

**Υπηρεσίες Διαχείρισης
MPEG Πληροφορίας
για Οπτικοακουστικές Εφαρμογές**

Γεωργία Κυριακάκη

Μεταπτυχιακή Διατριβή

Τμήμα Ηλεκτρονικών Μηχανικών και Μηχανικών Η/Υ

Πολυτεχνείο Κρήτης

Χανιά

Περίληψη

Τα σύγχρονα πληροφοριακά συστήματα διαχειρίζονται οπτικοακουστικό περιεχόμενο όπως βίντεο και ήχο, σε όλο και μεγαλύτερο βαθμό από ότι στο παρελθόν. Σημαντικά προβλήματα στη διαχείριση της οπτικοακουστικής πληροφορίας είναι η πρόσβαση στο υλικό βάσει του περιεχομένου του (content addressability), η ελεγχόμενη περιήγηση (browsing control) σε αυτό, η μεταβίβαση του υλικού μέσω διαβαθμισμένων υπηρεσιών και ανάλογα με το προφίλ των χρηστών και των συσκευών τους (quality of service support) και η ανταλλαγή περιεχομένου (content interchange) σε ετερογενή περιβάλλοντα για χρήση του με οποιοδήποτε τρόπο. Τα παραπάνω προβλήματα διαχείρισης είναι δύσκολο να αντιμετωπιστούν στην γενική τους μορφή και οι λύσεις που έχουν προταθεί μέχρι τώρα εστιάζουν μόνο σε συγκεκριμένα πεδία εφαρμογών. Συνεπώς οι διαφορετικές προτεινόμενες λύσεις για τη διαχείριση οπτικοακουστικής πληροφορίας δεν έχουν μια κοινά αποδεκτή βάση αντιμετώπισης, συνήθως ούτε για την ίδια εφαρμογή, με αποτέλεσμα να είναι δύσκολο να υποστηριχθούν στον πυρήνα των πληροφοριακών συστημάτων.

Έχοντας ως στόχο την αντιμετώπιση της οπτικοακουστικής πληροφορίας με ενιαίο και ευρέως αποδεκτό τρόπο, αναπτύχθηκαν πρόσφατα τα πρότυπα MPEG-2 και MPEG-4, για την κωδικοποίηση και μετάδοση των οπτικοακουστικών μέσων παρουσίασης. Η έμφαση και των δύο αυτών προτύπων είναι στην πολύ καλή συμπίεση της πληροφορίας, στη μετάδοσή της σε ροές πραγματικού χρόνου και στην διευκόλυνση της διαχείρισης του περιεχομένου των ροών. Ιδιαίτερα το πρότυπο MPEG-4, που αποτελεί το πλέον σύγχρονο πρότυπο στο συγκεκριμένο τομέα, προχωράει σε μεγάλο βάθος στην κωδικοποίηση του περιεχομένου των ροών παρουσίασης και στην σύνθεση πολλαπλών ροών σε μία παρουσίαση, καλύπτοντας πολύ μεγάλο εύρος δυνατοτήτων των τηλεπικοινωνιακών γραμμών και των συσκευών αποκωδικοποίησης.

Η παρούσα διατριβή εκμεταλλεύεται τα πρότυπα MPEG-2 και MPEG-4 με στόχο να δημιουργήσει τις βασικές υποδομές διαχείρισης οπτικοακουστικών μέσων σε ένα πληροφοριακό σύστημα, ούτως ώστε ανεξάρτητα από την εφαρμογή, να υπάρχει στον πυρήνα του πληροφοριακού συστήματος η βασική υποστήριξη για πρόσβαση στο περιεχόμενο, για ελεγχόμενη περιήγηση σε αυτό, για διαβαθμισμένη μεταβίβασή του σε συσκευές και χρήστες με διαφορετικά προφίλ και για ανταλλαγή του ανάμεσα σε διαφορετικά συστήματα διαχείρισης. Για να πετύχει το συγκεκριμένο στόχο, η διατριβή μοντελοποιεί το κωδικοποιημένο περιεχόμενο των οπτικοακουστικών μέσων όπως αυτό καθορίζεται από τα πρότυπα MPEG-2 και MPEG-4. Η μοντελοποίηση αυτή αποτελείται από

ένα επίπεδο αναπαράστασης που παρέχει τη δυνατότητα περιγραφής του περιεχομένου με ευρέως αποδεκτό, απλό και ευέλικτο τρόπο, κατάλληλο για διαχείριση και ανταλλαγή από τα πληροφοριακά συστήματα υποστήριξης οπτικοακουστικών εφαρμογών. Η αναπαράσταση του περιεχομένου γίνεται με την ανάπτυξη γλώσσας αναπαράστασης οπτικοακουστικής πληροφορίας, η οποία στηρίζεται στη γλώσσα XML του Παγκόσμιου Ιστού.

Η πρόσβαση στο οπτικοακουστικό περιεχόμενο γίνεται από σχεσιακές βάσεις δεδομένων. Οι βάσεις δεδομένων υποστηρίζουν την αποθήκευση της πληροφορίας περιεχομένου και την ανάκτησή της για την υλοποίηση της αυξημένης λειτουργικότητας (ανάκτηση βάσει περιεχομένου, ελεγχόμενη πλοήγηση, κλπ) των οπτικοακουστικών εφαρμογών. Το λογικό μοντέλο αποθήκευσης σε σχεσιακές βάσεις δεδομένων στηρίζεται στο μοντέλο αναπαράστασης που παρέχει η γλώσσα περιγραφής του περιεχομένου και επιπλέον υποστηρίζει ένα σύνολο από εργαλεία για την ευέλικτη διαχείριση του περιεχομένου είτε στη μορφή αποθήκευσης, είτε στη μορφή αναπαράστασης.

Τέλος, το βασικό σύνολο της λειτουργικότητας των οπτικοακουστικών εφαρμογών υποστηρίζεται στον πυρήνα των πληροφοριακών συστημάτων από μια ολοκληρωμένη αρχιτεκτονική εργαλείων διαχείρισης του MPEG περιεχομένου. Τα εργαλεία αυτά εκτός από την γλώσσα περιγραφής και τη βάση δεδομένων, περιλαμβάνουν και ένα σύνολο από συνδιαλλαγές ανάκτησης περιεχομένου. Η χρησιμότητα της ολοκληρωμένης υποδομής διαχείρισης οπτικοακουστικής πληροφορίας παρουσιάζεται με ένα παράδειγμα εκπομπής οπτικοακουστικού περιεχομένου με απαιτήσεις αυξημένης λειτουργικότητας.

Αφιέρωση

*Αφιερώνεται σε όλους τους συναδέλφους μου που βρέθηκαν στην ίδια θέση
και κατόρθωσαν με επιτυχία να αντεπεξέλθουν στις δυσκολίες των καιρών...*

Ευχαριστίες

Ευχαριστώ τον καθηγητή κ. Σταύρο Χριστοδουλάκη για την επίβλεψη και την πολύτιμη βοήθεια που μου παρείχε και είχε ως αποτέλεσμα την επιτυχή ολοκλήρωση της παρούσας διατριβής. Επίσης τον ευχαριστώ για την άψογη συνεργασία μας όλα τα χρόνια της παρουσίας μου στο Εργαστήριο Διανεμημένων Πληροφοριακών Συστημάτων και Εφαρμογών του Πολυτεχνείου Κρήτης.

Ευχαριστώ προκαταβολικά τους καθηγητές του τμήματος Ηλεκτρονικών και Μηχανικών Η/Υ κ. Π. Τριανταφύλλου και Μ. Κουμπάρκη για το χρόνο που διέθεσαν στην ανάγνωση της παρούσας διατριβής και για τη δική τους συμβολή στην εκπόνησή της.

Ιδιαίτερες ευχαριστίες αξίζουν σε όλους τους συναδέλφους μου από το Εργαστήριο Διανεμημένων Πληροφοριακών Συστημάτων και Εφαρμογών, οι οποίοι μου συμπαραστάθηκαν με οποιοδήποτε τρόπο καθ' όλη τη διάρκεια εκπόνησης της διατριβής.

Περιεχόμενα

Κεφάλαιο 1

Εισαγωγή	1
1.1 Εισαγωγή	1
1.2 Αναγκαιότητα	4
1.3 Επιμέρους στόχοι	8
1.4 Συνεισφορά της διατριβής	11
1.5 Δομή της διατριβής	12

Κεφάλαιο 2

Σχετική έρευνα	14
2.1 Εισαγωγή	14
2.2 Το πρότυπο MPEG-4	15
2.2.1 Οπτικά Αντικείμενα	17
2.2.1.1 Βίντεο	19
2.2.1.2 Πρόσωπα	21
2.2.1.3 Εικόνες και υφή	22
2.2.1.4 Πλέγματα	22
2.2.1.5 Προφίλ και επίπεδα οπτικών αντικειμένων	24
2.2.2 Ήχος	25
2.2.3 Περιγραφή Σκηνών	28
2.2.4 Περιγραφείς Αντικειμένων	32
2.2.5 Γλώσσα Συντακτικής Περιγραφής	35
2.2.6 Εφαρμογές του MPEG-4	36
2.3 Η γλώσσα XML	39
2.4 Διαχείριση οπτικοακουστικού περιεχομένου στο πρότυπο DAVIC	41
2.4.1 Μοντέλο οπτικοακουστικής πληροφορίας	41
2.4.2 Υπηρεσία περιεχομένου του DAVIC	42
2.5 Το σύστημα ΚΥΔΩΝΙΑ	45
2.5.1 Αρχιτεκτονική του συστήματος ΚΥΔΩΝΙΑ	46
2.5.2 Το πρότυπο DAVIC και το πρόγραμμα SICMA	46
2.5.2.1 Ο ΚΥΔΩΝΙΑ-D εξυπηρετητής πολυμέσων	49

Κεφάλαιο 3

Περιγραφή του MPEG-4 περιεχομένου	53
3.1 Εισαγωγή	53
3.2 Χρήση της AVDL	57
3.3 Χαρακτηριστικά της περιγραφής	61
3.4 Τύπος εγγράφων και προφίλ περιγραφής	63
3.5 Σημασιολογικές δομές της AVDL	66
3.6 Σύνθεση στοιχείων περιγραφής	68
3.7 Περιγραφή οπτικών αντικειμένων	71
3.7.1 Αντικείμενα Βίντεο	75
3.7.1.1 Βαθμίδα Αντικειμένου Βίντεο	76
3.7.1.2 Επίπεδο με μικρή επικεφαλίδα	82
3.7.1.3 Επίπεδο αντικειμένου βίντεο	85
3.7.1.4 Αντικείμενο τύπου sprite	92
3.7.1.5 Επίπεδο αντικειμένου κλιμάκωσης	95
3.7.1.6 Macroblock	97
3.7.2 Αντικείμενα Ακίνητης Υφής	100
3.7.3 Πλέγματα	101

3.7.4	Πρόσωπα	104
3.8	Περιγραφείς Αντικειμένων	105
3.9	Περιγραφή Σκηνής	111
3.9.1	Δομή της σκηνής	112
3.9.2	Κόμβοι σκηνής	114
3.9.3	Κατηγορίες των κόμβων της σκηνής	117
3.9.3.1	Κόμβος-κορυφή της σκηνής	119
3.9.3.2	Κοινοί 2Δ και 3Δ κόμβοι	121
3.9.3.3	Κόμβοι δύο διαστάσεων	126
3.9.3.4	Κόμβοι τριών διαστάσεων	128
3.9.3.5	Μικτοί κόμβοι	129
3.9.4	ROUTE	130
3.10	Ανακεφαλαίωση	130

Κεφάλαιο 4

Διαχείριση MPEG περιεχομένου σε εξυπηρετητή πολυμέσων 132

4.1	Εισαγωγή	132
4.2	Διαχείριση πολυμέσων στο σύστημα ΚΥΔΩΝΙΑ	134
4.3	Δεικτοδότηση MPEG περιεχομένου στο σύστημα ΚΥΔΩΝΙΑ	137
4.3.1	Δομή δεικτοδότησης MPEG βίντεο	138
4.3.2	Γρήγορη επισκόπηση βίντεο	140
4.3.3	Δεικτοδότηση συγχρονισμένου βίντεο και ήχου	142
4.3.4	Εναλλαγή λειτουργιών περιήγησης	143
4.3.5	Αρχιτεκτονική του συστήματος υποστήριξης λειτουργιών περιήγησης	144
4.4	Υποστήριξη της Υπηρεσίας Περιεχομένου στο σύστημα ΚΥΔΩΝΙΑ-D	146
4.5	Αναπαράσταση MPEG-2 βίντεο με την γλώσσα AVDL	151
4.6	Επέκταση των δομών δεικτοδότησης για τα δεδομένα του MPEG-4	158
4.7	Ανακεφαλαίωση	161

Κεφάλαιο 5

Βάση δεδομένων οπτικοακουστικού περιεχομένου 164

5.1	Εισαγωγή	164
5.2	Το μοντέλο οντοτήτων-συσχετίσεων	167
5.2.1	Επεκτάσεις του μοντέλου οντοτήτων-συσχετίσεων	170
5.3	Αποθήκευση περιγραφών σε βάση δεδομένων	172
5.3.1	Αντιστοίχιση της περιγραφής στο μοντέλο οντοτήτων-συσχετίσεων	173
5.3.2	Μοντέλο αποθήκευσης περιγραφών στη βάση δεδομένων	175
5.4	Ανάκτηση περιγραφών από τη βάση δεδομένων	180
5.5	Μοντέλο οντοτήτων-συσχετίσεων της AVDL	181
5.5.1	Μοντέλο οπτικών αντικειμένων του MPEG-4	182
5.5.2	Μοντέλο βάσης δεδομένων σκηνών του MPEG-4	186
5.5.3	Μοντέλο βάσης δεδομένων MPEG-2 βίντεο	188
5.6	Σχεσιακό μοντέλο δεδομένων	191
5.7	Παράδειγμα	194
5.8	Ανακεφαλαίωση	198

Κεφάλαιο 6

Υποστήριξη εφαρμογών οπτικοακουστικού περιεχομένου 200

6.1	Εισαγωγή	200
6.2	Υπηρεσίες περιήγησης και ελεγχόμενης παρουσίασης	204
6.3	Υπηρεσίες διαβαθμισμένης μεταβίβασης περιεχομένου	205

6.3.1	Προφίλ τερματικού	207
6.3.2	Προφίλ χρηστών	210
6.4	Υπηρεσίες ανάκτησης βάσει περιεχομένου	212
6.5	Υπηρεσίες ανταλλαγής περιεχομένου	213
6.6	Παράδειγμα εφαρμογής	214
6.6.1	Εκπαίδευση από απόσταση	214
6.6.1.1	Διαβαθμισμένη μεταβίβαση	223
6.6.1.2	Περιήγηση και ελεγχόμενη παρουσίαση	223
6.7	Ανακεφαλαίωση	224
Κεφάλαιο 7		
	Ανακεφαλαίωση	225

Πίνακας Σχημάτων

Σχήμα 2-1. Εργαλεία του MPEG-4	18
Σχήμα 2-2. Περιγραφείς Αντικειμένων	33
Σχήμα 2-3. Κλάσεις πληροφορίας στο πρότυπο DAVIC	42
Σχήμα 2-4. Αρχιτεκτονική του πληροφοριακού συστήματος ΚΥΔΩΝΙΑ	47
Σχήμα 2-5. Αρχιτεκτονική του συστήματος DAVIC	48
Σχήμα 2-6. Αρχιτεκτονική του εξυπηρετητή ΚΥΔΩΝΙΑ-D	50
Σχήμα 3-2. Δημιουργία της περιγραφής στη γλώσσα AVDL	57
Σχήμα 3-3. Χρήση της AVDL για την υποστήριξη λειτουργιών διαχείρισης οπτικοακουστικού περιεχομένου	58
Σχήμα 3-4. Χρήση της περιγραφής από τον προγραμματιστή εφαρμογής	59
Σχήμα 3-5. Διαχείριση της περιγραφής στην εφαρμογή	59
Σχήμα 3-6. Χρήση των οπτικοακουστικών δεδομένων και του περιεχομένου τους	60
Σχήμα 3-7. Δήλωση τύπου εγγράφου και στοιχείου - ρίζας	65
Σχήμα 3-8. Τύποι δεδομένων, περιορισμοί και συσχετίσεις στοιχείων	68
Σχήμα 3-9. Σύνθεση αυτόνομων περιγραφών από διαφορετικά έγγραφα	70
Σχήμα 3-10. Οπτικά Αντικείμενα	72
Σχήμα 3-11. Προφίλ και επίπεδα οπτικών Αντικειμένων	73
Σχήμα 3-12. Χαρακτηριστικά του σήματος βίντεο οπτικών αντικειμένων	74
Σχήμα 3-13. Αντικείμενα Βίντεο	75
Σχήμα 3-14. Βαθμίδα Αντικειμένου Βίντεο	76
Σχήμα 3-15. Βαθμίδα Βάσης	78
Σχήμα 3-16. Βαθμίδα Βάσης Βίντεο	78
Σχήμα 3-17. Βαθμίδα κλιμάκωσης	79
Σχήμα 3-18. Πολυπλοκότητα επιπέδων βίντεο	81
Σχήμα 3-19. Επίπεδο Βίντεο με μικρή επικεφαλίδα	83
Σχήμα 3-20. Ομάδα Επιπέδων Αντικειμένου Βίντεο	86
Σχήμα 3-21. Επίπεδο Αντικειμένου Βίντεο	86
Σχήμα 3-22. Επίπεδο Αντικειμένου Βίντεο τύπου I	88
Σχήμα 3-23. Επίπεδο Αντικειμένου Βίντεο τύπου P	90
Σχήμα 3-24. Επίπεδο Αντικειμένου Βίντεο τύπου B	91
Σχήμα 3-25. Αντικείμενο τύπου sprite	93
Σχήμα 3-26. Επίπεδο κλιμάκωσης	95
Σχήμα 3-27. Περιγραφή αντικειμένων ακίνητης υφής	100
Σχήμα 3-28. Περιγραφή πλεγμάτων	102
Σχήμα 3-28. Περιγραφή προσώπων	104
Σχήμα 3-29. Μηνύματα Περιγραφών Αντικειμένων	106
Σχήμα 3-30. Περιγραφέας Αντικειμένων και Αρχικός Περιγραφέας	107
Σχήμα 3-31. Περιγραφέας Στοιχειώδους Ροής	108
Σχήμα 3-32. Περιγραφείς παραμέτρων αποκωδικοποιητή	109
Σχήμα 3-33. Περιγραφή της πνευματικής ιδιοκτησίας του περιεχομένου	110
Σχήμα 3-34. Ποιότητα υπηρεσίας και περιγραφή άλλων μορφών περιεχομένου	111
Σχήμα 3-35. Δομή της ρίζας της σκηνής	113
Σχήμα 3-36. Περιγραφή των κόμβων πεδίου	115
Σχήμα 3-37. Σύνθετα πεδία κόμβων	115
Σχήμα 3-38. Στοιχεία βασικών τύπων πεδίων	116
Σχήμα 3-39. Κατηγορίες κόμβων σκηνής	119
Σχήμα 3-40. 2Δ βαθμίδα σκηνής	120
Σχήμα 3-41. 3Δ βαθμίδα σκηνής	121
Σχήμα 3-42. Κοινοί 2Δ και 3Δ κόμβοι σκηνών	122
Σχήμα 3-43. Κοινοί 2Δ και 3Δ κόμβοι ήχου	123
Σχήμα 3-44. Δυσδιάστατοι κόμβοι σκηνών	127
Σχήμα 3-45. ROUTEs	130
Σχήμα 4-1. Μοντελοποίηση αντικειμένων πολυμέσων στο σύστημα ΚΥΔΩΝΙΑ	135
Σχήμα 4-2. Μοντελοποίηση αντικειμένων βίντεο στο σύστημα ΚΥΔΩΝΙΑ	136
Σχήμα 4-3. Δεικτοδότηση ροών MPEG βίντεο	139
Σχήμα 4-4. Δεικτοδότηση συγχρονισμένων ροών MPEG-1 Systems	142
Σχήμα 4-5. Δομή των δεικτών ροών MPEG-1 και ροών γρήγορης επισκόπησης	144
Σχήμα 4-6. Διάγραμμα πακέτων της αρχιτεκτονικής υποστήριξης περιήγησης	145

Σχήμα 4-7. Διάγραμμα σειράς για την υποστήριξη των αιτήσεων περιήγησης στο σύστημα ΚΥΔΩΝΙΑ	146
Σχήμα 4-8. Μοντελοποίηση αντικειμένων του DAVIC στο σύστημα ΚΥΔΩΝΙΑ	147
Σχήμα 4-9. Η Υπηρεσία Περιεχομένου στον εξυπηρετητή ΚΥΔΩΝΙΑ-D	149
Σχήμα 4-10. Εγκατάσταση MPEG περιεχομένου στον εξυπηρετητή	150
Σχήμα 4-11. Αναπαράσταση σειρών MPEG-2 βίντεο με τη γλώσσα AVDL	152
Σχήμα 4-12. Προφίλ και επίπεδα σειρών MPEG-2 βίντεο στη γλώσσα AVDL	154
Σχήμα 4-13. Όροι δεικτοδότησης ρών MPEG-4 βίντεο	159
Σχήμα 4-14. Όροι δεικτοδότησης ρών MPEG-4 πλεγμάτων	159
Σχήμα 4-15. Όροι δεικτοδότησης ρών MPEG-4 προσώπων	160
Σχήμα 4-16. Όροι δεικτοδότησης σκηνών του MPEG-4	160
Σχήμα 4-17. Δομές δεικτοδότησης σκηνών και αντικειμένων του MPEG-4	161
Σχήμα 5-1. Στοιχεία του διαγράμματος οντοτήτων-συσχετίσεων	171
Σχήμα 5-2. Μοντέλο αποθήκευσης τύπων στοιχείων στη βάση δεδομένων	176
Σχήμα 5-3. Μοντέλο αποθήκευσης στοιχείων στη βάση δεδομένων	177
Σχήμα 5-4. Ιεραρχία οπτικοακουστικών αντικειμένων MPEG-4 και MPEG-2 στον μοντέλο βάσης δεδομένων	182
Σχήμα 5-5. Σειρά οπτικών αντικειμένων	183
Σχήμα 5-6. Οπτικά αντικείμενα	183
Σχήμα 5-7. Αντικείμενα βίντεο και βαθμίδες βίντεο	184
Σχήμα 5-8. Μοντέλο οντοτήτων-συσχετίσεων των βαθμίδων βίντεο	185
Σχήμα 5-9. Σχήμα των βαθμίδων βίντεο	186
Σχήμα 5-10. Επίπεδα αντικειμένων βίντεο	186
Σχήμα 5-11. Μοντέλο αναπαράστασης σκηνών	187
Σχήμα 5-12. Κόμβος σκηνής	187
Σχήμα 5-13. Γεγονότα σκηνών	188
Σχήμα 5-14. Μοντέλο οντοτήτων-συσχετίσεων σειράς MPEG-2 βίντεο	189
Σχήμα 5-15. Ιεραρχία εξειδίκευσης των δεδομένων της σειράς βίντεο	189
Σχήμα 5-16. Μοντέλο βαθμίδων κλιμάκωσης MPEG-2	190
Σχήμα 5-17. Μοντέλο εικόνων MPEG-2 βίντεο	191
Σχήμα 5-18. Σχεσιακό μοντέλο των τύπων στοιχείων της AVDL	193
Σχήμα 5-19. Σχεσιακό μοντέλο των στοιχείων και των εγγράφων της AVDL	193
Σχήμα 5-20. Σχεσιακό σχήμα οπτικών αντικειμένων του MPEG-4	194
Σχήμα 6-1. Αρχιτεκτονική υποστήριξης εφαρμογών οπτικοακουστικού περιεχομένου	201
Σχήμα 6-2. Υποστήριξη εφαρμογών διαχείρισης στατικού περιεχομένου	203
Σχήμα 6-3. Εφαρμογές διαχείρισης δυναμικού περιεχομένου	203
Σχήμα 6-4. Διάγραμμα υποστήριξης υπηρεσιών διαβαθμισμένης μεταβίβασης περιεχομένου	205
Σχήμα 6-5. Διάγραμμα υποστήριξης υπηρεσιών διαβαθμισμένης μεταβίβασης περιεχομένου	206
Σχήμα 6-6. Διάγραμμα υποστήριξης υπηρεσιών ανάκτησης βάσει περιεχομένου	213
Σχήμα 6-7. Σκηνή εφαρμογής εκπαίδευσης από απόσταση	215
Σχήμα 6-8. Λογική δομή της οπτικοακουστικής σκηνής της εφαρμογής εκπαίδευσης από απόσταση	216
Σχήμα 6-9. Περιγραφή του οπτικού αντικειμένου για την καθηγήτρια	218
Σχήμα 6-10. Περιγραφή του αντικειμένου της οπτικοακουστικής παρουσίασης	219
Σχήμα 6-11. Περιγραφή σκηνής της εφαρμογής εκπαίδευσης από απόσταση	222

Κεφάλαιο 1

Εισαγωγή

1.1 Εισαγωγή

Τα πολυμέσα κατέχουν πρωτεύοντα ρόλο στις ψηφιακές εφαρμογές λόγω των ελκυστικών χαρακτηριστικών τους. Τα σημαντικότερα από αυτά τα χαρακτηριστικά είναι ο πλούτος της πληροφορίας που περιέχουν, η ζωντάνια της παρουσίασης που παρέχουν, η δυνατότητα αλληλεπίδρασης που δίνουν στους χρήστες και η ποικιλία των τρόπων που μπορούν να χρησιμοποιηθούν. Ορισμένα από τα ποικίλα πεδία εφαρμογής τους είναι η ψυχαγωγία (παιχνίδια, μουσική, τηλεόραση, ταινίες, βίντεο κλπ), τα αλληλεπιδραστικά περιβάλλοντα (τεχνητή πραγματικότητα, εκπαίδευση, εγκυκλοπαίδειες, τηλεαγορές κλπ), οι συλλογές ή βιβλιοθήκες αρχειακού υλικού (ταινιοθήκες, βιντεοθήκες, φωτογραφικά αρχεία, κλπ) και η παραγωγή οπτικοακουστικού υλικού (στούντιο παραγωγής βίντεο, ταινιών και διαφήμισης, κλπ). Σε σημαντικό βαθμό έχει συντελέσει στην αύξηση της χρήσης των πολυμέσων η ευρεία εξάπλωση του Ίντερνετ, το οποίο έχει κατακτήσει χρήστες και επιχειρήσεις όλων των ειδών.

Η επιστημονική και τεχνολογική εξέλιξη που έχει συντελεστεί στον τομέα των ψηφιακών πολυμέσων τις τελευταίες δεκαετίες, έχει καλύψει θέματα όπως:

- η αποδοτική κωδικοποίηση και συμπίεση της ψηφιοποιημένης πληροφορίας με ευρέως αποδεκτά πρότυπα όπως τα πρότυπα κωδικοποίησης MPEG-1 και MPEG-2,
- η ανάπτυξη του προτύπου MPEG-4 για την κάλυψη ευρύτερου φάσματος δυνατοτήτων των δικτύων μετάδοσης και για την κωδικοποίηση της οπτικοακουστικής πληροφορίας με βάση το περιεχόμενό της,
- η αποδοτική αποθήκευση της πληροφορίας με τη χρήση αποθηκευτικών μέσων μεγάλης χωρητικότητας και ειδικών τεχνικών εκμετάλλευσης του αποθηκευτικού χώρου,
- η γρήγορη πρόσβαση στην αποθηκευμένη πληροφορία πολυμέσων με χρήση εξυπηρετητών αποθήκευσης και ανάκτησης,
- η μετάδοση της πληροφορίας σε πραγματικό χρόνο μέσα από δίκτυα υψηλών ταχυτήτων,
- η μετάδοση της πληροφορίας σε πολλούς απομακρυσμένους χρήστες με συστήματα ψηφιακής εκπομπής και εξυπηρετητές μετάδοσης κατ' απαίτηση,
- η ανάπτυξη μεθόδων διαχείρισης του περιεχομένου των πολυμέσων, με μεθόδους και πρότυπα αναπαράστασης πληροφορίας και
- η ανάκτηση των πολυμέσων με βάση το περιεχόμενό τους, με βελτίωση των τεχνολογιών που χρησιμοποιούν οι βάσεις δεδομένων και τα συστήματα ανάκτησης πληροφορίας.

Συγκεκριμένα, οι εξυπηρετητές πολυμέσων έδωσαν ώθηση σε εφαρμογές βίντεο κατ' απαίτηση (Video on Demand) για τη μετάδοση δεδομένων πάνω από δίκτυα υψηλών ταχυτήτων σε ετερογενείς χρήστες. Βασικό χαρακτηριστικό των εξυπηρετητών πολυμέσων είναι ότι οι περισσότεροι πλέον χρησιμοποιούν τα πρότυπα MPEG-1 ή MPEG-2, τα οποία παρέχουν πολύ αποδοτική συμπίεση βίντεο και ήχου, συγχρονισμένη μετάδοση και δυνατότητα υλοποίησης λειτουργιών VCR (pause, continue, rewind, fast-forward, fast-backward).

Από την άλλη, το Ίντερνετ έδωσε μεγάλη ώθηση στη διαχείριση εγγράφων με κύριο γνώμονα τη γλώσσα αναπαράστασης HTML. Η μεγάλη επιτυχία της οδήγησε πρόσφατα στην εξάπλωση μιας άλλης γενικότερης και ισχυρότερης γλώσσας εγγράφων, της XML, η οποία ουσιαστικά δίνει τη δυνατότητα αναπαράστασης ημιδομημένης πληροφορίας οποιουδήποτε είδους με κείμενο.

Επίσης, για το κείμενο έχουν αναπτυχθεί εδώ και καιρό σημαντικές μέθοδοι γρήγορης αναζήτησης και προσπέλασης με βάση το περιεχόμενό του, δηλ. τις λέξεις ή φράσεις που

περιέχει. Παρόμοιες μέθοδοι βάσει περιεχομένου έχουν αναπτυχθεί τα τελευταία χρόνια και για τα υπόλοιπα μέσα, κυρίως για τον ήχο και την εικόνα. Για το βίντεο, που είναι το μεγαλύτερο σε όγκο, το πλουσιότερο σε πληροφορία και το πιο πολύπλοκο από τα πολυμέσα, έχουν αναπτυχθεί εφαρμογές διαχείρισής του μόλις τα τελευταία χρόνια. Έχουν γίνει αρκετές θεωρητικές μελέτες των χαρακτηριστικών του, οι οποίες στοχεύουν στο να δημιουργήσουν μεθόδους εξαγωγής του περιεχομένου του από το ψηφιοποιημένο υλικό. Η πληροφορία αυτή έχει στόχο να χρησιμοποιηθεί ως ευρετήριο, περίληψη ή συνοπτική περιγραφή για να διευκολύνει τη γρήγορη αναζήτηση συγκεκριμένου περιεχομένου μέσα στο βίντεο. Στη διεθνή βιβλιογραφία έχουν εμφανιστεί τα τελευταία χρόνια προσπάθειες προσέγγισης του παραπάνω προβλήματος με τις εξής μεθόδους:

- Με αυτόματη ή ημιαυτόματη εξαγωγή περιεχομένου από το κείμενο, τις εικόνες, τον ήχο και το βίντεο. Η διαδικασία αυτή ονομάζεται ανάλυση (analysis) ή δεικτοδότηση (indexing) και χρησιμοποιεί αλγόριθμους εξαγωγής χαρακτηριστικών από το υλικό.
- Με ταυτοποίηση και αναπαράσταση μεταδεδομένων που σχετίζονται με το υλικό και περιγράφουν το περιεχόμενό του. Στο συγκεκριμένο τομέα έχουν αναπτυχθεί πρότυπα τα οποία διευκολύνουν την ανταλλαγή των μεταδεδομένων σε ετερογενή περιβάλλοντα.
- Με την ανάπτυξη του προτύπου MPEG-4 το οποίο παρέχει κωδικοποίηση οπτικοακουστικού υλικού βάσει του περιεχομένου του. Το πρότυπο MPEG-4 διακρίνει τα διαφορετικά αντικείμενα ή τμήματα του υλικού και τα κωδικοποιεί χωριστά, δίνοντας έτσι τη δυνατότητα απευθείας διαχείρισής τους.
- Με ανάπτυξη μεθόδων οργάνωσης του περιεχομένου και μεθόδων προσπέλασής του για την υποστήριξη αναζήτησης σε αυτό.

Παρόλα αυτά, τα πληροφοριακά συστήματα που υποστηρίζουν εφαρμογές πολυμέσων και ειδικότερα οπτικοακουστικών μέσων, εστιάζουν σε ένα περιορισμένο σύνολο εφαρμογών, στις οποίες παρέχουν διαχείριση της πληροφορίας και του περιεχομένου της και μεταβίβασή της στους τελικούς αποδέκτες. Η πληθώρα της πληροφορίας που περιέχεται στα οπτικοακουστικά μέσα και η ποικιλομορφία των εφαρμογών στις οποίες αυτά χρησιμοποιούνται, έχουν καταστήσει αδύνατη την αντιμετώπιση του προβλήματος της διαχείρισης της οπτικοακουστικής πληροφορίας στη γενική του μορφή, δηλαδή με ενιαίο τρόπο, ανεξάρτητα από την εφαρμογή και το περιβάλλον χρήσης. Η παρούσα διατριβή αντιμετωπίζει το πρόβλημα της ενιαίας διαχείρισης της πληροφορίας των οπτικοακουστικών μέσων στον πυρήνα των πληροφοριακών συστημάτων, με την ανάπτυξη των κατάλληλων υποδομών για την αναπαράσταση, ανάκτηση, ελεγχόμενη παρουσίαση, μεταβίβαση και ανταλλαγή αυτής της πληροφορίας.

Στην επόμενη ενότητα αναλύεται η αναγκαιότητα που αποτέλεσε το κίνητρο για την πραγματοποίηση της παρούσας διατριβής, ενώ στην ενότητα 1.3 παρουσιάζονται οι επιμέρους στόχοι που προέκυψαν. Στην ενότητα 1.4 γίνεται η παρουσίαση της συνεισφοράς της διατριβής.

1.2 Αναγκαιότητα

Η αναγκαιότητα της παρούσας διατριβής προέκυψε αρχικά από την ανάγκη για υποστήριξη εφαρμογών οπτικοακουστικών μέσων σε πραγματικά περιβάλλοντα και συγκεκριμένα στα πλαίσια του ευρωπαϊκού προγράμματος SICMA (Scaleable Interactive Continuous Media server and Applications). Ο στόχος του προγράμματος αυτού ήταν η υποστήριξη δύο εφαρμογών βίντεο κατ' απαίτηση και συγκεκριμένα για την παρουσίαση βίντεο σε μουσεία και για την εκπαίδευση από απόσταση με πολυμέσα σε πανεπιστήμια. Οι ανάγκες των παραπάνω εφαρμογών αφορούσαν:

- Στην μετάδοση και παρουσίαση πληροφορίας MPEG-2 βίντεο που αποτελούσε το υλικό των εφαρμογών. Το πρότυπο MPEG-2 επιλέχθηκε λόγω της καλής συμπίεσης, της διαβαθμισμένης κωδικοποίησης και της δυνατότητας τυχαίας προσπέλασης που παρέχει.
- Στην μεταβίβαση πληροφορίας σε ετερογενείς συσκευές και χρήστες με διαφορετικά προφίλ.
- Στον έλεγχο της παρουσίασης του βίντεο από τους χρήστες (επισκέπτες του μουσείου και φοιτητές) και στην ελεύθερη περιήγηση στο περιεχόμενό του.
- Στην ανταλλαγή περιεχομένου σε ετερογενή περιβάλλοντα χωρίς απαιτήσεις παρουσίασης σε πραγματικό χρόνο, για χρήση του με οποιονδήποτε τρόπο.
- Στην υποστήριξη του προτύπου DAVIC για την διαχείριση της οπτικοακουστικής πληροφορίας και των μεταδεδομένων που σχετίζονται με αυτήν σε εφαρμογές βίντεο κατ' απαίτηση.

Αργότερα προέκυψε η ανάγκη υποστήριξης νέων οπτικοακουστικών εφαρμογών που απαιτούσαν αυξημένη λειτουργικότητα όπως:

- Αλληλεπίδραση με το υλικό για έλεγχο της παρουσίασής του με νέους τρόπους, πέραν των κλασικών λειτουργιών πλοήγησης, όπως παύση και γρήγορη αναπαραγωγή προς τα εμπρός ή προς τα πίσω. Η αυξημένη λειτουργικότητα αλληλεπίδρασης με το οπτικοακουστικό υλικό στηρίζεται στο περιεχόμενό του, δηλαδή σε αυτό που ο άνθρωπος αντιλαμβάνεται όταν το βλέπει ή το ακούει.

- Αναζήτηση συγκεκριμένης πληροφορίας μέσα στο υλικό. Οι χρήστες έχουν ανάγκη από τρόπους να επιλέγουν εύκολα το υλικό που τους ενδιαφέρει ειδικά μέσα από τη μεγάλη ποσότητα και ποικιλία περιεχομένου που είναι χαρακτηριστικό των σύγχρονων εφαρμογών. Αυτό επιτυγχάνεται καθορίζοντας κριτήρια για το περιεχόμενο, δηλαδή με αναζήτηση υλικού με βάση προσωπικές προτιμήσεις και ενδιαφέροντα των χρηστών.
- Αυτόματο φιλτράρισμα της οπτικοακουστικής πληροφορίας όπου το σύστημα να λειτουργεί με έξυπνο τρόπο, εφαρμόζοντας το ίδιο τα κριτήρια του χρήστη για την επιλογή πληροφορίας κάνοντας έξυπνη ανάκτηση υλικού.
- Χρήση του υπάρχοντος υλικού για περαιτέρω παραγωγή. Η διαδικασία αυτή περιλαμβάνει επιλογή αλλά και επεξεργασία του υλικού και του περιεχομένου του για τη δημιουργία νέου.

Για την υποστήριξη ποικίλων οπτικοακουστικών εφαρμογών με διαφορετικά χαρακτηριστικά και απαιτήσεις στο ίδιο πληροφοριακό σύστημα, είναι αναγκαίο να αναπτυχθεί μια γενική υποδομή, η οποία να αντιμετωπίζει το πρόβλημα με ενιαίο τρόπο, καλύπτοντας τον βασικό πυρήνα της απαιτούμενης λειτουργικότητας και συγκεκριμένα:

- Τη χρήση ευρέως αποδεκτών προτύπων κωδικοποίησης και αναπαράστασης του οπτικοακουστικού υλικού,
- τον έλεγχο της παρουσίασης και περιήγησης στο υλικό (browsing control),
- την ανάκτηση της πληροφορίας βάσει του περιεχομένου της (content addressability),
- την υποστήριξη της μεταβίβασης του υλικού μέσω διαβαθμισμένων υπηρεσιών (quality of service support) και ανάλογα με το προφίλ των χρηστών και των συσκευών τους (user profiles) και
- την ανταλλαγή περιεχομένου με τη χρήση ευρέως αποδεκτών προτύπων ανταλλαγής εγγράφων του Παγκόσμιου Ιστού (content document interchange).

Ο παραπάνω πυρήνας λειτουργιών του πληροφοριακού συστήματος, στηρίζεται στη διαχείριση του οπτικοακουστικού υλικού και ιδιαίτερα του περιεχομένου του. Το περιεχόμενο των πολυμέσων ορίζεται ως η πληροφορία που περιέχουν και η οποία γίνεται αντιληπτή από τον άνθρωπο. Για παράδειγμα, το περιεχόμενο ενός εγγράφου είναι οι λέξεις, οι φράσεις, οι προτάσεις, οι παράγραφοι, οι επικεφαλίδες που περιέχει, αλλά και οι έννοιες που σχηματίζονται από τις λέξεις, προτάσεις, κλπ. Το περιεχόμενο μιας εικόνας είναι τα αντικείμενα που διακρίνουμε σε αυτήν, τα χαρακτηριστικά (χρώμα, σχήμα, θέση, κλπ) των αντικειμένων, αλλά και η κατάσταση ή το γεγονός που απεικονίζει. Το περιεχόμενο μιας ταινίας είναι οι σκηνές από τις οποίες αποτελείται, οι άνθρωποι και τα

αντικείμενα μέσα στις σκηνές, οι ενέργειές τους, η εξέλιξη των ενεργειών τους στο χρόνο αλλά και η πλοκή της ταινίας. Επίσης συχνά υπάρχει πληροφορία που δεν περιέχεται, αλλά συσχετίζεται άμεσα με το οπτικοακουστικό υλικό, όπως για παράδειγμα το όνομα του σκηνοθέτη και του παραγωγού, τα ονόματα των πρωταγωνιστών και η ημερομηνία παραγωγής μιας ταινίας. Αυτή η επιπλέον πληροφορία που σχετίζεται με τα πολυμέσα, ονομάζεται μεταδεδομένα. Στην παρούσα διατριβή θεωρούμε ότι η πληροφορία που βλέπουμε, ακούμε και γενικά αντιλαμβανόμαστε ότι υπάρχει στα ψηφιακά πολυμέσα, μαζί με τυχόν μεταδεδομένα που έχουν συσχετιστεί μαζί τους, αποτελούν συνολικά το *περιεχόμενο των πολυμέσων* και η διαχείρισή του αποτελεί αναγκαιότητα της παρούσας διατριβής. Η διαχείριση του οπτικοακουστικού περιεχομένου με ενιαίο τρόπο σε ένα πληροφοριακό σύστημα, στηρίζεται σε ένα σύνολο από μεθόδους που αφορούν:

- 1) Στην ταυτοποίηση του περιεχομένου των οπτικοακουστικών μέσων, δηλαδή στη διάκριση της πληροφορίας που θα αποτελέσει τον πυρήνα δεδομένων του πληροφοριακού συστήματος. Ο ενιαίος πυρήνας πληροφορίας εξασφαλίζει την ενιαία υποστήριξη των οπτικοακουστικών υπηρεσιών. Η πληροφορία αυτή είναι αναγκαίο να περιλαμβάνει το περιεχόμενο που αντιλαμβάνεται ο άνθρωπος και με το οποίο θα αλληλεπιδράσει ή το οποίο θα διαχειριστεί μέσα από την εφαρμογή, καθώς και το περιεχόμενο που θα επιτρέψει την υλοποίηση των υπηρεσιών διαβαθμισμένης και εξατομικευμένης (personalised) μεταβίβασης.
- 2) Στον καθορισμό αναπαράστασης για το περιεχόμενο. Η αναπαράσταση του περιεχομένου της οπτικοακουστικής πληροφορίας, πέραν της κωδικοποιημένης μορφής του, υποστηρίζει τη διαχείρισή του από το σύστημα για την υλοποίηση των επιθυμητών υπηρεσιών, καθώς και την ανταλλαγή ή μεταβίβασή του σε ετερογενή δικτυακά περιβάλλοντα.
- 3) Στην οργάνωση της αναπαράστασης του περιεχομένου από το πληροφοριακό σύστημα, ώστε να διευκολύνεται η ανάπτυξη των μεθόδων για την προσπέλασή του και επομένως για την υλοποίηση των επιθυμητών υπηρεσιών (έλεγχος παρουσίας, ανάκτηση βάσει περιεχομένου, διαβαθμισμένη και εξατομικευμένη μεταβίβαση).
- 4) Στην υποστήριξη μηχανισμών για την ανταλλαγή περιεχομένου σε περιβάλλοντα όπως αυτό του Παγκόσμιου Ιστού, σε όποια μορφή και αν βρίσκεται η οπτικοακουστική πληροφορία (κωδικοποιημένη, σε κάποια μορφή αναπαράστασης ή αποθηκευμένη σε βάση δεδομένων).
- 5) Στην περιγραφή της υποστήριξης των εφαρμογών στις οποίες στοχεύει το πληροφοριακό σύστημα ώστε να αποδεικνύεται η καταλληλότητα και η χρησιμότητά του.

Εφόσον η διατριβή εστιάζει στη διαχείριση του περιεχομένου των οπτικοακουστικών μέσων, με στόχο την υποστήριξη ενός συνόλου υπηρεσιών, είναι αναγκαίο να αντιμετωπίσει τα παραπάνω προβλήματα της ταυτοποίησης, αναπαράστασης, οργάνωσης και προσπέλασης του περιεχομένου. Η προσέγγιση που ακολουθείται στην παρούσα διατριβή για την ανάπτυξη της γενικής υποδομής διαχείρισης οπτικοακουστικής πληροφορίας είναι η εξής:

- 1) Η ταυτοποίηση και διάκριση του περιεχομένου των οπτικοακουστικών μέσων στηρίζεται στην κωδικοποίηση που παρέχουν τα πρότυπα MPEG-2 και MPEG-4.
- 2) Η αναπαράσταση του περιεχομένου στηρίζεται σε διεθνώς αποδεκτά πρότυπα που διευκολύνουν τη διαχείρισή του αλλά και την ανταλλαγή του σε ανοικτά και ετερογενή περιβάλλοντα.
- 3) Η οργάνωση και προσπέλαση του περιεχομένου στηρίζεται στην κατασκευή σχεσιακών βάσεων δεδομένων και στην υλοποίηση συνδιαλλαγών ανάκτησης δεδομένων.
- 4) Η υποστήριξη του πυρήνα των εφαρμογών, δηλαδή ο έλεγχος της παρουσίασης και η περιήγηση στο υλικό, η ανάκτηση βάσει περιεχομένου, η διαβαθμισμένη και εξατομικευμένη μεταβίβαση και η ανταλλαγή περιεχομένου, καθορίζονται πλήρως με γνώμονα τις παραπάνω υπηρεσίες διαχείρισης οπτικοακουστικής πληροφορίας.

Το πρότυπο MPEG-2 αποτελεί το πιο διαδεδομένο πρότυπο κωδικοποίησης και μετάδοσης βίντεο και ήχου, επομένως ήταν αναγκαίο να αποτελέσει μέρος της γενικής υποδομής υποστήριξης υπηρεσιών διαχείρισης οπτικοακουστικής πληροφορίας. Η κωδικοποίηση του MPEG-2 παρέχει πληροφορία για τα χαρακτηριστικά των οπτικοακουστικών ψηφιοσειρών και για τα ανεξάρτητα σημεία προσπέλασης μέσα σε αυτές. Τα ανεξάρτητα σημεία προσπέλασης δίνουν τη δυνατότητα περιήγησης στο περιεχόμενο των ψηφιοσειρών. Το πρότυπο MPEG-4 αποτελεί την τεχνολογία αιχμής στα πρότυπα κωδικοποίησης οπτικοακουστικών μέσων γιατί καλύπτει πολύ μεγάλο φάσμα δυνατοτήτων μετάδοσης. Επίσης προχωράει σε περισσότερο βάθος στην κωδικοποίηση του περιεχομένου απ' ό,τι το MPEG-2 και παρέχει πληροφορία για τα κωδικοποιημένα αντικείμενα, για τον τρόπο που συντίθενται σε σκηνές παρουσίασης και για τη δυνατότητα αλληλεπίδρασης με το περιεχόμενο.

Η αναπαράσταση της οπτικοακουστικής πληροφορίας στον πυρήνα του συστήματος στηρίζεται σε γλώσσα περιγραφής του περιεχομένου. Μια γλώσσα αναπαράστασης ή περιγραφής αποτελεί τον πλέον τυπικό τρόπο αναπαράστασης πληροφορίας και παρέχει ανεξαρτησία από τις εφαρμογές. Επίσης υπήρχε η ανάγκη κατασκευής γλώσσας που να είναι κατάλληλη και για τον Παγκόσμιο Ιστό, εφόσον αυτός αποτελεί το ευρύτερο και

πλέον διαδεδομένο περιβάλλον ανταλλαγής πληροφορίας. Η γλώσσα που αναπτύχθηκε στην παρούσα διατριβή αποτελεί εφαρμογή της μεταγλώσσας XML για την αναπαράσταση εγγράφων στον Παγκόσμιο Ιστό.

Η κατασκευή βάσεων δεδομένων είναι αναγκαία για την οργάνωση της πληροφορίας του περιεχομένου με τρόπο που να υποστηρίζεται η προσπέλαση και ανάκτησή της ανεξάρτητα από το περιβάλλον υποστήριξης των υπηρεσιών, τα δίκτυα μεταβίβασης, τα τερματικά, τις εφαρμογές και τους χρήστες.

Στην επόμενη ενότητα παρουσιάζονται οι επιμέρους στόχοι που προκύπτουν από την ανάγκη διαχείρισης του οπτικοακουστικού περιεχομένου σύμφωνα με την παραπάνω προσέγγιση.

1.3 Επιμέρους στόχοι

Οι επιμέρους στόχοι της διατριβής καθορίζονται με γνώμονα τις τρεις βασικές κατευθύνσεις που διακρίνονται στην προηγούμενη ενότητα, δηλαδή στην ταυτοποίηση του περιεχομένου, στην αναπαράστασή του στο πληροφοριακό σύστημα και στην οργάνωσή του για την υποστήριξη προσπέλασης και ανάκτησης. Όπως αναφέρθηκε στην προηγούμενη ενότητα, η ταυτοποίηση και διάκριση του περιεχομένου των οπτικοακουστικών μέσων στηρίζεται στην κωδικοποίηση βάσει περιεχομένου των προτύπων MPEG-2 και MPEG-4, εφόσον αποτελούν την τεχνολογία αιχμής στα πρότυπα κωδικοποίησης πολυμέσων. Όσον αφορά στην ταυτοποίηση και διάκριση του οπτικοακουστικού περιεχομένου καθορίστηκαν οι εξής επιμέρους στόχοι:

- Ταυτοποίηση εννοιολογικών τμημάτων του περιεχομένου, δηλαδή αντικειμένων με συγκεκριμένο και ξεχωριστό ρόλο μέσα στην πληροφορία. Τα αντικείμενα αυτά είναι σημαντικά για την εφαρμογή και το χρήστη, για τον οποίο αποτελούν το στόχο των ενεργειών του.
- Διάκριση αυτόνομων υποσυνόλων του σύνθετου και μεγάλου σε όγκο οπτικοακουστικού περιεχομένου. Είναι αναγκαίο να υπάρχει διάκριση των ανεξάρτητων τμημάτων του περιεχομένου, όπως προκύπτουν από τα κωδικοποιημένα δεδομένα και να μπορούν αυτά τα τμήματα να χρησιμοποιηθούν ως αυτόνομα υποσύνολα πληροφορίας.
- Ομαδοποίηση υποσυνόλων του περιεχομένου. Είναι αναγκαίο να μπορούν αυτόνομα τμήματα του περιεχομένου να χρησιμοποιηθούν ως ενιαίες οντότητες πληροφορίας για έναν κοινό σκοπό.

Όσον αφορά στη χρήση του προτύπου MPEG-4, που αποτελεί νέο πρότυπο, καθορίστηκαν οι εξής επιμέρους στόχοι:

- Μελέτη και κατανόηση της κωδικοποίησης του MPEG-4. Η κατανόηση αυτή έχει ως στόχο να συμπεριλάβει όλα τα τμήματα του προτύπου, ώστε να μπορεί να χρησιμοποιηθεί το περιεχόμενο ατομικών οπτικοακουστικών αντικειμένων, αλλά και σύνθετων παρουσιάσεων.
- Πλήρης εκμετάλλευση της πληροφορίας βάσει περιεχομένου που παρέχει το MPEG-4. Η πληροφορία που περιέχεται στην κωδικοποίηση του προτύπου αφορά στα χαρακτηριστικά των οπτικοακουστικών αντικειμένων και στον τρόπο παρουσιάσής τους. Για την υποστήριξη εφαρμογών με διαφορετικές απαιτήσεις, είναι αναγκαίο να είναι δυνατή η εκμετάλλευση οποιουδήποτε χαρακτηριστικού ή πληροφορίας που υπάρχει στην κωδικοποίηση.

Η αναπαράσταση του περιεχομένου των κωδικοποιημένων οπτικοακουστικών αντικειμένων που προκύπτει από την κατανόηση και χρήση των προτύπων κωδικοποίησης MPEG-2 και MPEG-4, πρέπει να πληρεί τις εξής επιμέρους στόχους:

- Απλότητα και ευκολία στη διαχείρισή της από το πληροφοριακό σύστημα, ώστε να μην αποτελεί ανασταλτικό παράγοντα στην υποστήριξη των οπτικοακουστικών υπηρεσιών.
- Ανταλλαγή περιεχομένου ανάμεσα σε ετερογενή συστήματα και κυρίως σε δικτυακά περιβάλλοντα όπως αυτό του Παγκοσμίου Ιστού. Ο στόχος επιτυγχάνεται με τη χρήση ευρέως αποδεκτών προτύπων αναπαράστασης πληροφορίας.
- Ανεξαρτησία από το υλικό, ώστε η αναπαράσταση του περιεχομένου να μπορεί να χρησιμοποιηθεί και αυτόνομη για την υποστήριξη των υπηρεσιών, εφόσον έχει απομονωθεί από το ίδιο το υλικό.
- Διατήρηση της σημασιολογίας του περιεχομένου των αρχικών δεδομένων, όπως αυτή περιέχεται στα χαρακτηριστικά, τη δομή αλλά και τη συσχέτισή τους. Ο στόχος αυτός εξασφαλίζει τη συνέπεια ανάμεσα στην κωδικοποιημένη πληροφορία και στην αναπαράστασή της στο πληροφοριακό σύστημα.
- Δυνατότητα εμπλουτισμού του περιεχομένου με επιπλέον πληροφορία ώστε να υπάρχει ευελιξία και προσαρμοστικότητα στις απαιτήσεις των κατασκευαστών, των εφαρμογών και τελικά των χρηστών.

Οι παραπάνω στόχοι αφορούν τελικά στην κατασκευή της γλώσσας αναπαράστασης του οπτικοακουστικού περιεχομένου. Η επιλογή της μεταγλώσσας XML για την ανάπτυξη της γλώσσας αναπαράστασης, εξασφαλίζει ως ένα βαθμό την τήρηση των παραπάνω στόχων,

εφόσον παρέχει απλότητα στη σύνταξη και ευκολία διαχείρισης, χρησιμοποιείται στον Παγκόσμιο Ιστό, είναι ανεξάρτητη από άλλα πρότυπα κωδικοποίησης και είναι επεκτάσιμη.

Η οργάνωση του περιεχομένου για αποδοτική διαχείρισή του από το πληροφοριακό σύστημα παροχής υπηρεσιών έχει ως επιμέρους στόχους:

- Να επιτρέπει και να διευκολύνει την αποθήκευση του περιεχομένου και τη διαχείριση αυτής. Ο στόχος αυτός είναι αναγκαίος για την αποδοτική διαχείριση του περιεχομένου και εξασφαλίζεται με την κατασκευή βάσεων δεδομένων.
- Να παρέχει ανεξαρτησία από τα ιδιαίτερα χαρακτηριστικά του συστήματος ώστε να μπορεί η οργάνωση του περιεχομένου να χρησιμοποιηθεί από οποιοδήποτε σύστημα. Ο στόχος αυτός εξασφαλίζεται με χρήση ευρέως αποδεκτών μοντέλων δεδομένων, όπως είναι το μοντέλο οντοτήτων-συσχετίσεων και το σχεσιακό μοντέλο για τον καθορισμό της οργάνωσης της πληροφορίας.
- Να διευκολύνει τη μετατροπή από την μορφή αναπαράστασης στη μορφή οργάνωσής του και αντίστροφα. Ο στόχος αυτός εξασφαλίζει την ανεξαρτησία των βάσεων δεδομένων από την κωδικοποιημένη πληροφορία και επιτρέπει τη χρήση του περιεχομένου με ευέλικτο τρόπο, δηλαδή είτε ως προς την αναπαράστασή του είτε ως προς την αποθηκευμένη μορφή του.
- Να υποστηρίζει την ανάκτηση της πληροφορίας του οπτικοακουστικού περιεχομένου. Η ανάκτηση του περιεχομένου αποτελεί αναγκαιότητα στην υποδομή υποστήριξης των υπηρεσιών περιήγησης, εξατομίκευσης και προσπέλασης βάσει περιεχομένου. Η ανάκτηση του περιεχομένου υποστηρίζεται από ερωτήσεις στη βάση δεδομένων.

Στόχος όλων των παραπάνω είναι η ανάπτυξη της μεθοδολογίας που υποστηρίζει τις βασικές λειτουργίες ενός ευρύ συνόλου οπτικοακουστικών εφαρμογών, οι οποίες στηρίζονται στις υπηρεσίες διαχείρισης της οπτικοακουστικής πληροφορίας. Επιμέρους στόχοι της διατριβής είναι επομένως:

- Ο καθορισμός της μεθοδολογίας που χρησιμοποιεί την αναπαράσταση και την οργάνωση του περιεχομένου για να υποστηρίξει τον έλεγχο της παρουσίασης και την περιήγηση στο υλικό.
- Ο καθορισμός των μεθόδων ανάκτησης της οπτικοακουστικής πληροφορίας βάσει του περιεχομένου της, με χρήση της αναπαράστασης και οργάνωσής του τελευταίου.
- Η περιγραφή της μεταβίβασης του υλικού μέσω διαβαθμισμένων υπηρεσιών και ανάλογα με το προφίλ των χρηστών και των συσκευών τους, όπως υποστηρίζονται από τις υπηρεσίες διαχείρισης της οπτικοακουστικής πληροφορίας.

- Ο καθορισμός των μεθόδων για την ανταλλαγή περιεχομένου ανεξάρτητα από τη μορφή αναπαράστασης ή οργάνωσής του.

Στην ενότητα που ακολουθεί περιγράφεται η συνεισφορά της διατριβής στο πεδίο των οπτικοακουστικών εφαρμογών και των υπηρεσιών και συστημάτων διαχείρισης οπτικοακουστικής πληροφορίας.

1.4 Συνεισφορά της διατριβής

Η συνεισφορά της διατριβής συνίσταται στην υποστήριξη υπηρεσιών διαχείρισης οπτικοακουστικής πληροφορίας, οι οποίες επιτρέπουν την υλοποίηση εφαρμογών όπως βίντεο κατ' απαίτηση, αποθήκευση και ανάκτηση βάσει περιεχομένου, ψηφιακές βιβλιοθήκες πολυμέσων κλπ. Η υποστήριξη αυτών των υπηρεσιών στα πλαίσια της παρούσας διατριβής, αποτελεί έναν γενικό πυρήνα λειτουργικότητας των συστημάτων διαχείρισης οπτικοακουστικής πληροφορίας, ο οποίος εξασφαλίζει ενιαίο τρόπο αντιμετώπισης των οπτικοακουστικών εφαρμογών. Αυτό αποτελεί συνεισφορά της διατριβής γιατί μέχρι τώρα τα πληροφοριακά συστήματα διαχείρισης οπτικοακουστικής πληροφορίας αντιμετώπιζαν το πρόβλημα με εξειδικευμένες λύσεις κατάλληλες μόνο για συγκεκριμένα υποσύνολα πληροφορίας και εφαρμογών. Η παρούσα διατριβή περιγράφει τη διαχείριση της οπτικοακουστικής πληροφορίας ανεξάρτητα από το είδος και τα ιδιαίτερα χαρακτηριστικά των εφαρμογών και μπορεί επομένως να υποστηρίξει πληθώρα διαφορετικών εφαρμογών με ομοιόμορφο τρόπο. Η διαχείριση αυτή στηρίζεται σε ένα βασικό σύνολο υπηρεσιών που καλύπτουν τις ανάγκες των εφαρμογών ως προς ελεγχόμενη παρουσίαση και περιήγηση στο περιεχόμενο, ανάκτηση βάσει περιεχομένου, εξατομικευμένα και διαβαθμισμένη μεταβίβαση και ανταλλαγή του περιεχομένου.

Σημαντικό μέρος της συνεισφοράς της διατριβής αποτελεί επίσης η χρήση του προτύπου MPEG-4, το οποίο αποτελεί την τεχνολογία αιχμής στα πρότυπα κωδικοποίησης οπτικοακουστικού περιεχομένου. Η συνεισφορά έγκειται στο ότι δεν έχει αναπτυχθεί μέχρι τώρα στη βιβλιογραφία σχετική εργασία, η οποία πέραν της χρήσης του προτύπου για την κωδικοποίηση και μετάδοση του υλικού, να έχει αναπτύξει μια μέθοδο για την εκμετάλλευση και χρήση της κωδικοποίησης βάσει περιεχομένου που παρέχει το πρότυπο. Η μέθοδος αυτή περιλαμβάνει την αναπαράσταση της κωδικοποιημένης πληροφορίας του MPEG-4 για την χρήση της στον πυρήνα υπηρεσιών πληροφοριακού συστήματος που υποστηρίζει ένα ευρύ φάσμα οπτικοακουστικών εφαρμογών και την οργάνωσή της σε σχεσιακές βάσεις δεδομένων για την υποστήριξη συνδιαλλαγών διαχείρισης και ανάκτησης.

Επίσης, όσον αφορά στη χρήση της μεταγλώσσας αναπαράστασης εγγράφων XML του Παγκόσμιου Ιστού για την κατασκευή της γλώσσα αναπαράστασης που περιγράφεται στην παρούσα διατριβή, είναι η πρώτη πλήρης προσπάθεια περιγραφής οπτικοακουστικού περιεχομένου με βάση την XML. Η τελευταία έχει χρησιμοποιηθεί έως τώρα για την αναπαράσταση με κείμενο άλλου είδους πληροφορίας. Η χρήση της XML για την περιγραφή της κωδικοποιημένης πληροφορίας των προτύπων MPEG-2 και MPEG-4 εξασφαλίζει την απλότητα της μορφής αναπαράστασης, την ευελιξία στην επέκταση του περιεχομένου με επιπλέον πληροφορία και την δυνατότητα ανταλλαγής περιεχομένου στον Παγκόσμιο Ιστό, ο οποίος αποτελεί την σημαντικότερη παγκόσμια πηγή πληροφορίας στις μέρες μας.

Η κατασκευή σχεσιακών βάσεων δεδομένων για την οργάνωση και αποθήκευση του οπτικοακουστικού περιεχομένου, μαζί με την γλώσσα αναπαράστασης και τις μεθοδολογίες μετάβασης από τη μια μορφή στην άλλη, δηλαδή από την περιγραφή στην αποθηκευμένη μορφή και αντίστροφα, αποτελούν επίσης συνεισφορά της διατριβής. Στο σημείο αυτό η συνεισφορά έγκειται στο γεγονός ότι υποστηρίζεται ένα σύνολο από εργαλεία που δίνουν πρόσθετη λειτουργικότητα σε μια βάση δεδομένων και σε μια γλώσσα αναπαράστασης, ταυτόχρονα. Η πρόσθετη αυτή λειτουργικότητα που παρέχεται από το πληροφοριακό σύστημα μέσω των συγκεκριμένων εργαλείων, παρέχει έναν αυτόματο τρόπο μετάβασης από μία μορφή του περιεχομένου σε μια άλλη, ανάλογα με τις ανάγκες των εφαρμογών. Μέχρι τώρα, οι εφαρμογές οπτικοακουστικού περιεχομένου εστίαζαν σε ένα μόνο τρόπο χρήσης της πληροφορίας αποκλειστικά για ανταλλαγή σε περιβάλλοντα όπως αυτό του Παγκοσμίου Ιστού ή αποκλειστικά για ανάκτηση από βάσεις δεδομένων, ενώ με τα εργαλεία μετάβασης από τη μια μορφή στην άλλη υποστηρίζεται η ελεύθερη εναλλαγή ανάμεσα σε μεταφορά και ανάκτηση βάσει περιεχομένου.

1.5 Δομή της διατριβής

Το δεύτερο κεφάλαιο της διατριβής αποτελεί σύντομη επισκόπηση της σχετικής έρευνας στον τομέα της διαχείρισης οπτικοακουστικής πληροφορίας. Η έμφαση είναι στην παρουσίαση του προτύπου MPEG-4 που αποτελεί νέο πρότυπο και χρησιμοποιήθηκε στην παρούσα διατριβή για την υποστήριξη ενός πυρήνα διαχείρισης οπτικοακουστικού περιεχομένου. Επίσης γίνεται σύντομη παρουσίαση της μεταγλώσσας XML που χρησιμοποιήθηκε για την κατασκευή γλώσσας αναπαράστασης περιεχομένου καθώς και δύο προτύπων περιγραφής μεταδεδομένων που χρησιμοποιούνται στα πλαίσια του DAVIC. Επίσης περιγράφεται συνοπτικά το σύστημα ΚΥΔΩΝΙΑ στα πλαίσια του προγράμματος SICMA, το οποίο χρησιμοποιήθηκε για την υλοποίηση μέρους της διατριβής.

Στο κεφάλαιο 3 παρουσιάζεται η αναπαράσταση των δεδομένων του προτύπου MPEG-2 και των μεταδεδομένων του προτύπου DAVIC. Επίσης περιγράφεται η υλοποίηση των υπηρεσιών περιήγησης στο περιεχόμενο MPEG-2 βίντεο και των υπηρεσιών εγκατάστασης και διαχείρισης περιεχομένου στα πλαίσια του DAVIC.

Στο κεφάλαιο 4 περιγράφεται η αναπαράσταση του περιεχομένου του MPEG-4 με τη γλώσσα οπτικοακουστικού περιεχομένου που αναπτύχθηκε στα πλαίσια της διατριβής.

Στο πέμπτο κεφάλαιο περιγράφεται το μοντέλο αποθήκευσης της αναπαράστασης του οπτικοακουστικού περιεχομένου σε σχεσιακή βάση δεδομένων. Το μοντέλο αποθήκευσης καθορίζεται με βάση το εννοιολογικό μοντέλο οντοτήτων-συσχετίσεων και με βάση το σχεσιακό μοντέλο δεδομένων. Επίσης παρουσιάζονται τα εργαλεία μετατροπής από τη μια μορφή αναπαράστασης του περιεχομένου στην άλλη, καθώς και οι συνδιαλλαγές υποστήριξης των βασικών υπηρεσιών διαχείρισης περιεχομένου.

Στο κεφάλαιο 6 περιγράφεται η υποστήριξη των ίδιων των εφαρμογών με χρήση της υποδομής διαχείρισης οπτικοακουστικού περιεχομένου που παρουσιάστηκε στο προηγούμενο μέρος της διατριβής.

Τέλος, στο κεφάλαιο 7 γίνεται ανακεφαλαίωση των επιτευγμάτων και της συνεισφοράς της διατριβής.

Κεφάλαιο 2

Σχετική έρευνα

Στο κεφάλαιο αυτό παρουσιάζονται συνοπτικά οι κυριότερες τεχνολογίες που χρησιμοποιήθηκαν από την παρούσα διατριβή με στόχο την υποστήριξη ενός πυρήνα υπηρεσιών διαχείρισης οπτικοακουστικού περιεχομένου. Οι τεχνολογίες αυτές περιλαμβάνουν ευρέως αποδεκτά πρότυπα κωδικοποίησης οπτικοακουστικών δεδομένων όπως το MPEG-4, πρότυπα αναπαράστασης πληροφορίας και συγκεκριμένα τη γλώσσα XML του Παγκόσμιου Ιστού, το πρότυπο υποστήριξης οπτικοακουστικών εφαρμογών DAVIC και το σύστημα ΚΥΔΩΝΙΑ που χρησιμοποιήθηκε ως πλατφόρμα για την υλοποίηση μέρους των οπτικοακουστικών υπηρεσιών.

2.1 Εισαγωγή

Στο κεφάλαιο αυτό παρουσιάζονται τα πρότυπα, οι τεχνολογίες και τα συστήματα που κατείχαν σημαντικό ρόλο στην πραγματοποίηση της παρούσας διατριβής. Στην ενότητα 2.2 παρουσιάζεται το πρότυπο MPEG-4, το οποίο ως νέο και πρωτοποριακό πρότυπο στην περιοχή της κωδικοποίησης οπτικοακουστικών μέσων, αποτέλεσε κύριο γνώμονα για την παρούσα διατριβή. Σημαντικό μέρος της συνεισφοράς της διατριβής αφορά στην κατανόηση, περιγραφή και χρήση της κωδικοποιημένης πληροφορίας του MPEG-4, με σκοπό η τελευταία να χρησιμοποιηθεί για την υποστήριξη υπηρεσιών διαχείρισης οπτικοακουστικού περιεχομένου. Συγκεκριμένα δίνεται έμφαση στα τμήματα του προτύπου που έχουν ιδιαίτερη σημασία για την ανάπτυξη της διατριβής και ειδικά στην αναπαράσταση των οπτικοακουστικών αντικειμένων, την περιγραφή και σύνθεση των σκηνών και τη συσχέτιση πληροφορίας με τα κωδικοποιημένα αντικείμενα.

Στην ενότητα 2.3 παρουσιάζεται συνοπτικά η γλώσσα XML του Παγκόσμιου Ιστού, η οποία χρησιμοποιήθηκε στην παρούσα διατριβή για την αναπαράσταση ή περιγραφή του MPEG περιεχομένου με ενιαίο, απλό και ευέλικτο τρόπο. Η αναπαράσταση του οπτικοακουστικού περιεχομένου με τη γλώσσα XML αποτελεί μέρος της γενικής υποδομής υποστήριξης οπτικοακουστικών υπηρεσιών σε πληροφοριακά συστήματα.

Στην ενότητα 2.4 παρουσιάζεται το πρότυπο DAVIC, το οποίο καθορίζει την υποστήριξη υπηρεσιών μετάδοσης πολυμέσων ανάμεσα σε ετερογενή συστήματα, με στόχο να επιτευχθεί η μέγιστη δυνατή διαλειτουργικότητα. Το πρώτο DAVIC χρησιμοποιήθηκε στην παρούσα διατριβή για την υποστήριξη υπηρεσιών διαχείρισης MPEG περιεχομένου στον εξυπηρετητή πολυμέσων ΚΥΔΩΝΙΑ. Τέλος, στην ενότητα 2.5 παρουσιάζεται το σύστημα ΚΥΔΩΝΙΑ που αποτέλεσε μέρος της πλατφόρμας υλοποίησης της παρούσας διατριβής.

2.2 Το πρότυπο MPEG-4

Το πρότυπο MPEG-4 ξεκίνησε το 1996 με στόχο να καλύψει τις εξελίξεις σε εφαρμογές πολυμέσων όπως η ψηφιακή τηλεόραση, τα αλληλεπιδραστικά γραφικά και ο Παγκόσμιος Ιστός. Η εξέλιξη στις παραπάνω εφαρμογές που αφορά στο εν λόγω πρότυπο, συνοψίζεται στα εξής χαρακτηριστικά:

- υπάρχει αυξανόμενη ανάγκη χρήσης τεχνητά παραγόμενου υλικού,
- υπάρχει ανάγκη συσχέτισης επιπλέον πληροφορίας με το ψηφιοποιημένο υλικό,
- το υλικό μεταδίδεται πάνω από εξελιγμένα δίκτυα μετάδοσης και τέλος
- υπάρχει έντονη ανάγκη για αλληλεπιδραστική κατανάλωση του υλικού.

Τα χαρακτηριστικά αυτά συνεπάγονται την ανάγκη για βελτίωση της αναπαράστασης και διαχείρισης του περιεχομένου, δηλαδή της πληροφορίας που περιέχει το οπτικοακουστικό υλικό των εφαρμογών. Στόχος του MPEG-4 είναι η συνένωση της παραγωγής, κατανομής και προσπέλασης του περιεχομένου και στα τρία είδη εφαρμογών που αναφέρονται παραπάνω. Στις εφαρμογές αυτές διακρίνονται τρεις βασικοί παράγοντες των οποίων οι ανάγκες θα καλυφθούν, οι *συγγραφείς*, οι *παροχείς δικτυακών υπηρεσιών* και οι *τελικοί χρήστες*. Οι συγγραφείς θα ωφεληθούν από το πρότυπο μέσω της δυνατότητας παραγωγής περιεχομένου που θα έχει μεγαλύτερη επαναχρησιμότητα και ευελιξία, ενώ οι τελικοί χρήστες θα διαθέτουν πλούσια λειτουργικότητα και υψηλό επίπεδο αλληλεπίδρασης με το περιεχόμενο. Οι παροχείς δικτυακών υπηρεσιών θα προσφέρουν διαφανή πληροφορία, η οποία θα μεταφράζεται στα κατάλληλα μηνύματα σήμανσης κάθε τύπου δικτύου.

Για να πετύχει τους στόχους του, το πρότυπο MPEG-4 καθορίζει εργαλεία για τις εξής λειτουργίες:

- Αναπαράσταση των οπτικοακουστικών αντικειμένων και των δομικών μονάδων του περιεχομένου τους. Τα οπτικοακουστικά αντικείμενα ονομάζονται *αντικείμενα μέσω* και μπορεί να είναι φυσικά ή συνθετικά. Οι δομικές μονάδες του περιεχομένου είναι υποαντικείμενα που διακρίνονται μέσα στο υλικό. Για παράδειγμα, μια σειρά βίντεο είναι ένα οπτικοακουστικό αντικείμενο, ενώ ένας άνθρωπος που εμφανίζεται στο βίντεο είναι επιμέρους δομική μονάδα.
- Περιγραφή της σύνθεσης αντικειμένων για την κατασκευή *οπτικοακουστικών σκηνών*.
- Πολυπλεξία και συγχρονισμός των δεδομένων για μεταφορά πάνω από δίκτυα, παρέχοντας *Ποιότητα Υπηρεσίας* (QoS) κατάλληλη για τη φύση κάθε αντικειμένου.
- *Αλληλεπίδραση* του χρήστη με την οπτικοακουστική σκηνή στο τερματικό.

Όσον αφορά στην αναπαράσταση και κωδικοποίηση των οπτικοακουστικών αντικειμένων, το πρότυπο καθορίζει ένα σύνολο από στοιχειώδη αντικείμενα μέσω που μπορούν να αναπαραστήσουν φυσικό αλλά και συνθετικό περιεχόμενο. Τα αντικείμενα αυτά είναι εικόνες, βίντεο, ήχος, κείμενο, γραφικά, ανθρώπινα πρόσωπα και συνθετικός ήχος και κωδικοποιούνται με ένα σύνολο από διαφορετικά εργαλεία κατάλληλα για μεγάλο εύρος εφαρμογών.

Στο τερματικό τα αντικείμενα μέσω συνθέτουν οπτικοακουστικές σκηνές οι οποίες παρουσιάζονται στο χρήστη και δέχονται τις ενέργειές του. Το MPEG-4 παρέχει λειτουργικότητα για την τοποθέτηση αντικειμένων στο σύστημα συντεταγμένων μιας σκηνής, παρέχει τη δυνατότητα εφαρμογής μετασχηματισμών για αλλαγή της γεωμετρικής και ακουστικής εμφάνισης των αντικειμένων μιας σκηνής, τη δυνατότητα ομαδοποίησης στοιχειωδών αντικειμένων σε σύνθετα αντικείμενα, τη δυνατότητα αλλαγής των ιδιοτήτων των αντικειμένων σύμφωνα με εισερχόμενα δεδομένα και τη δυνατότητα αλληλεπιδραστικής μεταβολής των σημείων θέασης και ακρόασης του χρήστη.

Η μετάδοση των αντικειμένων γίνεται μέσα σε *στοιχειώδεις ροές*, οι οποίες αντιστοιχίζονται σε μικρά αντικείμενα που ονομάζονται *περιγραφείς*. Με τους περιγραφείς επιτυγχάνεται ιεραρχική κωδικοποίηση των αντικειμένων και συσχέτιση μεταδεδομένων με αυτά. Κάθε ροή περιέχει πληροφορία για την Ποιότητα Υπηρεσίας με βάση την οποία πρέπει να μεταδοθεί. Ο συγχρονισμός των στοιχειωδών ροών επιτυγχάνεται με *χρονοσφραγίδες* που συσχετίζονται με ανεξάρτητες ατομικές μονάδες των αντικειμένων. Η μετάδοση καθορίζεται με βάση ένα επίπεδο συγχρονισμού και ένα επίπεδο μετάδοσης που αποτελείται από δύο επίπεδα πολυπλεξίας.

Η αλληλεπίδραση του χρήστη με μια οπτικοακουστική σκηνή αποτελείται από λειτουργίες που καθορίζονται από το συγγραφέα της σκηνής και μπορεί να περιλαμβάνουν την αλλαγή του σημείου θέασης ή ακρόασης, την αλλαγή της θέσης των αντικειμένων, την πυροδότηση μιας σειράς γεγονότων, την επιλογή γλώσσας, κλπ.

Στο πρότυπο δίνεται έμφαση και στην πληροφορία για τα πνευματικά δικαιώματα, η οποία μπορεί να κωδικοποιηθεί μέσα στα αντικείμενα μέσων. Η πληροφορία αυτή μπορεί να ανήκει σε διαφορετικά διεθνή συστήματα ή να είναι ένα ζευγάρι τιμών (ρόλος, όνομα), π.χ. («συγγραφέας», «John Smith»).

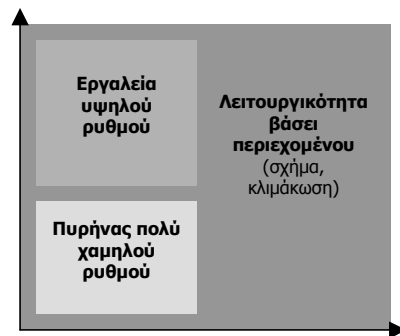
Όπως και τα προηγούμενα MPEG πρότυπα, το MPEG-4 χωρίζεται σε τμήματα. Τα κύρια τμήματά του είναι τα *Συστήματα* (Systems), ο *Ήχος* (Audio), τα *Οπτικά Αντικείμενα* (Visual) και το *Πλαίσιο Μετάδοσης Πολυμέσων* (DMIF). Τα τμήματα για την οπτική πληροφορία και τον ήχο, όπως αναφέρθηκε και παραπάνω, ορίζουν νέους ευέλικτους αλγόριθμους κωδικοποίησης φυσικών και συνθετικών αντικειμένων, ανθρώπινων προσώπων και τυχαίων τμημάτων εικόνων, φωνής, μουσικής και ήχου. Το τμήμα για τα Συστήματα, εκτός από το ότι καθορίζει το συγχρονισμό, την πολυπλεξία και την «πακετοποίηση» των ροών, ορίζει τη σύνθεση σκηνών από στοιχειώδη αντικείμενα και την διαχείριση των σκηνών στο τερματικό. Το Πλαίσιο Μετάδοσης Πολυμέσων (DMIF) αποτελεί ένα πρωτόκολλο συνόδων που αποκρύπτει τις λεπτομέρειες των χαμηλότερων δικτυακών επιπέδων.

2.2.1 Οπτικά Αντικείμενα

Το τμήμα του MPEG-4 για τα οπτικά αντικείμενα είναι αντίστοιχο των τμημάτων για την κωδικοποίηση βίντεο των προτύπων MPEG-1 και 2, στα οποία η έμφαση είχε δοθεί στην κωδικοποίηση σειρών από κινούμενες εικόνες με καλή ποιότητα και υψηλό βαθμό συμπίεσης. Οι εφαρμογές που στοχεύει να καλύψει το MPEG-4 είναι κυρίως αλληλεπιδραστικές, όπως πολυμέσα στο Ίντερνετ, βιντεοπαιχνίδια, διαπροσωπικές επικοινωνίες (βιντεοσυνεδριάσεις, κλπ), αλληλεπιδραστικά μέσα αποθήκευσης, ταχυδρόμηση πολυμέσων, δικτυακές βάσεις δεδομένων, απομακρυσμένα συστήματα άμεσης βοήθειας και ασύρματα πολυμέσα. Για αυτό το λόγο εκτός από την αποδοτική συμπίεση και την καλή ποιότητα κωδικοποίησης, το MPEG-4 εστιάζει στην *κωδικοποίηση με βάση το περιεχόμενο* των οπτικών αντικειμένων.

Επομένως, το τμήμα για τα οπτικά αντικείμενα λαμβάνει υπόψη του κυρίως δύο παράγοντες: το ρυθμό των δεδομένων και τη λειτουργικότητα βάσει του περιεχομένου. Όσον αφορά στο ρυθμό των δεδομένων, ορίζεται ένας πυρήνας εργαλείων για κωδικοποίηση του βίντεο σε πολύ χαμηλούς ρυθμούς, από 5 ως 64 Kbps, που ονομάζεται

Πυρήνας Βίντεο Πολύ Χαμηλού Ρυθμού, ή VLBV (Very Low Bitrate Video) Core. Ο πυρήνας αυτός καθορίζει την κωδικοποίηση συμβατικών, δηλαδή όχι βάσει του περιεχομένου τους, τετραγωνικών εικόνων για εφαρμογές επικοινωνίας και παρέχει τυχαία προσπέλαση και λειτουργικότητα VCR. Η ίδια λειτουργικότητα επεκτείνεται πέρα από το VLBV Core στην κωδικοποίηση υψηλότερων ρυθμών δεδομένων από 64 kbps ως 4 Mbps. Η λειτουργικότητα βάσει περιεχομένου θεωρείται πρόσθετη της κωδικοποίησης και επεκτείνει τις τεχνικές αναπαράστασης ορθογώνιων εικόνων, επιτρέποντας την ξεχωριστή κωδικοποίηση και αποκωδικοποίηση των αντικειμένων σε μια σκηνή. Στο παρακάτω σχήμα φαίνεται η ταξινόμηση των εργαλείων και αλγορίθμων του MPEG-4 για τα οπτικά αντικείμενα.



Σχήμα 2-1. Εργαλεία του MPEG-4

Τα εργαλεία κωδικοποίησης οπτικών αντικειμένων αποτελούνται από έναν πυρήνα κωδικοποίησης σε πολύ χαμηλούς ρυθμούς και ένα σύνολο εργαλείων πολύ υψηλού ρυθμού. Και τα δύο αυτά υποσύνολα συμπληρώνονται από τα εργαλεία κωδικοποίησης βάσει του περιεχομένου των ροών.

Η κωδικοποιημένη οπτική πληροφορία οργανώνεται μέσα στις ψηφιοσειρές σε βαθμίδες. Μια ψηφιοσειρά βίντεο περιλαμβάνει τη *βασική βαθμίδα* και καμία ή περισσότερες *βελτιωτικές βαθμίδες*. Οι βελτιωτικές βαθμίδες ονομάζονται και *βαθμίδες κλιμάκωσης* και παρέχουν κλιμακωτή αποκωδικοποίηση με παρόμοιο τρόπο όπως στα προηγούμενα MPEG πρότυπα. Τα δομικά στοιχεία της ψηφιοσειράς είναι τα οπτικά αντικείμενα. Η ψηφιοσειρά μπορεί να αποτελείται από ένα ή περισσότερα οπτικά αντικείμενα, καθένα από τα οποία περιέχει εκτός από το είδος (ή τύπο) του και τα κωδικοποιημένα δεδομένα, την ένδειξη για το Προφίλ και το Επίπεδο με το οποίο συμμορφώνεται. Το Προφίλ στα πρότυπα της οικογένειας MPEG καθορίζει ένα υποσύνολο των εργαλείων κωδικοποίησης, ενώ το Επίπεδο καθορίζει την πολυπλοκότητα των συγκεκριμένων εργαλείων. Ανάλογα με τον τύπο του αντικειμένου ορίζεται η δομή της ψηφιοσειράς που το περιέχει και είναι υποσύνολο μιας Οπτικής Ψηφιοσειράς.

Τα οπτικά αντικείμενα διακρίνονται σε φυσικά και συνθετικά. Τα φυσικά οπτικά αντικείμενα εκτός από βίντεο περιλαμβάνουν ακίνητες εικόνες και *υφή*. Για τα φυσικά οπτικά αντικείμενα παρέχεται συνοπτικά η εξής λειτουργικότητα:

- Συμπίεση εικόνων, βίντεο και υφής.
- Κωδικοποίηση αντικειμένων τυχαίου σχήματος εκτός από ορθογώνιες εικόνες.
- Τυχαία προσπέλαση σε όλα τα είδη των οπτικών αντικειμένων.
- Εκτεταμένη λειτουργικότητα διαχείρισης εικόνων και βίντεο.
- Κωδικοποίηση εικόνων και βίντεο με βάση το περιεχόμενό τους.
- Κλιμακωτή κωδικοποίηση υφής, εικόνων και βίντεο με βάση το περιεχόμενό τους.
- Χωρική, χρονική και ποιοτική κλιμάκωση.
- Σθεναρότητα και ανεκτικότητα σε λάθη.

Τα συνθετικά αντικείμενα του MPEG-4 είναι ένα υποσύνολο γραφικών που περιλαμβάνει ανθρώπινα πρόσωπα και υφή και αναπαρίστανται με τριγωνικά δυσδιάστατα πλέγματα.

2.2.1.1 Βίντεο

Τα αντικείμενα βίντεο αποτελούνται από μία ή περισσότερες *Βαθμίδες Αντικειμένου Βίντεο* (Video Object Layers ή VOL's). Κάθε Βαθμίδα Αντικειμένου Βίντεο μπορεί να είναι κωδικοποιημένη με *προοδευτική* (progressive) ή *πεπλεγμένη* (interlaced) κωδικοποίηση. Οι Βαθμίδες Αντικειμένων είναι δυνατόν να περιέχουν πληροφορία για τις διαστάσεις του αντικειμένου, οι οποίες εκφράζονται ως προς εικονοστοιχεία φωτεινότητας, για το ρυθμό δεδομένων, το ρυθμό των «εικόνων» του βίντεο, τη μορφή του χρώματος και τη μορφή του σχήματος της Βαθμίδας. Το σχήμα των αντικειμένων βίντεο μπορεί να είναι τετράγωνο, δισαδικό ή κλίμακας του γκρι. Για να επιτευχθεί η κωδικοποίηση βάσει του περιεχομένου ορίζεται η έννοια του *Επιπέδου Αντικειμένου Βίντεο* (Video Object Plane ή VOP), το οποίο παριστάνει μια τυχαίου σχήματος περιοχή της εικόνας και αντιστοιχεί σε κάποιο αντικείμενο ή περιοχή ενδιαφέροντος. Διαδοχικά VOP που αντιστοιχούν στο ίδιο φυσικό αντικείμενο ορίζουν ένα *Αντικείμενο Βίντεο* (Video Object), το οποίο κωδικοποιείται σε μια χωριστή Βαθμίδα (VOL), μαζί με την πληροφορία για το σχήμα, την υφή και την κίνηση των διαδοχικών Επιπέδων (VOP) της. Η πηγή δεδομένων για τα VOL μπορεί να δημιουργείται με τεχνολογίες όπως chroma-key, ή να παράγεται με τμηματοποίηση του αρχικού φυσικού βίντεο. Επομένως η έννοια του καρτέ (frame) έχει αντικατασταθεί στο βίντεο του MPEG-4 από την έννοια του Επιπέδου Αντικειμένου Βίντεο και υπάρχουν 4 είδη Επιπέδων Αντικειμένου Βίντεο:

- *Ενδοκωδικοποιημένα επίπεδα αντικειμένων βίντεο* (Intracoded VOP ή I-VOP), τα οποία κωδικοποιούνται ανεξάρτητα από άλλα VOP.
- *Προφητικά επίπεδα αντικειμένων* (Predictive ή P-VOP), τα οποία κωδικοποιούνται με πρόβλεψη κίνησης από ένα προηγούμενο VOP.
- *Προφητικά επίπεδα διπλής κατεύθυνσης* (Bidirectionally predictive ή B-VOP), τα οποία κωδικοποιούνται χρησιμοποιώντας πρόβλεψη κίνησης από ένα προηγούμενο ή και ένα επόμενο VOP.
- *Επίπεδα sprite* (Sprite VOP), τα οποία αντιστοιχούν σε μεγάλα στατικά αντικείμενα, όπως το φόντο εικόνων.

Τα VOP κατά αναλογία με τα καρέ στα πρότυπα MPEG-1 και 2 ομαδοποιούνται σε *Ομάδες από Επίπεδα Αντικειμένων Βίντεο* (Group of VOP's), οι οποίες αρχίζουν πάντα με ένα I-VOP και κωδικοποιούνται ανεξάρτητα από τις άλλες ομάδες (επόμενες ή προηγούμενες). Τα VOP με τη σειρά τους οργάνωνονται σε μη επικαλυπτόμενα macroblock, τα οποία αποτελούνται από χρωματική πληροφορία σε μορφή 4:2:0 και αποτελούν τη βάση για τα διανύσματα κίνησης. Όταν η μορφή του βίντεο είναι προοδευτική, τα macroblock φωτεινότητας προκύπτουν από συνεχόμενες γραμμές του σήματος και η οργάνωση αυτή ονομάζεται *κωδικοποίηση καρέ*. Όταν η μορφή του βίντεο είναι πεπλεγμένη, υπάρχουν δύο διαφορετικές δυνατότητες οργάνωσης των macroblock φωτεινότητας: από συνεχόμενες γραμμές όλης της εικόνας, δηλαδή γραμμές και από τα δύο *πεδία*, είτε από γραμμές μόνο του ενός εκ των δύο πεδίων. Η δεύτερη οργάνωση ονομάζεται *κωδικοποίηση πεδίου*. Τα macroblock για τη χρωματικότητα σχηματίζονται πάντα με κωδικοποίηση καρέ. Στα macroblock εφαρμόζεται διακριτός μετασχηματισμός συνημίτονου (Discrete Cosine Transform ή DCT).

Τα αντικείμενα τύπου sprite, όπως αναφέρθηκε και παραπάνω είναι μεγάλες στατικές εικόνες που συνήθως αντιστοιχούν στο φόντο μιας σειράς εικόνων και εξάγονται (διαχωρίζονται από τα άλλα αντικείμενα της σειράς) πριν γίνει κωδικοποίηση του βίντεο. Το sprite μεταδίδεται μόνο στο πρώτο καρέ της σειράς και στα υπόλοιπα καρέ εφαρμόζεται απόσβεση κίνησης και κωδικοποιούνται 8 παράμετροι που περιγράφουν την κίνησή του. Οι παράμετροι κίνησης των sprite αποτελούν το *συγγενή* (affine) μετασχηματισμό του sprite που μεταδόθηκε στο πρώτο καρέ. Στο δέκτη γίνεται σύνθεση του φόντου με τα υπόλοιπα αντικείμενα που κωδικοποιούνται ως προς το (τυχαίο) σχήμα τους. Σε εφαρμογές όπου η καθυστέρηση πρέπει να είναι μικρή, είναι δυνατόν να μεταδίδεται ένα sprite τμηματικά.

2.2.1.2 Πρόσωπα

Τα πρόσωπα ανήκουν στη κατηγορία των συνθετικών αντικειμένων που υποστηρίζει το MPEG-4. Τα *αντικείμενα προσώπων* αποτελούν αναπαραστάσεις ανθρώπινων προσώπων και είναι δομημένα έτσι ώστε να προβάλλουν κινήσεις ομιλίας και εκφράσεις. Η λειτουργικότητα που παρέχεται για τα πρόσωπα είναι:

- Ορισμός και κωδικοποίηση *Παραμέτρων Ορισμού Προσώπων* (Facial Definition Parameters – FDP’s). Οι παράμετροι αυτοί περιλαμβάνουν το σχήμα και την υφή που καλύπτει το πρόσωπο.
- Ορισμός και κωδικοποίηση *Παραμέτρων Κίνησης Προσώπων* (Facial Animation Parameters – FAP’s). Η κίνηση των προσώπων καθορίζεται από τις εκφράσεις, τα συναισθήματα και την προφορά.

Η κωδικοποίηση των προσώπων αρχίζει με κάποια ουδέτερη έκφραση, με βάση την οποία παίρνουν τιμές οι παράμετροι κίνησης. Το ουδέτερο πρόσωπο ορίζεται με συγκεκριμένες τιμές στις Παραμέτρους Ορισμού του Προσώπου και είναι έτοιμο για ζωγράφισμα στο τερματικό. Έχει επίσης τη δυνατότητα να λάβει από μία εισερχόμενη ψηφιοσειρά επιπλέον παραμέτρους ορισμού που θα το μετασχηματίσουν, καθώς και παραμέτρους κίνησης που θα καθορίσουν την έκφραση και την ομιλία του προσώπου. Το MPEG-4 δίνει επίσης τη δυνατότητα μετάδοσης ολόκληρου μοντέλου για πρόσωπα.

Οι παράμετροι κίνησης βασίζονται στη μελέτη της κίνησης των μυών του προσώπου και αναπαριστούν ένα πλήρες σύνολο βασικών ενεργειών και φυσικών εκφράσεων. Χωρίζονται σε δύο βασικές κατηγορίες, *οπτικά φωνήματα* (visual phonemes ή visemes) και *εκφράσεις* (expressions). Τα οπτικά φωνήματα επιτρέπουν την αναπαράσταση των κινήσεων κατά την ομιλία. Στο πρότυπο ορίζεται ένα σύνολο από στατικά οπτικά φωνήματα, τα οποία είναι εύκολα αναγνωρίσιμα. Οι εκφράσεις των προσώπων περιγράφονται και κωδικοποιούνται με κείμενο.

Οι μονάδες των παραμέτρων κίνησης ονομάζονται *Μονάδες Παραμέτρων Κίνησης Προσώπων* (Facial Animation Parameter Units – FAPU’s) και αντιστοιχούν σε κλάσματα των αποστάσεων ανάμεσα σε κάποια κύρια χαρακτηριστικά των προσώπων. Οι παράμετροι κίνησης δηλαδή ορίζουν τις αποστάσεις των παραμέτρων ορισμού (σημείων) των προσώπων με τρόπο ανεξάρτητο από το μοντέλο που χρησιμοποιείται στον αποκωδικοποιητή. Παραδείγματα μονάδων κίνησης είναι η απόσταση ανάμεσα στα μάτια, η απόσταση ανάμεσα σε μύτη και μάτι, η απόσταση στόματος και μύτης και το πλάτος του στόματος.

Τα πρόσωπα όπως και τα αντικείμενα βίντεο κωδικοποιούνται ως μια χρονική σειρά από *Επίπεδα Αντικειμένου Προσώπου* (Face Object Planes). Τα επίπεδα αυτά μπορούν να ομαδοποιηθούν σε χρονικές ομάδες που αποτελούνται από 16 συνεχόμενα επίπεδα προσώπου, οι οποίες ονομάζονται Ομάδες Επιπέδων Αντικειμένου Προσώπου. Στην πρώτη περίπτωση (κωδικοποίηση σε επίπεδα), η αποκωδικοποίηση βασίζεται σε καρέ, ενώ στη δεύτερη (κωδικοποίηση σε ομάδες επιπέδων), η αποκωδικοποίηση βασίζεται στο διακριτό μετασχηματισμό συνημίτονου.

2.2.1.3 Εικόνες και υφή

Για υφή και ακίνητες εικόνες χρησιμοποιείται ένας zerotree wavelet αλγόριθμος κωδικοποίησης, ο οποίος παρέχει καλή συμπίεση σε πολλές συχνότητες, έντεκα επίπεδα κλιμάκωσης και κωδικοποίηση τυχαίων σχημάτων. Τα είδη κλιμάκωσης που υποστηρίζονται είναι ως προς την ποιότητα, τη χωρική και τη χρονική ανάλυση. Η κωδικοποίηση του σχήματος τυχαίων αντικειμένων γίνεται για να παρέχονται αναπαραστάσεις βάσει του περιεχομένου κατάλληλες για εφαρμογές όπως βάσεις δεδομένων, παιχνίδια, κλπ. Γι' αυτό το σκοπό χρησιμοποιούνται οι *δυαδικοί πίνακες άλφα* (binary alpha planes), στους οποίους ορίζεται για κάθε εικονοστοιχείο αν αυτό ανήκει σε ένα αντικείμενο ή όχι.

2.2.1.4 Πλέγματα

Τα Αντικείμενα Πλεγμάτων ή *Πλέγματα* (Meshes) αποτελούν έναν από τους 4 τύπους οπτικών αντικειμένων και στο MPEG-4 είναι *δυσδιάστατα* και *τριγωνικά*. Βασικό πλεονέκτημα των δυσδιάστατων πλεγμάτων είναι ότι μπορούν να κατασκευαστούν από μία όψη των αντικειμένων και να μοντελοποιήσουν την κίνηση και το σχήμα των αντικειμένων του βίντεο σε ένα ενιαίο πλαίσιο. Τα τριγωνικά πλέγματα έχουν χρησιμοποιηθεί σε γραφικά για τη μοντελοποίηση τρισδιάστατων σχημάτων και αντικειμένων, ενώ τα δυσδιάστατα τριγωνικά πλέγματα μπορούν να θεωρηθούν προβολές τρισδιάστατων πλεγμάτων στο επίπεδο της εικόνας.

Το MPEG-4 υποστηρίζει πλέγματα δυναμικά, εμπρόσθια και βάσει περιεχομένου. Ένα δυσδιάστατο *δυναμικό πλέγμα* (dynamic mesh) ορίζεται ως μια χρονική σειρά θέσεων των κόμβων του ίδιου πλέγματος, δηλαδή κόμβων που έχουν την ίδια τοπολογία αλλά μεταβάλλεται η θέση τους με το χρόνο. Η αναπαράσταση των δυναμικών πλεγμάτων γίνεται με το αρχικό πλέγμα και με διανύσματα κίνησης των κόμβων του. *Εμπρόσθια πλέγματα* (forward tracking) ονομάζονται αυτά στα οποία οι κόμβοι του αρχικού πλέγματος παρακολουθούν τα μελλοντικά χαρακτηριστικά της εικόνας με διανύσματα κίνησης. Τα πλέγματα μπορούν να είναι κανονικά (για παράδειγμα να αποτελούνται από όμοια τρίγωνα που επαναλαμβάνονται το ένα δίπλα στο άλλο) ή βάσει του περιεχομένου

της εικόνας, στα οποία η δειγματοληψία κόμβων δεν είναι ομοιόμορφη αλλά γίνεται σε συγκεκριμένα σημεία των αντικειμένων (για παράδειγμα πλέγματα Delaunay, τα οποία καθορίζονται από τις θέσεις των κόμβων και τις ακμές).

Για την αντιστοίχιση της υψής από το πλέγμα αναφοράς σε κάποιο μελλοντικό πλέγμα, όπου φυσιολογικά το αρχικό «μωσαϊκό» έχει παραμορφωθεί από την κίνηση των κόμβων, χρησιμοποιείται παραμετρική μέθοδος, η οποία είναι συνάρτηση των διανυσμάτων κίνησης. Το MPEG-4 χρησιμοποιεί αντιστοίχιση τύπου affine, η οποία είναι γραμμική με χαμηλό υπολογιστικό κόστος και μπορεί να μοντελοποιήσει 6 βαθμούς ελευθερίας: μετατόπιση, περιστροφή, κλιμάκωση, ανάκλαση, αποκοπή και διατήρηση των ευθυγράμμων τμημάτων σε μια εικόνα. Οι βαθμοί αυτοί ταιριάζουν με τους βαθμούς ελευθερίας που δίνουν τα διανύσματα κίνησης των τριών κόμβων κάθε τριγώνου. Τα πλέγματα επομένως είναι αποδοτικά στην αναπαράσταση της συνεχούς κίνησης ήπιων παραμορφώσεων.

Η λειτουργικότητα που παρέχουν τα πλέγματα του MPEG-4 μπορεί να συνοψιστεί στα εξής:

- Διαχείριση των αντικειμένων του βίντεο και συγκεκριμένα:
 - *Επαυξημένη πραγματικότητα* (augmented reality), δηλαδή συγχώνευση τεχνητών εικόνων με πραγματικές για τη δημιουργία βελτιωμένης πληροφορίας για την παρουσίαση.
 - Μεταμόρφωση ή κίνηση συνθετικών αντικειμένων. Ένα φυσικό αντικείμενο του βίντεο μπορεί να αντικατασταθεί από ένα άλλο φυσικό αντικείμενο το οποίο μπορεί να εξαχθεί από ένα βίντεο ή από μια εικόνα με μεταμόρφωσή του, χρησιμοποιώντας την πληροφορία της κίνησης που υπάρχει για το αρχικό αντικείμενο.
 - Χωροχρονική παρεμβολή (spatiotemporal interpolation). Η μοντελοποίηση της κίνησης παρέχει σθεναρότερη παρεμβολή με βάση την κίνηση.
- Συμπίεση αντικειμένων του βίντεο. Η συμπίεση μπορεί να επιτευχθεί με τη μετάδοση πινάκων αντιστοίχισης υψής μόνο σε καρέ-κλειδιά.
- Δεικτοδότηση βάσει περιεχομένου. Η μοντελοποίηση με δυσδιάστατα πλέγματα επιτρέπει τη δημιουργία κινούμενων οπτικών στιγμιότυπων των αντικειμένων, τα οποία μπορούν να χρησιμοποιηθούν ως περιλήψεις (visual synopses).

Επίσης, η αναπαράσταση αντικειμένων με πλέγματα παρέχει ακριβή πληροφορία για την τροχιά τους, διευκολύνοντας έτσι την ανάκτηση με βάση την κίνηση. Επιπλέον παρέχει

καλύτερη αναπαράσταση και για το σχήμα των αντικειμένων, διευκολύνοντας την ανάκτηση αντικειμένων βάσει του σχήματός τους.

Στην ψηφιοσειρά του MPEG-4 τα πλέγματα οργανώνονται όπως και τα αντικείμενα βίντεο σε επίπεδα και μια σειρά από *Επίπεδα Αντικειμένων Πλεγμάτων* (Mesh Object Planes) αναπαριστά τους σχηματισμούς που πρέπει να εφαρμοστούν σε ένα επίπεδο αντικειμένου βίντεο ή υφής για να δημιουργεί ένα συνθετικό αντικείμενο βίντεο με κίνηση. Τα επίπεδα πλεγμάτων διακρίνονται σε δύο κατηγορίες: ενδοκωδικοποιημένα (intra-coded) και προφητικά (predictive). Η πρόβλεψη χρησιμοποιείται για τη θέση των κόμβων.

2.2.1.5 Προφίλ και επίπεδα οπτικών αντικειμένων

Τα προφίλ για τα οπτικά αντικείμενα είναι τα εξής: Simple (Απλό), Core (Βασικό), Main (Κύριο), Simple Scalable (Απλό Κλιμακωτό), 12-Bit, Basic Animated 2D Texture (Βασική κινούμενη 2D υφή), Animated 2D Mesh (Κινούμενο 2D πλέγμα), Simple Face (Απλό πρόσωπο), Simple Scalable Texture (Απλή κλιμακωτή υφή) και Core Scalable Texture (Βασική κλιμακωτή υφή). Τα προφίλ καθορίζονται ως προς την πολυπλοκότητα των εργαλείων κωδικοποίησης για τα οπτικά αντικείμενα. Στον παρακάτω πίνακα φαίνεται η σχέση των προφίλ με τα εργαλεία για τα οπτικά αντικείμενα.

Εργαλεία Οπτικών αντικειμένων	Προφίλ οπτικών αντικειμένων									
	Simple	Core	Main	Simple Scalable	12-Bit	Basic Anim. 2D Texture	Anim. 2D Mesh	Simple Face	Simple Scalable Texture	Core Scal. Texture
I-VOP	✓	✓	✓	✓	✓		✓			
P-VOP	✓	✓	✓	✓	✓		✓			
AC/DC Prediction	✓	✓	✓	✓	✓		✓			
Slice Resynchronization	✓	✓	✓	✓	✓		✓			
Data Partitioning	✓	✓	✓	✓	✓		✓			
Reversible VLC	✓	✓	✓	✓	✓		✓			
4MV, Unrestricted MV	✓	✓	✓	✓	✓		✓			
Binary Shape Coding		✓	✓		✓	✓	✓			
H.263/MPEG-2 Quantization Tables		✓	✓		✓		✓			
P-VOP based temporal scalability rect. shape		✓	✓	✓	✓		✓			
P-VOP based temporal scalability arbitrary shape		✓	✓		✓		✓			
B-VOP		✓	✓	✓			✓			
OBMC			✓							

Temporal scalability rectangular shape				√						
Temporal scalability arbitrary shape										
Spatial scalability rectangular shape				√						
Static spites			√							
Interlaced tools			√							
Grayscale Alpha shape coding			√							
4 to 12-bit pixel depth					√					
2D Dynamic Mesh with uniform topology						√	√			
2D Dynamic Mesh with Delaunay topology							√			
Facial Animation Parameters								√		
Scalable wavelet texture (rect., spatial & SNR scal.)						√	√			√
Scalable wavelet texture (spatial scalable)							√		√	√
Scalable wavelet texture (all tools, Shape Adaptive)						√	√			

Πίνακας 2-1. Προφίλ οπτικών αντικειμένων του MPEG-4

Οι στήλες του παραπάνω πίνακα καθορίζουν τα διαφορετικά προφίλ οπτικών αντικειμένων, ενώ οι γραμμές καθορίζουν όλα τα διαφορετικά εργαλεία κωδικοποίησης του προτύπου. Στο μέσο του πίνακα σημειώνονται τα εργαλεία που χρησιμοποιούνται από κάθε προφίλ.

2.2.2 Ήχος

Το MPEG-4 όσον αφορά τον ήχο καλύπτει ένα μεγάλο εύρος εφαρμογών, παρέχοντας εργαλεία αναπαράστασης από φωνή μέχρι πολυκαναλικό φυσικό και συνθετικό ήχο. Η αναπαράσταση συνθετικού ήχου μπορεί να εξαχθεί από κείμενο ή από περιγραφές οργάνων και παραμέτρους κωδικοποίησης για την αναπαράσταση ακουστικών εφέ. Οι αναπαραστάσεις που ορίζει το πρότυπο παρέχουν συμπίεση, αλλά και λειτουργίες όπως

κλιμάκωση και επεξεργασία των εφέ. Συγκεκριμένα τα είδη του ήχου και τα χαρακτηριστικά τους που υποστηρίζονται από το MPEG-4 είναι φωνή σε συχνότητες από 2 ως 24 kbps, κωδικοποιημένη κατάλληλα για επικοινωνίες, συνθετική φωνή με κλιμακωτούς κωδικοποιητές από κείμενο σε φωνή (TTS) από 200 bps ως 1.2 kbps, ήχος χαμηλού εύρους ζώνης (μουσική και φωνή για κανάλια μικρής συχνότητας), ήχος υψηλού εύρους ζώνης (μόνο και πολυκαναλικός ήχος ποιότητας εκπομπής), συνθετικός ήχος με χρήση ειδικής γλώσσας περιγραφής των μουσικών οργάνων, περιορισμένης πολυπλοκότητας συνθετικός ήχος. Η πρόσθετη λειτουργικότητα που ορίζεται για όλα τα παραπάνω περιλαμβάνει τα εξής:

- Έλεγχο της ταχύτητας του ήχου χωρίς να αλλάζει ο τόνος (pitch) κατά τη διάρκεια της αποκωδικοποίησης. Η λειτουργικότητα αυτή μπορεί να φανεί χρήσιμη στην υλοποίηση γρήγορης και αργής κίνησης (fast-forward και slow motion) και στην προσαρμογή του μήκους κάποιου ήχου σε συγκεκριμένο βίντεο.
- Έλεγχο του τόνου (pitch) χωρίς να αλλάζει η χρονική κλίμακα κατά την κωδικοποίηση /αποκωδικοποίηση. Είναι χρήσιμο για αλλαγή της φωνής σε εφαρμογές Karaoke ή για μετατροπή της φωνής. Η τεχνική αυτή εφαρμόζεται μόνο σε παραμετρικές και δομημένες μεθόδους κωδικοποίησης.
- Κλιμάκωση του ρυθμού για μετάδοση ή αποκωδικοποίηση μια ροής ήχου σε χαμηλότερο ρυθμό.
- Κλιμάκωση του εύρους ζώνης ώστε να μπορεί ένα τμήμα του φάσματος συχνότητας του ήχου να απορριφθεί κατά τη μετάδοση ή την αποκωδικοποίηση.
- Κλιμάκωση της πολυπλοκότητας του κωδικοποιητή για να μπορούν κωδικοποιητές διαφορετικής πολυπλοκότητας να παράγουν σωστές ψηφιοσειρές.
- Κλιμάκωση της πολυπλοκότητας του αποκωδικοποιητή για να μπορεί μια ψηφιοσειρά να αποκωδικοποιείται από αποκωδικοποιητές με διαφορετική πολυπλοκότητα.
- Σθεναρότητα σε σφάλματα για να αποφεύγει ο αποκωδικοποιητής την παραμόρφωση από τα λάθη μετάδοσης.
- Ακουστικά εφέ που παρέχουν τη δυνατότητα να γίνεται επεξεργασία αποκωδικοποιημένων ακουστικών σημάτων για λειτουργίες μίξης.

Ο φυσικός ήχος στο MPEG-4 κωδικοποιείται σε ρυθμούς από 2 ως 64 kbits/sec, ενώ όταν η κωδικοποίηση είναι μεταβλητού ρυθμού υποστηρίζονται και συχνότητες κάτω των 2 kbits/sec, για παράδειγμα με μέσο όρο 1.2 kbits/sec. Στους υψηλότερους ρυθμούς χρησιμοποιείται η κωδικοποίηση AAC του MPEG-2. Το πρότυπο ορίζει τη σύνταξη των

κωδικοποιημένων ψηφιοσειρών και τις αντίστοιχες διαδικασίες αποκωδικοποίησης ως ένα σύνολο από εργαλεία που συνενώνονται σε ένα κοινό πλαίσιο.

Η κωδικοποίηση φωνής υποστηρίζεται σε ρυθμούς από 2 ως 24 kbits/sec. Από 2 ως 4 kbits/sec χρησιμοποιείται το εργαλείο HVXC (Harmonic Vector eXcitation Coding) και από 4 ως 24 kbits/sec χρησιμοποιείται το εργαλείο CELP (Code Excited Linear Predictive). Η κωδικοποίηση μεταβλητού ρυθμού σε λιγότερο από 2 kbits/sec παρέχεται από το HVXC. Στην κωδικοποίηση CELP χρησιμοποιούνται δύο διαφορετικές συχνότητες δειγματοληψίας για την υποστήριξη φωνής στενής (narrowband) και ευρείας ζώνης (wideband) αντίστοιχα.

Για γενική κωδικοποίηση του ήχου σε ρυθμούς μεγαλύτερους των 6 kbits/sec εφαρμόζονται οι τεχνικές κωδικοποίησης μετασχηματισμού TwinVQ και AAC, όπου οι συχνότητες δειγματοληψίας των σημάτων αρχίζουν από 8 kHz. Επίσης αρχίζοντας από ένα σήμα χαμηλού ρυθμού και προσθέτοντας βελτιωτικά επίπεδα στο γενικό αποκωδικοποιητή ήχου, μπορεί να αυξηθεί η ποιότητα και το εύρος ζώνης του ήχου. Αντίστροφα με την κλιμάκωση, μια ψηφιοσειρά μπορεί να αποκωδικοποιηθεί ως ψηφιοσειρά χαμηλότερου ρυθμού. Η κλιμάκωση μπορεί να εφαρμοστεί μέσα σε ένα συγκεκριμένο εργαλείο κωδικοποίησης ή και ανάμεσα σε διαφορετικά εργαλεία του MPEG-4, όπως για παράδειγμα το βασικό επίπεδο να είναι CELP και το βελτιωτικό επίπεδο AAC.

Ο συνθετικός ήχος περιγράφεται με καλά ορισμένες δομές. Η συνθετική φωνή παράγεται από κείμενο με τον κατάλληλο αποκωδικοποιητή TTS (Text-To-Speech), ενώ υπάρχουν εργαλεία και για την αναπαράσταση πιο γενικού ήχου. Βασικό πλεονέκτημα του συνθετικού ήχου είναι ότι μπορεί να επιτευχθεί αποδεκτή ποιότητα σε πολύ χαμηλές συχνότητες.

Η συνθετική φωνή κωδικοποιείται σε ρυθμούς από 200 bits/sec ως 1,2 kbits/sec και επιτρέπει είσοδο με τη μορφή απλού κειμένου και κειμένου με προσωδιακές παραμέτρους, όπως το περίγραμμα οξύτητας (pitch contour), η διάρκεια των φωνημάτων, κλπ. Η συνθετική φωνή παρέχει και παραμέτρους για συγχρονισμό της φωνής με αντίστοιχη κίνηση προσώπων, διεθνείς γλώσσες για το κείμενο και διεθνή σύμβολα για τα φωνήματα. Αυτό που παρέχει το πρότυπο, ουσιαστικά είναι η διεπικοινωνία με τους μετατροπείς κειμένου σε φωνή και όχι η υλοποίηση κάποιου συγκεκριμένου μετατροπέα.

Τα εργαλεία συνθετικού ήχου (εκτός φωνής) παίρνουν ως είσοδο κάποια δομημένα δεδομένα και παράγουν ως έξοδο ήχο. Η αποκωδικοποίηση γίνεται με βάση μια ειδική γλώσσα σύνθεσης που ονομάζεται Ορχηστρική Γλώσσα Δομημένου Ήχου (Structured Audio Orchestra Language – SAOL). Αυτή η γλώσσα ορίζει ότι μια ορχήστρα αποτελείται από μουσικά όργανα που δημιουργούν και επεξεργάζονται δεδομένα. Τα όργανα αυτά δεν είναι σταθερά και καθορισμένα αλλά περιγράφονται μέσα στην ψηφιοσειρά που

μεταδίδεται στο τερματικό. Ένα όργανο αναπαρίσταται ως ένα δίκτυο από στοιχειώδεις λειτουργίες επεξεργασίας σήματος, οι οποίες μπορεί να προσομοιώνουν συγκεκριμένους ήχους, όπως ήχους μουσικών οργάνων. Το MPEG-4 ουσιαστικά δεν καθορίζει συγκεκριμένη μέθοδο σύνθεσης, αλλά μια μέθοδο περιγραφής της σύνθεσης. Ο έλεγχος της σύνθεσης γίνεται από «παρτιτούρες» ή «χειρόγραφα» που περιέχονται στην ψηφιοσειρά. Η παρτιτούρα είναι μια χρονική σειρά εντολών που βάζουν σε λειτουργία διάφορα όργανα σε καθορισμένες χρονικές στιγμές. Η περιγραφή της παρτιτούρας είναι σε μια γλώσσα που ονομάζεται SASL (Structured Audio Score Language) και εκτός από πληροφορία για τη δημιουργία νέων ήχων μπορεί να εκφράσει λειτουργίες τροποποίησης των υπάρχοντων ήχων. Για τον έλεγχο της ορχήστρας μπορεί επίσης να χρησιμοποιηθεί και το πρωτόκολλο MIDI.

Για εφαρμογές που δεν έχουν πολύ υψηλές απαιτήσεις ως προς το συνθετικό ήχο, αλλά και για τερματικά που δεν έχουν μεγάλες δυνατότητες, παρέχεται από το πρότυπο μια αναπαράσταση απλών συνθετικών ήχων και απλών φίλτρων. Η μορφή αυτή ονομάζεται «μορφή πινάκων κυματομορφών».

Τα εργαλεία για κωδικοποίηση σε συχνότητες από 6 ως 14 kbits/sec έχουν ελεγχθεί για εφαρμογές AM σε συνεργασία με την ομάδα NADIB (Narrow Band Digital Broadcasting). Έλεγχοι πιστοποίησης έχουν γίνει για τη φωνή στις συχνότητες 2, 4 (HVXC), 6, 8.3, 12 (CELP narrowband) και 18 (CELP wideband) kbits/sec. Τα συμπεράσματα από τα πειράματα έδειξαν ότι με ψηφιακές τεχνικές, όπως η κωδικοποίηση CELP που παρέχει το MPEG-4 και η κλιμάκωση, αυξάνεται η απόδοση σε σχέση με το υπάρχον σύστημα μετάδοσης AM.

2.2.3 Περιγραφή Σκηνών

Το MPEG-4 επικεντρώνεται σε συστήματα που ανταλλάσσουν πληροφορία με τη μορφή αλληλεπιδραστικών οπτικοακουστικών σκηνών. Οι οπτικοακουστικές σκηνές είναι οι πιο σύνθετες οντότητες MPEG-4 περιεχομένου, καθώς και οι πιο σύνθετες οντότητες παρουσίασης και αλληλεπίδρασης με το χρήστη. Μια *οπτικοακουστική σκηνή* του MPEG-4 αποτελείται από:

- Κωδικοποιημένα φυσικά ή συνθετικά οπτικοακουστικά αντικείμενα που ονομάζονται *αντικείμενα μέσων* (media objects).
- Κωδικοποιημένη πληροφορία για τη χωροχρονική θέση και την αλληλεπιδραστική συμπεριφορά των παραπάνω αντικειμένων. Η πληροφορία αυτή ονομάζεται *περιγραφή της σκηνής*.

- Κωδικοποιημένη πληροφορία για τη διαχείριση των ροών πληροφορίας, δηλαδή το συγχρονισμό, την ταυτοποίηση, την περιγραφή και τη συσχέτιση του περιεχομένου των ροών.

Το τερματικό πρέπει να έχει όλη την πληροφορία συνδυασμού των αντικειμένων, η οποία ονομάζεται περιγραφή της σκηνής. Η διαδικασία σύνθεσης της σκηνής τοποθετεί τα αντικείμενα στον ίδιο χώρο αναπαράστασης και το ζωγράφισμα (rendering) τα μετατρέπει στη μορφή που απαιτεί η συσκευή παρουσίασης. Όσον αφορά στη σύνθεση των σκηνών, το πρότυπο περιγράφει μόνο τη σύνταξη και τη σημασιολογία των ροών περιγραφής. Αυτή η αναπαράσταση ονομάζεται BIFS (BInary Format for Scenes). Η περιγραφή κωδικοποιείται ανεξάρτητα από τα αντικείμενα που αποτελούν μια σκηνή. Συνοπτικά η περιγραφή σκηνών παρέχει τις εξής δυνατότητες:

- Ιεραρχική μη στατική οργάνωση του περιεχομένου της σκηνής, όπου οι κόμβοι της ιεραρχίας αντιστοιχούν σε αντικείμενα μέσω και μπορούν να μεταβάλλονται δυναμικά.
- Χωροχρονική τοποθέτηση των αντικειμένων. Κάθε αντικείμενο έχει το δικό του τοπικό σύστημα συντεταγμένων και για να τοποθετηθεί στη σκηνή μετασχηματίζεται στο σύστημα συντεταγμένων της σκηνής.
- Επιλογή των χαρακτηριστικών των αντικειμένων. Οι κόμβοι μιας σκηνής έχουν χαρακτηριστικά τα οποία ελέγχουν τη συμπεριφορά τους. Το περιεχόμενο των αντικειμένων κωδικοποιείται σε διαφορετική ροή από τα χαρακτηριστικά.
- Αλληλεπίδραση με το χρήστη. Οι ενέργειες του χρήστη προκαλούν μεταβολές στα χαρακτηριστικά των κόμβων της σκηνής.

Γενικά, οι σκηνές του MPEG-4 έχουν τη δομή των σκηνών της γλώσσας VRML η οποία χρησιμοποιείται για την αναπαράσταση γραφικών στον Παγκόσμιο Ιστό. Η διαφορά των σκηνών του MPEG-4 με τις σκηνές της VRML είναι στην αναπαράσταση των ψηφιοσειρών, όπου το MPEG-4 χρησιμοποιεί δυαδική κωδικοποίηση, στην μετάδοση των σκηνών, ενώ η VRML δεν υποστηρίζει ροές πραγματικού χρόνου, στο είδος της συμπίεσης ανάμεσα στις δύο αναπαραστάσεις και τέλος στο ότι το MPEG-4 επιτρέπει τη δυναμική ενημέρωση των σκηνών.

Το BIFS είναι η δυαδική μορφή αναπαράστασης ενός προκαθορισμένου συνόλου αντικειμένων σκηνής καθώς και της συμπεριφοράς και των χωροχρονικών ιδιοτήτων τους. Το BIFS περιέχει τρόπο περιγραφής για τα χαρακτηριστικά των αντικειμένων μέσω (οπτικοακουστικές ιδιότητες), τη δομή της σκηνής, την προκαθορισμένη χωροχρονική συμπεριφορά των αντικειμένων και τη συμπεριφορά των αντικειμένων που προκαλείται από αλληλεπίδραση με το χρήστη.

Η γενική σύνταξη που παρέχει το BIFS είναι για την περιγραφή οπτικών και ακουστικών στοιχείων σε ένα τρισδιάστατο χώρο και στηρίζεται στη γλώσσα VRML. Η περιγραφή είναι ένα σύνολο κόμβων και κάθε αντικείμενο μέσω αναπαρίσταται από έναν ή περισσότερους κόμβους που ομαδοποιούνται. Κάθε κόμβος έχει ένα πεδίο που δηλώνει τον τύπο του. Τα αντικείμενα μέσω μπορεί να περιγράφονται πλήρως από το BIFS, ή να απαιτεί δεδομένα από μια στοιχειώδη ροή οπτικοακουστικών αντικειμένων. Οι κόμβοι αποτελούνται από πεδία και γεγονότα τα οποία περιγράφονται από τη VRML.

Οι μονάδες μέτρησης στην περιγραφή των σκηνών είναι μέτρα για την απόσταση, RGB για τα χρώματα, δευτερόλεπτα για το χρόνο και ακτίνια για τις γωνίες. Το δυσδιάστατο σύστημα συντεταγμένων έχει το κέντρο του στο κέντρο της περιοχής του rendering και επεκτείνεται κατά 1 μέτρο προς κάθε κατεύθυνση του άξονα x. Στον άξονα y το μήκος καθορίζεται από το aspect ratio έτσι ώστε η μονάδα να είναι ίδια και προς τις δύο κατευθύνσεις. Το τρισδιάστατο σύστημα συντεταγμένων είναι αυτό που περιγράφεται στο πρότυπο της VRML.

Τα δεδομένα περιγραφής σκηνών μεταδίδονται σε ειδικές στοιχειώδεις ροές, οι οποίες αποτελούνται από μονάδες προσπέλασης, έχουν μια χρονική βάση και περιέχουν χρονοσφραγίδες σύνθεσης, σε αναλογία με τις στοιχειώδεις ροές δεδομένων.

Επιπλέον της VRML το BIFS ορίζει δυσδιάστατες σκηνές, συνένωση δυσδιάστατης και τρισδιάστατης αναπαράστασης, περισσότερες δυνατότητες για τον ήχο, μοντέλο χρονισμού, πρωτόκολλο δυναμικής ενημέρωσης της σκηνής, πρωτόκολλο δυναμικής κίνησης της σκηνής (animation) και σύνταξη για δυαδική κωδικοποίηση της περιγραφής.

Οι τύποι των νέων χρονικά εξαρτώμενων κόμβων που προσθέτει το MPEG-4 στην περιγραφή των σκηνών είναι Ροές Κίνησης (Animation Streams), Πηγή Ήχου (Audio Source), Αντικείμενο Δυσδιάστατου Βίντεο (VideoObject2D) και Χρονικός Αισθητήρας Μέσων (MediaTimeSensor). Για παράδειγμα, όσον αφορά στον ήχο, η περιγραφή και σύνθεση των σκηνών έχει στόχο να καλύψει την αναπαράσταση του ακουστικού χώρου ιδεατών περιβαλλόντων, αλλά και παραδοσιακές εφαρμογές.

Μερικοί τύποι κόμβων φαίνονται στον παρακάτω πίνακα:

Κοινοί κόμβοι (2Δ,3Δ)	Κοινοί με VRML	2Δ κόμβοι	3Δ κόμβοι	Μικτοί 2Δ/3Δ
Ροή animation	Audio clip	2Δ φόντο	Σημείο ακρόασης	2Δ επίπεδο
Καθυστέρηση ήχου	Χρώμα	Κύκλος	Πρόσωπο	3Δ επίπεδο
Μίξη ήχων	Στυλ γραμματοσειράς	2Δ συντεταγμένα	Παράμετρος κίνησης προσώπου	Σύνθετη 2Δ υφή
Πηγή ήχου	Υφή εικόνας	2Δ καμπύλη	Φώνημα	Σύνθετη 3Δ υφή
Παράμ. κβαντοποίησης	Υφή ταινίας	Φόρμα διάταξης	Έκφραση	

Δυνατότητες τερματικού	Σχήμα	2D ομάδα	Παράμετρος ορισμού προσώπου	
	Κείμενο	2D εικόνα	Πλέγμα ορισμού προσώπου	
	Μετ/σμός κειμένου	Διάταξη	Φόντο	
		2D ήχος	Κώνος	
		2D μετ/σμός	Συντεταγμένη	
		2D αντικείμενο βίντεο	Κύλινδρος	
			Φως	
			Υλικό	
			Ήχος	
			Σημείο θέασης	

Πίνακας 2-2. Τύποι κόμβων σκηνής

Μια άλλη κατηγοριοποίηση των κόμβων του MPEG-4 μπορεί να γίνει με βάση τη χρήση τους όπως φαίνεται στον παρακάτω πίνακα. Οι βασικές κατηγορίες κόμβων είναι κόμβοι ομαδοποίησης, παρεμβολής για εφαρμογή κίνησης, αισθητήρες ενεργειών του χρήστη, γεωμετρικοί κόμβοι, κόμβοι-παιδιά άλλων κόμβων, κόμβοι μέσων, κόμβοι κίνησης προσώπων και σωμάτων, κλπ.

Κατηγορία κόμβων	Λειτουργικότητα	Κόμβοι
Κόμβοι Ομαδοποίησης	Περιέχουν ένα σύνολο από κόμβους-παιδιά.	Ομάδα, Φόρμα, 2D βαθμίδα, 2D μετασχηματισμός, 3D βαθμίδα, κόμβοι συνδυασμού ήχων (μίξη ήχων, καθυστέρηση ήχων), κλπ.
Κόμβοι Παρεμβολής	Εφαρμόζουν γραμμική παρεμβολή για την κίνηση καρτέ	Παρεμβολή χρώματος, 2D ή 3D θέσης, 2D ή 3D συντεταγμένης, προσανατολισμού, κλπ.
Κόμβοι Αισθητήρες	Ανιχνεύουν γεγονότα από το περιβάλλον και πυροδοτούν άλλα γεγονότα	Αισθητήρας ποντικίου, χρόνου, ορατότητας, κλπ.
Κόμβοι Γεωμετρίας	Αναπαριστούν γεωμετρικά αντικείμενα	Κύκλος, 2D καμπύλη, τετράγωνο, κείμενο, κώνος, κύλινδρος, σφαίρα, κλπ.
Προσαρτήσιμοι κόμβοι	Μόνος ένας κόμβος από κάθε είδος μπορεί να είναι ενεργός ανά πάσα στιγμή	2D και 3D φόντο, σημείο ακρόασης, σημείο θέασης, κλπ.
Κόμβοι-παιδιά	Παιδιά των κόμβων ομαδοποίησης. Περιέχει τις κατηγορίες: κόμβοι ομαδοποίησης, αισθητήρες, παρεμβολής και προσαρτήσιμοι	Επιπλέον κόμβοι: Ροή κίνησης, πρόσωπο, σχήμα, 2D ήχος, πηγή φωτός, ήχος.
Κόμβοι αντικειμένων μέσων	Επιτρέπουν συσχέτιση πολυμέσων με τη σκηνή	Ροή κίνησης, κομμάτι ήχου, πηγή ήχου, φόντο, υφή εικόνας, υφή ταινίας.
Πρόσωπα και Σώματα	Κόμβοι σχετικοί με κίνηση προσώπων και σωμάτων	Πρόσωπο, πλέγμα ορισμού προσώπου, μετασχηματισμοί προσώπων, παράμετροι ορισμού προσώπων, φωνήματα.
Διάφορα χαρακτηριστικά	Χαρακτηριστικά των κόμβων-παιδιών που αναπαρίστανται με συγκεκριμένους κόμβους	Εμφάνιση, χρώμα, στυλ γραμματοσειράς, 2D και 3D συντεταγμένη, 2D και 3D υλικό, συντεταγμένη υψός, μετασχ/μός υψός.

Κόμβοι-κορυφή της σκηνής		Ομάδα, 2Δ βαθμίδα, 3Δ βαθμίδα, διατεταγμένη ομάδα.
Κόμβοι διανομής	Κόμβοι για μεταφορά της σκηνής	Εντολές σκηνής και κίνηση σκηνής (BIFS-Command, BIFS-Anim)

Πίνακας 2-3. Κατηγοριοποίηση κόμβων σκηνής ανάλογα με τη λειτουργία τους

Το πρότυπο MPEG-4 ορίζει τη σύνταξη της δυαδικής κωδικοποίησης, καθώς και την διαδικασία αποκωδικοποίησης κάθε είδους κόμβου.

2.2.4 Περιγραφείς Αντικειμένων

Η περιγραφή και η σύνθεση των σκηνών είναι η βασικότερη λειτουργικότητα των Συστημάτων του MPEG-4. Οι σκηνές δεν συσχετίζονται απευθείας με τις στοιχειώδεις ροές οι οποίες μεταφέρουν τα δεδομένα, αλλά χρησιμοποιούν ένα ενδιάμεσο επίπεδο, τους *Περιγραφείς Αντικειμένων* (Object Descriptors). Οι περιγραφείς αυτοί διαχωρίζουν την δομή της σκηνής από τα ίδια τα δεδομένα, αλλά και από τα επίπεδα μεταφοράς και παρέχουν μια υψηλού επιπέδου αντιστοίχιση ανάμεσά τους.

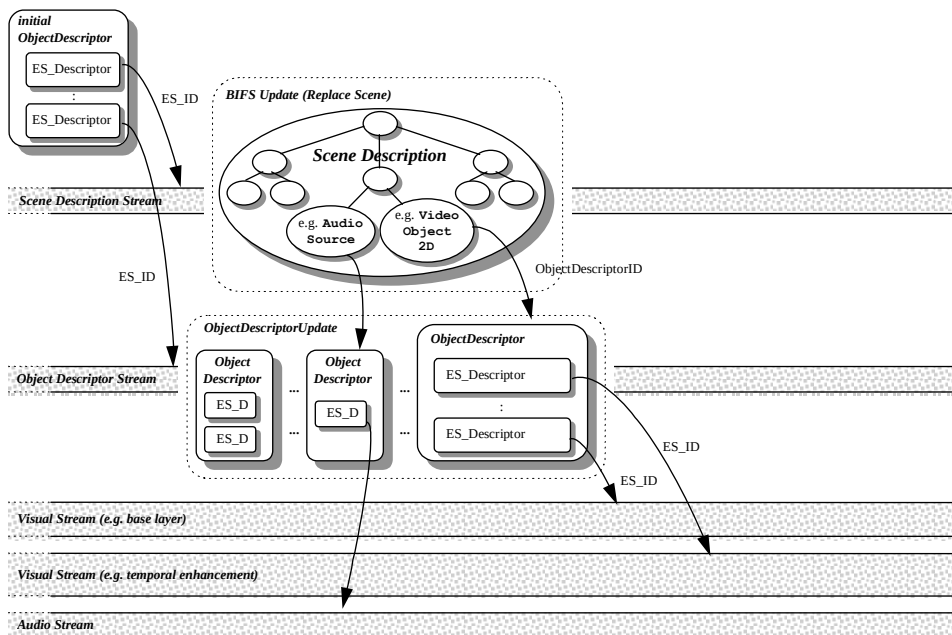
Σε κάθε αντικείμενο μέσων, το οποίο συσχετίζεται με μία ή περισσότερες στοιχειώδεις ροές, αντιστοιχεί ένας περιγραφέας. Οι περιγραφείς έχουν μια *μοναδική ταυτότητα* και αποτελούνται από ένα σύνολο *Περιγραφέντων Στοιχειωδών Ροών* (Elementary Stream Descriptors). Οι περιγραφείς στοιχειωδών ροών περιέχουν πληροφορία απαραίτητη για την αρχικοποίηση της αναπαραγωγής της στοιχειώδους ροής που περιέχει τα καθεαυτό δεδομένα, καθώς και για τα πνευματικά τους δικαιώματα.

Οι περιγραφείς αντικειμένων μεταφέρονται σε εδικές ροές, τις *ροές περιγραφέντων* (ροές περιγραφέντων αντικειμένων ή περιγραφέα σκηνής), οι οποίες επιτρέπουν λειτουργίες διαχείρισης των ίδιων των περιγραφέντων κατά τη διάρκεια της συνόδου. Οι λειτουργίες αυτές είναι ενημέρωση και διαγραφή και μπορούν να εφαρμοστούν και στα δομικά στοιχεία των περιγραφέντων αντικειμένων (δηλ. στους περιγραφείς στοιχειωδών ροών). Οι ενημερώσεις και διαγραφές είναι χρήσιμες για να παρακολουθούν τις αλλαγές στο ίδιο το περιεχόμενο. Οι ροές των περιγραφέντων δηλαδή αποτελούνται από μηνύματα, τα οποία περιέχουν τους ίδιους τους περιγραφείς και τις ενέργειες ενημέρωσής τους. Οι ενέργειες αυτές συνοδεύονται από μια χρονοσφραγίδα που δηλώνει τη χρονική στιγμή που πρέπει να συμβεί η ενέργεια. Οι μονάδες προσπέλασης μέσα στις ροές των περιγραφέντων είναι ομάδες μηνυμάτων που έχουν την ίδια χρονοσφραγίδα.

Επιπλέον χαρακτηριστικά των περιγραφέντων αντικειμένων που υλοποιούν τη συσχέτιση μεταξύ τους είναι τα εξής:

- Ένας περιγραφέας αντικειμένου (Object Descriptor) συσχετίζει στοιχειώδεις ροές με συμβατό τύπο (οπτικές, ήχου ή περιγραφή σκηνής) και δεν επιτρέπει την ανάμιξή τους. Αυτό σημαίνει ότι οι στοιχειώδεις ροές είναι πολλαπλές αναπαραστάσεις του ίδιου αντικειμένου. Οι στοιχειώδεις ροές διαφορετικού τύπου που μπορεί να αναφέρονται σε έναν περιγραφέα επιτρέπεται να είναι μόνο ροές OCI ή ροές Clock Reference.
- Οι στοιχειώδεις ροές που αναφέρονται στον ίδιο περιγραφέα μπορεί να εξαρτώνται η μία από την άλλη. Για παράδειγμα, τα επίπεδα βελτίωσης εξαρτώνται από το βασικό επίπεδο. Επίσης οι ροές εισόδου (επιστροφής) προς τον εξυπηρετητή εξαρτώνται πάντα από μία ροή εξόδου. Άλλο είδος εξάρτησης είναι η περιγραφή σκηνής να εξαρτάται από τη ροή που περιέχει τους περιγραφείς των αντικειμένων.

Η προσπέλαση του περιεχομένου σε κάθε σύνοδο ξεκινάει από τον περιγραφέα αντικειμένου που αντιστοιχεί στη ροή που περιέχει την περιγραφή της σκηνής. Ο περιγραφέας αυτός μεταδίδεται στο τερματικό κατά την αρχικοποίηση της συνόδου. Η όλη διαδικασία φαίνεται στο "Σχήμα 2-2. Περιγραφείς Αντικειμένων".



Σχήμα 2-2. Περιγραφείς Αντικειμένων

Η προσπέλαση της πληροφορίας ξεκινάει από ένα αρχικό περιγραφέα αντικειμένων, ο οποίος οδηγεί σε μια ροή που περιέχει την περιγραφή της σκηνής. Η τελευταία περιέχει τις ταυτότητες άλλων ροών που μπορεί να περιέχουν περιγραφείς ή στοιχειώδη αντικείμενα (βίντεο, ήχο, κλπ).

Όπως φαίνεται στο σχήμα, ο αρχικός περιγραφέας έχει έναν δείκτη (ταυτότητα ροής) σε μια ροή που περιέχει την περιγραφή της σκηνής και σε μια ροή που περιέχει τους περιγραφείς των αντικειμένων. Οι περιγραφείς των αντικειμένων με τη σειρά τους περιέχουν δείκτες στις στοιχειώδεις ροές που περιέχουν τα δεδομένα των αντικειμένων. Στο συγκεκριμένο παράδειγμα, μέσα στις ροές περιγραφής φαίνονται λειτουργίες ενημέρωσης, της περιγραφής της σκηνής και του περιγραφέα κάποιου αντικειμένου.

Για κάθε είδος περιγραφέα, το πρότυπο ορίζει τη σύνταξη του μέσα σε μια ροή. Στον παρακάτω πίνακα φαίνεται η σύνταξη και η σημασιολογία του περιγραφέα που μεταδίδεται στην αρχή της συνόδου, του περιγραφέα αντικειμένων του περιγραφέα στοιχειώδους ροής και του μηνύματος διαγραφής περιγραφέα στοιχειώδους ροής. Με παρόμοια σύνταξη στη γλώσσα συντακτικής περιγραφής, ορίζονται και τα υπόλοιπα είδη περιγραφέων.

Περιγραφέας Αντικειμένου
<pre>aligned (8) class ObjectDescriptor : bit(8) tag=ObjectDescrTag { bit(8) length; bit(10) ObjectDescriptorID; bit(1) URL_Flag; const bit(5) reserved = 0b1111.1; if (URL_Flag) { bit(8) URLString [length-2]; } ExtensionDescriptor extDescr[0 .. 255]; if (!URL_Flag) { OCI_Descriptor ociDescr[0..255]; ES_Descriptor esDescr[1..30]; } }</pre>
Περιγραφέας Στοιχειώδους Ροής
<pre>aligned (8) class ES_Descriptor : bit(8) tag=ES_DescrTag { bit(8) length; bit(16) ES_ID; bit(1) streamDependenceFlag; bit(1) URL_Flag; const bit(1) reserved=1; bit(5) streamPriority; if (streamDependencyFlag) bit(16) dependsOn_ES_ID; if (URL_Flag) bit(8) URLString[length-3-(streamDependenceFlag*2)]; ExtensionDescriptor extDescr[0 .. 255]; LanguageDescriptor langDescr[0 .. 1]; DecoderConfigDescriptor decConfigDescr; SLConfigDescriptor slConfigDescr; IPI_DescPointer ipiPtr[0 .. 1]; IP_IdentificationDataSet ipIDS[0 .. 1]; QoS_Descriptor qosDescr[0 .. 1]; }</pre>
Διαγραφή Περιγραφέα Στοιχειώδους Ροής
<pre>Aligned (8) class ES_DescriptorRemove : bit(8) tag= ES_DescrRemoveTag { bit(8) length; bit(10) objectDescriptorID; bit(5) streamCount; const bit(1) reserved=1; bit(16) ES_ID[streamCount];</pre>

Πίνακας 2-4. Παραδείγματα Σύνταξης Μηνυμάτων Περιγραφών

2.2.5 Γλώσσα Συντακτικής Περιγραφής

Το τμήμα αυτό περιγράφει τη δομή της ψηφιοσειράς του MPEG-4 μέσω μιας γλώσσας συντακτικής περιγραφής. Ο στόχος του προτύπου στο συγκεκριμένο πεδίο, ήταν η σύνταξη να είναι ανεξάρτητη από την αποκωδικοποίηση και το ζωγράφισμα της πληροφορίας και να περιγράφει τι περιέχεται μέσα στην ψηφιοσειρά, αλλά όχι πως αυτό θα εξαχθεί. Στο πρότυπο MPEG-1 και κατόπιν και στο MPEG-2, η δομή της ψηφιοσειράς καθοριζόταν με μια σύνταξη παρόμοια με αυτήν της γλώσσας C (pseudo-C), με βάση πίνακες τιμών και κείμενο με επεξηγήσεις. Αντίθετα, το MPEG-4 ορίζει μια γλώσσα που είναι επέκταση της C++ και χρησιμοποιείται όχι μόνο για την περιγραφή της δομής των ψηφιοσειρών, αλλά και για την περιγραφή των σκηνών.

Η γλώσσα αυτή ξεκίνησε με το όνομα *MPEG-4 Syntactic Description Language* (MSDL) με σκοπό να χρησιμοποιηθεί για την περιγραφή της μορφής της ψηφιοσειράς και να επιτρέψει το συνδυασμό των εργαλείων του MPEG-4 ως πλήρεις αλγορίθμους. Αρχικά ορίστηκε μια συγκεκριμένη γλώσσα "συντακτικής περιγραφής", η οποία υπέστη τροποποιήσεις και στη συνέχεια ονομάστηκε *Flavor* (Formal Language for Audio-Visual Object Representation) από τους εμπνευστές της, του πανεπιστημίου Columbia. Ο όρος MPEG-4 Syntactic Description Languages χρησιμοποιήθηκε έκτοτε για αναφορά σε όλα API των εργαλείων του MPEG-4, στη σύνθεση σκηνών, αλλά και στη γλώσσα που τα συνδέει όλα. Το συντακτικό μέρος της MSDL ονομάστηκε και MSDL-S (MSDL-Syntax). Τώρα χρησιμοποιείται ο όρος *Syntactic Description Language* (SDL) και αποτελεί το τμήμα 6 του προτύπου.

Η γενικότερη ιδέα του καθορισμού γλώσσας για τη σύνταξη της ψηφιακής ροής των δεδομένων είναι να είναι προγραμματίσιμοι οι πελάτες. Για παράδειγμα η γλώσσα Java επιτρέπει προγραμματίσιμους πελάτες, με την έννοια ότι ο κώδικας έρχεται και εκτελείται στον πελάτη. Η διαφορά όμως της Java με το περιβάλλον του MPEG-4 είναι ότι στο τελευταίο υπάρχει μια ροή δεδομένων που έρχεται στον πελάτη σε πραγματικό χρόνο.

Η γλώσσα που ορίζει το MPEG-4 και ονομάζεται SDL (Syntactic Description Language), στηρίζεται σε ένα αντικειμενοστρεφές πλαίσιο και χρησιμοποιεί το σύστημα τύπων της Java και της C++, το οποίο επεκτείνει για να οριστούν ποσότητες που προέρχονται από ψηφιοσειρές και για να οριστεί πως αυτές θα υποστούν συντακτική ανάλυση (parsing). Ο γενικός κανόνας που χρησιμοποιείται είναι μια μεταβλητή να αναλυθεί πριν προσπελαθεί. Αυτό ονομάζεται δηλωτική συντακτική ανάλυση. Στη δήλωση μιας κλάσης μπαίνει και η

πληροφορία για την ανάλυση. Αυτό σημαίνει ότι σε κάθε τερματικό πρέπει να υπάρχει ένας συντακτικός αναλυτής.

Οι βασικοί τύποι δεδομένων που ορίζει η γλώσσα συντακτικής περιγραφής του MPEG-4 είναι:

- Δυαδικά πεδία σταθερού μήκους που ονομάζονται και Κώδικες Σταθερού Μήκους (Fixed Length Codes).
- Δυαδικά πεδία μεταβλητού μήκους ή Παραμετρικοί Κώδικες Σταθερού Μήκους (parametric FLCs), των οποίων το μέγεθος καθορίζεται από την τιμή κάποιας άλλης παραμέτρου μέσα στην ψηφιοσειρά.
- Έμμεσα δυαδικά πεδία σταθερού μήκους, των οποίων η τιμή υπάρχει σε κάποιο πίνακα ή άλλη μεταβλητή.
- Έμμεσα δυαδικά πεδία μεταβλητού μήκους, δηλ. Κώδικες Huffman.

Εκτός από τους στοιχειώδεις τύπους ορίζονται και σύνθετοι τύποι, οι οποίοι ονομάζονται κλάσεις. Οι κλάσεις υποστηρίζουν κληρονομικότητα και ένα νέο προαιρετικό πεδίο που ονομάζεται `object_id` και αν οριστεί αποτελεί το διαχωρισμό ανάμεσα στα αντικείμενα μιας κλάσης και των υποκλάσεών της. Επίσης υποστηρίζεται και η έννοια των αφηρημένων κλάσεων. Η γλώσσα επίσης υποστηρίζει πίνακες, αριθμητικές και λογικές εκφράσεις, ροές ελέγχου ίδιες με της C++ (if, switch, for, while) και τον ενσωματωμένο τελετή `lengthOf()`. Η γλώσσα δεν ορίζει δομές για μεθόδους και συναρτήσεις. Στον Πίνακα 1 οι περιγραφείς αντικειμένων έχουν οριστεί με τη γλώσσα SDL.

Η SDL είναι τμήμα της MSDL. Η MSDL παρέχει το συνολικό περιβάλλον προγραμματισμού, δηλαδή ορίζει τα API για αποκωδικοποίηση και σύνθεση και περιέχει τα προγραμματιστικά στοιχεία για την υλοποίησή τους. Η SDL είναι ανεξάρτητο τμήμα της MSDL, δηλαδή δεν επηρεάζει τις δομές προγραμματισμού και αντίστροφα. Αυτό που είναι κοινό ανάμεσα στις δύο γλώσσες, είναι οι βασικές αρχές για τον ορισμό δομών δεδομένων και ιεραρχιών. Μέχρι τώρα στα πλαίσια του MPEG-4, έχουν εκφραστεί και τα Προφίλ του MPEG-2 στη γλώσσα SDL.

2.2.6 Εφαρμογές του MPEG-4

Όσον αφορά στις εφαρμογές του προτύπου, έχει επιλεγεί ένα σύνολο κατηγοριών, οι οποίες θα υποστηριχθούν λαμβάνοντας υπόψη κάποιες από τις απαιτήσεις τους στο σχεδιασμό των εργαλείων. Οι εφαρμογές που στοχεύει να διευκολύνει και να υποστηρίξει το MPEG-4, όπως παρουσιάζονται στο [10], ανήκουν στις παρακάτω κατηγορίες:

Εκπομπή (ευρεία μετάδοση)

Η εκπομπή αφορά στην παροχή υπηρεσιών σε ένα μονής κατεύθυνσης, σταθερό κανάλι πραγματικού χρόνου. Ο στόχος του MPEG-4 είναι να παρέχει στις εφαρμογές εκπομπής υψηλότερο εύρος ζώνης μέσω καλύτερης συμπίεσης αλλά και καλύτερη ποιότητα υπηρεσιών. Άλλες λειτουργίες που είναι επιθυμητές, είναι η προσπέλαση υπό συνθήκη, η συλλογή στατιστικών και η χρέωση. Οι απαιτήσεις όσον αφορά στη μετάδοση, είναι ενιαία αντιμετώπιση των διαφόρων τύπων καναλιών (π.χ. δορυφορικά, καλωδιακά, κλπ), πολυπλεξία πολλών προγραμμάτων μεταβλητού ρυθμού στο ίδιο κανάλι και σθεναρότητα σε λάθη.

Το περιεχόμενο των προγραμμάτων εκπομπής είναι οπτικοακουστικό ή σκέτο ακουστικό και μπορεί να απαιτεί κωδικοποίηση σε πραγματικό χρόνο αν πρόκειται για ζωντανή μετάδοση. Μαζί με το βασικό υλικό του προγράμματος είναι πολύ συχνά απαραίτητο να μεταδίδονται και άλλα δεδομένα, όπως κείμενο, λογότυπα και γραφικά.

Ανάλογα με την εφαρμογή εκπομπής, μπορεί να απαιτούνται υπηρεσίες όπως τυχαία προσπέλαση μέσα στο περιεχόμενο, υψηλή ποιότητα βίντεο, πεπλεγμένη ή προοδευτική σάρωση, λειτουργικότητα με βάση αντικείμενα, επιστρώματα με κείμενο και γραφικά, χρονικός και χωρικός συγχρονισμός αντικειμένων, ευελιξία και κλιμάκωση στην ποιότητα των αντικειμένων, συμβατότητα με το MPEG-2 και αλληλεπίδραση με το χρήστη.

Συνεργατική νοερή απεικόνιση σκηνών

Η συνεργατική νοερή απεικόνιση σκηνών (Collaborative Scene Visualization) περιλαμβάνει εφαρμογές συνεργασίας στηριζόμενης από τον υπολογιστή, στις οποίες οι ομάδες ατόμων που συνεργάζονται χρησιμοποιούν εργαλεία νοερής απεικόνισης για να εργαστούν σε έναν κοινό οπτικό χώρο. Η τάση στις συγκεκριμένες εφαρμογές είναι η παροχή Αυξημένης Πραγματικότητας (Augmented Reality), στην οποία συνυπάρχουν πραγματικά και συνθετικά αντικείμενα, όπως για παράδειγμα ένα πλάνο βίντεο από πραγματικό χώρο και αντικείμενα μέσα σε αυτόν που έχουν δημιουργηθεί με τον υπολογιστή.

Το MPEG-4 μπορεί να προσφέρει στις εφαρμογές αυτής της κατηγορίας υψηλή συμπίεση, κατάλληλη στάθμιση ανάμεσα σε ποιότητα και σε επιδόσεις (κλιμάκωση), αλληλεπίδραση με αντικείμενα που έχουν κωδικοποιηθεί ατομικά, σύνθεση αντικειμένων και αναπαράσταση προσώπων.

Αποθήκευση και ανάκτηση βάσει περιεχομένου

Ο όρος ανάκτηση βάσει περιεχομένου για το MPEG—4 έχει την έννοια της προσπέλασης βίντεο και ήχου με βάση χαρακτηριστικά που σχετίζονται με το περιεχόμενο του βίντεο. Αυτή η κατηγορία εφαρμογών είναι κατάλληλη για την οργάνωση βιβλιοθηκών με μεγάλες

συλλογές υλικού. Το πρότυπο έχει ως στόχο να διευκολύνει την τμηματοποίηση που συνήθως γίνεται στο υλικό και να παρέχει γρήγορη προσπέλαση στα τμήματά του, καθώς και να διευκολύνει την απάντηση σε ερωτήσεις αναζήτησης χωρίς να είναι απαραίτητη η αποκωδικοποίηση του υλικού.

Ψηφιακή εκπομπή ΑΜ

Ο στόχος της συγκεκριμένης εφαρμογής είναι η βελτίωση της ραδιοφωνικής εκπομπής ΑΜ με την προσθήκη ψηφιακών προγραμμάτων, συμβατών με την υπάρχουσα υποδομή, αλλά με πλεονεκτήματα όπως καλύτερη ποιότητα ήχου, δεδομένα συσχετιζόμενα με κάθε πρόγραμμα, εξωτερικές υπηρεσίες και πιο φιλική διαχείριση από το δέκτη. Η εφαρμογή απαιτεί κωδικοποίηση μονοφωνικού ή στερεοφωνικού ήχου, μίξη της πληροφορίας με άλλα δεδομένα και πολυπλεξία ενός ή περισσότερων προγραμμάτων.

Set Top Box ψηφιακής τηλεόρασης

Η ψηφιακή τηλεόραση έχει στόχο να παρέχει στους χρήστες σύνδεση των προγραμμάτων τηλεόρασης με πληροφορία στο Ίντερνετ, ηλεκτρονικό εμπόριο, αλληλεπιδραστικά παιχνίδια και βίντεο κατ' απαίτηση. Το MPEG-4 από τη φύση του είναι κατάλληλο για τη συγκεκριμένη εφαρμογή εφόσον παρέχει έλεγχο των ροών και της παρουσίας τους και σύνθεση και έλεγχο γραφικών. Επιπλέον σημαντικές λειτουργίες, όπως κανάλι επιστροφής, αλληλεπίδραση με το χρήστη, προσπέλαση υπό συνθήκη και προφίλ, μπορούν να υλοποιηθούν με τη φόρτωση προγραμμάτων που θα εκτελούνται στη συσκευή.

DVD

Τα DVD είναι κατάλληλα για εφαρμογές όπως αλληλεπιδραστικές ταινίες, οδηγούς, αυτομάθηση και παιχνίδια. Παραδείγματα σεναρίων αλληλεπιδραστικών ταινιών είναι η επιλογή της πλοκής, ή η απόρριψη σκηνών ακατάλληλων για τα παιδιά. Από την πλευρά του MPEG-4, αυτό σημαίνει αλλαγή της σύνθεσης των αντικειμένων κατά την αποκωδικοποίηση. Συνοπτικά η λειτουργικότητα που είναι χρήσιμη στις εφαρμογές DVD είναι αλληλεπίδραση με τη σύνθεση, συγχρονισμός αντικειμένων, αποδοτική κωδικοποίηση, τυχαία προσπέλαση, κλιμάκωση, βοηθητικά δεδομένα, κλπ.

Infotainment

Η κατηγορία αυτή περιλαμβάνει το συνδυασμό πληροφόρησης για συγκεκριμένα θέματα μαζί με διασκέδαση και απαιτεί την αλληλεπίδραση για τη ζήτηση επιπλέον πληροφορίας και την αλλαγή της σύνθεσης των αντικειμένων. Το MPEG-4 είναι ιδανικό για αυτού του είδους την εφαρμογή, εφόσον: α) υποστηρίζει πληθώρα τύπων πολυμέσων (όπως συνθετικούς), οι οποίοι μπορούν να συνδυαστούν σε ένα σενάριο παρουσίασης, β)

παρέχει συνοχή στη διαχείριση των αντικειμένων και την αλληλεπίδραση μεταξύ τους και γ) παρέχει πολύ καλή ποιότητα υπηρεσιών.

Κινητά πολυμέσα

Οι εφαρμογές κινητών πολυμέσων χαρακτηρίζονται από περιορισμούς στις δυνατότητες της ασύρματης επικοινωνίας και από την ποικιλία κινητών συσκευών και ασύρματων δικτύων. Το MPEG-4 υποστηρίζει εφαρμογές πολύ χαμηλού ρυθμού δεδομένων, ευελιξία στην πολυπλοκότητα αποκωδικοποίησης, αλληλεπίδραση με αντικείμενα και μείωση του εύρους ζώνης με την κωδικοποίηση παραμέτρων για τα πρόσωπα.

Επικοινωνίες πραγματικού χρόνου

Οι επικοινωνίες πραγματικού χρόνου περιλαμβάνουν είτε αλληλεπιδραστικές εφαρμογές επικοινωνίας διπλής κατεύθυνσης (π.χ. βιντεοτηλέφωνο), είτε εφαρμογές μονής κατεύθυνσης με αυστηρούς περιορισμούς καθυστέρησης. Κύρια χαρακτηριστικά αυτών των εφαρμογών είναι επίσης ο συγχρονισμός και η ποικιλία των μέσων μετάδοσης. Το MPEG-4 μπορεί να παρέχει σ' αυτές τις εφαρμογές καλή ποιότητα κωδικοποίησης, σθεναρότητα σε σφάλματα, συγχρονισμό, μικρή καθυστέρηση, χαμηλό ρυθμό δεδομένων και μικρή πολυπλοκότητα αποκωδικοποίησης.

Άλλες εφαρμογές του MPEG-4 είναι η Ροή βίντεο στο Ίντερνετ ή σε Ίντρανετ, η Παραγωγή στούντιο ή τηλεόρασης και η Εποπτεία (surveillance).

2.3 Η γλώσσα XML

Η XML (Extensible Markup Language) είναι μεταγλώσσα ανταλλαγής δεδομένων στον Παγκόσμιο Ιστό. Η XML είναι γραμμένη σε SGML, δηλαδή είναι υποσύνολο της SGML και στόχος της είναι να κάνει την SGML εύχρηστη στον Παγκόσμιο Ιστό. Σε σχέση με την HTML, η XML υποστηρίζει επεκτάσιμους τύπους εγγράφων, πιο πολύπλοκη δομή στα έγγραφα και μη υποχρεωτική γραμματική. Επίσης, εφόσον η XML είναι μεταγλώσσα χρησιμοποιείται για τον ορισμό άλλων γλωσσών και οι εφαρμογές της συνίστανται στην κατασκευή ενός *τύπου εγγράφων* με βάση τη γραμματική της XML.

Τα έγγραφα της XML αποτελούνται από μία σειρά *στοιχείων* (elements). Κάθε στοιχείο μπορεί με τη σειρά του να αποτελείται από άλλα στοιχεία, από χαρακτηριστικά (attributes) ή από κείμενο. Τα στοιχεία οριοθετούνται μέσα στο έγγραφο με ετικέτες (tags) αρχής και τέλους. Ο *Ορισμός Τύπου Εγγράφου* (DTD ή Document Type Definition) αποτελείται από *δηλώσεις* (markup declarations) για το τι μπορεί να περιέχει ένα έγγραφο. Οι δηλώσεις αυτές αποτελούνται από στοιχεία, *λίστες χαρακτηριστικών* (attribute lists) των στοιχείων,

οντότητες (entities), συμβολισμούς (notations) και εντολές επεξεργασίας (processing instructions).

Στη γλώσσα XML, η δήλωση των χαρακτηριστικών που περιέχονται σε τύπους στοιχείων έχει ως εξής:

Όνομα_χαρακτηριστικού_τύπος_χαρακτηριστικού_προκαθορισμένη_τιμή

Η γλώσσα παρέχει ένα προκαθορισμένο σύνολο τύπων χαρακτηριστικών: κείμενο (CDATA), ταυτότητα (ID), αναφορά σε ταυτότητα (IDREF), αναφορά σε ταυτότητες (IDREFS), οντότητα (ENTITY), οντότητες (ENTITIES), συμβολικό όνομα και συμβολικά ονόματα (NMTOKEN, NMTOKENS) και απαριθμημένο (enumerated) τύπο. Το κείμενο είναι οποιαδήποτε συμβολοσειρά χαρακτήρων που δεν περιέχει χαρακτήρα(ες) αρχής μιας συμβολικής δήλωσης. Ο τύπος ταυτότητα επιτρέπει σε στοιχεία του εγγράφου να έχουν ένα χαρακτηριστικό που είναι μοναδικό σε όλο το έγγραφο. Η αναφορά (και οι αναφορές) σε ταυτότητα(ες) επιτρέπουν την αναφορά σε ένα (ή περισσότερα) μοναδικά ταυτοποιημένα στοιχεία του εγγράφου. Δηλαδή οι τιμές τύπου IDREF πρέπει να ταιριάζουν με την τιμή ενός χαρακτηριστικού τύπου ID. Τα χαρακτηριστικά που έχουν τύπο οντότητα (ENTITY) πρέπει να έχουν τιμή το όνομα μίας μη αναλυμένης (unparsed) οντότητας που ορίζεται στον τύπο του εγγράφου. Οι μη αναλυμένες οντότητες μπορεί να περιέχουν κείμενο ή όχι και αν περιέχουν κείμενο αυτό μπορεί να μην είναι XML. Οι μη αναλυμένες οντότητες δηλώνονται ως εξής: "<!ENTITY Name ExternalID NDATA Name >". Δηλαδή σχετίζονται με ένα όνομα, μια εξωτερική ταυτότητα και ένα σημειογραφικό όνομα που ορίζεται μέσα στον τύπο του εγγράφου. Η εξωτερική ταυτότητα είναι αναγνωριστικό του συστήματος ή δημόσιο αναγνωριστικό. Το αναγνωριστικό συστήματος είναι ένας ομοιόμορφος δείκτης πόρων (URI) ενώ ένα δημόσιο αναγνωριστικό εκτός από ένα URI περιέχει μία συμβολοσειρά που μπορεί να χρησιμοποιηθεί για τη δημιουργία ενός εναλλακτικού URI. Στην ουσία ο αναλυτής του εγγράφου (και του Τύπου του) δεν ασχολείται με το περιεχόμενο του χαρακτηριστικού τύπου οντότητας, απλώς ενημερώνει την εφαρμογή για την ύπαρξη της οντότητας και τη συσχετιζόμενη σημειογραφία. Τα χαρακτηριστικά τύπου NMTOKEN, NMTOKENS έχουν τιμή κάποια συμβολοσειρά χαρακτήρων. Τα χαρακτηριστικά που έχουν απαριθμημένο τύπο είναι δύο ειδών: σημειογραφίας (NOTATION) ή απαρίθμησης (enumeration). Η τιμή ενός σημειογραφικού χαρακτηριστικού πρέπει να δηλώνεται στον Τύπο Εγγράφου, ενώ τα απαριθμημένα χαρακτηριστικά έχουν ένα σύνολο από αλφαριθμητικές τιμές που δηλώνονται μαζί με το χαρακτηριστικό. Η σημειογραφία ορίζεται με στοιχεία τύπου NOTATION ως εξής: "<!NOTATION Name (ExternalID | PublicID) >", επιτρέποντας έτσι στον επεξεργαστή της γλώσσας να καλέσει ή να επικοινωνήσει με άλλες εφαρμογές. Η προκαθορισμένη τιμή που

μπορεί να συνοδεύει ένα χαρακτηριστικό είναι τριών ειδών: η συμβολοσειρά "#REQUIRED", η συμβολοσειρά "#IMPLIED", μία σταθερή τιμή ή η συμβολοσειρά "#FIXED" συνοδευόμενη από μία τιμή σταθερή τιμή. Στην πρώτη περίπτωση δηλώνεται ότι το χαρακτηριστικό είναι απαραίτητο να ορίζεται και να παίρνει τιμή στα στιγμιότυπα, ενώ στη δεύτερη περίπτωση δεν είναι απαραίτητο. Η σταθερή τιμή αντιστοιχεί σε προκαθορισμένη τιμή για το χαρακτηριστικό, ενώ στην τελευταία περίπτωση η προκαθορισμένη τιμή είναι σταθερή και δεν μεταβάλλεται.

Ένα παράδειγμα εφαρμογής της XML είναι η γλώσσα VML (Vector Markup Language) η οποία αναπτύχθηκε για την κωδικοποίηση και την καθορισμό της παρουσίασης γραφικών με διανύσματα (vector graphics). Η VML αναπτύχθηκε για να παρέχει μια καλή αναπαράσταση των γραφικών με διανύσματα, για να διευκολύνει την ανταλλαγή δεδομένων αυτού του είδους ανάμεσα σε εφαρμογές, και για να διευκολύνει την επεξεργασία της αναπαράστασης των γραφικών από οποιαδήποτε εφαρμογή παρέχοντας τη δυνατότητα επεξεργασίας και με το χέρι.

2.4 Διαχείριση οπτικοακουστικού περιεχομένου στο πρότυπο DAVIC

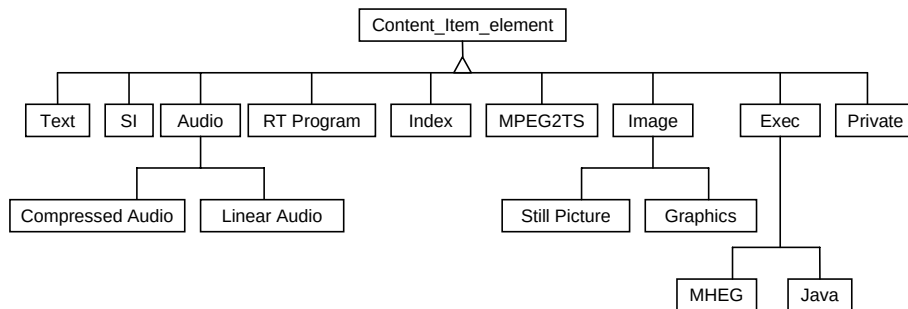
Το πρότυπο DAVIC καθορίζει ένα σύνολο από επίπεδα διαχείρισης της οπτικοακουστικής πληροφορίας εφαρμογών βίντεο κατ' απαίτηση. Τα επίπεδα αυτά αφορούν στην μετάδοση, διαχείριση και ανταλλαγή περιεχομένου ανάμεσα σε εξυπηρετητές πολυμέσων και απομακρυσμένα τερματικά. Στην παρούσα διατριβή, αναπτύχθηκε στα πλαίσια του εξυπηρετητή ΚΥΔΩΝΙΑ, ένα σύνολο από μεθόδους για τη διαχείριση της οπτικοακουστικής πληροφορίας εφαρμογών βίντεο κατ' απαίτηση. Οι μέθοδοι αυτές περιλαμβάνουν τη μοντελοποίηση, ανταλλαγή και δεικτοδότηση του οπτικοακουστικού περιεχομένου.

Στην επόμενη υποενότητα παρουσιάζεται το μοντέλο οπτικοακουστικής πληροφορίας και μεταδεδομένων για εφαρμογές βίντεο κατ' απαίτηση, ενώ στην υποενότητα 2.4.2 περιγράφεται η Υπηρεσία Περιεχομένου του DAVIC, η οποία αναπτύχθηκε για την υποστήριξη ανταλλαγής περιεχομένου ανάμεσα σε εξυπηρετητές και ιδιοκτήτες περιεχομένου.

2.4.1 Μοντέλο οπτικοακουστικής πληροφορίας

Στο πρότυπο DAVIC η οπτικοακουστική πληροφορία μοντελοποιείται από κλάσεις αντικειμένων, οι οποίες αναπαριστούν το περιεχόμενο καθώς και συνοδευτικά μεταδεδομένα. Οι κλάσεις αυτές μοντελοποιούν όλα τα είδη οπτικοακουστικής

πληροφορίας που είναι χρήσιμα για την υποστήριξη εφαρμογών βίντεο κατ' απαίτηση. Η ιεραρχία των κλάσεων πληροφορίας απεικονίζεται στο παρακάτω σχήμα.



Σχήμα 2-3. Κλάσεις πληροφορίας στο πρότυπο DAVIC

Η μοντελοποίηση της πληροφορίας στο DAVIC περιλαμβάνει ένα σύνολο από κλάσεις περιεχομένου, οι οποίες περιλαμβάνουν κείμενο, ήχο, ροές βίντεο, εικόνες, εκτελέσιμα προγράμματα και άλλα είδη περιεχομένου. Οι κλάσεις Compressed Audio, MPEG2TS και Still Picture στηρίζονται στην κωδικοποίηση του προτύπου MPEG-2.

Τα μεταδεδομένα περιγράφουν το περιεχόμενο που μεταφέρεται για εγκατάσταση σε κάποιον εξυπηρετητή και χρησιμεύουν στην οργάνωση και προσπέλασή του από τις εφαρμογές και ορίζονται για κάθε κλάση περιεχομένου. Τα μεταδεδομένα στο DAVIC αποτελούνται από πεδία με πρότυπη αναπαράσταση ώστε να είναι κοινά για όλους τους παροχείς περιεχομένου και εξυπηρετητές. Τα στοιχειώδη αντικείμενα πολυμέσων περιγράφονται από γενικά μεταδεδομένα όπως άδεια προσπέλασης, τύπο, όνομα, ιδιοκτήτη, μέγεθος και γλώσσες. Κάθε υποκλάση περιεχομένου περιγράφεται από συγκεκριμένα μεταδεδομένα που χρησιμοποιούνται για έλεγχο της προσπέλασης, για διαχείριση από τον εξυπηρετητή ή για περιγραφή του περιεχομένου.

2.4.2 Υπηρεσία περιεχομένου του DAVIC

Το πρότυπο DAVIC μεταξύ άλλων καθορίζει τον τρόπο διαχείρισης του περιεχομένου στο περιβάλλον των συστημάτων μετάδοσης οπτικοακουστικών μέσων. Η ανάγκη αυτή εμφανίζεται σε τρεις διαφορετικές περιπτώσεις: για διαχείριση του οπτικοακουστικού περιεχομένου από τον παροχέα ή ιδιοκτήτη του, εσωτερικά στον *Σύστημα Παροχής Υπηρεσιών* και ανάμεσα σε διαφορετικά Συστήματα Παροχής Υπηρεσιών. Ο *παροχέας του περιεχομένου* είναι ο ιδιοκτήτης των εφαρμογών που θα προσφερθούν σε καταναλωτές. Ο παροχέας υπηρεσιών στην ουσία δίνει τη δυνατότητα στον παροχέα του περιεχομένου να προσφέρει τις εφαρμογές σε τελικούς χρήστες παρέχοντας το σύστημα παράδοσής τους. Καθένας από αυτούς έχει συγκεκριμένες ανάγκες ως προς τη διαχείριση του περιεχομένου. Ο παροχέας του περιεχομένου έχει την ανάγκη να εγκαταστήσει περιεχόμενο σε μια

περιοχή ενός συστήματος εξυπηρέτησης, μετά την εγκατάστασή του να το διαχειρίζεται (ενημέρωση, διαγραφή) και να καθορίζει τις άδειες προσπέλασής του. Επομένως από την πλευρά του παροχέα περιεχομένου, η έμφαση δίνεται στην *αναπαράσταση του περιεχομένου*, η οποία έχει να κάνει με την πηγή από την οποία αυτό προέρχεται αλλά και με μετασχηματισμούς που θα εφαρμοστούν σ' αυτό, στην «συσκευασία» του περιεχομένου για να έτοιμο για εγκατάσταση σε έναν εξυπηρετητή και στη διεπικοινωνία για την απομακρυσμένη διαχείρισή του με διαλειτουργικό τρόπο.

Ο παροχέας υπηρεσιών ενεργεί ως ο μεσάζοντας που υλοποιεί τις ανάγκες του παροχέα περιεχομένου. Οι απαιτήσεις που τίθενται σε έναν παροχέα υπηρεσιών ως προς το περιεχόμενο είναι: υποστήριξη της διαδικασίας εγκατάστασης η οποία αρχικοποιείται από παροχείς περιεχομένου, αποδοτική διαχείριση του περιεχομένου για να παρέχεται η κατάλληλη ποιότητα υπηρεσιών στο χρήστη και εγγυήσεις για τον έλεγχο προσπέλασης.

Η προσέγγιση που ακολουθείται ως προς τη διαχείριση του περιεχομένου από τον ιδιοκτήτη του είναι ο ορισμός ενός συνόλου επιτρεπτών οπτικοακουστικών τύπων και η διεπικοινωνία για την εγκατάσταση και διαχείριση στοιχείων που ανήκουν σ' αυτούς τους τύπους. Η διαχείριση του περιεχομένου στο σύστημα εξυπηρέτησης γίνεται σε διαφορετικά επίπεδα: το εξωτερικό επίπεδο που εξάγει τη λειτουργικότητα διαχείρισης στον παροχέα του περιεχομένου και ένα σύνολο από εσωτερικά επίπεδα. Το εξωτερικό επίπεδο υποστηρίζει τη διεπικοινωνία (τύπους περιεχομένου, ιδιότητες των τύπων του περιεχομένου και διεπικοινωνία διαχείρισης) που χρησιμοποιεί ο παροχέας περιεχομένου. Τα εσωτερικά επίπεδα έχουν να κάνουν με την πιστοποίηση των προκαθορισμένων τύπων, καθώς και με τους εσωτερικούς τύπους, τις ιδιότητες και τις τιμές αυτών.

Το περιεχόμενο διακρίνεται σε αντικείμενα εφαρμογής επεξεργάσιμα από το τερματικό και αντικείμενα στον εξυπηρετητή. Τα αντικείμενα εφαρμογής είναι τμήματα ενός ή περισσότερων εφαρμογών οι οποίες είναι έτοιμες προς παράδοση σε καταναλωτές. Τα αντικείμενα αυτά μπορούν επιπλέον να διακριθούν σε ένα εκτελέσιμο για το τερματικό, ένα εκτελέσιμο για τον εξυπηρετητή και τα «αποθέματα» ή αντικείμενα δεδομένων της εφαρμογής. Από την άποψη της εφαρμογής τα αποθέματα είναι είτε αποθηκευμένα αντικείμενα ή ροές πραγματικού χρόνου.

Η διάκριση των αποθεμάτων σε ροές πραγματικού χρόνου και αποθηκευμένα αντικείμενα αναφέρεται στον τρόπο που μεταδίδονται το τερματικό από το σύστημα παροχής υπηρεσιών και παρουσιάζονται στον τελικό χρήστη. Οι ροές πραγματικού χρόνου καταναλώνονται την ώρα παράδοσης, ενώ τα αποθηκευμένα αντικείμενα παραμένουν στην τοπική αποθήκευση (κύρια μνήμη, λανθάνουσα μνήμη ή τοπικός δίσκος) του τερματικού

έτσι ώστε το εκτελέσιμο της εφαρμογής να έχει πλήρη έλεγχο της διαδικασίας κατανάλωσης.

Κάθε στοιχείο των αποθεμάτων ανήκει σε ένα σύνολο από προκαθορισμένους τύπους αντικειμένων πολυμέσων. Η γενική κλάση αντικειμένων πολυμέσων εξειδικεύεται σε στοιχεία «μονομέσων» και σύνθετα αντικείμενα πολυμέσων. Οι τύποι μονομέσων που υποστηρίζονται και για τους οποίους υπάρχει πρότυπη μορφή κωδικοποίησης είναι οι εξής: χαρακτήρες, πληροφορία για τη γλώσσα, πληροφορία για την υπηρεσία, κείμενο, γραμμικός ήχος, συμπιεσμένος ήχος, συμπιεσμένο βίντεο, συμπιεσμένες εικόνες και γραφικά. Τα σύνθετα αντικείμενα αποτελούνται από ένα ή περισσότερα στοιχειώδη αντικείμενα μονομέσων. Τα στοιχειώδη αντικείμενα μπορούν επίσης να αποτελούνται από μια σειρά αντικειμένων μονομέσων του ίδιου τύπου που παρουσιάζονται σειριακά στο χρήστη.

Όσον αφορά στην παράδοση των αντικειμένων, τα στοιχεία όλων των τύπων μονομέσων μπορούν να οριστούν ως αποθηκευμένα αντικείμενα αλλά μόνο ένα υποσύνολο από αυτά μπορούν να παραδοθούν ως ροές πραγματικού χρόνου. Οι ροές πραγματικού χρόνου μπορούν να αποτελούνται μόνο από MPEG-2 βίντεο, MPEG-1 ήχου, MPEG εικόνες ή γραφικά του DAVIC. Τα αποθηκευμένα αντικείμενα για να μεταδοθούν οργανώνονται σε πακέτα και μεταδίδονται ως Ιδιωτικές Ροές, οι οποίες αποτελούν επέκταση της δομής των PES που ορίζεται από το MPEG-2.

Η εγκατάσταση του περιεχομένου σε έναν εξυπηρετητή γίνεται σε μονάδες που ονομάζονται «συσκευασίες» περιεχομένου, οι οποίες περιέχουν ένα σύνολο από οπτικοακουστικά αντικείμενα. Η μορφοποίηση της συσκευασίας είναι για λόγους διαλειτουργικότητας κοινή για τον παροχέα υπηρεσιών και τον παροχέα περιεχομένου και ανεξάρτητη από το μέσο μετάδοσης. Η συσκευασία αυτή εκτός από τα πραγματικά αντικείμενα περιέχει και ένα σύνολο μεταδεδομένων που συσχετίζονται με αυτά και διευκολύνουν την όλη διαχείριση και χρήση του περιεχομένου. Τα μεταδεδομένα για το περιεχόμενο λειτουργούν ως περιγραφείς και ταξινομούνται σε δύο κατηγορίες: μεταδεδομένα της εφαρμογής και μεταδεδομένα διαχείρισης.

Τα μεταδεδομένα διαχείρισης προορίζονται για χρήση από τον παροχέα του περιεχομένου, επομένως και για εσωτερική χρήση από τον εξυπηρετητή και συγκεκριμένα στις διαδικασίες μεταφοράς, εγκατάστασης και πιστοποίησης. Τα μεταδεδομένα των εφαρμογών αφορούν στην πλοήγηση μέσα στο περιεχόμενο, στα πνευματικά δικαιώματα και στον έλεγχο πρόσβασης από τους χρήστες. Η διεπικοινωνία μεταφοράς του περιεχομένου ορίζεται στη γλώσσα IDL (Interface Definition Language) και σχετίζεται με τη

διεπικοινωνία Directory του προτύπου DSM-CC (Digital Storage Media – Command and Control).

2.5 Το σύστημα ΚΥΔΩΝΙΑ

Το σύστημα ΚΥΔΩΝΙΑ αναπτύχθηκε στο Εργαστήριο Διανεμημένων Συστημάτων και Εφαρμογών του Πολυτεχνείου Κρήτης και χρησιμοποιήθηκε στο ευρωπαϊκό πρόγραμμα SICMA (Scaleable Interactive Continuous Media Server and Application) για την υποστήριξη εφαρμογών Βίντεο κατ' Απαίτηση σε ένα περιβάλλον συμβατό με το πρότυπο DAVIC (Digital AudioVisual Council). Το σύστημα ΚΥΔΩΝΙΑ αποτελεί έναν εξυπηρετητή πολυμέσων που υποστηρίζει μετάδοση βίντεο, ήχου και άλλων αντικειμένων (εικόνες, κείμενο, κλπ) σε πραγματικό χρόνο. Η αρχιτεκτονική του αποτελείται από ένα υποσύστημα διαχείρισης της αποθήκευσης και μετάδοσης αντικειμένων ευαίσθητων σε χρονικές καθυστερήσεις και από έναν αντικειμενοστρεφή πυρήνα διαχείρισης των υπόλοιπων αντικειμένων του συστήματος. Η διαχείριση των πολυμέσων στο σύστημα ΚΥΔΩΝΙΑ, εκτός από την αποθήκευση και ανάκτησή τους με απαιτήσεις πραγματικού χρόνου, περιλαμβάνει τη μοντελοποίησή τους σύμφωνα με το αντικειμενοστρεφές παράδειγμα και την κατασκευή μιας αντικειμενοστρεφούς βάσης δεδομένων με τα μοντελοποιημένα αντικείμενα. Η μοντελοποίηση αυτή παρέχει μεγαλύτερη λειτουργικότητα στις εφαρμογές, όπως προσπέλαση των αντικειμένων μέσω της μοναδικής τους ταυτότητας, ταξινόμηση των αντικειμένων σε γνωστές κατηγορίες ή κλάσεις, συσχέτιση των αντικειμένων μεταξύ τους μέσω των σχέσεων όπως part-of και is-a, εμπλουτισμό των αντικειμένων με ιδιότητες και δυνατότητα επέκτασης των κλάσεων αντικειμένων και των ιδιοτήτων τους.

Στα πλαίσια του προγράμματος SICMA, το σύστημα ΚΥΔΩΝΙΑ επεκτάθηκε ώστε να γίνει συμβατό με το πρότυπο DAVIC, το οποίο καθορίζει τη λειτουργικότητα ενός εξυπηρετητή πολυμέσων σε ένα κατακευματισμένο περιβάλλον. Συγκεκριμένα το DAVIC καθορίζει πως θα εξασφαλιστεί η διαλειτουργικότητα ανάμεσα στον εξυπηρετητή πολυμέσων και την εφαρμογή, τον εξυπηρετητή και τον παροχέα περιεχομένου (και εφαρμογής) και ανάμεσα σε διαφορετικούς εξυπηρετητές. Σύμφωνα με το DAVIC, ένας εξυπηρετητής πολυμέσων παρέχει στο εξωτερικό του περιβάλλον ένα σύνολο από υπηρεσίες, όπως εγκατάσταση και διαχείριση περιεχομένου και εφαρμογών από τον ιδιοκτήτη τους, προσπέλαση των εφαρμογών και του περιεχομένου τους από απομακρυσμένα τερματικά, μετάδοση του περιεχομένου σε πραγματικό χρόνο, αλληλεπίδραση του χρήστη με το περιεχόμενο και προνόμια προσπέλασης των υπηρεσιών του συστήματος.

Στην ενότητα που ακολουθεί περιγράφεται συνοπτικά η αρχιτεκτονική του συστήματος ΚΥΔΩΝΙΑ, ενώ στην ενότητα 2.5.2 παρουσιάζονται οι στόχοι του προγράμματος SICMA και η μετάβαση του συστήματος ΚΥΔΩΝΙΑ προς μία αρχιτεκτονική συμβατή με το DAVIC, η οποία υποστηρίζει υπηρεσίες διαχείρισης οπτικοακουστικού περιεχομένου.

2.5.1 Αρχιτεκτονική του συστήματος ΚΥΔΩΝΙΑ

Το σύστημα ΚΥΔΩΝΙΑ έχει αρχιτεκτονική πελάτη-εξυπηρετητή και αποτελείται από τρία βασικά υποσυστήματα όπως εμφανίζονται στο Σχήμα 2–4. Το υποσύστημα υποστήριξης του πελάτη που εμφανίζεται στο υψηλότερο επίπεδο της αρχιτεκτονικής, αναπτύχθηκε με στόχο να υποστηρίξει την ανάπτυξη και την εκτέλεση των εφαρμογών πολυμέσων που θα έτρεχαν πάνω στον εξυπηρετητή. Το υποσύστημα αυτό αναλαμβάνει την δημιουργία και διαχείριση των μηνυμάτων που ανταλλάσσονται ανάμεσα σε πελάτη και εξυπηρετητή, διαχειρίζεται τα αντικείμενα και υποστηρίζει λειτουργίες προσωρινής αποθήκευσης δεδομένων. Επίσης εξάγει στις εφαρμογές την λειτουργικότητα που παρέχει ο εξυπηρετητής για διαχείριση του αντικειμενοστρεφούς σχήματος της βάσης δεδομένων, δημιουργία δεικτών και υποβολή ερωτήσεων και πλοήγηση στο περιεχόμενο των πολυμέσων.

Το Υποσύστημα Διαχείρισης Αντικειμένων και Συνόδων (Object and Session Management System ή OSM) εκτός από την αποκωδικοποίηση και δρομολόγηση των εισερχόμενων μηνυμάτων, αναλαμβάνει την παροχή λειτουργικότητας όπως τη διαχείριση των αντικειμένων πολυμέσων στην κύρια μνήμη του εξυπηρετητή, τη δεικτοδότηση του περιεχομένου των πολυμέσων και την υποστήριξη ερωτήσεων ανάκτησης περιεχομένου.

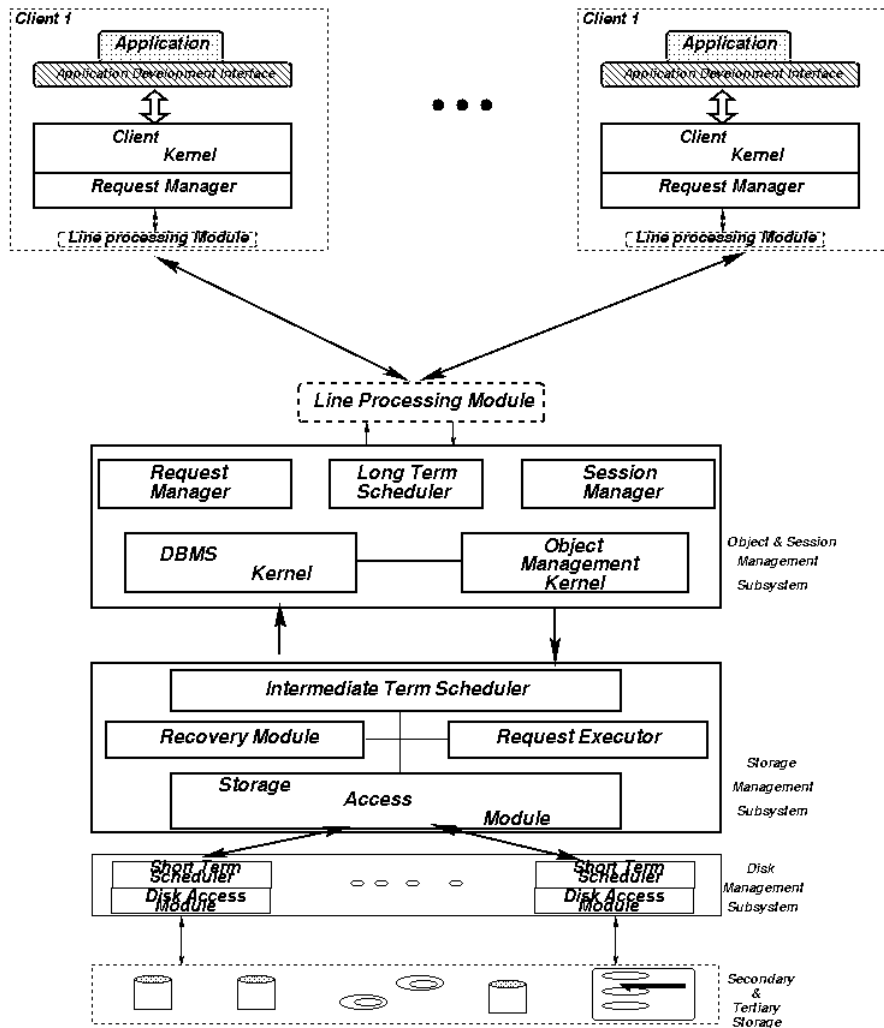
Τέλος, η αποθήκευση των δεδομένων στην δευτερεύουσα και τριτεύουσα μνήμη αναλαμβάνεται από το υποσύστημα Διαχείρισης Αποθήκευσης που βρίσκεται στο κατώτερο επίπεδο της αρχιτεκτονικής.

2.5.2 Το πρότυπο DAVIC και το πρόγραμμα SICMA

Τα τελευταία χρόνια υπάρχει παγκόσμιο ενδιαφέρον για την κατασκευή υποδομών που θα επιτρέπουν την κατανομή της ψηφιακής πληροφορίας σε ένα σύνολο από εξυπηρετητές που θα την παρέχουν σε χρήστες κατ' απαίτηση και μέσω δικτύων υψηλής χωρητικότητας.

Οι τελικοί χρήστες στους οποίους απευθύνονται αυτού του είδους οι υποδομές είναι σπίτια εφοδιασμένα με συσκευές ανάκτησης ή Set Top Units (STUs) και η πληροφορία στην οποία θα έχουν πρόσβαση θα είναι κυρίως πληροφορία πολυμέσων (κείμενο, εικόνες, video και ήχος). Οι εξυπηρετητές πολυμέσων αποτελούν βασικό στοιχείο αυτών των υποδομών εφόσον προσφέρουν αποθήκευση και ανάκτηση μεγάλου όγκου πληροφορίας

πολυμέσων σε ευρύ πεδίο εφαρμογών όπως βίντεο κατ' απαίτηση, τηλεγορές και νέα κατ' απαίτηση.



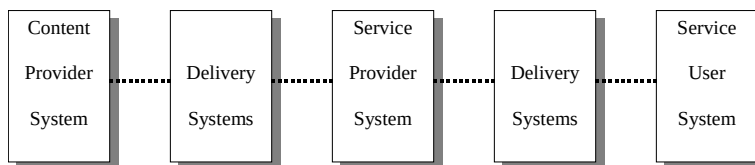
Σχήμα 2-4. Αρχιτεκτονική του πληροφοριακού συστήματος ΚΥΔΩΝΙΑ

Η αρχιτεκτονική του πληροφοριακού συστήματος διαχείρισης πολυμέσων ΚΥΔΩΝΙΑ, αποτελείται από 4 επίπεδα από κάτω προς τα πάνω: διαχείριση δίσκων, διαχείριση δευτερεύουσας αποθήκευσης, διαχείριση αντικειμένων και συνόδων και υποσύστημα εφαρμογής.

Το DAVIC είναι ένας μη κερδοσκοπικός οργανισμός που είχε στόχο να συγκεντρώσει τις προσπάθειες μεγάλων οργανισμών και εταιρειών που κινούνταν στα πεδία των τηλεπικοινωνιών, της καλωδιακής τηλεόρασης, των εξυπηρετητών και των συσκευών

λήψης και αναπαραγωγής ψηφιακών πολυμέσων για την καθιέρωση ενός προτύπου που θα εξασφάλιζε διαλειτουργικότητα ανάμεσα στα διαφορετικά τμήματα των υποδομών εκπομπής οπτικοακουστικής πληροφορίας. Το πρότυπο DAVIC που αναπτύχθηκε στα πλαίσια αυτής της προσπάθειας περιγράφει τη λειτουργικότητα και τα πρωτόκολλα που πρέπει να υποστηρίζονται από κάθε τμήμα της υποδομής ώστε να επιτυγχάνεται ανάμεσά τους η επιθυμητή διαλειτουργικότητα. Όσον αφορά στον πληροφοριακό εξυπηρετητή το DAVIC απαιτεί να είναι μεγάλης κλίμακας, ικανός να ανταποκρίνεται σε μεγάλο αριθμό αιτήσεων χωρίς να μειώνεται η απόδοσή του και να διαχειρίζεται το περιεχόμενο πολυμέσων με αποδοτικό τρόπο, ώστε να υποστηρίζει τις απαιτήσεις πραγματικού χρόνου των δεδομένων πολυμέσων, όπως την ευαισθησία του βίντεο σε χρονικές καθυστερήσεις.

Η συνολική αρχιτεκτονική των υπηρεσιών του προτύπου DAVIC ορίζεται μέσω στοιχείων εξυπηρέτησης (service elements), συστημάτων ζεύξης και επαφής (interfaces), ροών πληροφορίας (information flows) και πρωτοκόλλων επικοινωνίας (protocol stacks). Στο παρακάτω σχήμα φαίνονται τα επιμέρους υποσυστήματα από τα οποία αποτελείται ένα πληροφορικό σύστημα συμβατό με το πρότυπο DAVIC. Το Σύστημα Παροχής Περιεχομένου (Content Provider System) είναι υπεύθυνο για την είσοδο δεδομένων και πληροφορίας. Το Σύστημα Παροχής Υπηρεσιών (Service Provider System) αποτελεί τον κεντρικό εξυπηρετητή της υποδομής, ο οποίος παρέχει υπηρεσίες μετάδοσης οπτικοακουστικής πληροφορίας στο Σύστημα Εξυπηρέτησης του Χρήστη (Service User System). Τα επιμέρους συστήματα ενώνονται μεταξύ τους με Συστήματα Μεταφοράς δεδομένων (Delivery Systems). Ιδιαίτερη σημασία δίνεται από τα στοιχεία του DAVIC στην μετάδοση οπτικοακουστικής πληροφορίας κωδικοποιημένης σύμφωνα με το πρότυπο MPEG-2 και στη μεταφορά της σύμφωνα με το πρωτόκολλο MPEG-2 Transport Stream. Το πρότυπο MPEG-2 επιλέχθηκε την εποχή που ξεκίνησε το DAVIC ως το πλέον σύγχρονο και αποδοτικό πρότυπο κωδικοποίησης και μεταφοράς βίντεο και ήχου.



Σχήμα 2-5. Αρχιτεκτονική του συστήματος DAVIC

Ένα σύστημα υποστήριξης οπτικοακουστικών εφαρμογών που είναι συμβατό με το πρότυπο DAVIC αποτελείται από το Σύστημα Παροχής Περιεχομένου, το Σύστημα Παροχής Υπηρεσιών, το Σύστημα του Χρήστη και ενδιάμεσα Συστήματα Παράδοσης πληροφορίας.

Ο στόχος του ερευνητικού ευρωπαϊκού προγράμματος SICMA, στο οποίο συμμετείχε το Εργαστήριο Διανεμημένων Πληροφοριακών Συστημάτων και Εφαρμογών του Πολυτεχνείου Κρήτης, ήταν η ανάπτυξη ενός μεγάλης κλίμακας πληροφοριακού εξυπηρετητή πολυμέσων συμβατού με το πρότυπο DAVIC. Ο εξυπηρετητής είχε στόχο να αποθηκεύει, να διαχειρίζεται και να μεταδίδει δεδομένα πολυμέσων σε ένα μεγάλο αριθμό πελατών με τη μέγιστη δυνατή διαλειτουργικότητα ως προς άλλα στοιχεία της υποδομής μετάδοσης πολυμέσων, για την υποστήριξη ενός μεγάλου φάσματος εφαρμογών. Στην ενότητα που ακολουθεί περιγράφεται η αρχιτεκτονική και τα χαρακτηριστικά του ΚΥΔΩΝΙΑ-D εξυπηρετητή πολυμέσων, ο οποίος αναπτύχθηκε στα πλαίσια του προγράμματος SICMA από το Εργαστήριο Διανεμημένων Πληροφοριακών Συστημάτων και Εφαρμογών.

2.5.2.1 Ο ΚΥΔΩΝΙΑ-D εξυπηρετητής πολυμέσων

Στα πλαίσια του ευρωπαϊκού προγράμματος SICMA, έγιναν οι απαραίτητες ενέργειες και μετατροπές στον εξυπηρετητή ΚΥΔΩΝΙΑ, ώστε να ακολουθήσει το πρότυπο DAVIC. Το αποτέλεσμα αυτής της προσπάθειας ήταν η ανάπτυξη του εξυπηρετητή ΚΥΔΩΝΙΑ-D. Επομένως, σύμφωνα με το DAVIC ο εξυπηρετητής ΚΥΔΩΝΙΑ-D έχει αντικειμενοστρεφή αρχιτεκτονική και παρέχει ένα σύνολο από υπηρεσίες που υποστηρίζουν τις εφαρμογές πολυμέσων:

- Την Υπηρεσία Εγκατάστασης Περιεχομένου, η οποία χρησιμοποιείται από τους παροχείς ή ιδιοκτήτες της πληροφορίας για να εγκαταστήσουν νέες εφαρμογές και περιεχόμενο στον εξυπηρετητή και να ενημερώσουν το ήδη υπάρχον υλικό.
- Την Υπηρεσία Επιλογής Υπηρεσιών που επιτρέπει σε πελάτες να εντοπίσουν και να επιλέξουν κάποια από τις διαθέσιμες υπηρεσίες του εξυπηρετητή.
- Την Υπηρεσία Διαχείρισης Ροών Πληροφορίας, η οποία μεταδίδει και διαχειρίζεται τις συνεχείς ροές πληροφορίας από τον εξυπηρετητή στον τελικό πελάτη.

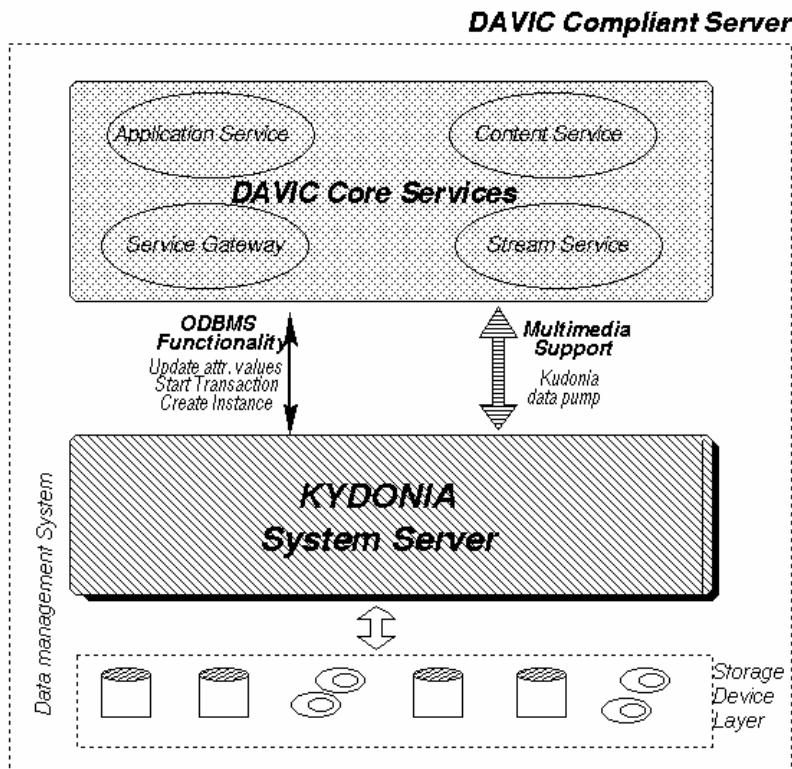
Η ανάπτυξη της παραπάνω λειτουργικότητας σε ένα σύστημα όπως το ΚΥΔΩΝΙΑ παρέχει σημαντικά πλεονεκτήματα στους τομείς της διαχείρισης αντικειμένων πολυμέσων, της οργάνωσης και επέκτασης της λειτουργικότητας του συστήματος, της πλοήγησης και συσχέτισης αντικειμένων και τέλος στην ανάπτυξη επιπρόσθετων μηχανισμών υποστήριξης όπως η ασφάλεια, τιμολόγηση και ερωτήσεις. Τα πλεονεκτήματα αυτά οφείλονται στην αντικειμενοστρεφή βάση δεδομένων του συστήματος.

Ο εξυπηρετητής του προγράμματος SICMA χρησιμοποιεί το σύστημα ΚΥΔΩΝΙΑ. Ο ρόλος του συστήματος ΚΥΔΩΝΙΑ αφορά:

- στην διαχείριση της αποθήκευσης αντικειμένων πολυμέσων

- στην παροχή συνεχών ροών πραγματικού χρόνου μέσω μηχανισμών κατάλληλης δρομολόγησης των αιτήσεων των πελατών
- στην υποστήριξη λειτουργικότητας βάσεων δεδομένων πολυμέσων, όπως δημιουργία αντικειμένων πολυμέσων και αντικειμένων που αναπαριστούν το περιεχόμενο των πολυμέσων
- επαρκή διαχείριση MPEG-2 βίντεο και HTML εγγράφων
- υποστήριξη μεθόδων προσπέλασης για δεδομένα βίντεο και κειμένου
- υποστήριξη ανάκτησης πολυμέσων βάσει του περιεχομένου τους

Ο εξυπηρετητής του SICMA υποστηρίζει ένα μεγάλο αριθμό πελατών χρησιμοποιώντας μια αρχιτεκτονική από επεξεργαστικά τμήματα, συσκευές αποθήκευσης και εξωτερικά τμήματα διεπικοινωνίας.



Σχήμα 2-6. Αρχιτεκτονική του εξυπηρετητή KYΔΩΝΙΑ-D

Η αρχιτεκτονική του συστήματος KYΔΩΝΙΑ επεκτάθηκε με έναν πυρήνα βασικών στοιχείων υποστήριξης των υπηρεσιών του DAVIC, αποτελώντας έτσι τον εξυπηρετητή KYΔΩΝΙΑ-D.

Ο εξυπηρετητής του SICMA μπορεί επίσης να μεταφέρει δεδομένα βίντεο και ήχου κωδικοποιημένα σε MPEG-2 μορφή και να μεταδώσει ροές πληροφορίας πάνω από ATM (Asynchronous Transfer Mode) δίκτυα υψηλών ταχυτήτων. Άλλα χαρακτηριστικά και υπηρεσίες που παρέχονται είναι η μεταφορά ηλεκτρονικών αρχείων και εκτελέσιμων προγραμμάτων από τον εξυπηρετητή στη συσκευή που αποτελεί τον τελικό αποδέκτη.

Τα βασικά στοιχεία υπηρεσίας (Core Service elements) αναπτύχθηκαν πάνω στον εξυπηρετητή βάσης πολυμέσων ΚΥΔΩΝΙΑ και αποτελούν το κέλυφος με το οποίο επιτυγχάνεται η συμβατότητα με το πρότυπο DAVIC καταλήγοντας στον αναμορφωμένο ΚΥΔΩΝΙΑ-D εξυπηρετητή.

Η προσπάθεια και οι εργασίες που έλαβαν χώρα ώστε ο εξυπηρετητής ΚΥΔΩΝΙΑ να γίνει συμβατός με το πρότυπο DAVIC, ενίσχυσαν τον εξυπηρετητή με καινούργια χαρακτηριστικά όπως νέα πρωτόκολλα και συστήματα ζεύξης, ενώ παράλληλα οδήγησαν στην ενσωμάτωση στο λογισμικό του εξυπηρετητή νέων μηχανισμών για την αποδοτικότερη διαχείριση των δεδομένων πολυμέσων, ώστε να ανταποκριθεί στις απαιτήσεις ποιότητας υπηρεσιών (Quality Of Service) που απορρέουν από το πρότυπο. Με βάση την αρχιτεκτονική του DAVIC για τον εξυπηρετητή έχουν αναπτυχθεί τα ακόλουθα βασικά στοιχεία υπηρεσιών (Core Service Elements): Η υπηρεσία περιεχομένου (Content Service), η οποία αναλαμβάνει την εγκατάσταση νέων εφαρμογών μέσα στον εξυπηρετητή, η υπηρεσία εισόδου (Service Gateway) που αναλαμβάνει την οργάνωση των υπηρεσιών (service domain) και παρέχει την δυνατότητα στους εξωτερικούς πελάτες να ανακαλύψουν και να χρησιμοποιήσουν διαθέσιμες υπηρεσίες και η υπηρεσία ροής (Stream Service) η οποία οδηγεί την πρόσβαση σε χρονικά ευαίσθητα αντικείμενα και επιτρέπει στους πελάτες να έχουν έλεγχο της ροής των δεδομένων. Τα συστήματα ζεύξης (APIs) των στοιχείων υπηρεσίας είναι κυρίως βασισμένα στο DSM-CC User to User specification. Τα δεδομένα συνεχόμενης ροής μεταδίδονται μέσω του MPEG-2 Transport πρωτοκόλλου, ενώ για την μεταφορά των κλήσεων των υπηρεσιών από τους πελάτες ακολουθείται η αρχιτεκτονική CORBA και χρησιμοποιείται το Orbix 2.0.1 προϊόν.

Ο απώτερος στόχος του DAVIC είναι να αυξήσει την διαλειτουργικότητα και συνεπώς η εσωτερική υλοποίηση του εξυπηρετητή δεν καθορίζεται. Έτσι αυξάνεται η ευελιξία καθώς είναι υπευθυνότητα του εξυπηρετητή να οργανώσει και να διαχειριστεί τα δεδομένα πολυμέσων με τέτοιο τρόπο ώστε να ικανοποιήσει καλύτερα τις απαιτήσεις στην ροή δεδομένων που επιβάλλονται από τα ανώτερα επίπεδα. Το σύστημα ΚΥΔΩΝΙΑ ενεργεί ως διαχειριστής δεδομένων (data manager) μέσα στον DAVIC εξυπηρετητή και χρησιμοποιείται από τα στοιχεία υπηρεσίας για την δημιουργία, οργάνωση και ανάκτηση πληροφορίας πολυμέσων.

Παράλληλα με την ανάπτυξη λειτουργικότητας συμβατής με το DAVIC, πάνω στον εξυπηρετητή ΚΥΔΩΝΙΑ στα πλαίσια του προγράμματος SICMA, έγινε ουσιαστική έρευνα, πειραματισμοί και ανάπτυξη πάνω στον εξυπηρετητή αποθήκευσης που γίνεται τελικά ο αποδέκτης όλου του φόρτου για αποθήκευση και ανάκτηση δεδομένων. Κύριος στόχος ήταν η αύξηση της απόδοσης εξυπηρέτησης δεδομένων συνεχόμενης ροής (streams) με τον σχεδιασμό της αρχιτεκτονικής και την κατάλληλη χρήση τεχνικών ανάπτυξης λογισμικού που θα εξασφάλιζαν την βελτιστοποίηση των πόρων του συστήματος με απόρροια την μεγιστοποίηση της απόδοσης αλλά και θα παρείχαν τις απαραίτητες εγγυήσεις ποιότητας εξυπηρέτησης. Τα αποτελέσματα αυτής της προσπάθειας περιγράφονται στην τρέχουσα διατριβή.

Κεφάλαιο 3

Περιγραφή του MPEG-4 περιεχομένου

Στο κεφάλαιο αυτό παρουσιάζεται η περιγραφή του περιεχομένου του MPEG-4 με τη γλώσσα αναπαράστασης οπτικοακουστικού περιεχομένου AVDL (Audio Visual Description Language) που αναπτύχθηκε στα πλαίσια της διατριβής. Η γλώσσα AVDL αποτελεί εφαρμογή της μεταγλώσσας XML για την αναπαράσταση εγγράφων με οποιοδήποτε περιεχόμενο στον Παγκόσμιο Ιστό. Τα χαρακτηριστικά της γλώσσας είναι ότι υποστηρίζει ημιδομημένες περιγραφές με σύνθετο και ιεραρχικά οργανωμένο περιεχόμενο και εκφράζει τις σημασιολογικές δομές των κωδικοποιημένων δεδομένων του MPEG-4. Επίσης υποστηρίζει διαβαθμισμένες περιγραφές του περιεχομένου μέσα από προφίλ που επιτρέπουν την χρήση της περιγραφής ανάλογα με τις ανάγκες των εφαρμογών. Η περιγραφή των δεδομένων του MPEG-4 με την AVDL είναι πλήρης και περιλαμβάνει την περιγραφή των οπτικών αντικειμένων, των σκηνών και των περιγραφών αντικειμένων. Επίσης υποστηρίζει τη σύνθεση αυτόνομων περιγραφών από διαφορετικά έγγραφα στον Παγκόσμιο Ιστό.

3.1 Εισαγωγή

Οι πλέον σύγχρονες εφαρμογές οπτικοακουστικών μέσων απαιτούν αυξημένη λειτουργικότητα, όπως υψηλό βαθμό αλληλεπίδρασης του χρήστη με το υλικό, ποικίλους τρόπους διαχείρισης των πολυμέσων και εμπλουτισμό του υλικού με περισσότερη πληροφορία. Αυτού του είδους οι εφαρμογές απαιτούν διαχείριση του υλικού με βάση το περιεχόμενό του, όπως αυτό γίνεται αντιληπτό από το χρήστη ή την εφαρμογή. Ενώ έχουν αναπτυχθεί μέθοδοι και συστήματα εξαγωγής πληροφορίας από τις εικόνες, το βίντεο και τον ήχο, δεν υπάρχει ενιαίος τρόπος αναπαράστασής της, γεγονός που αποτελεί

ανασταλτικό παράγοντα στην εξάπλωση των εφαρμογών βάσει περιεχομένου. Στο κεφάλαιο 2 είδαμε ότι το πρότυπο MPEG-4 στοχεύει να καλύψει τις απαιτήσεις νέων εφαρμογών οπτικοακουστικών μέσων παρέχοντας κωδικοποίηση του υλικού με βάση το περιεχόμενό του και αλληλεπίδραση με το χρήστη.

Παρόλα αυτά, το MPEG-4 δεν καθορίζει τον τρόπο που μπορούν να υποστηριχθούν εφαρμογές στις οποίες υπάρχει ανάγκη για διαχείριση περιεχομένου πέραν αυτής που παρέχει η κωδικοποίηση των αντικειμένων με τα εργαλεία του. Επίσης μεγάλο μέρος των εφαρμογών του MPEG-4 και των οπτικοακουστικών εφαρμογών γενικότερα, αφορούν στον Παγκόσμιο Ιστό. Το κυριότερο χαρακτηριστικό όμως του Παγκόσμιου Ιστού είναι ότι έχουν ήδη καθιερωθεί πρότυπα αναπαράστασης εγγράφων που παρέχουν πλεονεκτήματα όπως απευθείας διαχείριση του περιεχομένου των εγγράφων (εφόσον είναι κείμενο) και δυνατότητα ανταλλαγής του ανεξάρτητα από το είδος των εφαρμογών ή των συστημάτων. Τα πρότυπα αυτά είναι πολύ δημοφιλή και αποτελούν αναπόσπαστο μέρος των εφαρμογών και επομένως είναι δύσκολο να εγκαταλειφθούν ή ακόμα και να αγνοηθούν.

Επομένως συμπεραίνουμε ότι υπάρχει η ανάγκη υποστήριξης από τα πληροφοριακά συστήματα πολυμέσων, οπτικοακουστικών εφαρμογών με αυξημένη λειτουργικότητα, η οποία να στηρίζεται στο περιεχόμενο του υλικού και η υποστήριξη αυτή πρέπει να παρέχεται από τον πυρήνα του συστήματος με ενιαίο τρόπο για όλες τις εφαρμογές. Το πρότυπο MPEG-4 παρέχει κωδικοποίηση βάσει περιεχομένου, άρα δίνει τη δυνατότητα διαχείρισης και περαιτέρω επεξεργασίας του περιεχομένου. Στο κεφάλαιο αυτό παρουσιάζουμε το πλαίσιο που αναπτύχθηκε στην παρούσα διατριβή για την εκμετάλλευση της πληροφορίας που παρέχει το MPEG-4 σχετικά με το περιεχόμενο των οπτικοακουστικών μέσων και για τη χρήση της πληροφορίας αυτής σε εφαρμογές τεχνολογίας αιχμής, συμπεριλαμβανομένων των εφαρμογών του Παγκόσμιου Ιστού. Το πλαίσιο αυτό αναπτύχθηκε με στόχο την υποστήριξη των εξής αναγκών των εφαρμογών διαχείρισης οπτικοακουστικών μέσων:

- Έλεγχος της αλληλεπίδρασης και της παρουσίασης του υλικού. Οι σύγχρονες εφαρμογές απαιτούν υψηλό βαθμό αλληλεπίδρασης με το χρήστη και περιήγηση στο οπτικοακουστικό περιεχόμενο. Ο έλεγχος της αλληλεπίδρασης καθιστά το σύστημα ικανό να αποκριθεί σε πολύπλοκες ενέργειες των χρηστών και η περιήγηση δίνει στο χρήστη ελευθερία στην επισκόπηση του περιεχομένου.
- Εξατομίκευση των εφαρμογών. Οι χρήστες έχουν την ανάγκη να "βλέπουν" το υλικό προσαρμοσμένο στα δικά τους ενδιαφέροντα και προτιμήσεις. Η εξατομίκευση αυτή

περιλαμβάνει επιλογή του τρόπου παρουσίασης, αλλά και επιλογή του ίδιου του υλικού που θα παρουσιαστεί.

- Μεταβίβαση υλικού και τρόπου παρουσίασης σύμφωνα με τα χαρακτηριστικά του δικτύου ή της συσκευής παρουσίασης, δηλαδή του περιβάλλοντος της εφαρμογής.
- Ανταλλαγή του περιεχομένου σε ετερογενή περιβάλλοντα ανεξάρτητα από τον τρόπο διαχείρισής του από το πληροφοριακό σύστημα.

Η επίτευξη των παραπάνω στόχων πραγματοποιείται στην παρούσα διατριβή με την *περιγραφή του περιεχομένου* του κωδικοποιημένου οπτικοακουστικού υλικού του MPEG-4. Περιγραφή του περιεχομένου των πολυμέσων σημαίνει κατασκευή ενός υψηλού επιπέδου αναπαράστασης το οποίο συλλαμβάνει, εκφράζει ή και εμπλουτίζει τη δομή και τη σημασιολογία των δεδομένων. Ο πιο διαδεδομένος τρόπος αναπαράστασης περιγραφών επιτυγχάνεται με τη χρήση *γλωσσών περιγραφής*. Οι γλώσσες περιγραφής είναι ουσιαστικά γλώσσες αναπαράστασης οι οποίες αφορούν στην αναπαράσταση συγκεκριμένου είδους πληροφορίας, η οποία δεν αποτελεί το καθαυτό υλικό, αλλά είναι *μεταπληροφορία* ή *μεταδεδομένα* για το υλικό. Το ίδιο το υλικό μπορεί να αναπαρίσταται με διαφορετικό τρόπο και συγκεκριμένα στην περίπτωση των οπτικοακουστικών μέσων είναι κωδικοποιημένο και συμπιεσμένο. Η έννοια μεταδεδομένα έχει ήδη χρησιμοποιηθεί σε εξυπηρετητές πολυμέσων (όπως στο πρότυπο DAVIC), αλλά η αναπαράστασή τους συνήθως δε γινόταν με γλώσσες. Τα τελευταία χρόνια αναπτύχθηκαν γλώσσες αναπαράστασης μεταδεδομένων για έγγραφα του Παγκόσμιου Ιστού, ενώ η πρώτη προσπάθεια χρήσης γλωσσών περιγραφής για οπτικοακουστικό περιεχόμενο ξεκίνησε με το πρότυπο MPEG-7, αλλά δεν έχει ολοκληρωθεί ακόμα.

Όπως αναφέρεται στο κεφάλαιο 2, στο πρότυπο MPEG-4 η σύνταξη των ψηφιοσειρών δεδομένων εκφράζεται ήδη σε μια γλώσσα προγραμματισμού κατάλληλη για δυαδική κωδικοποίηση. Αυτή η γλώσσα ονομάζεται SDL και εκφράζει τη μορφή των κωδικοποιημένων δεδομένων όπως θα φτάσουν στον αποκωδικοποιητή. Εκτός από τη σύνταξη των ψηφιοσειρών, η SDL κωδικοποιεί και τις οπτικοακουστικές σκηνές που συντίθενται από ένα σύνολο αντικειμένων. Όμως η γλώσσα SDL είναι χαμηλού επιπέδου, κατάλληλη κυρίως για κωδικοποίηση και δεν μπορεί να εκφράσει τη σημασιολογία των δεδομένων σε υψηλό επίπεδο. Μια αναπαράσταση υψηλότερου επιπέδου είναι απαραίτητη για τη χρήση του περιεχομένου σε εφαρμογές που δε στοχεύουν μόνο στην αναπαραγωγή περιεχομένου MPEG-4, αλλά στην υποστήριξη αυξημένης λειτουργικότητας.

Στόχος της διατριβής ήταν να χρησιμοποιηθεί μια γλώσσα περιγραφής του περιεχομένου του MPEG-4, η οποία εκτός των άλλων να είναι κατάλληλη και για τον Παγκόσμιο Ιστό, ώστε να επιτευχθεί ενοποίηση του προτύπου MPEG-4, το οποίο αποτελεί πρότυπο

κωδικοποίησης οπτικοακουστικού περιεχομένου, με γνωστά πρότυπα αναπαράστασης εγγράφων του Παγκόσμιου Ιστού. Η γλώσσα HTML, η οποία αποτελεί την πιο δημοφιλή γλώσσα αναπαράστασης εγγράφων του Παγκόσμιου Ιστού, μπορεί να θεωρηθεί γλώσσα περιγραφής με την έννοια ότι περιγράφει τη δομή των εγγράφων, το περιεχόμενο των δομικών στοιχείων, αλλά και το μη δομημένο περιεχόμενο που μπορεί να αποτελεί τμήμα του εγγράφου μαζί με τη δομή. Το πρόβλημα με την HTML είναι ότι εστιάζει σε έγγραφα που περιέχουν κείμενο το οποίο οργανώνεται σε παραγράφους και άλλα δομικά στοιχεία που είναι χρήσιμα μόνο για την παρουσίαση και ανταλλαγή κειμένου.

Εν τούτοις η XML, η οποία αποτελεί μεταγλώσσα ορισμού γλωσσών αναπαράστασης ημιδομημένων εγγράφων και ανήκει στην ίδια οικογένεια με την HTML, είναι κατάλληλη για την αναπαράσταση εγγράφων με οποιοδήποτε περιεχόμενο. Επίσης, η απλότητα και η γενικότητά της την καθιστούν ιδανική για χρήση σε οποιοδήποτε περιβάλλον διαχείρισης οπτικοακουστικού περιεχομένου. Επειδή η γλώσσα XML χρησιμοποιεί κείμενο επιτρέπει την άμεση επεξεργασία της πληροφορίας που αναπαριστά και είναι κατάλληλη για ανταλλαγή δεδομένων ανάμεσα σε εφαρμογές του Παγκόσμιου Ιστού, αλλά και σε άλλες εφαρμογές. Λόγω των πλεονεκτημάτων που προσφέρει, η μεταγλώσσα XML επιλέχθηκε ως η καταλληλότερη για την περιγραφή των δεδομένων του MPEG-4. Δηλαδή η XML χρησιμοποιήθηκε στην παρούσα διατριβή για τον ορισμό μιας γλώσσας περιγραφής (ή αναπαράστασης) του περιεχομένου των δεδομένων του MPEG-4.

Επομένως, στο κεφάλαιο αυτό παρουσιάζουμε την αναπαράσταση του περιεχομένου οπτικοακουστικών δεδομένων, τα οποία είναι κωδικοποιημένα σύμφωνα με το πρότυπο MPEG-4, χρησιμοποιώντας μια γλώσσα περιγραφής που στηρίζεται στην XML. Τη γλώσσα αυτή ονομάζουμε *Γλώσσα Περιγραφής Οπτικοακουστικού Περιεχομένου* (Audio Visual Content Description Language ή AVDL). Η αναπαράσταση του περιεχομένου των δεδομένων σύμφωνα με την AVDL ονομάζεται *Περιγραφή του MPEG-4* (MPEG-4 Description).

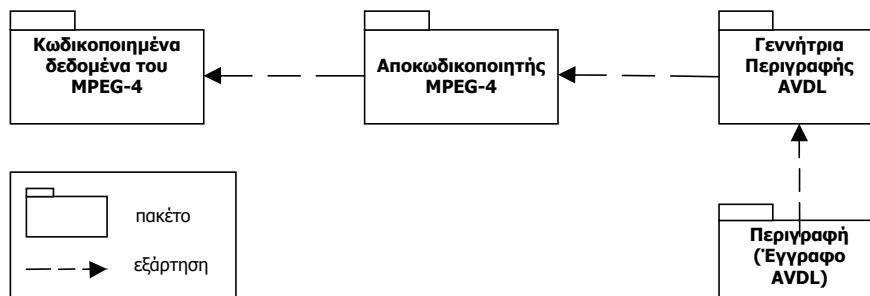
Ο στόχος της γλώσσας AVDL είναι να παρέχει μια ευρέως αποδεκτή, ευέλικτη και απλή αναπαράσταση του κωδικοποιημένου περιεχομένου του MPEG-4, ούτως ώστε να μπορεί αυτό να χρησιμοποιηθεί στον πυρήνα ενός πληροφοριακού συστήματος για την υποστήριξη των βασικών υπηρεσιών οπτικοακουστικών εφαρμογών

Στην ενότητα 3.2 παρουσιάζεται η χρήση της περιγραφής σε περιβάλλοντα χρήσης και σε συστήματα υποστήριξης εφαρμογών. Στην ενότητα 3.3 παρουσιάζονται με λεπτομέρεια τα χαρακτηριστικά της περιγραφής του περιεχομένου του MPEG-4 και στις επόμενες ενότητες γίνεται αναλυτική παρουσίαση γλώσσας της περιγραφής του MPEG-4.

3.2 Χρήση της AVDL

Η Γλώσσα Περιγραφής Οπτικοακουστικού Περιεχομένου (AVDL) χρησιμοποιείται για τον ορισμό της περιγραφής δεδομένων που είναι κωδικοποιημένα σύμφωνα με το πρότυπο MPEG-4. Η περιγραφή έχει σκοπό να υποκαταστήσει και να εμπλουτίσει την κωδικοποιημένη μορφή των δεδομένων σε περιβάλλοντα και εφαρμογές όπου απαιτείται ευκολότερη διαχείριση του περιεχομένου και πλουσιότερη λειτουργικότητα, όπως αλληλεπίδραση, εξατομίκευση και ανάκτηση.

Η περιγραφή των αντικειμένων του MPEG-4 χρησιμοποιείται είτε από το πληροφοριακό σύστημα που υποστηρίζει τις εφαρμογές, είτε από το ίδιο το σύστημα που τρέχει τις εφαρμογές. Για να χρησιμοποιηθεί η περιγραφή σε ένα σύστημα, πρώτα κατασκευάζεται με βάση τα κωδικοποιημένα οπτικοακουστικά δεδομένα. Αυτό επιτυγχάνεται με τη χρήση αποκωδικοποιητή του MPEG-4, ο οποίος παρέχει πληροφορία στη γεννήτρια της περιγραφής για την κατασκευή της. Η διαδικασία αυτή φαίνεται στο παρακάτω διάγραμμα, το οποίο εκφράζεται με τα *διαγράμματα πακέτων* (package diagrams) της γλώσσας UML (Unified Modeling Language). Τα πακέτα (packages) συμβολίζουν τμήματα του συστήματος σε μια αντικειμενοστρεφή αρχιτεκτονική, ενώ οι διακεκομμένες γραμμές που τα συνδέουν εκφράζουν τις εξαρτήσεις μεταξύ τους. Ουσιαστικά τα πακέτα στη γλώσσα UML προέρχονται από ομαδοποιήσεις κλάσεων υλοποίησης του συστήματος. Στην συγκεκριμένη περίπτωση, οι κλάσεις που υλοποιούν τα πακέτα αποτελούν λεπτομέρεια που δεν εμφανίζεται στο διάγραμμα.

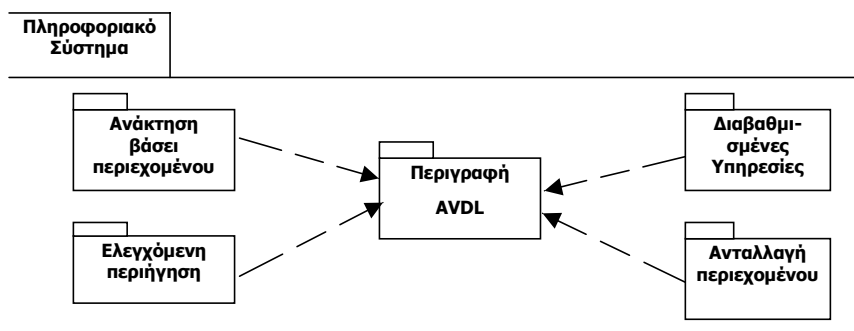


Σχήμα 3-2. Δημιουργία της περιγραφής στη γλώσσα AVDL

Το διάγραμμα πακέτων δείχνει τις εξαρτήσεις ανάμεσα στα τμήματα της αρχιτεκτονικής του συστήματος που είναι υπεύθυνα για την κατασκευή της περιγραφής των δεδομένων σύμφωνα με τη γλώσσα AVDL. Από δεξιά προς τα αριστερά του διαγράμματος, η Περιγραφή δημιουργείται από τη γεννήτρια περιγραφής, η οποία με τη σειρά της εξαρτάται από την αποκωδικοποίηση των δεδομένων. Η Περιγραφή, η Γεννήτρια Περιγραφής, ο Αποκωδικοποιητής του MPEG-4 και τα ίδια τα κωδικοποιημένα δεδομένα,

μοντελοποιούνται ως πακέτα (ομάδες) κλάσεων. Οι κλάσεις υλοποίησης των πακέτων δε φαίνονται στο διάγραμμα.

Το επόμενο στάδιο μετά τη δημιουργία της περιγραφής, είναι να χρησιμοποιηθεί για να επιτευχθεί η επιθυμητή λειτουργικότητα. Στο σημείο αυτό διακρίνονται δύο περιπτώσεις χρήσης: στον πυρήνα του πληροφοριακού συστήματος ή στην ίδια την εφαρμογή. Η χρήση της περιγραφής στον πυρήνα του συστήματος διαχείρισης οπτικοακουστικού περιεχομένου αποτελεί στόχο της διατριβής, ο οποίος θα υποστηριχθεί από ένα σύνολο εργαλείων που αυτοματοποιούν την εν λόγω διαδικασία και καθιστούν τη χρήση της περιγραφής διαφανή για την εφαρμογή και το χρήστη. Η επίτευξη αυτού του στόχου επιτυγχάνεται και στα επόμενα κεφάλαια της διατριβής και παρουσιάζεται ολοκληρωμένη στο Κεφάλαιο 6. Στο σχήμα που ακολουθεί απεικονίζεται η χρήση της περιγραφής στον πυρήνα του πληροφοριακού συστήματος για την υποστήριξη των βασικών λειτουργιών διαχείρισης οπτικοακουστικού περιεχομένου με διαφανή τρόπο.

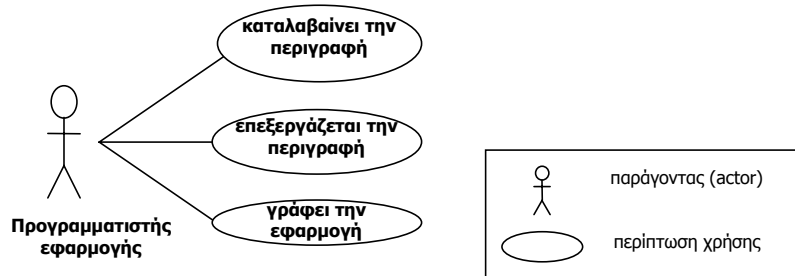


Σχήμα 3-3. Χρήση της AVDL για την υποστήριξης λειτουργιών διαχείρισης οπτικοακουστικού περιεχομένου

Το πληροφοριακό σύστημα διαχείρισης οπτικοακουστικού περιεχομένου, περιέχει στον πυρήνα του το οπτικοακουστικό περιεχόμενο στη μορφή της γλώσσας AVDL. Η αναπαράσταση αυτή χρησιμοποιείται για την υποστήριξη των τεσσάρων βασικών λειτουργιών διαχείρισης που παριστάνονται ως πακέτα του συστήματος: ανάκτηση βάσεις περιεχομένου, ελεγχόμενη περιήγηση, διαβαθμισμένες υπηρεσίες μετάδοσης και ανταλλαγή περιεχομένου.

Στην περίπτωση όπου η περιγραφή χρησιμοποιείται και από την ίδια την εφαρμογή, είναι απαραίτητο ο προγραμματιστής της εφαρμογής να καταλαβαίνει και να γνωρίζει πως να χρησιμοποιήσει τη γλώσσα XML ή να έχει στη διάθεσή του ένα εργαλείο διαχείρισης XML εγγράφων. Όπως έχει ήδη αναφερθεί, η απλότητα της XML και η αναπαράστασή της με κείμενο διευκολύνουν ακόμα και την απευθείας διαχείριση της περιγραφής από τον προγραμματιστή, σε αντίθεση με τη γλώσσα SDL του MPEG-4 και τη μορφή BIFS

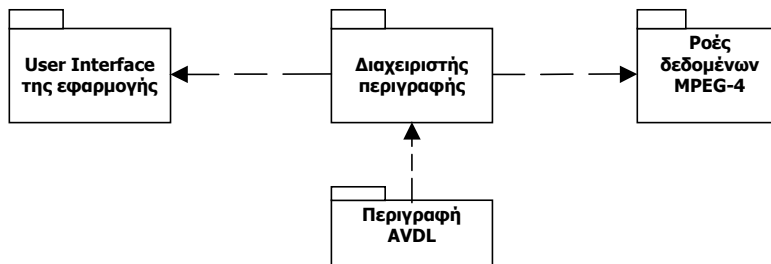
(αναπαράσταση των σκηνών). Η σχέση του προγραμματιστή εφαρμογής με την περιγραφή φαίνεται στο παρακάτω *διαγράμματος περίπτωσης χρήσης* (use case diagram) που εκφράζεται με τη γλώσσα UML. Τα διαγράμματα περιπτώσεων χρήσης γενικά μοντελοποιούν το ρόλο χρηστών ή εξωτερικών παραγόντων σε ένα σύστημα.



Σχήμα 3-4. Χρήση της περιγραφής από τον προγραμματιστή εφαρμογής

Το διάγραμμα αυτό δείχνει τη συσχέτιση του προγραμματιστή της εφαρμογής με την περιγραφή του περιεχομένου μέσω περιπτώσεων χρήσης (use cases). Η κατανόηση της εφαρμογής και η κατασκευή της αποτελούν χρήσεις που δεν υλοποιούνται στο σύστημα, αλλά παρουσιάζονται για λόγους ομοιομορφίας. Η επεξεργασία της περιγραφής έχει σημασία για το σύστημα γιατί εμφανίζεται και σε άλλες περιπτώσεις. Για παράδειγμα η αλληλεπίδραση με το υλικό που προκαλεί μεταβολές στην περιγραφή του υπονοεί τη δυνατότητα επεξεργασίας της (ενημέρωσης, κλπ).

Η εφαρμογή αποτελείται από διαφορετικά πακέτα (ή τμήματα), όπως το σύστημα διεπαφής (user interface) και ο διαχειριστής της περιγραφής. Οι ενέργειες του χρήστη μεταβάλλουν την κατάσταση του διαχειριστή της περιγραφής, επομένως υπάρχει εξάρτηση ανάμεσα στο σύστημα διεπαφής της εφαρμογής και τον διαχειριστή της περιγραφής, ο οποίος με τη σειρά του προκαλεί μεταβολές στην ίδια την περιγραφή.



Σχήμα 3-5. Διαχείριση της περιγραφής στην εφαρμογή

Στο διάγραμμα φαίνεται η εξάρτηση της περιγραφής από άλλα πακέτα του συστήματος. Κατά τη διάρκεια εκτέλεσης της εφαρμογής υπάρχουν αντικείμενα που υλοποιούν το διαχειριστή της περιγραφής και επιτρέπουν την επεξεργασία της. Ο διαχειριστής αυτός εξαρτάται από ενέργειες που προέρχονται από το σύστημα διεπαφής (user interface) της

εφαρμογής αλλά και από τα ίδια τα δεδομένα, τα οποία είναι δυνατόν να περιέχουν δυναμικές λειτουργίες της παρουσιάσής τους, άρα και της περιγραφής τους.

Ο χρήστης των οπτικοακουστικών εφαρμογών έρχεται σε επαφή ή χρησιμοποιεί τα οπτικοακουστικά δεδομένα με τρεις διαφορετικούς τρόπους: τα "αισθάνεται", δηλαδή τα βλέπει ή τα ακούει, αλληλεπιδρά μαζί τους μεταβάλλοντας τον τρόπο που τα "αισθάνεται" και τέλος, επιλέγει τα δεδομένα που τον ενδιαφέρουν. Η σχέση του χρήστη με το οπτικοακουστικό υλικό φαίνεται στο αριστερό μέρος του παρακάτω διαγράμματος περίπτωσης χρήσης. Το δεξί μέρος του διαγράμματος απεικονίζει τις περιπτώσεις χρήσης του περιεχομένου των οπτικοακουστικών δεδομένων, όπως ως περιεχόμενο θεωρούμε την πληροφορία που εξάγεται από τα δεδομένα ή εμπλουτίζει τα δεδομένα και αναπαρίσταται στην εφαρμογή. Το περιεχόμενο των οπτικοακουστικών μέσων είναι δυνατόν να είναι και το ίδιο οπτικοακουστικής φύσης (ακόμα και στην περίπτωση που είναι κείμενο, ο χρήστης το διαβάζει). Στις εφαρμογές επομένως όπου χρησιμοποιείται (αναπαρίσταται και διαχειρίζεται) το περιεχόμενο των οπτικοακουστικών μέσων, ο χρήστης έχει τη δυνατότητα να το "αισθανθεί" όταν αυτό του παρουσιάζεται, να αλληλεπιδράσει μαζί του και να επιλέξει συγκεκριμένο περιεχόμενο. Το διάγραμμα αυτό είναι αρκετά γενικό και περιλαμβάνει όλες τις κατηγορίες εφαρμογών. Σε ορισμένες εφαρμογές είναι δυνατόν να εφαρμόζεται ένα υποσύνολο των περιπτώσεων χρήσης του περιεχομένου, συμπεριλαμβανόμενης της περίπτωσης όπου ο χρήστης δεν έρχεται καθόλου σε επαφή με το περιεχόμενο.



Σχήμα 3-6. Χρήση των οπτικοακουστικών δεδομένων και του περιεχομένου τους

Σε όλες τις οπτικοακουστικές εφαρμογές ο χρήστης έχει στη διάθεσή του τρεις τρόπους για να διαχειριστεί το υλικό αλλά και το περιεχόμενό του: να το "δει" μέσα από την παρουσίασή του, να αλληλεπιδράσει με την τελευταία και να ανακτήσει υλικό και περιεχόμενο που να ταιριάζει στις ατομικές του ανάγκες.

Η διαχείριση του περιεχομένου πρέπει να είναι τέτοια ώστε να επιτρέπει στο χρήστη την εναλλαγή ανάμεσα σε υλικό και περιεχόμενο ούτως ώστε οι δύο αυτές διαφορετικές

αναπαραστάσεις της πληροφορίας να μπορούν να χρησιμοποιηθούν συμπληρωματικά η μία προς την άλλη. Η Γλώσσα Περιγραφής Οπτικοακουστικού Περιεχομένου που αναπτύχθηκε στα πλαίσια της διατριβής, παρέχει εύκολο τρόπο μεταβίβασης, ανταλλαγής και παρουσίασης του περιεχομένου μέσα από τον Παγκόσμιο Ιστό.

3.3 Χαρακτηριστικά της περιγραφής

Η περιγραφή του περιεχομένου του MPEG-4 που παρουσιάζεται στο κεφάλαιο αυτό έχει τα εξής γενικά χαρακτηριστικά:

- Εκμεταλλεύεται πλήρως τη δομή αλλά και τη σημασιολογία των δεδομένων του MPEG-4, αποτελεί δηλαδή πλήρη περιγραφή των κωδικοποιημένων δεδομένων. Η δομή περιέχει τη διάκριση των διακριτών αντικειμένων, τα χαρακτηριστικά τους και το σύνθετο περιεχόμενό τους. Η σημασιολογία των αντικειμένων περιλαμβάνει τη συσχέτιση μεταξύ τους.
- Εκφράζεται με τη χρήση γνωστών και ευρέως αποδεκτών προτύπων αναπαράστασης και συγκεκριμένα της XML.
- Είναι σε μορφή κατάλληλη για αποθήκευση αλλά και διαχείριση από διάφορα συστήματα και εφαρμογές, εφόσον είναι ανεξάρτητη της υλοποίησης των συστημάτων.
- Διευκολύνει την υλοποίηση λειτουργικότητας με βάση το περιεχόμενο.

Συγκεκριμένα το περιεχόμενο του MPEG-4 έχει τα εξής χαρακτηριστικά που είναι απαραίτητο να αναπαρασταθούν στην περιγραφή με τη γλώσσα AVDL:

- Σύνθετη δομή. Τα δεδομένα είναι κωδικοποιημένα με βάση το περιεχόμενό τους, δηλαδή με βάση τα αντικείμενα που περιέχουν. Τα αντικείμενα αυτά περιέχουν ένα σύνολο από δυαδικά πεδία ή χαρακτηριστικά αλλά είναι δυνατόν να εμπεριέχουν και άλλα αντικείμενα. Επομένως και στην περιγραφή των οπτικοακουστικών αντικειμένων πρέπει να υποστηρίζονται χαρακτηριστικά και σύνθετα (ή φωλιασμένα) στοιχεία.
- Ιεραρχική δομή. Οι σκηνές του MPEG-4 έχουν ιεραρχική δομή, αποτελούμενη από κόμβους οι οποίοι έχουν παιδιά άλλους κόμβους.
- Τύποι δεδομένων και περιορισμοί στις τιμές των δεδομένων. Πολλά από τα πεδία που αποτελούν τα κωδικοποιημένα αντικείμενα έχουν συγκεκριμένη σημασιολογία που καθορίζεται από περιορισμούς στο είδος και το εύρος των τιμών που δέχονται. Το είδος των τιμών αποτελεί τον τύπο των δεδομένων, ενώ το σύνολο τιμών εκφράζεται

ως περιορισμός. Η περιγραφή είναι αναγκαίο να μπορεί να εκφράσει τέτοιου είδους περιορισμούς στα χαρακτηριστικά των δομικών στοιχείων των εγγράφων.

- Περιορισμοί στον πληθάρισμο αντικειμένων και χαρακτηριστικών. Τα αντικείμενα και τα πεδία που αποτελούν άλλα αντικείμενα συχνά χαρακτηρίζονται από έναν πληθάρισμο όσον αφορά τη σχέση "περιέχει" ή "είναι τμήμα του" που υπάρχει μεταξύ τους.

Όπως αναφέρεται και στο κεφάλαιο 2 (ενότητα 2.3), τα κυριότερα συστατικά ενός XML εγγράφου, μέσω των οποίων είναι αναγκαίο να οριστεί η περιγραφή των δεδομένων του MPEG-4 είναι τα εξής:

- Η Δήλωση Τύπου Εγγράφου (Document Type Declaration) η οποία καθορίζει ποιος είναι, που βρίσκεται, από τι αποτελείται πως θα γίνει επεξεργασία του τύπου του εγγράφου.
- Ο Ορισμός Τύπου Εγγράφου (Document Type Definition) που αποτελεί τον τύπο του εγγράφου, δηλαδή το πρότυπο σύμφωνα με το οποίο θα φτιαχτεί το έγγραφο. Ο Τύπος Εγγράφου αποτελείται από δηλώσεις τύπων δομικών στοιχείων. Δηλαδή τα δομικά στοιχεία ή απλά στοιχεία του εγγράφου, είναι στιγμιότυπα κάποιου τύπου δομικών στοιχείων που ορίζεται στον τύπο του εγγράφου και αποτελούν τα βασικά συστατικά στοιχεία των εγγράφων. Ο Τύπος Εγγράφου στην ουσία αποτελεί τη γραμματική της γλώσσας AVDL.
- Στιγμιότυπα στοιχείων που περιέχουν και τα δεδομένα του εγγράφου.
- Άλλες δηλώσεις, όπως δηλώσεις για "υπό συνθήκη" (conditional) τμήματα και οντότητες (entities) των εγγράφων.

Αναλυτικά η μεταγλώσσα XML και τα εργαλεία που παρέχει για την αναπαράσταση πληροφορίας χρησιμοποιούνται για τον ορισμό της περιγραφής των δεδομένων του MPEG-4, ως εξής:

- Περιγράφονται οι σύνθετες οντότητες των ψηφιοσειρών του MPEG-4. Παραδείγματα τέτοιων οντοτήτων είναι τα κωδικοποιημένα αντικείμενα αλλά και οι σκηνές. Για την περιγραφή των σύνθετων οντοτήτων χρησιμοποιούνται τα δομικά στοιχεία των εγγράφων της XML (elements).
- Εκφράζεται ο τύπος των κωδικοποιημένων οντοτήτων. Αυτό επιτυγχάνεται αντιστοιχίζοντας τους δυαδικούς κώδικες έναρξης (start codes) των κωδικοποιημένων αντικειμένων ή τμημάτων αυτών με λεκτική περιγραφή του τύπου τους, η οποία αποτελεί μέρος του τύπου στοιχείων που τα αναπαριστά.

- Τα πεδία των οπτικοακουστικών αντικειμένων αναπαρίστανται ως χαρακτηριστικά των στοιχείων. Ο τύπος των χαρακτηριστικών καθορίζεται από τη σημασιολογία και το πεδίο τιμών των δυαδικών πεδίων που παριστάνουν και εκφράζεται με τους τύπους δεδομένων της XML, αλλά και άλλους εξειδικευμένους τύπους δεδομένων.
- Περιγράφεται η σημασιολογία των στοιχείων που η ύπαρξή τους εξαρτάται από κάποια συνθήκη, δηλαδή από την τιμή κάποιου άλλου χαρακτηριστικού.
- Παρέχεται μηχανισμός υλοποίησης των προφίλ, δηλαδή χρήσης της περιγραφής σε ένα έγγραφο είτε ολόκληρης είτε ως προς ένα μέρος της. Αυτό επιτυγχάνεται με τμήματα υπό συνθήκη του τύπου των εγγράφων που είναι δυνατόν να μη συμπεριληφθούν στα έγγραφα.

Τα πλεονεκτήματα της κατασκευής της Γλώσσας Περιγραφής Οπτικοακουστικού Περιεχομένου σύμφωνα με την παραπάνω προσέγγιση είναι ότι:

- Διακρίνονται μόνο τα αντικείμενα που είναι σημαντικά για την περιγραφή και αναπαράσταση του περιεχομένου και τα οποία θα χρησιμοποιηθούν για την υποστήριξη των βασικών υπηρεσιών διαχείρισης. Οι πολύ χαμηλού επιπέδου κώδικες που δεν είναι χρήσιμοι εκτός της διαδικασίας αποκωδικοποίησης δεν περιλαμβάνονται στην περιγραφή.
- Παρέχεται επεκτασιμότητα στην περιγραφή με χρήση των μηχανισμών της XML. Έτσι η περιγραφή μπορεί να επεκταθεί με νέα χαρακτηριστικά για τα στοιχεία, αλλά και με νέα στοιχεία.
- Υπάρχει ευελιξία, εφόσον η περιγραφή μπορεί να χρησιμοποιηθεί είτε ολόκληρη είτε ως προς ένα μέρος της, επιλέγοντας το κατάλληλο προφίλ περιγραφής.

Για την περιγραφή του οπτικοακουστικού περιεχομένου με τη γλώσσα AVDL που στηρίζεται στη μεταγλώσσα XML, κατασκευάζεται ένας τύπος εγγράφων, ο οποίος ορίζει την αναπαράσταση των οπτικοακουστικών δεδομένων. Στην επόμενη ενότητα παρουσιάζεται με λεπτομέρεια ο τύπος των εγγράφων που περιγράφει τα οπτικοακουστικά δεδομένα του MPEG-4. Με βάση αυτόν τον τύπο εγγράφων κατασκευάζονται έγγραφα-στιγμιότυπα του που περιέχουν την περιγραφή οπτικοακουστικών δεδομένων .

3.4 Τύπος εγγράφων και προφίλ περιγραφής

Ο τύπος ενός XML εγγράφου αποτελείται από το σύνολο των δηλώσεων που μπορούν να χρησιμοποιηθούν μέσα στο έγγραφο για τον ορισμό (στιγμιοποίηση) ημιδομημένων στοιχείων. Για κάθε στοιχείο δηλαδή που ορίζεται σε ένα XML έγγραφο είναι απαραίτητο

να υπάρχει η δήλωση του τύπου του, δηλαδή η γραμματική του, καθώς και η συσχέτιση του με τους υπόλοιπους τύπους στοιχείων του εγγράφου. Η *Δήλωση Τύπου Εγγράφου* (Document Type Declaration) είναι το τμήμα των εγγράφων που περιέχει τη γραμματική τους. Ο τύπος ενός εγγράφου μπορεί να αποτελείται από εσωτερικές και εξωτερικές δηλώσεις. Οι εξωτερικές δηλώσεις καθορίζονται στη Δήλωση Τύπου με ένα *Ομοιόμορφο Δείκτη Πόρων* (URI) που "δείχνει" το έγγραφο που περιέχει τις δηλώσεις, κάπου στον Παγκόσμιο Ιστό. Όλες οι δηλώσεις μαζί (εσωτερικές και μη) συνιστούν τον *Τύπο Εγγράφου* (Document Type Definition ή DTD). Το όνομα του Τύπου Εγγράφου δηλώνεται στην αρχή των εγγράφων και συμπίπτει με το όνομα του στοιχείου που αποτελεί τη ρίζα του εγγράφου στην ιεραρχική του δομή. Επομένως το στοιχείο ρίζα είναι το πρώτο στοιχείο που δηλώνεται στον Τύπο Εγγράφου και ορίζεται στο έγγραφο. Στον Τύπο Εγγράφου η δήλωση του στοιχείου-ρίζα είναι η ρίζα των δηλώσεων όλων των άλλων στοιχείων και στο έγγραφο, το στοιχείο ρίζα περιέχει όλα τα άλλα στοιχεία του εγγράφου με την τρόπο που καθορίζεται στον Τύπο του.

Για τα οπτικοακουστικά δεδομένα των προτύπων MPEG, ο τύπος των εγγράφων ονομάζεται avdl από τα αρχικά της γλώσσας AVDL και η δήλωση του τύπου που γίνεται στην αρχή των εγγράφων, αποτελείται από το URL όπου βρίσκεται το έγγραφο που περιέχει τον Τύπο του Εγγράφου (Document Type Definition). Η σύνταξη της γλώσσας XML για τη δήλωση του τύπου εγγράφου φαίνεται στο σχήμα (Σχήμα 3–7. Δήλωση τύπου εγγράφου και στοιχείου - ρίζας) παρακάτω.

Το στοιχείο ρίζα ενός XML Document Type Definition είναι αυτό που περιέχει όλα τα άλλα δομικά στοιχεία της περιγραφής. Στη συγκεκριμένη περιγραφή όπως ήδη αναφέρθηκε έχει το όνομα avdl και η δομή του είναι η ρίζα της ιεραρχίας των οντοτήτων του MPEG-4 και του MPEG-2. Το στοιχείο αυτό αποτελείται από τα στοιχεία mpeg4 και mpeg2 που αποτελούν τη ρίζα της ιεραρχίας για κάθε ένα από τα πρότυπα. Επομένως το στοιχείο mpeg4 μπορεί να περιέχει μία ή περισσότερες από τις βασικές οντότητες που κωδικοποιεί το πρότυπο MPEG-4, δηλαδή οπτικά αντικείμενα, ακουστικά αντικείμενα, την περιγραφή σκηνών ή περιγραφείς αντικειμένων. Ο διαχωρισμός των οντοτήτων προκύπτει από το ότι αυτές κωδικοποιούνται χωριστά στο πρότυπο και μεταδίδονται σε διαφορετικές στοιχειώδεις ροές. Στο παρακάτω σχήμα φαίνεται η σύνταξη του στοιχείου ρίζα, το οποίο αποτελείται από ένα σύνολο αντικειμένων του MPEG-4, τα οποία με τη σειρά τους ανήκουν σε έναν από τους 4 τύπους στοιχείων.

```
<!-- Δήλωση τύπου εγγράφου και στοιχείου-ρίζας των εγγράφων -->
<!DOCTYPE avdl SYSTEM "http://www.music.tuc.gr/~georgia/AVDL/avdl.dtd">
<avdl>
</avdl>

<!-- Ορισμός του στοιχείου ρίζα στον Τύπο Εγγράφου -->
```

```

<!ELEMENT avdl (mpeg4 | mpeg2) >

<!-- Προφίλ περιγραφής του MPEG-4 -->
<!ENTITY % MPEG4.description "IGNORE">

<!-- Τμήμα που περιλαμβάνει το προφίλ του MPEG-4 υπό συνθήκη -->
<![ %MPEG4.description; [

<!ELEMENT mpeg4 (MPEG4Object*, DescriptionId*)>

<!ELEMENT MPEG4Object (Visual | Scene | ObjectDescriptor)>

<!-- Μηχανισμός επιλογής από τον τύπο του εγγράφου, ενός υποσυνόλου από
τους 4 βασικούς τύπους στοιχείων που αποτελούν το στοιχείο ρίζα, για να
χρησιμοποιηθούν στο έγγραφο -->

<!-- Δηλώσεις στον Τύπο Εγγράφου -->
<!ENTITY % MPEG4Visual.description "IGNORE">
<!ENTITY % MPEG4Scene.description "IGNORE">
<!ENTITY % MPEG4ObjectDescriptor.description "IGNORE">

<!-- Ο τύπος του στοιχείου Visual δηλώνεται μέσα σε ένα υπό συνθήκη τμήμα
στον Τύπο Εγγράφου. Η συνθήκη είναι η τιμή της οντότητας
MPEG4Visual.description -->

<![ %MPEG4Visual.description; [
    <!ELEMENT Visual ...>
]]>

<!-- Δήλωση στο έγγραφο για την χρήση του στοιχείου Visual.description
από τον Τύπο Εγγράφου -->

<!ENTITY % MPEG4Visual.description "INCLUDE">

<!-- Έγγραφο που χρησιμοποιεί το προφίλ περιγραφής οπτικών αντικειμένων
του MPEG-4 -->
<!DOCTYPE avdl SYSTEM "http://www.music.tuc.gr/~georgia/AVDL/avdl.dtd">
<!ENTITY % MPEG4.description "INCLUDE">
<!ENTITY % MPEG4Visual.description "INCLUDE">
<avdl>
<mpeg4>
<Visual>
</Visual>
</mpeg4>
</avdl>

```

Σχήμα 3-7. Δήλωση τύπου εγγράφου και στοιχείου - ρίζας

Ο συμβολισμός που ακολουθείται στην συγκεκριμένη περιγραφή είναι ότι τα ονόματα των δομικών στοιχείων (elements) αποτελούνται από λέξεις των οποίων τα αρχικά είναι κεφαλαία. Τα ονόματα των χαρακτηριστικών αποτελούνται από μικρά αρχικά και ανάμεσα σε διαφορετικές λέξεις του ονόματος ενός χαρακτηριστικού υπάρχει ο χαρακτήρας `_'.

Στην περιγραφή αυτή υπάρχει η ευελιξία να συμπεριληφθούν ή να αφαιρεθούν από συγκεκριμένα έγγραφα, κάποια από τα τμήματά που ορίζονται στον Τύπο Εγγράφου, δηλαδή η χρήση *Προφίλ Περιγραφής*. Αυτός ο μηχανισμός είναι χρήσιμος όταν ένα έγγραφο περιέχει στοιχεία μόνο ενός από τους 4 βασικούς τύπους, για παράδειγμα την

περιγραφή ενός μόνο οπτικού αντικειμένου. Η αφαίρεση όλων των άλλων τύπων στοιχείων (ακουστικά, σκηνές και περιγραφείς αντικειμένων), επιταχύνει σε μεγάλο βαθμό τη συντακτική ανάλυση του εγγράφου εφόσον ο Τύπος Εγγράφου είναι αρκετά μικρότερος.

Το παραπάνω επιτυγχάνεται με ειδικές δηλώσεις που υποστηρίζει η XML για το αν ένα συγκεκριμένο τμήμα του DTD θα συμπεριληφθεί ή θα αγνοηθεί κατά τη συντακτική ανάλυση. Το πρώτο στάδιο αυτού του μηχανισμού είναι στο Document Type Definition, δηλαδή στη γραμματική της γλώσσας AVDL, να οριστούν τα υπό συνθήκη τμήματα δηλώσεων MPEG4Visual.description, MPEG4Scene.description και MPEG4ObjectDescriptor.description. Κατόπιν καθένα από τα τέσσερα αυτά τμήματα δηλώσεων του DTD θα περιέχεται μέσα σε ένα υπό συνθήκη τμήμα (conditional section) του Τύπου Εγγράφου.

Σε περίπτωση που ο συγγραφέας της περιγραφής θελήσει να χρησιμοποιήσει κάποιο τμήμα της, τότε πρέπει να δηλώσει ότι το τμήμα που την περιέχει θα χρησιμοποιηθεί, μέσα στο ίδιο το έγγραφο όπως φαίνεται στο παραπάνω, όπου δηλώνεται ότι θα συμπεριληφθεί μόνο ο τύπος στοιχείων MPEG-4 Visual. Η δήλωση αυτή υπερισχύει της πρώτης που υπάρχει στον Τύπο Εγγράφου και καθορίζει ότι η γραμματική του αντίστοιχου τμήματος θα συμπεριληφθεί στη συντακτική ανάλυση της περιγραφής, δηλαδή θα είναι συντακτικό λάθος αν μέσα στο έγγραφο επιτρέπονται στοιχεία εκτός του τύπου Visual.

Εκτός από τα στοιχεία της AVDL που αφορούν σε ένα από τα δύο είδη οπτικοακουστικού περιεχομένου MPEG-4 και MPEG-2, υπάρχουν στοιχεία και χαρακτηριστικά που είναι κοινά και για τα δύο και ορίζονται στον Τύπο Εγγράφου της γλώσσας και συγκεκριμένα στο τμήμα των κοινών δηλώσεων που δεν περιέχονται σε κάποιο από τα υπό συνθήκη τμήματα. Οι κοινές αυτές δηλώσεις διακρίνονται από τα υπόλοιπα στοιχεία στις επόμενες που αφορούν στην παρουσίαση της γλώσσας AVDL.

3.5 Σημασιολογικές δομές της AVDL

Σε αυτήν την ενότητα περιγράφονται οι σημασιολογικές δομές της AVDL, οι οποίες αποτελούνται από τύπους δεδομένων για τα χαρακτηριστικά των στοιχείων, περιορισμούς στις τιμές των χαρακτηριστικών και συσχετίσεις ανάμεσα στους τύπους στοιχείων.

Οι τύποι δεδομένων της AVDL βοηθούν στο να εκφραστεί η σημασιολογία των κωδικοποιημένων οπτικοακουστικών αντικειμένων. Η ανάγκη για την ύπαρξη τύπων δεδομένων προέρχεται από τη φύση των κωδικοποιημένων δεδομένων του MPEG-4, τα οποία περιέχουν πληροφορία διαφόρων ειδών. Οι τύποι δεδομένων που είναι ανάγκη να οριστούν από την AVDL είναι:

- Αριθμητικά δεδομένα. Τα αριθμητικά δεδομένα αποτελούν μεγάλο μέρος των επιμέρους χαρακτηριστικών των αντικειμένων και είναι κατά κανόνα ακέραιοι αριθμοί, δηλ. κώδικες.
- Αλφαριθμητικά δεδομένα (ή συμβολοσειρές). Οι συμβολοσειρές χρησιμοποιούνται σε τιμές χαρακτηριστικών που έχουν φραστική αναπαράσταση. Στη συγκεκριμένη περιγραφή χρησιμοποιούνται αλφαριθμητικά δεδομένα και σε περιπτώσεις όπου η κωδικοποίηση χρησιμοποιεί αριθμητικά, τα οποία αντιπροσωπεύουν συμβολικές έννοιες όταν αποκωδικοποιηθούν και αναλυθούν συντακτικά.
- Λογικά δεδομένα που έχουν τιμή 0 ή 1 ή αλήθεια/ ψέμα (TRUE/FALSE).
- Ταυτότητες αντικειμένων, οι οποίες χρησιμοποιούνται ως σημεία αναφοράς από άλλα αντικείμενα.

Εκτός από τον καθορισμό του τύπου των δεδομένων της AVDL, είναι συχνά απαραίτητο να εκφραστούν περιορισμοί στις τιμές ή τη δομή τους. Οι περιορισμοί στους οποίους είναι δυνατόν να υφίστανται τα δεδομένα του MPEG-4 είναι οι εξής:

- Σύνολο προκαθορισμένων τιμών αριθμητικών ή αλφαριθμητικών δεδομένων. Η τιμή ενός χαρακτηριστικού μπορεί να ανήκει σε ένα προκαθορισμένο σύνολο τιμών που περιγράφεται στον Τύπο Εγγράφου μαζί με το χαρακτηριστικό.
- Περιορισμοί στο εύρος τιμών αριθμητικών δεδομένων. Η συνηθέστερη περίπτωση είναι η τιμή ενός αριθμητικού κώδικα να περιορίζεται σε ένα διάστημα τιμών από μία τιμή έως μία άλλη. Το διάστημα αυτό μπορεί να είναι άπειρο ή και ανοικτό στο κάτω ή το άνω άκρο του.
- **Κανονική έκφραση αλφαριθμητικών δεδομένων.** Ένα χαρακτηριστικό που έχει τύπο συμβολοσειρά, δηλαδή κείμενο, είναι δυνατόν να υπόκειται σε μία κανονική έκφραση που καθορίζει τη μορφή της τιμής του.

Comment [Γ.Κ.1]: Να γίνει παρουσίαση των περιορισμών αναλυτικά και πως υποστηρίζονται...

Εκτός από περιορισμούς στις τιμές των δεδομένων, τα στοιχεία της AVDL οργανώνονται σε ιεραρχίες, όπου ένα στοιχείο "περιέχεται" μέσα σε ένα άλλο. Η συσχέτιση αυτή ανάμεσα στα στοιχεία συχνά έχει τη σημασιολογία της εξειδίκευσης (ή γενίκευσης) και είναι αναγκαίο να μπορεί να διαχωρίζεται από τη συσχέτιση "περιέχεται", ειδικά σε περιβάλλοντα όπου η περιγραφή θα αποθηκευτεί σε βάσεις δεδομένων, όπου τα ευρέως αποδεκτά μοντέλα ενθαρρύνουν τη δημιουργία ιεραρχιών εξειδίκευσης.

```
<!-- Κοινές δηλώσεις για όλα τα προφίλ της AVDL -->
<!-- Τύπος δεδομένων κείμενο -->
<!ENTITY % text "CDATA">
```

```

<!-- Τύπος δεδομένων ακέραιος -->
<!ENTITY % integer "CDATA">

<!-- Τύπος δεδομένων boolean -->
<!ENTITY % boolean "(TRUE|FALSE)">

<!-- Τύπος δεδομένων ταυτότητα -->
<!ENTITY % id_type "ID">

<!-- Τύπος δεδομένων αναφορά σε στοιχείο -->
<!ENTITY % ref_type "IDREF">

<!-- Τύπος δεδομένων ταυτότητα εγγράφου -->
<!ENTITY % docid_type "ENTITY">

<!-- Τύπος δεδομένων αναφορά σε έγγραφο -->
<!ENTITY % docref_type "ENTITY">

<!-- Εξειδίκευση ανάμεσα σε δύο τύπους στοιχείων με το χαρακτηριστικό
specializes -->
<!ENTITY % specializes "specializes CDATA %fixed;">

<!-- Οντότητες γενικής χρήσης -->
<!ENTITY % implied "IMPLIED">
<!ENTITY % required "REQUIRED">
<!ENTITY % fixed "FIXED">

```

Σχήμα 3-8. Τύποι δεδομένων, περιορισμοί και συσχετίσεις στοιχείων

Στο παραπάνω σχήμα παρουσιάζεται η σύνταξη της AVDL για την αναπαράσταση των τύπων δεδομένων που χρησιμοποιούνται στην περιγραφή του οπτικοακουστικού περιεχομένου, των περιορισμών στις τιμές τους και των συσχετίσεων εξειδίκευσης ή γενίκευσης. Οι δηλώσεις αυτές είναι κοινές για όλα τα προφίλ της γλώσσας.

3.6 Σύνθεση στοιχείων περιγραφής

Η περιγραφή του οπτικοακουστικού περιεχομένου του MPEG-4 που παρουσιάζεται στο παρόν κεφάλαιο της διατριβής, λειτουργεί ως ένα επίπεδο αναπαράστασης για τα δεδομένα του MPEG-4. Η αναπαράσταση αυτή αποτελεί σημαντικό μέρος της γενικής υποδομής υποστήριξης υπηρεσιών διαχείρισης οπτικοακουστικής πληροφορίας και χρησιμοποιείται για την υποστήριξη της αυξημένης λειτουργικότητας των εφαρμογών. Στο επίπεδο της περιγραφής βασικός στόχος είναι να διατηρηθούν τα κύρια χαρακτηριστικά των κωδικοποιημένων αντικειμένων, αλλά και οι συσχετίσεις μεταξύ τους, οι οποίες επιτυγχάνουν τη σύνθεσή τους σε σκηνές. Η ανάγκη συσχέτισης αντικειμένων στο MPEG-4, προκύπτει από το γεγονός ότι τα κωδικοποιημένα αντικείμενα μεταφέρονται σε διαφορετικές ροές γιατί καθένα έχει τις δικές του απαιτήσεις μεταφοράς. Η συσχέτιση αυτή επιτυγχάνεται στο πρότυπο MPEG-4 με τη χρήση περιγραφών αντικειμένων και ταυτοτήτων. Η σύνθεση των σκηνών υλοποιείται σε τρία στάδια:

- Η σκηνή αρχίζει από έναν περιγραφέα που καθορίζει την ταυτότητα της ροής η οποία περιέχει την περιγραφή της σκηνής.
- Οι κωδικοποιημένες οπτικοακουστικές σκηνές όταν θέλουν να αναφερθούν σε αντικείμενα μέσω των χρησιμοποιούν ταυτότητες περιγραφών αντικειμένων.
- Οι περιγραφείς αντικειμένων περιέχουν ταυτότητες ροών δεδομένων που μεταφέρουν τα αντικείμενα μέσω των.

Σύμφωνα με τα παραπάνω, ένας από τους στόχους της γλώσσας AVDL είναι να υποστηρίξει τη σύνθεση περιγραφών μεμονωμένων αντικειμένων, όπως αυτές προκύπτουν από τα κωδικοποιημένα δεδομένα, σε ολοκληρωμένες περιγραφές σκηνών. Οι αυτόνομες μεμονωμένες περιγραφές διαφορετικών αντικειμένων πρέπει επομένως να διατηρούν τη συσχέτιση των αντικειμένων στο επίπεδο της σκηνής, όπως αυτή υπήρχε στα κωδικοποιημένα δεδομένα.

Επιπλέον, η AVDL έχει ως στόχο να παρέχει ανεξαρτησία της αναπαράστασης του περιεχομένου από τα ίδια τα δεδομένα και να επιτρέψει τη σύνθεση εγγράφων με τυχαίο περιεχόμενο και οποιοδήποτε συνδυασμό περιγραφών αντικειμένων. Συνεπώς ένα αντικείμενο της γλώσσας είναι αναγκαίο να χρησιμοποιείται σε πολλά διαφορετικά έγγραφα ή σε πολλά σημεία μέσα στο ίδιο έγγραφο. Άρα ο διαχωρισμός και η ευέλικτη σύνθεση των αντικειμένων στο επίπεδο της περιγραφής είναι χρήσιμα σε περιπτώσεις όπως στη σύνθεση περιεχομένου, για την κατασκευή νέων σκηνών από υπάρχοντα αντικείμενα και στη χρήση των ίδιων περιγραφών από πολλές εφαρμογές.

Στην γλώσσα AVDL της διατριβής, το επίπεδο συσχέτισης των αντικειμένων του MPEG-4 διατηρείται αυτούσιο όπως στα κωδικοποιημένα δεδομένα. Αυτό επιτυγχάνεται διατηρώντας τις ταυτότητες των αντικειμένων του MPEG-4 στα στοιχεία περιγραφής. Στο επόμενο στάδιο γίνεται η συσχέτιση των στοιχείων περιγραφής μέσω των ταυτοτήτων τους. Στοιχεία που περιέχονται στο ίδιο έγγραφο συσχετίζονται με το μηχανισμό IDREF της XML. Εφόσον όμως υποστηρίζονται αυτόνομες περιγραφές, η συσχέτιση ταυτοτήτων μπορεί να είναι εξωτερική, δηλαδή ανάμεσα σε διαφορετικά έγγραφα. Αυτό επιτυγχάνεται με έναν μηχανισμό της γλώσσας AVDL, ο οποίος φαίνεται στο παρακάτω σχήμα. Ορίζεται το στοιχείο ExternalId το οποίο περιέχει την ταυτότητα ενός εξωτερικού στοιχείου που θα χρησιμοποιηθεί στο παρόν έγγραφο και το στοιχείο ExternalRef το οποίο παριστάνει την αναφορά σε ταυτότητα εξωτερικών εγγράφων ή στοιχείων. Οπότε ορίζονται δύο επιπλέον στοιχεία (DocRef και ElementRef) που αντιστοιχούν στους δύο τύπους αναφορών. Στο ίδιο σχήμα φαίνεται και ένα παράδειγμα εγγράφου που χρησιμοποιεί εξωτερική αναφορά στην ταυτότητα OD1 κάποιου αντικειμένου που περιέχεται στο έγγραφο doc1.AVDL σε κάποιο άλλο σημείο στον Παγκόσμιο Ιστό.

```

<!-- Κοινές δηλώσεις για όλα τα προφίλ της AVDL -->

<!-- Ταυτότητα εξωτερικού αντικειμένου -->
<!ELEMENT ExternalId (DocRef?)>
<ATTLIST ExternalId
    id      %id_type;    %implied;>

<!-- Αναφορά σε ταυτότητα αντικειμένου -->
<!ELEMENT ExternalRef (DocRef | ElementRef)>

<!-- Αναφορά σε εξωτερική ταυτότητα εγγράφου -->
<ATTLIST DocRef
    doc_ref      %docref_type;          %implied;>

<!-- Αναφορά σε εξωτερική ταυτότητα στοιχείου -->
<!ELEMENT ElementRef (DocRef)>
<ATTLIST ElementRef
    element_ref      %ref_type;    %implied;>

<!-- Παράδειγμα εγγράφου που περιέχει αναφορά σε ταυτότητα εξωτερικού
στοιχείου -->

<!-- A) Έγγραφο doc1.AVDL που περιέχει το στοιχείο με ταυτότητα ID10 -->

<!DOCTYPE avdl SYSTEM "http://www.music.tuc.gr/~georgia/AVDL/avdl.dtd">
<avdl>
    ...
    <... element_id="ID10"> ... </...>
    ...
</avdl>

<!-- B) Έγγραφο doc2.AVDL που αναφέρεται στο εξωτερικό στοιχείο ID10 -->

<!DOCTYPE avdl SYSTEM "http://www.music.tuc.gr/~georgia/AVDL/avdl.dtd">

<!-- Οντότητα που ορίζει το URL του εγγράφου doc1.avdl, στο οποίο
περιέχεται το στοιχείο ID10 -->

<ENTITY ExternalId1 PUBLIC
"http://www.music.tuc.gr/~georgia/AVDL/doc1.AVDL" >

<avdl>
    <ExternalId id=ID10"></ExternalId>
    ...
    <ExternalRef>
        <ElementRef element_ref="ID10">
            <DocRef doc_ref=ExternalId1> </DocRef>
        </ElementRef>
    </ExternalRef>
    ...
</avdl>

```

Σχήμα 3-9. Σύνθεση αυτόνομων περιγραφών από διαφορετικά έγγραφα

Εκτός από τον διαχωρισμό των αντικειμένων, η περιγραφή ήδη επιτρέπει την ομαδοποίηση αντικειμένων διαφορετικού τύπου μέσα στο ίδιο έγγραφο. Αυτό φαίνεται στο Σχήμα 3-7 όπου το στοιχείο ρίζα της περιγραφής μπορεί να περιέχει περισσότερα από ένα αντικείμενα διαφορετικού τύπου. Με το μηχανισμό που παρουσιάζεται σε αυτήν την ενότητα, η συσχέτιση των αντικειμένων διαφορετικού τύπου σε μία σκηνή μπορεί να γίνει ανεξάρτητα από το που βρίσκονται ή σε ποιο έγγραφο περιέχονται.

3.7 Περιγραφή οπτικών αντικειμένων

Τα οπτικά αντικείμενα του MPEG-4 περιγράφονται γιατί αποτελούν βασικό συστατικό του κωδικοποιημένου οπτικοακουστικού περιεχομένου, άρα και βασικό συστατικό της πληροφορίας που διαχειρίζεται η υποδομή υποστήριξης οπτικοακουστικών εφαρμογών. Η αναπαράστασή τους στη γλώσσα AVDL περιλαμβάνει την μοντελοποίηση της δομής και των χαρακτηριστικών τους, όπως αυτά περιέχονται στην κωδικοποιημένη πληροφορία. Το περιεχόμενο των οπτικών αντικειμένων στη γλώσσα AVDL, χρησιμοποιείται για την υποστήριξη διαβαθμισμένων υπηρεσιών μετάδοσης με βάση το προφίλ χρηστών ή συσκευών, καθώς και για ανάκτηση βάσει περιεχομένου. Στο υπόλοιπο της ενότητας 3.7 , παρουσιάζεται η αναπαράσταση των οπτικών αντικειμένων του MPEG-4 με τη γλώσσα AVDL.

Τα κωδικοποιημένα οπτικά αντικείμενα περιέχονται σε μια σειρά μαζί με άλλα δεδομένα. Η σειρά οπτικών αντικειμένων περιέχει ένα ή περισσότερα οπτικά αντικείμενα διαφόρων τύπων. Όπως φαίνεται στο παρακάτω σχήμα η οπτική πληροφορία αναπαρίσταται από τον τύπο στοιχείου Visual που αποτελείται από ένα στοιχείο τύπου VisualObjectSequence. Αυτό με τη σειρά του αποτελείται από 0 ή περισσότερα στοιχεία τύπου VisualObject.

Εκτός από οπτικά αντικείμενα μια σειρά αντικειμένων μπορεί να περιέχει και δεδομένα που δεν αποκωδικοποιούνται και ονομάζονται *δεδομένα χρήστη* (User Data). Επομένως ο τύπος στοιχείου UserData αποτελεί δεδομένα που παρέχονται από το χρήστη. Τα στοιχεία αυτού του τύπου χρησιμοποιούνται σε πολλά διαφορετικά σημεία της περιγραφής και μπορούν να θεωρηθούν δεδομένα κειμένου.

```
<!-- Στοιχείο που περιέχει οπτική πληροφορία -->
<!ELEMENT Visual (VisualObjectSequence)>

<!-- Σειρά οπτικών αντικειμένων -->
<!ELEMENT VisualObjectSequence (UserData*, VisualObject*)>

<!-- Δεδομένα χρήστη -->
<!ELEMENT UserData (#PCDATA)>

<!-- Τύπος στοιχείου οπτικού αντικειμένου -->
<!ELEMENT VisualObject
(MPEG4VisualProfileAndLevel, VersionAndPriority?, UserData*,
((VideoSignalType, VideoObject) |
 (VideoSignalType, StillTextureObject) |
 MeshObject | FaceObject)) >

<!-- Χαρακτηριστικά οπτικών αντικειμένων -->
<!ATTLIST VisualObject
    type (videoID | stillTextureID | meshID | faceID) %required;>

<!-- Έκδοση και προτεραιότητα οπτικών αντικειμένων -->
<!ELEMENT VersionAndPriority EMPTY>
<!ATTLIST VersionAndPriority
```

```

version      %int1_15; %implied;
priority     %int1_7;  %implied;>

<!ENTITY % int1_15 "(1|2|3|4|5|6|7|8|9|10|11|12|13|14|15)">
<!ENTITY % int1_7  "(1|2|3|4|5|6|7)">

```

Σχήμα 3–10. Οπτικά Αντικείμενα

Τα οπτικά αντικείμενα που αναπαρίστανται με στοιχεία του τύπου VisualObject περιέχουν την ένδειξη για το προφίλ και το επίπεδό τους, την έκδοση και την προτεραιότητά τους, που περιγράφονται με διαφορετικούς τύπους στοιχείων. Επίσης ανήκουν σε έναν από τους τέσσερις βασικούς τύπους οπτικών αντικειμένων: βίντεο, ακίνητη υφή, πλέγμα και πρόσωπο, όπως φαίνεται παραπάνω. Ο τύπος των οπτικών αντικειμένων αναπαρίστανται ως χαρακτηριστικό που παίρνει τιμές ένα σύνολο από συγκεκριμένες συμβολοσειρές και είναι απαραίτητο να παρέχεται στα στιγμιότυπα της περιγραφής.

Η έκδοση και η προτεραιότητα των οπτικών αντικειμένων αναπαρίστανται ως ένα στοιχείο που δεν περιέχει επιπλέον δομικά στοιχεία, αλλά δύο χαρακτηριστικά που λαμβάνουν αριθμητικές τιμές από 1 ως 15 και από 1 ως 7 αντίστοιχα. Το εύρος των τιμών των χαρακτηριστικών αναπαρίστανται με παραμετρικές οντότητες απαριθμημένων τύπων.

```

<!-- Προφίλ και επίπεδα οπτικών αντικειμένων MPEG-4 -->
<!ELEMENT MPEG4VisualProfileAndLevel
      (MPEG4VisualProfile, MPEG4Visual.Profile@Level)>

<!ELEMENT MPEG4VisualProfile
      (Simple | Core | Main | SimpleScaleable | 12Bit |
       BasicAnimated2DTexture | Animated2DMesh |
       SimpleFace | SimpleScaleableTexture |
       CoreScalableTexture )

<!-- Χαρακτηριστικά του απλού προφίλ -->
<!ENTITY % profile.simple.attr "
      IVOP          %boolean; %fixed; &quot; TRUE &quot;
      PVOP          %boolean; %fixed; &quot; TRUE &quot;
      ACDC_prediction %boolean; %fixed; &quot; TRUE &quot;
      slice_resync    %boolean; %fixed; &quot; TRUE &quot;
      data_partitioning %boolean; %fixed; &quot; TRUE &quot;
      reversible_VLC  %boolean; %fixed; &quot; TRUE &quot;
      4MV_Unrestricted %boolean; %fixed; &quot; TRUE &quot; ">

<!-- Απλό προφίλ -->
<!ATTLIST Simple
      %profile.simple.attr; >

<!-- Χαρακτηριστικά του βασικού προφίλ -->
<!ENTITY% %profile.core.attr "
      binary_shape_coding          %boolean; %fixed; &quot; TRUE &quot;
      MPEG2_quant_tables           %boolean; %fixed; &quot; TRUE &quot;
      PVOP_temp_scal_rect_shape    %boolean; %fixed; &quot; TRUE &quot;
      PVOP_temp_scal_arbitr_shape  %boolean; %fixed; &quot; TRUE &quot;
      BVOP                         %boolean; %fixed; &quot; TRUE &quot; ">

<!-- Βασικό προφίλ -->
<!ATTLIST Core

```

```

        %profile.simple.attr;
        %profile.core.attr; >

<!-- Κύριο προφίλ -->
<!ATTLIST Main
        %profile.simple.attr;
        %profile.core.attr;
        obmc                %boolean; %fixed; "TRUE"
        static_sprites      %boolean; %fixed; "TRUE"
        interlaced_tools    %boolean; %fixed; "TRUE"
        grayscale_alpha_shape %boolean; %fixed; "TRUE" >

<!-- Προφίλ σε Επίπεδο -->

<!ELEMENT MPEG4Visual.Profile@Level
        ( 12Bit@Level2 | Main@Level3 | Main@Level2 | Main@Level1 |
          Core@Level2 | Core@Level3 | SimpleScaleable@Level2 |
          SimpleScaleable@Level1 |
          Simple@Level1 | Simple@Level2 | Simple@Level3 )>

<!ATTLIST Main@Level2
        typical_session_size      %fixed; "CIF"
        max_total_num_of_objects  %fixed; "8"
        max_number_per_type       %fixed; "8"
        max_MBs_per_sec           %fixed; "23760"
        max_bitrate               %fixed; "2"
        max_enhancement_layers_per_object %fixed; "1" >

```

Σχήμα 3-11. Προφίλ και επίπεδα οπτικών Αντικειμένων

Το στοιχείο VideoSignalType όπως φαίνεται στο "Σχήμα 3-11. Προφίλ και επίπεδα οπτικών Αντικειμένων", περιέχεται σε αντικείμενα τύπου βίντεο ή ακίνητης υφής και περιγράφει τα χαρακτηριστικά του σήματος πριν την κωδικοποίηση. Αποτελείται από τη μορφή και το εύρος του σήματος βίντεο και είναι δυνατόν να περιέχει και πληροφορία για το χρώμα. Η μορφή του σήματος πριν την κωδικοποίηση ανήκει σε κάποιο από τα γνωστά πρότυπα σήματος βίντεο, δηλαδή Component, PAL, NTSC, SECAM, MAC. Υπάρχει επίσης η δυνατότητα η μορφή του σήματος να περιγραφεί ως ακαθόριστη. Το χρώμα περιγράφεται από τους συντελεστές του RGB, τα οπτοηλεκτρονικά χαρακτηριστικά του σήματος και τους συντελεστές υπολογισμού της χρωματότητας και φωτεινότητας από το σύστημα RGB. Η πληροφορία αυτή αναπαρίσταται με διαφορετικούς τύπους στοιχείων όπως φαίνεται στο παρακάτω σχήμα.

```

<!-- Χαρακτηριστικά οπτικών σημάτων βίντεο και ακίνητης υφής -->
<!ELEMENT VideoSignalType
        (VideoFormat, VideoRange, MPEG4ColorDescription?)>

<!-- Εύρος φωτεινότητας και χρωματότητας του βίντεο -->
<!ATTLIST VideoRange
        video_range ( Y16-35CbCr-112+112 | Y0-255CbCr-128+127 ) %implied;>

<!-- Περιγραφή των χαρακτηριστικών του χρώματος -->
<!ELEMENT MPEG4ColorDescription
        (MPEG4ColorPrimaries, TransferCharacteristics, MatrixCoefficients)>

<!-- Συστήματα περιγραφής των βασικών χρωματικών στοιχείων -->

```

```

<!ELEMENT MPEG4ColorPrimaries
  (Unspecified | ITU-R_BT.709 | ITU-R_BT.470-2_System_M |
   ITU-R_BT.470-2_System_BG | SMPTE_170M | SMPTE_240M | Generic_film)>

<!-- Κοινές δηλώσεις για όλα τα προφίλ της AVDL -->

<!-- Μορφή του σήματος βίντεο πριν την κωδικοποίηση -->
<!ATTLIST VideoFormat
  video_format (Component|PAL|NTSC|SECAM|MAC|Unspecified) %implied; >

<!-- Στοιχείο ακαθόριστου περιεχομένου -->
<ELEMENT Unspecified ANY>

<!-- Βασικά χαρακτηριστικά των χρωμάτων -->
<!ENTITY % primary "primary CDATA %fixed;">
<!ENTITY % green "green CDATA %fixed;">
<!ENTITY % blue "blue CDATA %fixed;">
<!ENTITY % red "red CDATA %fixed;">
<!ENTITY % whiteD65 "whiteD65 CDATA %fixed;">
<!ENTITY % whiteC "whiteC CDATA %fixed;">

<!-- Χαρακτηριστικά του προτύπου ITU-R_BT.709 -->
<!ENTITY % common.ITU-R.709 "
%primary; "x y"
  %green; "0,300 0,600"
  %blue; "0,150 0,060"
  %red; "0,640 0,330" ">

```

Σχήμα 3-12.Χαρακτηριστικά του σήματος βίντεο οπτικών αντικειμένων

Το στοιχείο VideoRange δείχνει το επίπεδο του μαύρου (C_b) και το εύρος της φωτεινότητας (Y) και της χρωματότητας (C_r) που εξαρτώνται από τον αριθμό των bits για το βίντεο. Η τιμή του χαρακτηριστικού έχει δύο τιμές (0 ή 1 στο πρότυπο) και κάθε μία αντιστοιχεί σε ένα σύνολο εξισώσεων που υπολογίζουν τις τιμές των Y , C_b και C_r με παραμέτρους τον αριθμό των bits και κάποιες αναλογικές τιμές. Στην περιγραφή οι τιμές 0 και 1 έχουν αντικατασταθεί από συμβολικές εκφράσεις της σημασιολογίας των δύο συστημάτων. Για παράδειγμα, η τιμή 0 έχει αντικατασταθεί από την τιμή "Y16-35CbCr-112+112" που υπονοεί ότι η φωτεινότητα παίρνει τιμές από 16 ως 35, ενώ η χρωματότητα από -112 ως +112.

Το στοιχείο "TransferCharacteristics" περιγράφει τα οπτοηλεκτρονικά χαρακτηριστικά μεταφοράς της εικόνας πηγής και "MatrixCoefficients" είναι οι συντελεστές για την εξαγωγή φωτεινότητας και χρωματότητας από το σύστημα R,G,B. Και τα τρία αυτά χαρακτηριστικά ορίζονται ως προς κάποιο πρότυπο σύστημα με τρόπο ανάλογο με αυτόν που περιγράφεται το στοιχείο "ColorPrimaries". Η λεπτομερής περιγραφή των στοιχείων αυτών παρατίθεται στο παράρτημα Α.

3.7.1 Αντικείμενα Βίντεο

Τα αντικείμενα τύπου βίντεο αποτελούν εξειδικευμένο τύπο οπτικών αντικειμένων που αναπαρίστανται με τον τύπο στοιχείου "VideoObject". Η σχέση αυτή του τύπου των αντικειμένων βίντεο με τον τύπο των οπτικών αντικειμένων δηλώνεται με το χαρακτηριστικό "specializes" που έχει την τιμή του γενικού τύπου, δηλαδή "VisualObject". Η έννοια της εξειδίκευσης τύπων στοιχείων που χρησιμοποιείται εδώ δεν είναι ταυτόσημη με αυτήν που χρησιμοποιείται στο αντικειμενοστρεφές παράδειγμα για τη δημιουργία υποκλάσεων και ιεραρχιών κληρονομικότητας. Στην παρούσα περιγραφή η εξειδίκευση έχει μόνο εννοιολογική σημασία, η οποία προκύπτει από το ίδιο το πρότυπο MPEG-4. Δηλαδή, στο πρότυπο αναφέρεται ότι η δομή της σειράς οπτικών αντικειμένων είναι ένα σύνολο από οπτικά αντικείμενα, καθένα από τα οποία μπορεί να ανήκει σε ένα από τους τέσσερις βασικούς τύπους (βίντεο, ακίνητη υφή, πλέγμα, πρόσωπο). Η σχέση ότι ένα οπτικό αντικείμενο αποτελείται από ένα αντικείμενο βίντεο έχει ήδη εκφραστεί στη δομή του οπτικού αντικειμένου. Τα αντικείμενα βίντεο με το χαρακτηριστικό "specializes" δηλώνουν ότι είναι είδη οπτικών αντικειμένων.

Εκτός από το να δηλώνουν ότι αποτελούν εξειδίκευση του τύπου στοιχείου "VisualObject", τα αντικείμενα τύπου βίντεο είναι δυνατόν να αναφέρονται εκπεφρασμένα στο οπτικό αντικείμενο που τα περιέχει, μέσω της ταυτότητας του τελευταίου, όπως δηλώνει το χαρακτηριστικό "partof_VisualObject". Επίσης είναι δυνατόν να έχουν δική τους μοναδική ταυτότητα. Η ταυτότητα αυτή είναι προαιρετική εφόσον δεν επιβάλλεται από το πρότυπο ως μηχανισμός αναφοράς στο αντικείμενο βίντεο από άλλα αντικείμενα.

```
<!-- Δομή και χαρακτηριστικά των αντικειμένων βίντεο -->
<!ELEMENT VideoObject
    (UserData*, VideoObjectLayer*)>

<!ATTLIST VideoObject
    %specializes;          "VisualObject"
    partof_VisualObject    %ref_type; %implied;
    id                     %id_type; %implied; >
```

Σχήμα 3-13.Αντικείμενα Βίντεο

Όπως φαίνεται στο "Σχήμα 3-13.Αντικείμενα Βίντεο", τα αντικείμενα τύπου βίντεο αποτελούνται από ένα σύνολο Βαθμίδων Αντικειμένου Βίντεο (Video Object Layers). Οι Βαθμίδες Αντικειμένου Βίντεο υποστηρίζουν κλιμάκωση στην κωδικοποίηση. Ένα οπτικό αντικείμενο μπορεί δηλαδή να αντιστοιχεί σε μία ή περισσότερες βαθμίδες, μία από τις οποίες είναι η βαθμίδα βάσης και οι υπόλοιπες αποτελούν βελτίωση της βάσης, περιέχοντας επιπλέον πληροφορία για τον τρόπο με τον οποίο θα επιτευχθεί η βελτίωση. Η περιγραφή των βαθμίδων γίνεται στην επόμενη ενότητα.

3.7.1.1 Βαθμίδα Αντικειμένου Βίντεο

Οι Βαθμίδες Αντικειμένου Βίντεο είναι δύο ειδών: συμβατές με προηγούμενα πρότυπα και κανονικές βαθμίδες του MPEG-4. Διακρίνονται μεταξύ τους από τη δομή και το περιεχόμενό τους και κυρίως από το μέγεθος και τη λεπτομέρεια της επικεφαλίδας που περιέχουν στην αρχή, η οποία αποκωδικοποιείται πριν από τα υπόλοιπα δεδομένα της βαθμίδας. Το βίντεο με μικρή επικεφαλίδα είναι συμβατό ως προς το πρότυπο H.263 και περιγράφεται ως ξεχωριστό στοιχείο που ονομάζεται "VideoPlaneWithShortHeader". Η Βαθμίδα Αντικειμένου Βίντεο με λεπτομερή επικεφαλίδα αποτελείται από ένα επίπεδο βάσης και κανένα ή περισσότερα επίπεδα κλιμάκωσης, όπως φαίνεται παρακάτω.

```
<!-- Δομή της Βαθμίδας Αντικειμένου Βίντεο -->
<!ELEMENT VideoObjectLayer
    ((BaseLayer, ScalabilityLayer*) | VideoPlaneWithShortHeader) >

<!ENTITY % VOL.Common.attr "
    obmc_disable           %boolean; %implied;
    quantization_type      (FIRST_INVERSE | SECOND_INVERSE) %implied;
    error_resilient_disable %boolean; %implied;
    data_partitioned       %boolean; %implied;
    reversible_vlc         %boolean; %implied;" >

<!-- Χαρακτηριστικά των βαθμίδων του βίντεο -->
<!ATTLIST VideoObjectLayer
    id                %id_type; %required;
    partof_video_object %ref_type; %implied;
    shape             (rectangular | binary | binary_only |
    grayscale) %implied;
    VOP_time_increment_resolution %integer; %implied;
    fixed_VOP_rate      %boolean; %implied;
    random_accessible_vol %boolean; %implied;>
```

Σχήμα 3-14. Βαθμίδα Αντικειμένου Βίντεο

Χαρακτηριστικά της βαθμίδας είναι η ταυτότητά της και η ταυτότητα του αντικειμένου βίντεο στο οποίο περιέχεται. Κάθε βαθμίδα βίντεο έχει μοναδική ταυτότητα για να μπορεί μια βελτιωτική βαθμίδα να αναφέρεται σε συγκεκριμένη βαθμίδα βάσης. Επίσης κάθε βαθμίδα περιέχει το είδος του σχήματός της, πληροφορία για την υποδιαίρεση του χρόνου, καθώς και για το ρυθμό και το είδος των εικόνων από τις οποίες αποτελείται. Για την εξασφάλιση μοναδικότητας στην ταυτότητα της βαθμίδας και για να είναι δυνατή η αναφορά σε αυτήν από άλλα στοιχεία χρησιμοποιείται ο τύπος δεδομένων ID που ορίζεται από τη γλώσσα XML.

Το σχήμα της βαθμίδας μπορεί να είναι τετράγωνο, δυαδικό, μόνο δυαδικό ή κλίμακας του γκρι και περιγράφεται από το χαρακτηριστικό "shape" που έχει τιμές συμβολικές αναπαραστάσεις του είδους του σχήματος. Οι Βαθμίδες Αντικειμένου Βίντεο όπως περιγράφεται παρακάτω, αποτελούνται από ένα σύνολο Επιπέδων Αντικειμένου Βίντεο (Video Object Planes). Το χαρακτηριστικό "VOP_time_increment_resolution" δείχνει τον

αριθμό των υποδιαστημάτων που διαιρείται η χρονική μονάδα του ενός δευτερολέπτου, ενώ το χαρακτηριστικό "fixed_VOP_rate" δηλώνει αν όλα τα Επίπεδα Αντικειμένου Βίντεο (VOP) της βαθμίδας έχουν σταθερό ρυθμό εικόνων. Το χαρακτηριστικό "random_accessible_voi" δείχνει αν κάθε VOP μπορεί να αποκωδικοποιηθεί ατομικά ή όχι, δηλαδή αν είναι I-VOP ή όχι. Η ένδειξη αυτή δεν χρησιμοποιείται στην αποκωδικοποίηση, αλλά για τη διευκόλυνση λειτουργιών τυχαίας προσπέλασης.

Όλες οι Βαθμίδες Βίντεο (βάσης, κλιμάκωσης και μικρής επικεφαλίδας) έχουν ένα σύνολο από κοινά χαρακτηριστικά που περιγράφονται στην οντότητα "VOL.Common.attr". Στη βαθμίδα με μικρή επικεφαλίδα τα χαρακτηριστικά αυτά έχουν σταθερές, προκαθορισμένες τιμές. Γι' αυτό το λόγο στην παρούσα περιγραφή περιέχονται στη δομή κάθε είδους βαθμίδας χωριστά. Το χαρακτηριστικό "obmc_disable (overlapped block motion compensation disable)" δείχνει αν επιτρέπεται ή όχι η απόσβεση κίνησης σε επικαλυπτόμενα τμήματα. Το "quantization_type" δείχνει ποιος από δύο δυνατούς τύπους κβαντοποίησης χρησιμοποιείται. Αν ο τύπος είναι FIRST_INVERSE, χρησιμοποιείται η πρώτη αντίστροφη μέθοδος κβαντοποίησης και ορίζεται και το στοιχείο "FirstInverseQuantization", το οποίο περιέχει δύο πίνακες κβαντοποίησης. Αν ο τύπος έχει τιμή SECOND_INVERSE, χρησιμοποιείται η δεύτερη αντίστροφη μέθοδος κβαντοποίησης για την αντίστροφη κβαντοποίηση των συντελεστών DCT. Το χαρακτηριστικό "data_partitioned" υπάρχει αν το "error_resilient_disable" είναι FALSE και δηλώνει ότι τα δεδομένα των macroblock του βίντεο είναι τμηματοποιημένα, δηλαδή τα διανύσματα κίνησης είναι χωριστά από τα δεδομένα υψής (δηλ. τους συντελεστές DCT). Το "reversible_vlc" αντίστοιχα ορίζεται αν το "data_partitioned" είναι TRUE και δηλώνει ότι για την αποκωδικοποίηση των συντελεστών DCT χρησιμοποιούνται οι πίνακες αντίστροφου μεταβλητού μήκους που ορίζονται από το πρότυπο.

Τα στοιχεία που περιγράφουν τη βασική βαθμίδα και τις βαθμίδες κλιμάκωσης έχουν παρόμοια δομή. Περιέχουν πληροφορία για την έκδοση και την προτεραιότητα, παραμέτρους ελέγχου που είναι όμοιες και για τα δύο είδη βαθμίδων καθώς και ένα σύνολο από επίπεδα αντικειμένου βίντεο. Η διαφορά τους είναι στην περιγραφή του σχήματος της βαθμίδας όπως φαίνεται και παρακάτω.

```
<!-- Δομή της βαθμίδας βάσης -->
<!ELEMENT BaseLayer
  (VersionAndPriority?,
   VOLControlParameters?,
   BaseLayerShape?,
   UserData*,
   VideoObjectPlane?,
   (GroupOfVideoObjectPlane?, VideoObjectPlane)* ) >

<ATTLIST BaseLayer
  %specializes                                "VideoObjectLayer"
```

Comment [Γ.Κ.2]: Τα χαρακτηριστικά "data_partitioned" και "reversible_vlc" εξαρτώνται από τις τιμές άλλων χαρακτηριστικών και δεν έχει εκφραστεί. Θα μπορούσε να γίνει με attributes: dependency1 CDATA "data_partitioned exists when error_resilient_disable=FALSE" %fixed; ή depends.attr.data_partitioned CDATA "exists when error_resilient_disable=FALSE" %fixed;

```
partof_VideoObjectLayer      %ref_type; %implied;
%VOL.Common.attr; >
```

Σχήμα 3-15. Βαθμίδα Βάσης

Το στοιχείο VersionAndPriority της βαθμίδας που έχει οριστεί σε προηγούμενη ενότητα περιέχει τον αριθμό έκδοσης και την προτεραιότητα της βαθμίδας. Το στοιχείο VOLControlParameters περιέχει παραμέτρους ελέγχου της βαθμίδας. Οι παράμετροι αυτές είναι ο ρυθμός του βίντεο (bit_rate) που ορίζεται σε μονάδες των 400 bits/sec, το chroma_format που είναι σταθερά με τιμή 4:2:0 και το δίτιμο χαρακτηριστικό low_delay που όταν έχει την τιμή TRUE, σημαίνει ότι δεν περιέχονται B-VOPs στο VOL και κατά την αποκωδικοποίηση δεν υπάρχει η καθυστέρηση αναδιάταξης των επιπέδων (VOPs) της βαθμίδας.

Το σχήμα της βασικής βαθμίδας μπορεί να είναι τετράγωνο, να επιτρέπει τη μετάδοση παραμέτρων για αντικείμενα τύπου Sprite, να περιέχει την ακρίβεια του βίντεο, τον τύπο κβαντοποίησης, καθώς και παραμέτρους για την πολυπλοκότητα της πρόβλεψης.

```
<!-- Παράμετροι ελέγχου των Βαθμίδων του Βίντεο -->
<!ATTLIST VOLControlParameters
    aspect_ratio_info %integer; %implied;
    VOP_rate_code     %integer; %implied;
    bit_rate          %integer; %implied;
    vbv_buffer_size   %integer; %implied;
    chroma_format      %fixed; "4:2:0"
    low_delay          %boolean; %implied; >

<!-- Δομή και χαρακτηριστικά του σχήματος της βαθμίδας βάσης -->
<!ELEMENT BaseLayerShape
    (VOLRectangularShape?,
    SpriteEnable?,
    Not8BitPrecision?,
    FirstInverseQuantization?,
    DefineVOPComplexityEstimationHeader?) >
```

Σχήμα 3-16. Βαθμίδα Βάσης Βίντεο

Οι βαθμίδες κλιμάκωσης έχουν παρόμοια δομή με τις βασικές βαθμίδες, αλλά διαφορετικό περιεχόμενο. Περιέχουν χαρακτηριστικά που αφορούν στην περιγραφή της κλιμάκωσης. Ref_layer_id είναι η ταυτότητα της βαθμίδας που θα χρησιμοποιηθεί ως πρόβλεψη για τη βαθμίδα κλιμάκωσης. Το χαρακτηριστικό ref_layer_sampling_direct δείχνει αν η βαθμίδα αναφοράς έχει υψηλότερη ή ίδια ή χαμηλότερη ανάλυση από τη βελτιωτική βαθμίδα. Το χαρακτηριστικό enhancement_type δηλώνει αν η συγκεκριμένη βαθμίδα αποτελεί κλιμάκωση για ένα τμήμα του επιπέδου αναφοράς ή για όλη την περιοχή της βαθμίδας αναφοράς.

```
<!-- Δομή των βαθμίδων κλιμάκωσης -->
<!ELEMENT ScalabilityLayer
```



```

        (VersionAndPriority?,
         VOLControlParameters?,
         ScalabilityLayerShape,
         UserData*,
         VideoObjectPlane?,
         (GroupOfVideoObjectPlane?, VideoObjectPlane)* ) >

<!-- Χαρακτηριστικά των βαθμίδων κλιμάκωσης -->
<!ATTLIST ScalabilityLayer
    %specializes                "VideoObjectLayer"
    partof_VideoObjectLayer    %ref_type; %implied;
    ref_layer_id                %ref_type; %required;
    enhancement_type            (PARTIAL | ENTIRE) %implied;
    ref_layer_sampling_direct   (SAME_OR_LOWER | HIGHER) %implied;
    %VOL.Common.attr; >

<!-- Δομή και χαρακτηριστικά του σχήματος των βαθμίδων κλιμάκωσης -->
<!ELEMENT ScalabilityLayerShape
    (VOLRectangularShape?,
     SpriteEnable?,
     Not8BitPrecision?,
     FirstInverseQuantization?,
     DefineVOPComplexityEstimationHeader?) >

<!ATTLIST ScalabilityLayerShape
    hor_sampling_factor_numerator    %integer; %implied;
    hor_sampling_factor_denominator  %integer; %implied;
    vert_sampling_factor_numerator    %integer; %implied;
    vert_sampling_factor_denominator %integer; %implied;>

```

Σχήμα 3-17. Βαθμίδα κλιμάκωσης

Το σχήμα των βαθμίδων κλιμάκωσης μπορεί να είναι τετράγωνο, να επιτρέπει τη μετάδοση παραμέτρων για αντικείμενα τύπου Sprite, να περιέχει την ακρίβεια του βίντεο, τον τύπο κβαντοποίησης, καθώς και παραμέτρους για την πολυπλοκότητα της πρόβλεψης. Τα κλάσματα οριζόντιας και κάθετης δειγματοληψίας που περιγράφονται ως χαρακτηριστικά του σχήματος της βαθμίδας αφορούν στη χωρική κλιμάκωση (spatial scalability).

Το στοιχείο VOLRectangularShape της Βαθμίδας υπάρχει αν το σχήμα της βαθμίδας είναι ορθογώνιο, δηλ. το χαρακτηριστικό shape του στοιχείου VideoObjectLayer, έχει την τιμή rectangular. Η εξάρτηση αυτή δηλώνεται με το χαρακτηριστικό "exists_when" του στοιχείου. Η τιμή του χαρακτηριστικού είναι η συμβολική περιγραφή της συνθήκης από την οποία εξαρτάται το στοιχείο VOLRectangularShape. Το πλάτος και το ύψος του τετράγωνου σχήματος ορίζονται στην ψηφιοσειρά σε μονάδες των pixels.

Το στοιχείο SpriteEnable σημαίνει την ύπαρξη Sprite στη βαθμίδα του βίντεο και περιγράφεται στη σελίδα 91, μαζί με τα υπόλοιπα στοιχεία που περιγράφουν τα αντικείμενα τύπου sprite.

```

<!ENTITY % RectangularShape.attr "
width %integer; %implied;
height %integer; %implied;" >

<!ATTLIST VOLRectangularShape

```

```

        partof_VideoObjectLayer %ref_type; %implied;
        exists_when CDATA %fixed; "VideoObjectLayer.attr.shape=rectangular"
% RectangularShape.attr; >

<!--ATTLIST Not8BitPrecision
bits_per_pixel      (4|5|6|7|8|9|10|11|12)    %implied;
quantization_precision (3|4|5|6|7|8|9)        %implied;
no_gray_quant_update  (FIXED_QUANTIZER | UPDATEABLE_QUANTIZER)
%implied; >

```

Η ύπαρξη του στοιχείου Not8BitPrecision σημαίνει ότι το βίντεο δεν έχει ακρίβεια 8 διφία ανά εικονοστοιχείο (bits per pixel). Ο αριθμός των διφίων ανά εικονοστοιχείο παίρνει τιμές από 4 ως 12 και επιτρέπει σε κάθε βαθμίδα να έχει τη δική της ακρίβεια, ανεξάρτητα από τις υπόλοιπες βαθμίδες της σειράς του βίντεο. Η ακρίβεια κβαντοποίησης περιγράφεται με το quantization_precision και δηλώνει τον αριθμό των διφίων που χρησιμοποιούνται για τις παραμέτρους κβαντοποίησης. Όταν το βίντεο κωδικοποιείται με 8 bits per pixel, η ακρίβεια κβαντοποίησης θεωρείται ότι είναι 5. Το χαρακτηριστικό no_gray_quant_update υπάρχει αν το σχήμα της βαθμίδας έχει την τιμή "grayscale" και παίρνει δύο δυνατές τιμές. Η τιμή "FIXED_QUANTIZER" σημαίνει ότι χρησιμοποιείται σταθερός κβαντιστής για τα δεδομένα κλίμακας του γκρι, ενώ η τιμή "UPDATEABLE_QUANTISER" δηλώνει ότι ο κβαντιστής επαναυπολογίζεται σε κάθε macroblock.

Αν ο τύπος κβαντοποίησης της Βαθμίδας Βίντεο είναι 1 περιγράφεται το στοιχείο QuantizationType1, που δηλώνει ότι χρησιμοποιείται η πρώτη αντίστροφη μέθοδος κβαντοποίησης για τους συντελεστές DCT. Σε αυτήν χρησιμοποιούνται δύο πίνακες, ένας για τα intra blocks και ένας για τα non-intra blocks. Οι πίνακες περιέχουν 2 ως 64 αριθμούς των 8 bits.

```

<!-- Πρώτος τύπος κβαντοποίησης των βαθμίδων του βίντεο -->

<!--ELEMENT FirstInverseQuantization
(IntraQuantMat?,
NonIntraQuantMat?,
(IntraQuantMatGrayscale?,
NonIntraQuantMatGrayscale?))>

<!--ATTLIST FirstInverseQuantization
partof_VideoObjectLayer %ref_type; %implied;
exists_when CDATA %fixed;
"VideoObjectLayer.attr.quantization_type=TYPE1" >

<!--ELEMENT IntraQuantMat (#PCDATA)>
<!--ELEMENT NonIntraQuantMat (#PCDATA)>
<!--ELEMENT IntraQuantMatGrayscale (#PCDATA)>
<!--ELEMENT NonIntraQuantMatGrayscale (#PCDATA)>

```

Το στοιχείο DefineVOPComplexityEstimationHeader της Βαθμίδας Βίντεο, περιέχει τον τύπο της μεθόδου εκτίμησης, που μπορεί να έχει τιμές από 0 ως 3. Όταν η μέθοδος εκτίμησης (estimation_method) έχει την τιμή 0, ορίζονται από το πρότυπο επιπλέον παράμετροι. Στην πρώτη έκδοση του προτύπου MPEG-4 προτείνεται μόνο μία μέθοδος εκτίμησης.

Το στοιχείο ShapeComplexityEstimationDisable δηλώνει ότι δεν ορίζονται παράμετροι για το σχήμα. Περιέχει ένα σύνολο από πεδία που καθορίζουν παραμέτρους για το τι θα μεταδοθεί στον αποκωδικοποιητή. Το πεδίο opaque έχει τιμή 1 για να επιτραπεί η μετάδοση του ποσοστού των τμημάτων φωτεινότητας και χρωματότητας που κωδικοποιούνται με opaque mode . Ομοίως τα υπόλοιπα χαρακτηριστικά αναφέρονται σε τρόπους κωδικοποίησης των τμημάτων φωτεινότητας και χρωματότητας. Το πεδίο transparent επιτρέπει τη μετάδοση του ποσοστού των blocks φωτεινότητας και χρωματότητας που είναι κωδικοποιημένα με διαφανή τρόπο. Τα χαρακτηριστικά intra_cae και inter_cae επιτρέπουν τη μετάδοση του ποσοστού των τμημάτων που έχουν κωδικοποιηθεί με τις μεθόδους IntraCAE και InterCAE αντίστοιχα. CAE (Context-based Arithmetic Encoding) είναι η μέθοδος που χρησιμοποιείται για την κωδικοποίηση του δυαδικού σχήματος. Παρόμοια, το χαρακτηριστικό no_update επιτρέπει τη μετάδοση του ποσοστού των τμημάτων που έχουν κωδικοποιηθεί με τρόπο "no update". Τα πεδία αυτά τελικά θα μεταδοθούν στο στοιχείο MotionShapeTexture (σελ. 64, όπου περιγράφεται η δομή του επιπέδου αντικειμένου βίντεο).

```
<!-- Πολυπλοκότητα των Επιπέδων Βίντεο -->
<!ELEMENT DefineVOPComplexityEstimationHeader
  (EstimationMethod0?)>

<!-- ATTLIST DefineVOPComplexityEstimationHeader
  partof_VideoObjectLayer %ref_type;          %implied;
  estimation_method (0|1|2|3) %implied; >

<!ELEMENT EstimationMethod0
  (ShapeComplexityEstimationDisable?,
  TextureComplexityEstimationSet1?,
  TextureComplexityEstimationSet2?,
  MotionCompensationComplexity?)?>

<!-- ATTLIST EstimationMethod0
  exists_when CDATA %fixed;
  "DefineVOPComplexityEstimationHeader.attr.estimation_method=0">

<!-- ATTLIST ShapeComplexityEstimationDisable
  opaque %boolean; %implied;
  transparent %boolean; %implied;
  intra_cae %boolean; %implied;
  inter_cae %boolean; %implied;
  no_update %boolean; %implied;
  up_sampling %boolean; %implied; >
```

Σχήμα 3-18. Πολυπλοκότητα επιπέδων βίντεο

Το στοιχείο TextureComplexityEstimationSet1 επιτρέπει τη χρήση του πρώτου συνόλου παραμέτρων για την υφή. Οι παράμετροι αυτοί είναι το ποσοστό των τμημάτων φωτεινότητας και χρωματότητας που είναι ενδοκωδικοποιημένα (Intra, Intra+Q), των τμημάτων που είναι κωδικοποιημένα με τη μέθοδο Inter ή Inter+Q, των τμημάτων που

είναι κωδικοποιημένα με τη μέθοδο Inter4V και το ποσοστό των τμημάτων που δεν είναι κωδικοποιημένα.

Το δεύτερο σύνολο παραμέτρων για την πολυπλοκότητα της υφής περιέχει μεγέθη για το ποσοστό των συντελεστών DCT στο μέγιστο αριθμό των συντελεστών στα κωδικοποιημένα τμήματα, το ποσοστό των 8x1 DCT γραμμών στο μέγιστο αριθμό τέτοιων γραμμών, το μέσο αριθμό VLC (Variable Length Code) συμβόλων ανά macroblock και το μέσο αριθμό ψηφίων για κάθε σύμβολο.

```
<!-- TextureComplexityEstimationSet1
      intra_blocks      %boolean; %implied;
      inter_blocks      %boolean; %implied;
      inter4v_blocks    %boolean; %implied;
      not_coded_blocks  %boolean; %implied; >

<!-- TextureComplexityEstimationSet2
      dct_coefs`        %boolean; %implied;
      dct_lines         %boolean; %implied;
      vlc_symbols       %boolean; %implied;
      vlc_bits          %boolean; %implied; >

<!-- MotionCompensationComplexity
      apm               %boolean; %implied;
      npm               %boolean; %implied;
      interpolate_mc_q  %boolean; %implied;
      forw_back_mc_q    %boolean; %implied;
      halfpel2          %boolean; %implied;
      halfpel4          %boolean; %implied; >
```

Οι παράμετροι πολυπλοκότητας της εκτίμησης κίνησης αποτελούνται από το ποσοστό των τμημάτων στα οποία έχει χρησιμοποιηθεί η πρόβλεψη APM (Advanced Prediction Mode) μέσα στο επίπεδο αντικειμένου βίντεο, το ποσοστό των τμημάτων στα οποία έχει χρησιμοποιηθεί κανονική πρόβλεψη (Normal Prediction Mode – NPM), το ποσοστό των τμημάτων στα οποία έχει χρησιμοποιηθεί παρεμβολή (interpolation) και των τμημάτων στα οποία έχει χρησιμοποιηθεί πρόβλεψη. Τα χαρακτηριστικά halfpel2 και halfpel4 επιτρέπουν τη μετάδοση του ποσοστού των τμημάτων που έχουν προβλεφθεί με διάνυσμα του μισού pixel σε μία ή σε δύο διαστάσεις αντίστοιχα.

3.7.1.2 Επίπεδο με μικρή επικεφαλίδα

Το στοιχείο VideoPlaneWithShortHeader παρέχει συμβατότητα με προηγούμενα πρότυπα κωδικοποίησης βίντεο. Η δομή και το περιεχόμενό του είναι αρκετά απλά. Αποτελείται από ένα στοιχείο που περιγράφει τη μορφή του σήματος πηγής και από ένα σύνολο βαθμίδων που περιέχουν ομάδες από blocks δεδομένων. Τα χαρακτηριστικά του είναι ένα σύνολο παραμέτρων με προκαθορισμένες τιμές, όπως φαίνεται και στο παρακάτω σχήμα.

```
<!-- Επίπεδο Βίντεο με μικρή επικεφαλίδα -->
<!-- ELEMENT VideoPlaneWithShortHeader
```

```

(SourceFormat, GOBLayer*) >
<!-- ATTLIST VideoPlaneWithShortHeader
    %specializes          "VideoObjectLayer"
    partof_VideoObjectLayer %ref_type; %implied;

    temporal_reference      %integer; %implied;
    document_camera_indicator %boolean; %implied;

    full_picture_freeze_release %boolean; %implied;
    picture_coding_type      ( I | P ) %implied;

<!-- κοινά χαρακτηριστικά όλων των τύπων βαθμίδων -->
    video_object_layer_shape %fixed; "rectangular"
    odbmc_disable            %fixed; TRUE
    quant_type              %fixed; SECOND_INVERSE
    error_resilient_disable %fixed; TRUE
    data_partitioned        %fixed; FALSE
    reversible_vlc          %fixed; FALSE

    block_count             %fixed; 6
    vop_rounding_type       %fixed; 0
    vop_fcode_forward       %fixed; 1
    vop_coded               %fixed; TRUE
    interlaced              %fixed; FALSE
    complexity_estimation_disable %fixed; TRUE
    use_intra_dc_vlc        %fixed; FALSE
    scalability             %fixed; FALSE
    not_8_bit               %fixed; FALSE
    bits_per_pixel          %fixed; 8
    vop_quant               CDATA %implied;
    psupp                   CDATA %implied; >

```

Σχήμα 3-19. Επίπεδο Βίντεο με μικρή επικεφαλίδα

Το χαρακτηριστικό "temporal_reference" μεταβάλλεται ανάμεσα στο σύνολο των επιπέδων με μικρή επικεφαλίδα που αποτελούν το βίντεο. Η τιμή του εξαρτάται από την τιμή του πεδίου στο προηγούμενο επίπεδο με σύντομη επικεφαλίδα και προκύπτει αυξάνοντας την προηγούμενη τιμή κατά 1 συν τον αριθμό των μη μεταδιδόμενων εικόνων από την προηγούμενη μεταδιδόμενη εικόνα. Το χαρακτηριστικό "document_camera_indicator" δεν έχει επίδραση στην αποκωδικοποίηση του VOP. Δηλώνει αν το περιεχόμενο του VOP πηγάει από document camera ή graphic αναπαράσταση, αντί για φυσικό βίντεο.

Το χαρακτηριστικό "full_picture_freeze_release" δηλώνει αν η συνέχιση της ανανέωσης των εικόνων πρέπει να ενεργοποιηθεί στην περίπτωση που το περιεχόμενο του βίντεο έχει για κάποιο λόγο παγώσει, για παράδειγμα εξαιτίας σφαλμάτων στη μετάδοση. Το πεδίο αυτό επίσης δεν έχει επίδραση στη διαδικασία αποκωδικοποίησης. Επίσης το βίντεο με μικρή επικεφαλίδα αποτελείται από εικόνες 2 ειδών: I ή P όπως φαίνεται από το χαρακτηριστικό "VOP_coding_type".

Όσον αφορά στα χαρακτηριστικά που περιέχονται σε όλους τους τύπους βαθμίδων βίντεο, το σχήμα του βίντεο με μικρή επικεφαλίδα είναι τετράγωνο, δεν επιτρέπεται η απόσβεση

κίνησης σε επικαλυπτόμενα τμήματα, ο τύπος κβαντοποίησης είναι δευτέρα αντίστροφη και τα δεδομένα του βίντεο δεν είναι τμηματοποιημένα.

Η τιμή του χαρακτηριστικού "VOP_rounding_type" είναι 0 δηλώνοντας ότι η παράμετρος που χρησιμοποιείται για παρεμβολή της τιμής των εικονοστοιχείων των επιπέδων τύπου P είναι 0. Για την αποκωδικοποίηση των διανυσμάτων κίνησης χρησιμοποιείται η παράμετρος "VOP_fcode_forward" με σταθερή τιμή 1. Η τιμή του χαρακτηριστικού "VOP_coded" δηλώνει ότι ακολουθούν και άλλα δεδομένα για αποκωδικοποίηση του επιπέδου. Επίσης το βίντεο δεν περιέχει πεπλεγμένες εικόνες, αλλά μόνο προοδευτικά κωδικοποιημένες. Στην αρχή κάθε επιπέδου βίντεο δεν υπάρχει επικεφαλίδα εκτίμησης της πολυπλοκότητας του επιπέδου και δεν επιτρέπεται η εναλλαγή ανάμεσα σε δύο διαφορετικούς κώδικες VLC για την αποκωδικοποίηση των συντελεστών Intra DC (χαρακτηριστικό "use_intra_dc_vlc").

Το στοιχείο "SourceFormat" περιγράφει τις διαστάσεις του σήματος πηγής ως προς τα γνωστά πρότυπα SubQCIF, QCIF, CIF, 4CIF, 16CIF. Για το πρώτο, οι διαστάσεις του βίντεο είναι 128x96, ο αριθμός των ομάδων από block είναι 6 σε κάθε Επίπεδο Αντικειμένου και ο αριθμός των macroblock σε κάθε ομάδα blocks είναι 8.

Το βίντεο με σύντομη επικεφαλίδα περιέχει έναν αριθμό από στοιχεία τύπου GOBLayer, ή βαθμίδες ομάδων από blocks. Ο αριθμός αυτός είναι χαρακτηριστικό του στοιχείου SourceFormat και για τις διάφορες μορφές όπως ορίζονται παραπάνω μπορεί να παίρνει τις τιμές {6,11,22,88,352}. Τα στοιχεία τύπου GOBLayer μπορούν να περιέχουν μια επικεφαλίδα και έναν αριθμό από macroblock.

Comment [Γ.Κ.3]: ????????

```
<!ELEMENT SourceFormat (SubQCIF | QCIF | CIF | 4CIF | 16CIF )>
<!ATTLIST SubQCIF
    vop_width           %integer; %fixed; '128'
    vop_height          %integer; %fixed; '96'
    gobs_in_vop         %integer% %fixed; '6'
    macroblocks_in_gob  %integer; %fixed; '8' >

<!ELEMENT GOBLayer (GOBHeader?, Macroblock*) >
<!ATTLIST GOBHeader
    gob_number          CDATA %implied;
    gob_frame_id        CDATA %implied;
    quant_scale         CDATA %implied; >
```

Το χαρακτηριστικό gob_number δείχνει ποια ομάδα από blocks μέσα στο επίπεδο είναι η συγκεκριμένη. Οι ομάδες αυτές κωδικοποιούνται σε αύξουσα σειρά στο βίντεο με σύντομη επικεφαλίδα. Το χαρακτηριστικό gob_frame_id υπάρχει για αντιμετώπιση της περίπτωσης να έχει χαθεί η επικεφαλίδα του επιπέδου βίντεο. Όλες οι ομάδες από blocks που ανήκουν στο ίδιο επίπεδο έχουν την ίδια τιμή στο χαρακτηριστικό gob_frame_id.

3.7.1.3 Επίπεδο αντικειμένου βίντεο

Η Βαθμίδα Αντικειμένου Βίντεο με μεγάλη επικεφαλίδα αποτελείται από μια σειρά από Επίπεδα Αντικειμένου Βίντεο (VideoObjectPlane) όπως περιγράφεται στη σελίδα 76. Τα επίπεδα αυτά αντικαθιστούν την έννοια των καρτέ ή εικόνων που υπήρχε στα προηγούμενα MPEG πρότυπα. Στην ψηφιοσειρά του βίντεο τα Επίπεδα Αντικειμένων Βίντεο σχηματίζουν ομάδες (GroupOfVideoObjectPlane) με κοινή επικεφαλίδα. Η επικεφαλίδα αυτή περιέχει το χρόνο παρουσίασης (time_code) του πρώτου επιπέδου της ομάδας. Ο χρόνος παρουσίασης αποτελείται από ώρες, λεπτά και δευτερόλεπτα όπως φαίνεται στο παρακάτω σχήμα.

Το χαρακτηριστικό "closed_gov" δηλώνει το είδος της πρόβλεψης που έχει χρησιμοποιηθεί στα πρώτα συνεχόμενα B-VOP της ομάδας και χρησιμοποιείται σε περιπτώσεις όπου γίνεται editing του βίντεο σε συνδυασμό με το χαρακτηριστικό "broken_link". Η τιμή 1 για το "closed_gov" δηλώνει ότι έχει χρησιμοποιηθεί προς τα πίσω πρόβλεψη ή ενδοκωδικοποίηση των επιπέδων τύπου B. Αν κατά την επέμβαση (editing) στο βίντεο έχουν αφαιρεθεί προηγούμενες εικόνες, η τιμή του "broken_link" γίνεται 1, έτσι ώστε τα B-επίπεδα που ακολουθούν να μην παρουσιαστούν.

Το επίπεδο αντικειμένου βίντεο δομείται ανάλογα με το είδος του σχήματός του. Συγκεκριμένα διακρίνεται η ειδική περίπτωση όπου το σχήμα είναι μόνο δυαδικό, δηλ. η τιμή του χαρακτηριστικού shape της βαθμίδας του βίντεο είναι binary_only. Τα υπόλοιπα Επίπεδα σχήματος ορθογώνιου, γκρι ή δυαδικού έχουν ομοιόμορφη δομή και χαρακτηριστικά. Τα χαρακτηριστικά του επιπέδου αντικειμένου βίντεο είναι ο τύπος του, δηλαδή αν είναι ενδοκωδικοποιημένο (τύπου I), αν περιέχει πρόβλεψη (τύπου P ή B) ή αν είναι sprite. Επίσης έχει χαρακτηριστικό την ταυτότητα της βαθμίδας που το περιέχει και είναι δυνατόν να έχει και δική του ταυτότητα.

```
<!-- Ομάδα Επιπέδων Αντικειμένου Βίντεο -->
<ELEMENT GroupOfVideoObjectPlane (UserData*, TimeCode)>

<!ATTLIST GroupOfVideoObjectPlane
    partof_VideoObjectLayer %ref_type;          %implied;
    refers_to_VOPs          IDREFS             %implied;
    closed_gov               (0|1) %implied;
    broken_link              (0|1) %implied; >

<!-- Κοινές δηλώσεις της AVDL για τα οπτικά αντικείμενα MPEG-4 και MPEG-2 -->
<!ATTLIST TimeCode
    time_code_hours          %hour;             %implied;
    time_code_minutes        %minutes;          %implied;
    time_code_seconds        %seconds;          %implied; >

<!ENTITY    % hours
" (0|1|2|3|4|5|6|7|8|9|10|11|12|13|14|15|16|17|18|19|20|21|22|23) " >

<!ENTITY    % minutes
```

```
"(0|1|2|3|4|5|6|7|8|9|10|11|12|13|14|15|
16|17|18|19|20|21|22|23|24|25|26|27|28|
29|30|31|32|33|34|35|36|37|38|39|40|41|
42|43|44|45|46|47|48|49|50|51|52|53|54|55|56|57|58|59)" >

<!ENTITY % seconds "%minutes;">
```

Σχήμα 3-20. Ομάδα Επιπέδων Αντικειμένου Βίντεο

Το στοιχείο ModuloTimeBase αναπαριστά την τοπική χρονική βάση σε δευτερόλεπτα. Για τα I και P Επίπεδα αναπαριστά τον αριθμό των δευτερολέπτων από το σημείο συγχρονισμού του προηγούμενου I ή P Επιπέδου. Για τα B Επίπεδα Αντικειμένου Βίντεο δείχνει τον αριθμό των δευτερολέπτων επίσης από το προηγούμενο I ή P Επίπεδο.

Το στοιχείο "VOPCoded" δηλώνει ότι υπάρχουν επιπλέον δεδομένα για το κωδικοποιημένο επίπεδο. Η ύπαρξη αυτών των δεδομένων δεν είναι υποχρεωτική και αν δεν υπάρχουν το Επίπεδο έχει συγκεκριμένο ρόλο: α) αν το σχήμα του επιπέδου είναι δυαδικό ή μόνο δυαδικό, τότε το ανακατασκευασμένο άλφα επίπεδο του είναι εντελώς διαφανές και β) αν το σχήμα του αντικειμένου είναι τετράγωνο, τότε το αντίστοιχο τετραγωνικό άλφα επίπεδο που έχει το ίδιο μέγεθος με τη συνιστώσα της φωτεινότητας είναι εντελώς διαφανές. Στην τελευταία περίπτωση, αν δεν υπάρχει άλφα επίπεδο κατά την αποκωδικοποίηση και σύνθεση του αντικειμένου, τότε το ανακατασκευασμένο Επίπεδο Αντικειμένου Βίντεο "γемίζεται" με το περιεχόμενο του αμέσως προηγούμενου Επιπέδου για το οποίο υπάρχουν επιπλέον κωδικοποιημένα δεδομένα.

```
<!-- Επίπεδο Αντικειμένου Βίντεο -->
<!ELEMENT VideoObjectPlane (ModuloTimeBase*, VOPCoded?)>

<!-- ATTLIST VideoObjectPlane -->
<ATTLIST VideoObjectPlane
    partof_VideoObjectLayer %ref_type;          %implied;
    id %id_type;                                %implied;
    VOP_coding_type (I|P|B|S) %implied;
    VOP_time_increment %integer; %implied;
    VOP_coded (FALSE | TRUE) %implied;>

<!-- Τοπική χρονική βάση -->
<!-- ATTLIST ModuloTimeBase -->
<ATTLIST ModuloTimeBase
    modulo_time_base %integer; %implied;>

<!-- ELEMENT VOPCoded -->
<!-- (ReadVOPComplexityEstimationHeader?, -->
<!-- (IVOP | PVOP | BVOP | Sprite))>
```

Σχήμα 3-21. Επίπεδο Αντικειμένου Βίντεο

Το στοιχείο "VOPCoded" όπως φαίνεται στο παραπάνω σχήμα είναι δυνατόν να περιέχει μια επικεφαλίδα για την εκτίμηση της πολυπλοκότητάς του και ένα αντικείμενο συγκεκριμένου τύπου κωδικοποίησης, δηλαδή ένα Επίπεδο τύπου I, P, B ή sprite.

Το στοιχείο "ReadVOPComplexityEstimationHeader" που περιέχεται στο επίπεδο αντικειμένου βίντεο ανεξάρτητα από τον τύπο του, ορίζει χαρακτηριστικά που εξαρτώνται από τη μέθοδο εκτίμησης της κίνησης καθώς και τον τύπο πρόβλεψης. Τα χαρακτηριστικά αυτά είναι παράμετροι εκτίμησης της πολυπλοκότητας και ορίζονται ως ποσοστιαία μεγέθη στο σύνολο των blocks από τα οποία αποτελείται το επίπεδο. Η ύπαρξη των χαρακτηριστικών αυτών εξαρτάται από τις τιμές των πεδίων του στοιχείου "DefineVOPComplexityEstimationHeader" που ορίστηκε στο "Σχήμα 3-18. Πολυπλοκότητα επιπέδων βίντεο". Οι παράμετροι για τον τύπο πρόβλεψης 0 φαίνονται στο παρακάτω σχήμα. Αν το χαρακτηριστικό " DefineVOPComplexityEstimationHeader.opaque" έχει τιμή TRUE, τότε περιέχεται το χαρακτηριστικό "dcecs_opaque" μέσα στο στοιχείο "ReadVOPComplexityEstimationHeader".

```
<!-- Πολυπλοκότητα Επιπέδων Βίντεο -->

<!ELEMENT ReadVOPComplexityEstimationHeader
(Estimation0PredictionParams?) >

<!ELEMENT Estimation0PredictionParams
(Prediction0Params | Prediction1Params |
 Prediction2Params | Prediction3Params) >

<!ELEMENT Prediction0Params EMPTY>
<!-- ATTLIST Prediction0Params
    dcecs_opaque           %percent; %implied;
    dcecs_transparent      %percent; %implied;
    dcecs_intra_cae        %percent; %implied;
    dcecs_inter_cae        %percent; %implied;
    dcecs_no_update        %percent; %implied;
    dcecs_upsampling       %percent; %implied;
    dcecs_intra_blocks      %percent; %implied;
    dcecs_not_coded_blocks  %percent; %implied;
    dcecs_dct_coefs         %percent; %implied;
    dcecs_dct_lines         %percent; %implied;
    dcecs_vlc_symbols       %percent; %implied;
    dcecs_vlc_bits          %percent; %implied;
-->
```

Αντικείμενα βίντεο τύπου I

Τα Επίπεδα Αντικειμένου Βίντεο τύπου I διακρίνονται ανάλογα με το σχήμα τους. Αν το σχήμα δεν είναι τετράγωνο, δηλαδή είναι δυαδικό ή κλίμακας του γκρι, περιγράφεται με τα στοιχεία "INotRectangularshape" και "VOPConstantAlpha" το άλφα επίπεδο του αντικειμένου. Η τιμή του στοιχείου VOPConstantAlpha αναπαριστά την κλίμακα που θα εφαρμοστεί στα δυαδικά και γκρι σχήματα κατά την προ-επεξεργασία της αποκωδικοποίησης. Ο τύπος κλιμάκωσης της τιμής των εικονοστοιχείων έχει ως εξής:

$$\text{Κλιμακωμένο_εικονοστοιχείο} = (\text{αρχικό_εικονοστοιχείο} * (\text{κλίμακα}+1))/256$$

Το χαρακτηριστικό "conv_ratio" του στοιχείου "INotRectangularshape" δηλώνει αν ο λόγος μετατροπής είναι ή όχι κωδικοποιημένος μέσα στη βαθμίδα των macroblock. Στη δεύτερη περίπτωση έχει τη σταθερή τιμή 1, όπως φαίνεται και στο σχήμα. Το

χαρακτηριστικό αυτό αποτελεί τον παράγοντα *upsampling* που εφαρμόζεται μετά την αποκωδικοποίηση του σχήματος.

Το στοιχείο "InotBinaryOnlyShape" περιγράφει τα επίπεδα τύπου I με σχήμα που δεν είναι "μόνο δυαδικό" και περιέχει στοιχεία για την περιγραφή πεπλεγμένου (interlaced) βίντεο, βίντεο κλίμακας γκρι και στοιχεία που αφορούν στην ύπαρξη ή όχι κλιμάκωσης. Περιέχει το χαρακτηριστικό "VOP_quant" που έχει την τιμή που θα χρησιμοποιηθεί για αποκβαντοποίηση του επόμενου macroblock και το χαρακτηριστικό "intra_dc_vlc" που επιτρέπει στον αποκωδικοποιητή να αλλάξει τη μέθοδο αποκωδικοποίησης των δυαδικών κωδικών μεταβλητού μεγέθους για τους συντελεστές DCT. Το χαρακτηριστικό αυτό λαμβάνει τιμές από 0 ως 7, όπου 0 σημαίνει ότι χρησιμοποιείται η μέθοδος "Intra DC VLC" για όλο το Επίπεδο, η τιμή 7 σημαίνει ότι χρησιμοποιείται η μέθοδος "Intra AC VLC" για όλο το επίπεδο και οι ενδιάμεσες τιμές επιτρέπουν αλλαγή της μεθόδου από DC σε AC.

```
<!-- Επίπεδο Αντικειμένου Τύπου I -->
<ELEMENT IVOP (INotRectangularShape?,
    (INotBinaryOnlyShape | BinaryOnlyShape))>

<!ATTLIST IVOP
    %specializes;                "VideoObjectPlane"
    part of_VideoObjectPlane     %ref_type; %implied;>

<!-- Μη τετραγωνικό σχήμα Επιπέδου τύπου I -->
<ELEMENT INotRectangularShape (VOPConstantAlpha?)>

<!ATTLIST INotRectangularShape
    conv_ratio (CODED | 1) %implied; >

<!ATTLIST VOPConstantAlpha
    vop_constant_alpha %integer; %implied;>

<!-- Μη "δυαδικό μόνο" σχήμα Επιπέδου τύπου I -->
<ELEMENT INotBinaryOnlyShape
    (Interlaced?, Grayscale?, (INoScalability | Scalability) )>

<!ATTLIST INotBinaryOnlyShape
    VOP_quant           %integer; %implied;
    intra_dc_vlc_thr    (0|1|2|3|4|5|6|7) %implied;>
```

Σχήμα 3-22. Επίπεδο Αντικειμένου Βίντεο τύπου I

Το στοιχείο Interlaced περιγράφει τα χαρακτηριστικά του βίντεο με πεπλεγμένη κωδικοποίηση. Αν αυτό το στοιχείο δεν ορίζεται στη δομή του επιπέδου, τότε υπονοείται ότι η κωδικοποίηση του βίντεο είναι προοδευτική. Το χαρακτηριστικό *top_field_first* δείχνει αν το πεδίο που περιέχει την πρώτη γραμμή των εικόνων, θα είναι και το πρώτο που θα παρουσιαστεί ή όχι. Αν έχει την τιμή 0, τότε το κάτω πεδίο θα παρουσιαστεί πρώτο. Το δίτιμο πεδίο *alternate_vertical_scan* όταν έχει τιμή 1 δείχνει ότι τα πεπλεγμένα VOP χρησιμοποιούν εναλλασσόμενη κάθετη σάρωση. Επίσης, αν το σχήμα του

αντικειμένου είναι κλίμακας του γκρι (grayscale), τότε ορίζεται η τιμή "VOP_alpha_quant" του κβαντιστή για τα δεδομένα.

Όπως φαίνεται στο παρακάτω σχήμα, ένα επίπεδο τύπου I που δεν έχει μόνο δυαδικό σχήμα είναι δυνατόν να περιέχει κλιμάκωση. Τα επίπεδα με μόνο δυαδικό σχήμα δεν περιέχουν κλιμάκωση. Στα Επίπεδα τύπου I χωρίς κλιμάκωση περιέχονται δεδομένα για την Υφή Κινούμενου Σχήματος (MotionShapeTexture). Το στοιχείο "MotionShapeTexture" περιγράφεται αναλυτικά παρακάτω. Το στοιχείο "Resync" αποτελείται από την επικεφαλίδα του πακέτου βίντεο και τα δεδομένα της υφής του κινούμενου σχήματος.

```
<!-- Πεπλεγμένο Επίπεδο Αντικειμένου Βίντεο -->
<!ATTLIST Interlaced
    top_field_first          %boolean; %implied;
    alternate_vertical_scan %boolean; %implied;>

<!-- Τιμή κβαντιστή για αντικείμενα κλίμακας του γκρι -->
<!ATTLIST Grayscale
    VOP_alpha_quant    CDATA %implied; >

<!-- Επίπεδο Αντικειμένου Βίντεο τύπου I χωρίς κλιμάκωση -->

<!ELEMENT INoScalability
    (IErrorResilientEnable | MotionShapeTexture)>

<!ELEMENT IErrorResilientEnable
    (MotionShapeTexture, IResync*)>

<!ELEMENT IResync
    (IVideoPacketHeader, MotionShapeTexture)>

<!ELEMENT IVideoPacketHeader
    (CommonVideoPacketHeader, IVPHEExtension?)>

<!-- Κοινά χαρακτηριστικά για όλες τις κεφαλίδες πακέτων βίντεο -->
<!ATTLIST CommonVideoPacketHeader
    macroblock_number    %integer; %implied;
    quant_scale          %integer; %implied; >

<!ELEMENT IVPHEExtension (CommonVPHEExtension)>

<!-- Κοινή επέκταση για όλες τις κεφαλίδες πακέτων βίντεο -->
<!ELEMENT CommonVPHEExtension (ModuloTimeBase*)>

<!ATTLIST CommonVPHEExtension
    VOP_time_increment    %integer; %implied;
    VOP_coding_type       (I|P|B|S) %implied;
    intra_dc_vlc_thr      %integer; %implied; >
```

Macroblock_number είναι ο αριθμός του macroblock μέσα στο επίπεδο αντικειμένου βίντεο. Το πάνω αριστερά macroblock του επιπέδου έχει τον αριθμό 0 και ο αριθμός αυξάνεται προς τα δεξιά και προς τα κάτω.

```
<!-- Δυαδικό μόνο σχήμα Επιπέδου τύπου I -->
<!ELEMENT BinaryOnlyShape
    (BinaryOnlyErrorResilientEnable | MotionShapeTexture)>

<!ELEMENT BinaryOnlyErrorResilientEnable
```

```

(MotionShapeTexture, BinaryOnlyResync*)>
<!-- ELEMENT BinaryOnlyResync
      (BinaryOnlyVideoPacketHeader, MotionShapeTexture)>
<!-- ELEMENT BinaryOnlyVideoPacketHeader (CommonVPHEExtension?)>
<!-- ATTLIST BinaryOnlyVideoPacketHeader
      macroblock_number CDATA %implied;>

```

Αντικείμενα βίντεο τύπου P ή B

Το στοιχείο που περιγράφει το μη ορθογώνιο σχήμα των αντικειμένων τύπου P ή B περιέχει τις διαστάσεις και τη θέση του ορθογωνίου που περικλείει το αντικείμενο τυχαίου σχήματος. Τα χαρακτηριστικά `horizontal_mc_spatial_ref` και `vertical_mc_spatial_ref` ορίζονται ως η οριζόντια και η κατακόρυφη θέση του πάνω αριστερά σημείου του αντίστοιχα και χρησιμοποιούνται για αποκωδικοποίηση και σύνθεση της εικόνας.

```

<!-- Επίπεδο αντικειμένου βίντεο τύπου P -->
<!-- ELEMENT PPOP
      (POrBNotRectangular?, (PNotBinaryOnlyShape | BinaryOnlyShape),
<!-- ELEMENT POrBNotRectangular (ShapeCoding?) >
<!-- ATTLIST POrBNotRectangular
      width                %integer; %implied;
      height               %integer; %implied;
      horizontal_mc_spatial_ref %integer; %implied;
      vertical_mc_spatial_ref %integer; %implied;
      change_conv_ratio_disable %boolean; %implied; >
<!-- ATTLIST ShapeCoding
      type                ( INTRA | INTER ) %implied;>
<!-- ELEMENT PNotBinaryOnlyShape
      (Interlaced?,
      Grayscale?,
      VOPFCODEForward,
      (PNoScalability | Scalability))>
<!-- ATTLIST PNotBinaryOnlyShape
      VOP_rounding_type (0|1) %implied;
      VOP_quant         CDATA %implied;
      intra_dc_vlc_thr  (0|1|2|3|4|5|6|7) %implied;>

```

Σχήμα 3-23. Επίπεδο Αντικειμένου Βίντεο τύπου P

Το χαρακτηριστικό `VOP_rounding_type` περιέχεται στο επίπεδο αντικειμένου όταν το σχήμα δεν είναι `binary_only` και η κωδικοποίηση του VOP (`VOP_coding_type`) έχει τιμή 'P'. Δηλώνει την τιμή της παραμέτρου που χρησιμοποιείται για παρεμβολή της τιμής των εικονοστοιχείων στην πρόβλεψη κίνησης ως P-VOP.

```

<!-- Επίπεδο τύπου P χωρίς κλιμάκωση -->
<!-- ELEMENT PNoScalability
      (PErrorResilientEnable | MotionShapeTexture)>
<!-- ELEMENT PErrorResilientEnable (MotionShapeTexture, PResync*)>

```

```

<!ELEMENT PResync (PVideoPacketHeader, MotionShapeTexture)>

<!ELEMENT PVideoPacketHeader
  (CommonVideoPacketHeader, PVPHEExtension?)>

<!ELEMENT PVPHEExtension (VOPFcodeForward)>

<!-- Επίπεδο Αντικειμένου Βίντεο τύπου B -->
<!ELEMENT BVOP
  (POrBNotRectangular?,
   (BNotBinaryOnlyShape | BinaryOnlyShape))>

<!ELEMENT BNotBinaryOnlyShape
  (Interlaced?,
   Grayscale?,
   VOPFcodeForward,
   VOPFcodeBackward,
   (BNoScalability | Scalability))>

<!ELEMENT VOPFcodeForward EMPTY>
<!-- ATTLIST VOPFcodeForward
  vop_fcode_forward (1|2|3|4|5|6|7) %implied; -->

<!ELEMENT VOPFcodeBackward EMPTY>
<!-- ATTLIST VOPFcodeBackward
  vop_fcode_backward (1|2|3|4|5|6|7) %implied; -->

<!--Επίπεδο τύπου B χωρίς κλιμάκωση -->
<!ELEMENT BNoScalability
  (BErrorResilientEnable | MotionShapeTexture)>

<!ELEMENT BErrorResilientEnable (MotionShapeTexture, BResync*)>

<!ELEMENT Bresync (BVideoPacketHeader, MotionShapeTexture)>

<!ELEMENT BVideoPacketHeader
  (CommonVideoPacketHeader, BVPHEExtension?)>

<!ELEMENT BVPHEExtension (PVPHEExtension, VOPFcodeBackward)>

```

Comment [Γ.Κ.4]: Τι σημαίνουν αυτές οι τιμές;

Σχήμα 3–24. Επίπεδο Αντικειμένου Βίντεο τύπου B

Στα αντικείμενα τύπου P και B καθορίζονται παράμετροι για την αποκωδικοποίηση των διανυσμάτων κίνησης. Αν ο τύπος του επιπέδου δεν είναι I (δηλ. είναι P ή B), ορίζεται η παράμετρος για εμπρόσθια αποκωδικοποίηση, ενώ αν είναι τύπου B ορίζεται παράμετρος και για την προς τα πίσω εξάρτηση των εικόνων.

Υφή κινούμενου σχήματος

Η υφή κινούμενου σχήματος περιέχεται σε όλα τα είδη αντικειμένων βίντεο αλλά και στα αντικείμενα κλιμάκωσης. Η δομή του στοιχείου "MotionShapeTexture" αποτελείται από δύο τύπους στοιχείων τμηματοποιημένη και συνδυασμένη υφή. Η τμηματοποιημένη υφή κινούμενου σχήματος περιγράφεται από το στοιχείο "DataPartitioningMST" και επιτρέπεται όταν το σχήμα της βαθμίδας δεν είναι "μόνο δυαδικό" και ο τύπος του επιπέδου αντικειμένου βίντεο δεν είναι B. Στις δύο αυτές περιπτώσεις η υφή είναι υποχρεωτικά συνδυασμένη και αποτελείται από ένα σύνολο από macroblock.

```

<!-- Υφή κινούμενου σχήματος -->
<!ELEMENT MotionShapeTexture
(DataPartitioningMST | CombinedMST) >

<!-- Συνδυασμένη υφή κινούμενου σχήματος -->
<!ELEMENT CombinedMST (Macroblock*) >

<!-- Τμηματοποιημένη υφή κινούμενου σχήματος -->
<!ELEMENT DataPartitioningMST
(DataPartitionedIVOP | DataPartitionedPVOP | BVOPCombinedMST) >

<!ELEMENT DataPartitionedIVOP
((NotRectangularShapeParams?, NonTransparentMB?)*,
MBInVideoPacket*, Block* ) >

<!ELEMENT NotRectangularShapeParams (IntraCAE?)>
<ATTLIST NotRectangularShapeParams
binary_alpha_block_type ( MVDs0NoUpdate| MVDsNon0NoUpdate |
transparent | opaque | intraCAE | MVDs0interCAE | VDsNon0interCAE )
%implied; >

<!ELEMENT MBInVideoPacket (ACPredFlag, CBFY)>

```

3.7.1.4 Αντικείμενο τύπου sprite

Τα αντικείμενα τύπου sprite περιέχονται μέσα στο επίπεδο αντικειμένου βίντεο και αποτελούν έναν από τους 4 βασικούς τύπους των αντικειμένων βίντεο. Το στοιχείο "SpriteEnable" εμφανίζεται πριν από το καθεαυτό αντικείμενο τύπου Sprite μέσα σε βασική βαθμίδα ή βαθμίδα κλιμάκωσης (Σχήμα 3–16. Βαθμίδα Βάσης και Σχήμα 3–17. Βαθμίδα κλιμάκωσης), ενώ τα αντικείμενα τύπου Sprite περιέχονται στα Επίπεδα Αντικειμένου Βίντεο (Σχήμα 3–21. Επίπεδο Αντικειμένου Βίντεο).

```

<!-- Στοιχείο παραμέτρων για αντικείμενα τύπου sprite -->
<!ELEMENT SpriteEnable EMPTY>

<ATTLIST SpriteEnable
sprite_width           %integer; %implied;
sprite_height          %integer; %implied;
sprite_left_coordinate %integer; %implied;
sprite_top_coordinate  %integer; %implied;
no_of_sprite_warping_points (0|1|2|3|4) %implied;
sprite_warping_accuracy_denom (2 | 4 | 8 | 16 ) %implied;
sprite_brightness_change (0|1) %implied;
low_latency_sprite_enable (0|1) %implied; >

<!-- Αντικείμενο τύπου sprite -->
<!ELEMENT Sprite
(SpriteTrajectory?,
BrightnessChangeFactor?,
NotStopLowLatencySprite?) >

<ATTLIST Sprite
sprite_transmit_mode ( STOP | PIECE | UPDATE | PAUSE )
%implied;>

<!-- Τροχιά αντικειμένου sprite -->
<!ELEMENT SpriteTrajectory (WarpingMotionVectorCode*)>
<!ELEMENT WarpingMotionVectorCode (#PCDATA) >

```

```

<!ELEMENT BrightnessChangeFactor EMPTY>
<!-- ATTLIST BrightnessChangeFactor
      size %integer; %implied;
      code %integer; %implied; -->
<!-- Τρόπος μετάδοσης του αντικειμένου -->
<!ELEMENT NotStopLowLatencySprite (DecodeSpritePiece*)>

```

Comment [Γ.Κ.5]: μπορεί να μη χρειάζεται αλλά δεν το έχω βρει στο πρότυπο

Σχήμα 3-25. Αντικείμενο τύπου sprite

Οι παράμετροι περιγραφής των αντικειμένων τύπου sprite, οι οποίες αποτελούν χαρακτηριστικά του στοιχείου "SpriteEnable", περιέχουν τις διαστάσεις (μήκος και πλάτος) του αντικειμένου, τις συντεταγμένες του καθώς και άλλες παραμέτρους κωδικοποίησης. Οι διαστάσεις του αντικειμένου καθορίζονται σε μονάδες των macroblock. Ο αριθμός των σημείων παραμόρφωσης (no_of_sprite_warping points) του sprite υπονοεί τη συνάρτησης παραμόρφωσης και αν είναι 0, τότε το sprite είναι στατικό (stationary). Αν είναι 1, τότε χρησιμοποιείται translation, αν είναι 2 ή 3 affine transformation και αν είναι 4 χρησιμοποιείται μετασχηματισμός προοπτικής (perspective transformation). Το χαρακτηριστικό "sprite_warping_accuracy_denom" παριστάνει την ακρίβεια κβαντοποίησης των διανυσμάτων κίνησης στην παραμόρφωση και είναι κλάσμα του ενός pixel. Επίσης στις παραμέτρους περιγράφεται αν μεταβάλλεται η φωτεινότητα του sprite με το χαρακτηριστικό "sprite_brightness_change" και αν αυτό είναι χαμηλής καθυστέρησης (latency) ή βασικού τύπου.

Το καθεαυτό στοιχείο τύπου "Sprite" αποτελείται από στοιχεία που περιγράφουν την τροχιά του, την αλλαγή της φωτεινότητάς του, καθώς και στοιχεία που περιέχουν τα δεδομένα υψής του στατικού αντικειμένου. Η τροχιά των αντικειμένων τύπου sprite αποτελείται από ένα σύνολο κωδίκων που αντιπροσωπεύουν τα διανύσματα κίνησης. Το στοιχείο "BrightnessChangeFactor" υπάρχει στη δομή του Sprite αν αλλάζει η φωτεινότητα του αντικειμένου κατά τη διάρκεια της παραμόρφωσης (warping) και συνδέεται με την τιμή του δίτιμου χαρακτηριστικού "sprite_brightness_change" του στοιχείου "SpriteEnable". Η τιμή του παράγοντα αυτού είναι κάποιος κώδικας.

Το χαρακτηριστικό "sprite_transmit_mode" του αντικειμένου τύπου sprite, παίρνει τις τιμές "STOP", "PIECE", "UPDATE" ή "PAUSE". Κατά την αρχικοποίηση της βαθμίδας του βίντεο, η τιμή αυτή γίνεται "PIECE". Όταν όλα τα τμήματα του αντικειμένου και οι ενημερώσεις αυτών, σταλούν για ολόκληρη τη βαθμίδα, τότε το χαρακτηριστικό παίρνει την τιμή "STOP". Όταν στέλνεται ένα τμήμα του αντικειμένου παίρνει την τιμή "PIECE", ενώ όταν ενημερώνεται ένα τμήμα του αντικειμένου παίρνει την τιμή "UPDATE". Όταν σταλούν όλα τα τμήματα του αντικειμένου και οι ενημερώσεις αυτών, τότε το χαρακτηριστικό λαμβάνει την τιμή "PAUSE".

Το στοιχείο "NotStopLowLatencySprite" περιγράφει τον τρόπο μετάδοσης του sprite όταν αυτός δεν έχει την τιμή "STOP". Το στοιχείο "DecodeSpritePiece" επομένως εξαρτάται από την τιμή του χαρακτηριστικού "sprite_transmit_mode". Αν η τιμή αυτή είναι "PIECE" ή "UPDATE", τότε υπάρχει και το στοιχείο "DecodeSpritePiece", το οποίο περιγράφει ή ενημερώνει ένα τμήμα του αντικειμένου, όπως φαίνεται στο παρακάτω σχήμα. Τα χαρακτηριστικά του είναι οι διαστάσεις και η θέση του συγκεκριμένου τμήματος που στέλνεται ή ενημερώνεται, τα οποία ορίζονται σε μονάδες των macroblock. Η θέση του αντικειμένου ορίζεται ως προς το πάνω (yoffset) αριστερό σημείο (xoffset) του αντικειμένου και χρησιμοποιείται από τον αποκωδικοποιητή για τοποθέτηση του τμήματος στον προσωρινό αποταμιευτή που περιέχει το αντικείμενο. Όλα τα στατικά τμήματα των αντικειμένων τύπου sprite κωδικοποιούνται σύμφωνα με τη σύνταξη των I-VOP, ενώ τα τμήματα ενημέρωσης χρησιμοποιούν τη σύνταξη των P-VOP. Τα τμήματα ενημέρωσης περιγράφουν δηλαδή τη διαφορά ανάμεσα στην αρχική υφή του αντικειμένου και σε αυτήν του ανακατασκευασμένου αντικειμένου που προκύπτει με συναρμολόγηση όλων των τμημάτων του sprite.

Όπως φαίνεται παρακάτω, η δομή του τμήματος και του τμήματος ενημέρωσης είναι ίδια, δηλαδή ένας αριθμός (πλάτος επί ύψος) από macroblock. Η διαφορά είναι ότι αν ένα τμήμα έχει ήδη σταλεί στον κωδικοποιητή, τότε δεν περιγράφεται ξανά μέσα στο sprite. Η κωδικοποίηση των τμημάτων γίνεται με σάρωση των macroblock σε σειρά raster.

Comment [Γ.Κ.6]: είναι σωστή αυτή η έκφραση;

```
<!-- Αποκωδικοποίηση ή ενημέρωση τμήματος του sprite -->
<!ELEMENT DecodeSpritePiece
  (SpriteShapeTexture)>

<!ATTLIST DecodeSpritePiece
  quant           %integer;    %implied;
  width_in_mbs    %integer;    %implied;
  heights_in_mbs  %integer;    %implied;
  xoffset         %integer;    %implied;
  yoffset         %integer;    %implied;>

<!ELEMENT SpriteShapeTexture
  (SpritePiece | SpriteUpdate)>

<!ELEMENT SpritePiece (Macroblock*)>
<!ELEMENT SpriteUpdate (Macroblock*)>

<!ELEMENT NotRectangularSprite
  (VOPConstantAlpha?, ScalabilitySprite?) >
<!ATTLIST NotRectangularSprite
  change_conv_ratio_disable %boolean; %implied; >

<!ELEMENT ScalabilitySprite
  (BackgroundComposition)>
```


3.7.1.5 Επίπεδο αντικειμένου κλιμάκωσης

Όπως ήδη αναφέρθηκε οι βαθμίδες του βίντεο μπορεί να είναι βάσης ή κλιμάκωσης. Κάθε επίπεδο αντικειμένου βίντεο είναι δυνατόν να περιέχει χαρακτηριστικά κλιμάκωσης. Τα στοιχεία των αντικειμένων βίντεο που αποτελούν μέρος των βαθμίδων κλιμάκωσης αναφέρονται σε ένα τμήμα ή σε ολόκληρη την περιοχή ενός επιπέδου αντικειμένου βίντεο. Αν αναφέρονται σε υποπεριοχή του αντικειμένου αναφοράς, η βελτίωση ονομάζεται τύπου "1" και περιγράφεται με το στοιχείο "PartialEnhancement" όπως φαίνεται παρακάτω. Το στοιχείο "RefSelectCode" περιγράφει το αντικείμενο αναφοράς για τα αντικείμενα τύπου P και B. Στα αντικείμενα κλιμάκωσης περιέχεται και η κωδικοποιημένη κινούμενη υφή (MotionShapeTexture).

```
<!-- Κλιμάκωση των αντικειμένων βίντεο -->
<!ELEMENT Scalability
  (ReferenceVOP,
   PartialEnhancement?,
   MotionShapeTexture)>

<!ELEMENT ReferenceVOP
  (PReferenceVOP | BReferenceVOP)>
```

Σχήμα 3-26. Επίπεδο κλιμάκωσης

Το στοιχείο "ReferenceVOP" που καθορίζει το επίπεδο αναφοράς (πρόβλεψης) της κλιμάκωσης διαχωρίζεται σε αντικείμενα τύπου P και B. Στα αντικείμενα τύπου B όπου η εξάρτηση είναι προς δύο κατευθύνσεις, χρησιμοποιούνται δύο αντικείμενα αναφοράς. Οι δυνατές τιμές του αντικειμένου αναφοράς για P-VOP είναι οι εξής:

- πρόβλεψη από το πιο πρόσφατα αποκωδικοποιημένο βελτιωτικό επίπεδο (VOP) της ίδιας βαθμίδας
- πρόβλεψη από το πιο πρόσφατο επίπεδο (VOP) σε σειρά προβολής της βαθμίδας αναφοράς
- πρόβλεψη από το επόμενο VOP σε σειρά προβολής της βαθμίδας αναφοράς
- πρόβλεψη από το χρονικά ταυτόσημο VOP της βαθμίδας αναφοράς, οπότε δεν περιέχονται διανύσματα κίνησης

Για τα B-VOP όπου η πρόβλεψη μπορεί να είναι προς τα μπρος αλλά και προς τα πίσω, οι δυνατές τιμές των VOP αναφοράς είναι:

- εμπρόσθια πρόβλεψη από το πιο πρόσφατα αποκωδικοποιημένο βελτιωτικό επίπεδο (VOP) της ίδιας βαθμίδας και πίσω πρόβλεψη από το χρονικά ταυτόσημο VOP της βαθμίδας αναφοράς

- εμπρόσθια πρόβλεψη από το πιο πρόσφατα αποκωδικοποιημένο βελτιωτικό VOP της ίδιας βαθμίδας και πίσω πρόβλεψη από το πιο πρόσφατο VOP ως προς τη σειρά παρουσίασης, που ανήκει στη βαθμίδα αναφοράς
- εμπρόσθια πρόβλεψη από το πιο πρόσφατα αποκωδικοποιημένο βελτιωτικό VOP της ίδιας βαθμίδας και πίσω πρόβλεψη από το επόμενο VOP σε σειρά παρουσίασης που ανήκει στη βαθμίδα αναφοράς
- εμπρόσθια πρόβλεψη από το επόμενο VOP της βαθμίδας αναφοράς σε σειρά παρουσίασης και πίσω πρόβλεψη από το επόμενο VOP σε σειρά παρουσίασης που ανήκει στη βαθμίδα αναφοράς

Οι τιμές αυτές του είδους της πρόβλεψης περιγράφονται στην παρούσα διατριβή με συμβολοσειρές ενώ στο πρότυπο είναι δυαδικοί κώδικες.

Το στοιχείο "PartialEnhancement" μπορεί να αποτελείται από πληροφορία για το σχήμα της υποπεριοχής στην οποία αναφέρεται. Το εμπρόσθιο και το πίσω σχήμα χρησιμοποιούνται στη διαδικασία σύνθεσης του υπόβαθρου. Το πίσω σχήμα είναι το σχήμα του βελτιωμένου αντικειμένου στο επόμενο VOP σε σειρά παρουσίασης, το οποίο ανήκει στο επίπεδο αναφοράς. Το μπροστινό σχήμα είναι το σχήμα του βελτιωμένου αντικειμένου στο προηγούμενο VOP σε σειρά παρουσίασης, που ανήκει στο επίπεδο αναφοράς.

```
<!ELEMENT PReferenceVOP EMPTY>
<!ATTLIST PReferenceVOP
    ref_code ( MOST_RECENT_ENHANCEMENT_VOP |
               MOST_RECENT_REFLAYER_DISPLAY_VOP |
               NEXT_REFLAYER_DISPLAY_VOP |
               COINCIDENT_REFLAYER_VOP ) %implied; >

<!ELEMENT BReferenceVOP EMPTY>
<!ATTLIST BReferenceVOP
    ref_code
        (MOST_RECENT_ENHANCEMENT_VOP_COINCIDENT_REFLAYER_VOP |
         MOST_RECENT_ENHANCEMENT_VOP_MOST_RECENT_REFLAYER_DISPLAY_VOP |
         MOST_RECENT_ENHANCEMENT_VOP_NEXT_REFLAYER_DISPLAY_VOP |
         MOST_RECENT_REFLAYER_DISPLAY_VOP_NEXT_REFLAYER_DISPLAY_VOP)
    %implied; >

<!ELEMENT PartialEnhancement
    (BackgroundComposition, LoadBackwardShape?)>

<!ELEMENT BackgroundComposition EMPTY>
<!ATTLIST BackgroundComposition
    background_composition (0|1) %implied; >
```

Το στοιχείο BackgroundComposition περιγράφει τη μέθοδο σύνθεσης του υπόβαθρου και υπάρχει στη δομή μόνο όταν το επίπεδο είναι βελτιωτικό επίπεδο κλιμάκωσης.

Η ύπαρξη του στοιχείου LoadBackwardShape στη δομή του PartialEnhancement υπονοεί ότι το πίσω σχήμα του προηγούμενου VOP στην ίδια βαθμίδα αντιγράφεται στο μπροστινό

σχήμα του τρέχοντος VOP και το πίσω σχήμα του τρέχοντος VOP αποκωδικοποιείται από την ψηφιοσειρά. Όταν το στοιχείο αυτό δεν υπάρχει, τότε το μπροστινό σχήμα του προηγούμενου VOP αντιγράφεται στο μπροστινό σχήμα του τρέχοντος και το πίσω σχήμα του προηγούμενου VOP στην ίδια βαθμίδα, αντιγράφεται στο πίσω σχήμα του τρέχοντος VOP. Το στοιχείο αυτό υπάρχει στη δομή αν το χαρακτηριστικό background_composition είναι 1 και το VOP_coded είναι 0 ή το VOP είναι το πρώτο της τρέχουσας βαθμίδας. Τα χαρακτηριστικά πλάτος, ύψος και συντεταγμένες του πίσω σχήματος καθορίζουν το ορθογώνιο που το περιέχει.

Το στοιχείο ForwardShape υπονοεί ότι θα αποκωδικοποιηθεί το εμπρόσθιο σχήμα, το οποίο ορίζεται από το ορθογώνιο που το περικλείει.

```
<!ELEMENT LoadBackwardShape
    (BabType+,
      (ConvRatio?, ScanType, BinaryArithmeticCode)+,
      ForwardShape?) >
<!ATTLIST LoadBackwardShape
    backward_shape_width          CDATA %implied;
    backward_shape_height         CDATA %implied;
    backward_shape_horizontal_mc_spatial_ref CDATA %implied;
    backward_shape_vertical_mc_spatial_ref  CDATA %implied; >
<!ELEMENT ForwardShape
    (BabType+, (ConvRatio?, ScanType, BinaryArithmeticCode)*) >
<!ATTLIST ForwardShape
    forward_shape_width          CDATA %implied;
    forward_shape_height         CDATA %implied;
    forward_shape_horizontal_mc_spatial_ref CDATA %implied;
    forward_shape_vertical_mc_spatial_ref  CDATA %implied; >
```

3.7.1.6 Macroblock

Τα macroblock διακρίνονται ανάλογα με τον τύπο του VOP που τα περιέχει, σε I, P, B, κλιμάκωσης και κλίμακας του γκρι.

```
<!ELEMENT Macroblock
    ((NotBMacroblock | BPartialEnhancementMb)?,
     GrayscaleNonTransparentMb?)>

<!ELEMENT NotBMacroblock
    (NotRectangularNotSpriteMb?, NotBinaryOnlyMb?) >

<!ELEMENT NotRectangularNotSpriteMb
    (MbBinaryShapingCoding)>

<!ELEMENT MbBinaryShapingCoding
    (PorBVOP?, BabType4?)>
<!ATTLIST MbBinaryShapingCoding
    bab_type          CDATA %implied; >

<!ELEMENT PorBVOP (BabType1or6?)>
<!ELEMENT BabType1or6 EMPTY>
<!ATTLIST BabType1or6
    mvds_x          CDATA %implied;
    mvds_y          CDATA %implied; >
```

Αν το derived_mbtype είναι 3 ή 4 υπάρχει το DCTType. Αν είναι P VOP και το derived_mbtype είναι 0 ή 1, ή αν είναι B VOP και το derived_mbtype είναι 1 υπάρχει το στοιχείο PVOPMBType0Or1_BVOPMBType1.

```
<!ELEMENT NotBinaryOnlyMb (NonTransparentMb?)>
<!ELEMENT NonTransparentMb (CodedMb | NotCodedMb)>
<!ELEMENT CodedMb (NotShortHeaderMbType3or4?, NotStuffingMb?)>
<!ATTLIST CodedMb
    type
    chrominance_block_pattern >
<!ATTLIST NotShortHeaderMbType3or4
    ac_pred_flag (0|1) %implied; >
<!ELEMENT NotStuffingMb
    (MbType1or4?,
    InterlacedInformation?,
    NotScalabilityNotSpriteMb?,
    Block*) >
<!ATTLIST NotStuffingMb
    luminance_pattern CDATA %implied; >
<!ELEMENT MbType1or4 EMPTY>
<!ATTLIST MbType1or4
    dquant ("1" | "-2" | 1 | 2 ) %implied; >
<!ELEMENT InterlacedInformation
    (DCTType?, PVOPMBType0Or1_BVOPMBType1?) >
<!ELEMENT DCTType EMPTY>
<!ATTLIST DCTType
    dc_type ( 0|1 ) %implied; >
```

Η τιμή του στοιχείου DCTType δείχνει αν το macroblock είναι frame DCT (τιμή 0) ή field DCT (τιμή 1). Αυτό το στοιχείο υπάρχει μόνο αν το interlaced έχει την τιμή 1 και το macroblock είναι κωδικοποιημένο ή ενδο-κωδικοποιημένο (intra-coded).

Το στοιχείο FieldPrediction υπάρχει αν το macroblock είναι κωδικοποιημένο με field motion vectors, δηλαδή η τιμή του χαρακτηριστικού field_prediction είναι 1. Αλλιώς χρησιμοποιείται frame prediction 16x16 ή 8x8. Το στοιχείο αυτό υπάρχει μόνο αν είναι interlaced και το derived_mbtype είναι 0 ή 1 σε P ή BVOP.

```
<!ELEMENT PVOPMBType0Or1_BVOPMBType1
    (FieldPrediction?)>
<!ATTLIST PVOPMBType0Or1_BVOPMBType1
    field_prediction (0|1) %implied; >
<!ELEMENT FieldPrediction
    (PVOP_BVOPMBType1, BVOPMBType1?)>
```

Το χαρακτηριστικό forward_top_field_reference δείχνει το πεδίο αναφοράς για την εμπρόσθια απόσβεση κίνησης του πάνω πεδίου, ενώ αντίστοιχα το forward_bottom_field_reference δείχνει το πεδίο αναφοράς για το κάτω πεδίο.

Στα VOP τύπου B ορίζονται και τα πεδία αναφοράς για την προς τα πίσω απόσβεση κίνησης.

```
<!--ELEMENT PVOP_BVOPMBType1 EMPTY>
<!--ATTLIST PVOP_BVOPMBType1
    forward_top_field_reference    (0|1) %implied;
    forward_bottom_field_reference (0|1) %implied;>

<!--ELEMENT BVOPMBType1 EMPTY>
<!--ATTLIST BVOPMBType1
    backward_top_field_reference    (0|1) %implied;
    backward_bottom_field_reference (0|1) %implied;>

<!--ELEMENT NotScalabilityNotSpriteMb
    ( MbType0or1 | MbType2 )>
<!--ELEMENT MbType0or1
    ( ForwardMV, PredictionMV?)>
<!--ELEMENT PredictionMV (ForwardMV)>
<!--ELEMENT MbType2 (ForwardMV*)>

<!--ELEMENT NotCodedMb EMPTY>
```

Το στοιχείο BPartialEnhancementMb υπάρχει στη δομή του macroblock όταν η βαθμίδα είναι κλιμάκωσης, ο τύπος αναφοράς είναι COINCIDENT_REF_LAYER_VOP για τα επίπεδα τύπου P ή (MOST_RECENT_REFLAYER_DISPLAY_VOP, NEXT_REFLAYER_DISPLAY_VOP) για τα επίπεδα τύπου B αντίστοιχα και η κλιμάκωση είναι για τμήμα του αντικειμένου.

```
<!--ELEMENT BPartialEnhancementMb
    (NotRectangularPEMb?, NotBinaryOnlyPEMb?)

<!--ELEMENT NotRectangularPEMb
    (MbBinaryShapingCoding)>

<!--ELEMENT NotBinaryOnlyPEMb
    (NonTransparentScalMb?) >

<!--ELEMENT NonTransparentPEMb
    (BTypeScalMb?)>

<!--ATTLIST BTypeScalMb
    type          CDATA %implied;
    block_pattern CDATA %implied; >

<!--ELEMENT GrayscaleNonTransparentMb
    (IVOPorPVOPICodedMb | PorBMb) >

<!--ELEMENT IVOPorPVOPICodedMb (NotCODA_I?)>
<!--ELEMENT NotCODA_I (AlphaBlock+)>
<!--ATTLIST NotCODA_I
    ac_pred_flag_alpha    (0|1) %implied;
    CBPA                  CDATA %implied;>

<!--ATTLIST DirectMV
    horizontal_mv_data    CDATA %implied;
    vertical_mv_data      CDATA %implied; >

<!--ELEMENT ForwardMV (Not1FCodeNot0HorData?) >
<!--ATTLIST ForwardMV
    horizontal_mv_data    CDATA %implied;
    vertical_mv_data      CDATA %implied; >
```

```
<!ATTLIST NotIFCodeNot0HorData
    horizontal_mv_residual CDATA %implied; >
```

3.7.2 Αντικείμενα Ακίνητης Υφής

Τα αντικείμενα υφής έχουν μοναδική ταυτότητα. Το χαρακτηριστικό `wavelet_filter_type` δείχνει την αριθμητική ακρίβεια που χρησιμοποιείται για τη διάσπαση των `wavelet` και μπορεί να είναι ακέραιος ή διπλός αριθμός κινητής υποδιαστολής. Το χαρακτηριστικό `wavelet_download` δείχνει αν θα χρησιμοποιηθούν τα προκαθορισμένα φίλτρα, ή αν τα φίλτρα μεταδίδονται μέσα στην ψηφιοσειρά.

Το χαρακτηριστικό `scan_direction` δείχνει την κατεύθυνση σάρωσης των AC συντελεστών. Υπάρχουν 3 τρόποι κβαντοποίησης που καθορίζουν τη σημασία αυτού του χαρακτηριστικού: `single`, `multi` και `bi-level`. Στην πρώτη και δεύτερη περίπτωση, αν η κατεύθυνση είναι 0, τότε η σάρωση γίνεται προς το βάθος του δέντρου, ενώ αν είναι 1, τότε σαρώνονται στην υποζώνη (`subband`). Σε `bi-level` κβαντοποίηση, σαρώνονται με δύο τρόπους ανάλογα με την τιμή της κατεύθυνσης: σε `bitplanes` με `subband` σε κάθε `bitplane` (τιμή 0) ή από το χαμηλό επίπεδο διάσπασης προς το υψηλό επίπεδο διάσπασης (τιμή 1).

Το σχήμα ενός αντικειμένου υφής περιγράφεται με το χαρακτηριστικό `texture_object_layer_shape_type` και μπορεί να είναι τετράγωνο ή δισαδικό.

```
<!ELEMENT StillTextureObject
    (DownloadWaveletFilters?,
    (RectangularTextureShape | BinaryTextureShape),
    YWaveletDecode, UWaveletDecode, VWaveletDecode,
    (SingleQuantizationType | MultiQuantizationType |
    BilevelQuantizationType))

<ATTLIST StillTextureObject
    texture_object_id      %id_type; %implied;
    wavelet_filter_type    ("INTEGER" | "DOUBLE_FLOAT")
    wavelet_download      ("DEFAULT_FILTERS" |
    "SPECIFIED_IN_BITSTREAM")
    wavelet_stuffing_quantization_type CDATA %implied;
    scan_direction        ( 0 | 1 ) %implied;
    start_code_enable      ( 0 | 1 ) %implied;
    wavelet_stuffing        %fixed; "111"
    texture_object_layer_shape_type ( rectangular | binary ) %implied;
    quantization_type      ( single | multi | bilevel ) %implied;>
```

Comment [Γ.Κ.7]: στο πρότυπο σελ. 46 η ψηφιοσειρά και σελ. 81 η σημασία

Σχήμα 3-27. Περιγραφή αντικειμένων ακίνητης υφής

Η ύπαρξη του στοιχείου `DownloadWaveletFilters` σημαίνει ότι τα φίλτρα μεταδίδονται μέσα στην ψηφιοσειρά ή θα χρησιμοποιηθούν τα προκαθορισμένα φίλτρα. Αν το σχήμα της υφής είναι τετράγωνο, περιγράφεται το πλάτος και το ύψος του εμφανούς τμήματος της φωτεινότητας. Οι διαστάσεις αυτές υπονοούν αριθμό εικονοστοιχείων. Τα αντικείμενα υφής δισαδικού σχήματος περιγράφεται από τη θέση και τις διαστάσεις των ορθογωνίων

που τα περικλείουν. Η οριζόντια και η κάθετη θέση αφορούν στην πάνω αριστερή γωνία του περικλειόντος ορθογωνίου. Τα δυαδικά τμήματα από τα οποία αποτελείται το σχήμα είναι τριών ειδών: διαφανή, αδιαφανή ή κωδικοποιημένα με intra-mode CAE (Context-based Arithmetic Encoding).

```
<!--ATTLIST RectangularTextureShape
        texture_object_layer_width      %integer;
        texture_object_layer_height     %integer;>

<!--ELEMENT BinaryTextureShape (ShapeObjectDecoding)>
<!--ATTLIST BinaryTextureShape
        horizontal_ref      %integer;
        vertical_ref        %integer;
        object_width        %integer;
        object_height       %integer; >

<!--ELEMENT ShapeObjectDecoding
        (STOConstantAlphaValue?, BinaryAlphaBlock*)
<!--ATTLIST ShapeObjectDecoding
        change_conv_ratio_disable      (0|1) %implied;
        STO_constant_alpha              (0|1) %implied;>

<!--ELEMENT BinaryAlphaBlock (IntraCAEBAB?)>
<!--ATTLIST BinaryAlphaBlock
        coding_type (transparent | opaque | intraCAE ) %implied;>

<!--ELEMENT IntraCAEBAB (ConvRatio?, BinaryArithmeticDecode)>
<!--ATTLIST IntraCAEBAB
        scan_type (transposed | not_transposed) %implied;>

<!--ATTLIST ConvRatio ( 1 |2 |4 ) %implied;>
```

Το στοιχείο ConvRatio περιέχει τον παράγοντα υπο-δειγματοληψίας του 16x16 δυαδικού τμήματος, με τον οποίο πρέπει να γίνει upsample το αποκωδικοποιημένο τμήμα. Ο τύπος σάρωσης (χαρακτηριστικό scan_type), δείχνει αν το δυαδικό τμήμα είχε μετατεθεί ή όχι πριν από την κωδικοποίηση. Αν είχε μετατεθεί, τότε ο αποκωδικοποιητής πρέπει να το επαναφέρει στην αρχική του μορφή.

3.7.3 Πλέγματα

Τα αντικείμενα πλεγμάτων αποτελούνται από ένα ή περισσότερα επίπεδα αντικειμένων πλεγμάτων. Κάθε επίπεδο αποτελείται από μια κεφαλίδα και τα δεδομένα του επιπέδου. Η κεφαλίδα περιγράφει αν το πλέγμα είναι ενδο-κωδικοποιημένο ή όχι και περιέχει μια κεφαλίδα που περιγράφει χρονικές παραμέτρους.

```
<!--ELEMENT MeshObject (MeshObjectPlane*)>

<!--ELEMENT MeshObjectPlane
        (MeshObjectPlaneHeader, MeshObjectPlaneData)>

<!--ELEMENT MeshObjectPlaneHeader (TemporalHeader)>
<!--ATTLIST MeshObjectPlaneHeader
        is_intra      %boolean; %implied;
        mesh_mask     %boolean; %implied;>
```

```

<!ELEMENT TemporalHeader (IsIntra?, DecodeSkipFrames?)>
<!--ATTLIST TemporalHeader
      skip_frames %boolean; %implied;-->

<!ELEMENT IsIntra (DecodeFrameRate?, TimeCode?)>
<!--ATTLIST IsIntra
      is_frame_rate      (0|1) %implied;
      is_time_code       (0|1) %implied;-->

<!ELEMENT DecodeFrameRate EMPTY>
<!--ATTLIST DecodeFrameRate
      frame_rate %integer;
      seconds    %integer;
      frequency_offset (0|1) %implied;-->

<!ELEMENT TimeCode EMPTY>
<!--ATTLIST TimeCode
      hours %integer; %implied;
      minutes %integer; %implied;
      seconds %integer; %implied;-->

<!ELEMENT DecodeSkipFrames (NumberToSkip)>
<!--ELEMENT NumberToSkip EMPTY>
<!--ATTLIST NumberToSkip
      num_to_skip %integer;-->

```

Σχήμα 3-28. Περιγραφή πλεγματών

Αν το επίπεδο αντικειμένου πλέγματος είναι ενδο-κωδικοποιημένο, η χρονική κεφαλίδα περιγράφει το ρυθμό των εικόνων και το χρόνο του επιπέδου στη σειρά παρουσίασης. Ο ρυθμός των εικόνων αποτελείται από δύο αριθμούς, το ρυθμό αναφοράς και ένα κλάσμα του ρυθμού αναφοράς. Ο τελικός ρυθμός υπολογίζεται με την πράξη "ρυθμός εικόνων = frame_rate+seconds/16". Το χαρακτηριστικό "frequency_offset" δείχνει αν ο ρυθμός περιέχει μετατόπιση της συχνότητας 1000/1001 όπως καθορίζεται από το πρότυπο NTSC, δηλαδή ο ρυθμός υπολογίζεται με τον "τύπο ρυθμός εικόνων = (1000/1001) * (frame_rate+ seconds/16)". Ο χρόνος παρουσίασης του επιπέδου πλέγματος αποτελείται από τρία μέρη: ώρες, λεπτά και δευτερόλεπτα.

Η χρονική επικεφαλίδα μπορεί να περιέχει και τον αριθμό των εικόνων που θα παραλειφθούν κατά την παρουσίαση.

Πληροφορία για το αντικείμενο πλέγματος περιέχεται μέσα στα δεδομένα, όταν η μάσκα του πλέγματος έχει την τιμή 1. Επιπλέον η μάσκα διαχωρίζει τα πλέγματα σε ενδοκωδικοποιημένα ή μη ενδοκωδικοποιημένα. Τα ενδοκωδικοποιημένα επίπεδα αντικειμένων πλεγματών περιγράφονται μέσω της γεωμετρίας τους ενώ τα μη ενδοκωδικοποιημένα μέσω της κίνησής τους, όπως φαίνεται και παρακάτω.

```

<!ELEMENT MeshObjectPlaneData (MeshMask1?)>
<!--ELEMENT MeshMask1 (IsIntraMeshObjectPlane | IsNotIntraMeshObjectPlane)>
<!--ELEMENT IsIntraMeshObjectPlane (MeshGeometry)>
<!--ELEMENT IsNotIntraMeshObjectPlane (MeshMotion)>

```



```

<!ELEMENT MeshGeometry (UniformMesh | DelaunayMesh)>
<!--ATTLIST MeshGeometry
      mesh_type (uniform |Delaunay) %implied;-->

<!ELEMENT UniformMesh EMPTY>
<!--ATTLIST UniformMesh
      nr_of_mesh_nodes_hor      %integer;
      nr_of_mesh_nodes_vert     %integer;
      mesh_rect_size_hor        %integer;
      mesh_rect_size_vert       %integer;
      triangle_split_code       (TOPLEFT_TO_RIGHTBOTTOM |
                                BOTTOMLEFT_TO_TOPRIGHT |
                                ALTERNATE_TOPLEFT_TO_RIGHTBOTTOM_BOTTOMLEFT_TO_TOPRIGHT |
                                ALTERNATE_BOTTOMLEFT_TO_TOPRIGHT_TOPLEFT_TO_RIGHTBOTTOM) >

```

Η γεωμετρία του πλέγματος μπορεί να είναι ομοιόμορφη ή Delaunay. Η ομοιόμορφη γεωμετρία περιγράφεται από τον αριθμό των κόμβων ανά γραμμή και ανά στήλη του πλέγματος (nr_of_mesh_nodes_hor, nr_of_mesh_nodes_vert). Το ομοιόμορφο πλέγμα όπως έχει ήδη αναφερθεί στο Κεφάλαιο 2, αποτελείται από όμοια τρίγωνα. Τα χαρακτηριστικά mesh_rect_size_hor και mesh_rect_size_vert περιγράφουν το πλάτος και το ύψος ενός ορθογωνίου που περιέχει δύο τρίγωνα του πλέγματος, σε μονάδες του μισού εικονοστοιχείου. Το χαρακτηριστικό triangle_split_code, περιγράφει τον τρόπο με τον οποίο χωρίζεται ένα ορθογώνιο του πλέγματος σε δύο τρίγωνα. Υπάρχουν τέσσερις τρόποι για να γίνει αυτό: από την πάνω αριστερή γωνία μέχρι την κάτω δεξιά γωνία του ορθογωνίου, από την κάτω αριστερή μέχρι την πάνω δεξιά γωνία αλληλοδιαδοχικά και με τους δύο παραπάνω τρόπους.

Το πλέγμα Delaunay περιγράφεται από το συνολικό αριθμό κόμβων που περιέχει, τον αριθμό των κόμβων που υπάρχουν πάνω στα όρια του πλέγματος και τις συντεταγμένες (node0_x, node0_y) του πρώτου οριακού κόμβου ως προς τοπικό σύστημα συντεταγμένων και σε μονάδες των μισών εικονοστοιχείων. Οι συντεταγμένες των υπολοίπων κόμβων καθορίζονται ως προς τη διαφορά τους με τις συντεταγμένες του προηγούμενου κωδικοποιημένου κόμβου (delta_x και delta_y).

```

<!ELEMENT DelaunayMesh (MeshNodeCoordDiff*)>
<!--ATTLIST DelaunayMesh
      nr_of_mesh_nodes          %integer;
      nr_of_boundary_nodes      %integer;
      node0_x                   %integer;
      node0_y                   %integer; >

<!--ATTLIST MeshNodeCoordDiff
      delta_x                   %integer;
      delta_y                   %integer;-->

<!ELEMENT MeshMotion (NodeMotionVector*)>
<!--ATTLIST MeshMotion
      motion_range_code (MIN32_TO_31 | MIN64_TO_63 | MIN128_TO_127) >

<!ELEMENT NodeMotionVector (NonZeroMotionVectorDiff?)>

```

```
<!ELEMENT NonZeroMotionVectorDiff EMPTY>
<!--ATTLIST NonZeroMotionVectorDiff
      delta_mv_x_res      %integer;
      delta_mv_y_res      %integer;-->
```

Η κίνηση του πλέγματος περιγράφεται ως προς τα διανύσματα κίνησης των κόμβων, για τα οποία καθορίζεται το εύρος των τιμών τους με το χαρακτηριστικό `motion_range_code`. Για κάποιο κόμβο το διάνυσμα κίνησής του μπορεί να είναι 0, οπότε δεν κωδικοποιείται.

3.7.4 Πρόσωπα

Τα αντικείμενα προσώπων, όπως και όλα τα οπτικά αντικείμενα, αποτελούνται από επίπεδα αντικειμένων. Τα επίπεδα αντικειμένων προσώπων αποτελούνται από μια κεφαλίδα και τα δεδομένα του επιπέδου. Η κεφαλίδα καθορίζει αν το επίπεδο προσώπου είναι ενδοκωδικοποιημένο ή όχι, καθώς και τη μάσκα και περιέχει μια χρονική κεφαλίδα, όπως και τα πλέγματα (ενότητα 3.7.3). Η μάσκα δείχνει αν μέσα στο επίπεδο υπάρχουν δεδομένα Παραμέτρων Κίνησης Προσώπων (FAP).

Αν μαζί με το πρόσωπο καθορίζονται οι παράμετροι κίνησής του, διακρίνεται σε δύο τύπους, ενδοκωδικοποιημένο ή μη. Το ενδοκωδικοποιημένο πρόσωπο περιέχει το είδος της κωδικοποίησης (`face_object_coding_type`), που είναι είτε με πρόβλεψη ή DCT.

```
<!ELEMENT FaceObject (FaceObjectPlane*)>
<!--ELEMENT FaceObjectPlane
      (FaceObjectPlaneHeader, FaceObjectPlaneData)>
<!--ELEMENT FaceObjectPlaneHeader (TemporalHeader)>
<!--ATTLIST FaceObjectPlaneHeader
      is_intra      (0|1) %implied;
      face_object_mask (unused | FAPpresent) %implied;-->

<!--ELEMENT FaceObjectPlaneData (FAPPresent?)>
<!--ELEMENT FAPPresent (IsIntraFace | IsNotIntraFace)>
<!--ELEMENT IsIntraFace (FAPGroupMask*, (PredictiveFBA | DCTFBA))>
<!--ATTLIST IsIntraFace
      fap_quant      %integer;
      fap_mask_type (NO_MASK_NOR_FAP |
      GROUP_MASK | FAP) %implied;
      face_object_coding_type (predictive |DCT ) %implied; >
<!--ELEMENT FAPGroupMask EMPTY>
<!--ATTLIST FAPGroupMask
      fap_group_mask %integer;-->
<!--ELEMENT PredictiveFBA (DecodeNewMinmax, DecodeIFAP)>
<!--ATTLIST PredictiveFBA
      is_i_new_max (0|1) %implied;
      is_i_new_min (0|1) %implied;
      is_p_new_max (0|1) %implied;
      is_p_new_min (0|1) %implied;-->
<!--ELEMENT DCTFBA (DecodeISegment)>

<!--ELEMENT IsNotIntraFace (DecodePFAP | DecodePSegment)>
```

Σχήμα 3-29. Περιγραφή προσώπων

3.8 Περιγραφείς Αντικειμένων

Όπως αναφέρεται και στην ενότητα "2.2.4.Περιγραφείς Αντικειμένων", οι περιγραφείς αυτοί αποτελούν τον τρόπο συσχέτισης των σκηνών με τα οπτικοακουστικά αντικείμενα, τα οποία κωδικοποιούνται και μεταφέρονται σε ξεχωριστές ροές. Στο πρότυπο MPEG-4 οι Περιγραφείς Αντικειμένων αναπαρίστανται ως αντικείμενα της γλώσσας MSDL (MPEG-4 Systems Description Language). Στην παρούσα ενότητα περιγράφεται η αναπαράστασή τους με ομοιόμορφο τρόπο με τα οπτικά αντικείμενα, όπως έγινε στην ενότητα 3.7 , δηλαδή με τη γλώσσα XML.

Οι Περιγραφείς Αντικειμένων και τα συστατικά τους στοιχεία μεταφέρονται μέσα σε δομές που ονομάζονται *Μηνύματα Περιγραφέων Αντικειμένων* (Object Descriptor Messages). Στο Σχήμα 3–30 παρακάτω παρουσιάζονται τα είδη των μηνυμάτων που αφορούν στους Περιγραφείς Αντικειμένων. Για κάθε τύπο μηνύματος υπάρχει ένας τύπος στοιχείου που αποτελεί εξειδίκευση του στοιχείου "ObjectDescriptorMessage". Οι τύποι των μηνυμάτων είναι α) ενημέρωση περιγραφέα, β) διαγραφή περιγραφέα, γ) διαγραφή όλων των περιγραφέων της ροής, δ) ενημέρωση περιγραφέα στοιχειώδους ροής και ε) διαγραφή περιγραφέα στοιχειώδους ροής.

```
<!-- Τύποι Μηνυμάτων Περιγραφέων Αντικειμένων -->
<!ELEMENT ObjectDescriptorMessage
    (ObjectDescriptorUpdate |
     ObjectDescriptorRemove | ObjectDescriptorRemoveAll |
     ES_DescriptorUpdate | ES_DescriptorRemove)

<!-- Ενημέρωση Περιγραφέα Αντικειμένου -->
<!ELEMENT ObjectDescriptorUpdate (ObjectDescriptor+)>

<!ATTLIST ObjectDescriptorUpdate
    %specializes;          "ObjectDescriptorMessage"
    length                 %integer; %implied;>

<!-- Διαγραφή Περιγραφέα Αντικειμένου -->
<!ELEMENT ObjectDescriptorRemove (ObjectDescriptorId+)>

<!ATTLIST ObjectDescriptorRemove
    %specializes;          "ObjectDescriptorMessage"
    length                 %integer; %implied;>

<!ATTLIST ObjectDescriptorId
    object_descriptor_id   %ref_type;          %implied;>

<!ATTLIST ObjectDescriptorRemoveAll
    %specializes;          "ObjectDescriptorMessage">

<!-- Ενημέρωση Περιγραφέα Στοιχειώδους Ροής -->
<!ELEMENT ES_DescriptorUpdate (ESDescriptor+)>

<!ATTLIST ES_DescriptorUpdate
    %specializes          "ObjectDescriptorMessage"
    length                %integer;          %implied;
    object_descriptor_id  %ref_type;          %implied;>
```

```

<!-- Διαγραφή Περιγραφέα Στοιχειώδους Ροής -->
<!ELEMENT ES_DescriptorRemove      (ESDescriptorId*)>

<!-- Ταυτότητα Περιγραφέα Στοιχειώδους Ροής -->
<!-- ATTTLIST ES_DescriptorRemove
      %specializes      "ObjectDescriptorMessage"
      object_descriptor_id  %ref_type;      %implied;
      stream_count        %integer;      %implied;>

<!-- Ταυτότητα Περιγραφέα Στοιχειώδους Ροής -->
<!-- ATTTLIST ESDescriptorId
      ES_descriptor_id    %ref_type;      %implied;>

```

Σχήμα 3-30. Μηνύματα Περιγραφών Αντικειμένων

Το μήνυμα ενημέρωσης Περιγραφέα Αντικειμένων που αναπαρίσταται με το στοιχείο "ObjectDescriptorUpdate" όπως φαίνεται στο παραπάνω σχήμα, περιέχει ένα ή περισσότερα στοιχεία τύπου "ObjectDescriptor". Τα στοιχεία αυτά είναι περιγραφείς που προστίθενται ή ενημερώνονται. Τα συστατικά ενός υπάρχοντος περιγραφέα δεν μεταβάλλονται με ένα μήνυμα ενημέρωσης. Η λειτουργία αυτή μπορεί να επιτευχθεί με προσθήκη ή διαγραφή περιγραφέων στοιχειωδών ροών στον ή από τον περιγραφέα αντικειμένου. Το στοιχείο διαγραφής περιγραφέων ("ObjectDescriptorRemove") καθιστά ανενεργούς ένα σύνολο από περιγραφείς αντικειμένων έτσι ώστε τα αντικείμενα μέσω των που συσχετίζονται με αυτούς δεν αναφέρονται πλέον στις στοιχειώδεις ροές που υπήρχαν στους διαγραφέντες περιγραφείς. Το στοιχείο "ObjectDescriptorRemoveAll" διαγράφει ή καθιστά μη διαθέσιμους όλους τους περιγραφείς που βρίσκονται στην εμβέλεια της ροής.

Το στοιχείο ενημέρωσης Περιγραφέων Στοιχειωδών ροών έχει παρόμοιο περιεχόμενο με το μήνυμα ενημέρωσης περιγραφέων αντικειμένων, δηλαδή περιέχει ένα ή περισσότερα στοιχεία τύπου "ES_Descriptor". Το στοιχείο αυτό αφορά σε έναν Περιγραφέα Αντικειμένου, στον οποίον θα προστεθούν αναφορές σε περιγραφείς στοιχειωδών ροών ή του οποίου οι αναφορές σε περιγραφείς στοιχειωδών ροών θα ενημερωθούν. Για να μεταβληθούν τα χαρακτηριστικά μιας στοιχειώδους ροής είναι απαραίτητο να διαγραφεί πρώτα ο περιγραφέας της και να αντικατασταθεί από ένα νέο. Η διαγραφή αναφορών σε στοιχειώδεις ροές γίνεται με το στοιχείο τύπου "ES_DescriptorRemove".

Στο σχήμα που ακολουθεί ορίζεται η δομή του περιγραφέα αντικειμένων, ο οποίος αποτελείται από τρία διαφορετικά μέρη: i) την ταυτότητα του περιγραφέα, ii) την απομακρυσμένη θέση του ή ένα σύνολο από περιγραφείς στοιχειωδών ροών και iii) ένα σύνολο από προαιρετικά στοιχεία που υποστηρίζουν μελλοντική επέκταση των περιγραφέων καθώς και ομοιόμορφη μεταφορά των ιδιωτικών δεδομένων.

```

<!-- Περιγραφέας αντικειμένων -->
<!ELEMENT ObjectDescriptor
  ((Url | (OCI_Descriptor*, ES_Descriptor+)),
  ExtensionDescriptor)>

```

```

<!ATTLIST ObjectDescriptor
    object_descriptor_id      %id_type;          %implied;
    oci_descriptor_count      %integer;          %implied;
    es_descriptor_count       %integer;          %implied; >

<!-- Αρχικός Περιγραφέας αντικειμένων -->
<!ELEMENT InitialObjectDescriptor
    ((Url | (ProfileObjectDescriptor, OCI_Descriptor*, ES_Descriptor+)),
    ExtensionDescriptor*)>

<!ATTLIST InitialObjectDescriptor
    object_descriptor_id      %id_type;          %implied;>

<!-- Περιγραφέας του προφίλ -->
<!ELEMENT ProfileObjectDescriptor
    (SceneProfile, AudioProfile, VisualProfile, GraphicsProfile)>

<!ATTLIST ProfileObjectDescriptor
    scene_profile      (systems_profile | user_private |
        no_profile | no_capability) %implied;
    audio_profile      (audio_profile | user_private |
        no_profile | no_capability) %implied;
    visual_profile     (visual_profile | user_private |
        no_profile | no_capability) %implied;
    graphics_profile   (graphics_profile | user_private |
        no_profile | no_capability) %implied;>

```

Σχήμα 3-31. Περιγραφέας Αντικειμένων και Αρχικός Περιγραφέας

Οι περιγραφείς στοιχειωδών ροών που περιέχονται μέσα στον περιγραφέα αντικειμένων περιέχουν παραμέτρους για τις στοιχειώδεις ροές που συσχετίζονται με τα αντικείμενα μέσων. Οι περιγραφείς στοιχειωδών ροών είναι δυνατό να συνοδεύονται και από ένα σύνολο στοιχείων τύπου "OCI_Descriptor". Τα στοιχεία τύπου "OCI_Descriptor" καθώς και ο περιγραφέας επέκτασης "ExtensionDescriptor" περιγράφονται παρακάτω.

Στο παραπάνω σχήμα εκτός από τον Περιγραφέα Αντικειμένων καθορίζεται και ο Αρχικός Περιγραφέας αντικειμένων, ο οποίος αποτελεί παραλλαγή του Περιγραφέα αντικειμένων και χρησιμοποιείται για την αρχική προσπέλαση του MPEG-4 περιεχομένου. Η βασική διαφορά του με τον περιγραφέα αντικειμένου είναι ότι περιέχει το προφίλ του περιεχομένου που σχετίζεται με τον συγκεκριμένο περιγραφέα.

Ο περιγραφέας στοιχειώδους ροής που φαίνεται στο παρακάτω σχήμα περιέχει πληροφορία που σχετίζεται με μία συγκεκριμένη στοιχειώδη ροή και αποτελείται από τρία βασικά τμήματα. Το πρώτο τμήμα περιέχει την ταυτότητα της στοιχειώδους ροής "ES_ID" που θα χρησιμοποιηθεί για αναφορά σε αυτήν, ένα μηχανισμό ομαδοποίησης στοιχειωδών ροών με τον ίδια περιγραφέα και ένα προαιρετικό URL που δείχνει στη θέση της συγχρονισμένα πακεταρισμένης (SL-packetized) στοιχειώδους ροής. Η ομαδοποίηση ροών γίνεται με το στοιχείο "DependsOnES", το οποίο καθορίζει την εξάρτηση αυτής της στοιχειώδους ροής από κάποια άλλη.

```

<!-- Περιγραφέας Στοιχειώδους Ροής -->

<!ELEMENT ES_Descriptor
  (DependsOnES?,
   Url?,
   ExtensionDescriptor*,
   LanguageDescriptor*,
   DecoderConfigDescriptor, SLConfigDescriptor,
   (IPIDescPointer | IPIIdentificationDataSet+)?,
   QoSDescriptor?)>

<!ATTLIST ES_Descriptor
  length          %integer;      %implied;
  ES_id           %id_type;       %implied;
  stream_dependence %boolean;     %implied;
  url_flag        %boolean;       %implied;
  num_urls        %integer;       %implied;
  stream_priority %integer;       %implied;>

<!-- Εξάρτηση στοιχειώδους ροής από μία άλλη -->

<!ATTLIST DependsOnES
  ES_id          %ref_type;       %required;>

```

Σχήμα 3–32. Περιγραφέας Στοιχειώδους Ροής

Το δεύτερο τμήμα του περιγραφέα αποτελείται από ένα σύνολο προαιρετικών περιγραφών επέκτασης, ενώ το τρίτο μέρος αποτελείται από τα στοιχεία "DecoderConfigDescriptor", "SLConfigDescriptor", "IPI_Descriptor" και "QoSDecsriptor", οι οποίοι περιέχουν παραμέτρους και απαιτήσεις της στοιχειώδους ροής που αντιστοιχεί στον περιγραφέα.

Το στοιχείο "DecoderConfigDescriptor" περιέχει τον τύπο της ροής, το μέγεθος σε bytes του προσωρινού αποταμιευτή που είναι απαραίτητος για την αποκωδικοποίηση της ροής, το μέγιστο ρυθμό δεδομένων της ροής που εμφανίζεται σε χρονικό παράθυρο του ενός δευτερολέπτου και το μέσο ρυθμό της ροής. Επίσης αποτελείται το στοιχείο "ObjectProfileIndication" που περιγράφει το προφίλ του αντικειμένου της στοιχειώδους ροής και το στοιχείο "DecoderSpecificInfo" που επιτρέπει τον καθορισμό επιπλέον πληροφορίας για συγκεκριμένο αποκωδικοποιητή, όπως φαίνεται στο Σχήμα 3–33.

```

<!-- Περιγραφέας του τύπου του αποκωδικοποιητή και των πόρων που
χρειάζεται για την αποκωδικοποίηση της στοιχειώδους ροής -->

<!ELEMENT DecoderConfigDescriptor
  (ObjectProfileIndication, DecoderSpecificInfo)>

<!ATTLIST DecoderConfigDescriptor
  length          %integer;      %implied;
  stream_type      (object_descriptor | clock_reference |
                   scene_description | visual | audio | mpeg7 |
                   object_content_info | user_private) %implied;
  up_stream        %boolean;      "FALSE" %implied;
  buffer_sizeDB    %integer;      %implied;
  max_bitrate      %integer;      %implied;
  avg_bitrate      %integer;      %implied; >

```

```

<!-- Προφίλ του αντικειμένου που περιέχεται στη στοιχειώδη ροή -->
<!ATTLIST ObjectProfileIndication
  object_profile_indication      (simple_scene | 2D_scene | VRML_scene
| audio_scene |
  complete_scene | simple_visual | core_visual | main_visual |
  simple_scalable_visual | 12_Bit_visual |
  basic_anim_2D_texture | anim_2D_mesh | simple_face |
  simple_scalable_texture | core_scalable_texture |

  AAC_main_audio | AAC_LC_audio | TF_audio |
  TF_main_scalable_audio | TF_LC_scalable_audio |
  Twin_VQ_core_audio | CELP_audio | HVXC_audio |
  HILN_audio | TTSI_audio | main_synthetic_audio |
  wavetable_synthesis_audio |

  simple_mpeg2_visual | main_mpeg2_visual | SNR_mpeg2_visual |
  spatial_mpeg2_visual | high_mpeg2_visual | 422_mpeg2_visual |
  7_mpeg2_audio | 3_mpeg2_audio |
  mpeg1_visual | mpeg1_audio |

  no_profile | user_private)>
<!-- Περιγραφέας πληροφορίας για συγκεκριμένο αποκωδικοποιητή μέσω -->
<!ATTLIST DecoderSpecificInfo
  specific_info CDATA %implied;>

```

Σχήμα 3-33. Περιγραφείς παραμέτρων αποκωδικοποιητή

Η επέκταση των περιγραφέων υλοποιείται με το στοιχείο "ExtensionDescriptor" και επιτρέπει τον ορισμό πρόσθετων περιγραφέων κυρίως για μελλοντική χρήση.

Το πρότυπο MPEG-4 επιτρέπει τον καθορισμό στοιχείων που περιγράφουν την πνευματική ιδιοκτησία του περιεχομένου. Επομένως ο περιγραφέας της στοιχειώδους ροής είναι δυνατόν να περιέχει, είτε έναν δείκτη σε μια άλλη ροή που περιέχει την πληροφορία αυτή είτε εκπεφρασμένα μέσα στο σώμα του, στοιχεία για τα πνευματικά δικαιώματα. Το στοιχείο "IP_IdentificationDataSet" χρησιμοποιείται για περιγραφή της ταυτοποίησης του περιεχομένου. Περιγράφουν τον τύπο του περιεχομένου, ο οποίος όπως φαίνεται και στο επόμενο σχήμα μπορεί να είναι από οπτικοακουστικός έως βιβλίο ή μουσικό κομμάτι, κλπ. Επίσης καθορίζει μέσω ενός διεθνούς κωδικού (ContentIdentifier) την παγκόσμια ταυτότητα του περιεχομένου.

```

<!-- Επέκταση του περιγραφέα στοιχειώδους ροής -->
<!ATTLIST ExtensionDescriptor
  length          %integer;      %implied;
  description_data %text;         %implied;>
<!-- Δείκτης στη στοιχειώδη ροή που περιέχει τα στοιχεία ταυτοποίησης του περιεχομένου -->
<!ATTLIST IPIDescPointer
  length          %text;          %implied;
  IPI_ES_id       %ref_type;      %implied;>

```

```

<!-- Περιγραφέας ταυτοποίησης και πνευματικής ιδιοκτησίας του
περιεχομένου -->

<!ELEMENT IP_IdentificationDataSet
(ContentIdentifier?, SupplContentIdentifier*)>

<!-- ATTLIST IP_IdentificationDataSet
      suppl_content_identifier          %boolean; %implied;
      suppl_content_identifier_count    %integer; %implied;
      content_type (audio_visual | book | serial | text |
                    item_or_contribution | sheet_music |
                    sound_rec_or_music_video |
                    still_picture | musical_work | others) %implied;>

<!-- ATTLIST ContentIdentifier
      type (ISAN | ISBN | ISSN | SICI | BICI | ISMN | ISRC |
            ISWC_T | ISWC_L | SPIFF | DOI) %implied;
      length %integer; %implied;
      content_identifier %text; %implied;>

<!-- ATTLIST SupplContentIdentifier
      language_code %text; %implied;
      title_length %integer; %implied;
      title %text; %implied;
      value_length %integer; %implied;
      value %text; %implied;>

```

Σχήμα 3-34. Περιγραφή της πνευματικής ιδιοκτησίας του περιεχομένου

Ένας περιγραφέας στοιχειώδους ροής είναι δυνατόν να περιέχει πληροφορία και για την ποιότητα υπηρεσιών του περιεχομένου, μέσω του περιγραφέα "QoSDescriptor", όπως φαίνεται στο παρακάτω σχήμα. Η ποιότητα υπηρεσιών περιλαμβάνει τις απαιτήσεις της ροής από το κανάλι μεταφοράς και την "κυκλοφορία" που θα δημιουργήσει η ροή στο κανάλι κατά τη μεταφορά της. Τα πιστοποιητικά της ποιότητας υπηρεσιών αποτελούνται από τον τύπο τους, τη μονάδα μέτρησης, το μήκος και την τιμή τους. Οι διαφορετικοί τύποι μέτρων που υποστηρίζονται είναι: η μέγιστη καθυστέρηση από άκρη σε άκρη καθυστέρηση της ροής, η προτιμητέα μέγιστη καθυστέρηση, η επιτρεπτή πιθανότητα απώλειας μιας μονάδας προσπέλασης, ο μέγιστος αριθμός συνεχόμενων απωλειών μονάδων προσπέλασης, το μέγιστο και το μέσο μέγεθος μιας μονάδας προσπέλασης και ο μέγιστος ρυθμός άφιξης των μονάδων προσπέλασης.

```

<!-- Περιγραφέας ποιότητας υπηρεσιών -->

<!ELEMENT QosDescriptor (QosQualifier*)>

<!-- ATTLIST QosDescriptor
      length %integer; %implied;
      predefined_profile (custom | other) %implied;
      qos_qualifier_count %integer; %implied;>

<!-- Πιστοποιητικά ποιότητας υπηρεσιών -->
<!-- ATTLIST QosQualifier
      metric_type (MAX_DELAY | PREF_MAX_DELAY | LOSS_PROB |

```



```

MAX_GAP_LOSS | MAX_AU_SIZE | AVG_AU_SIZE |
metric_unit    MAX_AU_RATE | user_defined ) %implied;
               (microseconds | fraction | integer_au | bytes |
               au_per_seconds ) %implied;
length         %integer;    %implied;
value          %text; %implied;>

<!-- Περιγραφέας άλλων μορφών περιεχομένου -->

<!ATTLIST RegistrationDescriptor
    %extends;                "ExtensionDescriptor"
    length                   %integer;    %implied;
    format_identifier        %text;       %implied;
    additional_identification_info %text; %implied;>

```

Σχήμα 3-35. Ποιότητα υπηρεσίας και περιγραφή άλλων μορφών περιεχομένου

Επίσης με τον περιγραφέα "RegistrationDescriptor" επιτρέπεται η περιγραφή της μορφής άλλου περιεχομένου που δεν καθορίζεται από το πρότυπο MPEG-4 αλλά μεταφέρεται σε στοιχειώδεις ροές αυτού. Το περιεχόμενο αυτό περιέχεται στο τμήμα private data των στοιχειωδών ροών του MPEG-4.

3.9 Περιγραφή Σκηνής

Όπως αναφέρθηκε στο κεφάλαιο 2, η περιγραφή της σκηνής γίνεται στο πρότυπο MPEG-4 με το εργαλείο BIFS (Binary Format for Scene Interchange) του τμήματος Systems. Όπως προκύπτει από το πρότυπο, η περιγραφή των σκηνών του MPEG-4 έχει τις εξής απαιτήσεις:

- Περιγραφή της ιεραρχικής δομής της σκηνής. Όπως έχει ήδη αναφερθεί η σκηνή αποτελεί ένα δέντρο από κόμβους σκηνής.
- Περιγραφή της δυναμικής συμπεριφοράς των σκηνών, εφόσον οι σκηνές του MPEG-4 δεν είναι στατικές.
- Περιγραφή των μετασχηματισμών των οπτικοακουστικών αντικειμένων από το τοπικό τους σύστημα συντεταγμένων στο σύστημα συντεταγμένων της σκηνής. Τα αντικείμενα που συμμετέχουν σε μια σύνθετη σκηνή κωδικοποιούνται χωριστά σε ένα ατομικό σύστημα συντεταγμένων και κατά η παρουσίασή τους μαζί με άλλα αντικείμενα, δηλαδή μέσα σε μια σκηνή πρέπει να γίνει σε ένα ενιαίο σύστημα συντεταγμένων. Στο εργαλείο BIFS, ο μετασχηματισμός είναι κόμβος-πατέρας του κόμβου που περιέχει το ίδιο το αντικείμενο.
- Περιγραφή των ιδιοτήτων των οπτικοακουστικών αντικειμένων οι οποίες αφορούν στον έλεγχο της θέσης και της συμπεριφοράς τους. Οι ιδιότητες αυτές είναι κατεξοχήν

πληροφορία που αφορά στην τοποθέτηση των αντικειμένων μέσα σε μια σκηνή και όχι στο περιεχόμενό τους.

- Συσχέτιση των αντικειμένων σκηνής με τα αντικείμενα που περιέχουν τα πραγματικά δεδομένα. Στο MPEG-4 η περιγραφή της σκηνής δεν περιέχει τα ίδια τα αντικείμενα και το περιεχόμενό τους, αλλά αναφορές σε αυτά. Τα αντικείμενα μέσω των μεταδίδονται στο τερματικό χωριστά.
- Περιγραφή της αλληλεπίδρασης του χρήστη με τη σκηνή. Η αλληλεπίδραση αυτή λαμβάνει χώρα μόνο στο τερματικό και δεν αφορά στον εξυπηρετητή (ή στην "άλλη άκρη" του συστήματος). Η αλληλεπίδραση του χρήστη με τη σκηνή συνεπάγεται δυναμική τροποποίηση των ιδιοτήτων των αντικειμένων της σκηνής.
- Διάκριση των σκηνών που περιέχουν μόνο ένα είδος αντικειμένων και των σκηνών με μεικτό περιεχόμενο: μόνο ακουστικές σκηνές, δυσδιάστατες σκηνές, τρισδιάστατες σκηνές και 3 τρόπους ανάμιξης του περιεχομένου των σκηνών: α) πλήρεις 2Δ και 3Δ σκηνές διαβαθμισμένες σε ένα 2Δ χώρο με βάθος, β) 2Δ και 3Δ σκηνές που χρησιμοποιούνται ως πίνακες υφής για 2Δ ή 3Δ αντικείμενα και γ) 2Δ σκηνές που παρουσιάζονται στο τοπικό x-y επίπεδο του τοπικού συστήματος συντεταγμένων μιας 3Δ σκηνής.
- Περιγραφή υποπεριοχών του συστήματος συντεταγμένων της σκηνής και περιγραφή ομαδοποίησης αντικειμένων που ανήκουν στο ίδιο τμήμα του χώρου της σκηνής.
- Υποστήριξη ταυτοποίησης των κόμβων της σκηνής για να είναι δυνατή η αναφορά μεταξύ τους.

Η περιγραφή της σκηνής στο XML DTD που παρουσιάζεται στην αναφορά έγινε με βάση την περιγραφή των κόμβων του BIFS σε MSDL, όπως ορίζεται στα Συστήματα του MPEG4. Η δομή της σκηνής περιέχει όλους τους κόμβους και κατόπιν τα ROUTEs.

3.9.1 Δομή της σκηνής

Στην παρούσα ενότητα αρχικά περιγράφεται η δομή της σκηνής και των κόμβων που την αποτελούν και κατόπιν περιγράφεται η διαδικασία αποστολής εντολών στη σκηνή, εφαρμογής κίνησης και αλληλεπίδρασης με τη σκηνή. Η δομή των κόμβων περιγράφεται στην ενότητα 3.9.2, ενώ οι διαφορετικοί τύποι κόμβων που υποστηρίζονται περιγράφονται αναλυτικά στην ενότητα 0. Τα στοιχεία τύπου ROUTE περιγράφουν τη δυναμική συμπεριφορά της σκηνής και παρουσιάζονται στην ενότητα 3.9.4.

Η σκηνή στο MPEG-4 μπορεί να είναι δύο τύπων "BIFSCCommand" ή "BIFSAAnim", δηλαδή να περιέχει εντολές τροποποίησης της σκηνής ή να περιέχει ένα σύνολο από βήματα για την κίνηση κάποιου από τα στοιχεία της σκηνής. Κατά κανόνα μια νέα σκηνή ξεκινάει με μια εντολή ενημέρωσης σκηνής, δηλαδή είναι τύπου "BIFSCCommand". Μέσα σε έναν κόμβο τύπου "SceneConfig" περιγράφονται χαρακτηριστικά που αφορούν στον τύπο της σκηνής, όπως είναι η μονάδα μέτρησης των διαστάσεων (εικονοστοιχεία ή μέτρα) και το μέγεθος της σκηνής τύπου "BIFSCCommand", ή η μάσκα της σκηνής τύπου "BIFSAAnim".

Η σκηνή επομένως περιέχει ένα στοιχείο με παραμέτρους ρύθμισης του αποκωδικοποιητή (SceneConfig), έναν κόμβο-ρίζα της ιεραρχίας των κόμβων (SceneTopNode) και ένα σύνολο από στοιχεία τύπου ROUTE. Οι κόμβοι περιέχουν πεδία τιμών που ανήκουν σε έναν από τους βασικούς τύπους δεδομένων. Στο παρακάτω σχήμα φαίνονται τα αρχικά επίπεδα της ιεραρχίας της σκηνής.

```
<!-- Ρίζα της δομής της σκηνής του MPEG-4 -->
<!ELEMENT Scene
      (SceneConfig, SceneTopNode, ROUTE*) >

<!-- Στοιχείο διαμόρφωσης της σκηνής -->
<!-- Στοιχείο διαμόρφωσης σκηνής εντολών -->
<!-- Στοιχείο διαμόρφωσης σκηνών κίνησης -->
<!-- Νέοι και επαναχρησιμοποιούμενοι κόμβοι σκηνής -->
<!-- Περιγραφή των νέων κόμβων της σκηνής -->
```

Comment [Γ.Κ.8]: εξαρτάται από το scene_type παραπάνω

Comment [Γ.Κ.9]: δεν έχουν χρησιμοποιηθεί ποτέ στην περιγραφή

Σχήμα 3-36. Δομή της ρίζας της σκηνής

Ένας κόμβος μπορεί να είναι νέος (στοιχείο "NewNode") ή να χρησιμοποιεί ένα ήδη ορισμένο κόμβο (στοιχείο "UseNode"), οπότε περιέχει την ταυτότητα του κόμβου που χρησιμοποιεί. Οι νέοι κόμβοι, αν έχουν σκοπό να χρησιμοποιηθούν σε μετέπειτα ενημερώσεις της σκηνής πρέπει να αποκτήσουν ταυτότητα, για να είναι δυνατή η αναφορά σε αυτούς. Η γλώσσα VRML χρησιμοποιεί γι' αυτό το σκοπό το μηχανισμό DEF, ο οποίος καθορίζει στοιχεία της σκηνής που μπορούν να αναφέρονται από άλλα στοιχεία. Στην παρούσα περιγραφή, διατηρούμε τα ονόματα DEF και USE για να ονομάσουμε τα χαρακτηριστικά που χρησιμοποιούνται για αναφορά στα στοιχεία και αυτά που αποτελούν αναφορά.

Η περιγραφή των κόμβων μπορεί να γίνει με 2 τρόπους: με πεδία ή με τυποποιημένα στοιχεία. Η περιγραφή με πεδία (fields) είναι αυτή που χρησιμοποιεί το BIFS. Κάθε κόμβος περιέχει ένα χαρακτηριστικό που περιγράφει τον τύπο του και ένα σύνολο από πεδία κάποιου βασικού τύπου δεδομένων. Ο αριθμός, το είδος και η σειρά των πεδίων υπονοούνται από το είδος του κόμβου. Στο δεύτερο είδος περιγραφής, ο τύπος των κόμβων καθορίζει την περαιτέρω δομή τους, η οποία αποτελείται από άλλους τύπους κόμβων. Ο τύπος αυτός είναι ένας από τους περίπου 100 δυνατούς τύπους κόμβων που μπορεί να περιέχει η σκηνή.

3.9.2 Κόμβοι σκηνής

Οι κόμβοι της σκηνής στη μορφή BIFS γενικά αποτελούνται από πεδία (Fields) που περιγράφουν τη δομή, τα χαρακτηριστικά αλλά και τη συμπεριφορά τους. Οι κόμβοι γενικά είναι δύο ειδών: μονού πεδίου (single field) ή σύνθετοι, δηλαδή πολλαπλών πεδίων (multiple field). Κάθε ένα από αυτά τα πεδία μπορεί να είναι και το ίδιο κόμβος ή κάποιο από πεδία βασικού τύπου που περιγράφονται παρακάτω. Το στοιχείο SFNode είναι κόμβος γενικού τύπου και περιέχει πεδία (Fields), τα οποία μπορεί να είναι απλά στοιχεία ή λίστες από απλούς τύπους όπως Bool, Float κλπ ή άλλου είδους κόμβοι.

Η καθεαυτό περιγραφή του κόμβου είναι δύο ειδών, MaskNodeDescription και ListNodeDescription. Η κωδικοποίηση των κόμβων με μάσκα, καθορίζει για κάθε πεδίο τύπου DEF του κόμβου, αν παρέχεται η τιμή του μέσα στην κωδικοποιημένη ψηφιοσειρά. Ο τύπος του πεδίου είναι κωδικοποιημένος στην κατάλληλη σειρά ώστε να είναι δυνατή βάσει αυτού η αποκωδικοποίησή του. Όταν η περιγραφή του κόμβου είναι ListNodeDescription, κάθε πεδίο του κόμβου είναι δυνατόν να αναφερθεί ευθέως μέσω της ταυτότητάς του. Η συντακτική ανάλυση του πεδίου εξαρτάται από το είδος του κόμβου. Όπως φαίνεται και παρακάτω, και τα δύο είδη περιγραφής των κόμβων αποτελούνται από ένα σύνολο πεδίων, καθένα από τα οποία μπορεί με τη σειρά του να είναι κόμβος.

```

<!-- Στοιχείο τύπου κόμβου πεδίου -->
<!ELEMENT   FieldNode (FieldNodeDescription)>

<!ELEMENT   FieldNodeDescription
            (MaskNodeDescription | ListNodeDescription) >

<!ELEMENT   MaskNodeDescription (Field*)>

<!ELEMENT   ListNodeDescription ((FieldRef, Field)*)>

```

Σχήμα 3-37. Περιγραφή των κόμβων πεδίου

Τα πεδία είναι δύο ειδών, μονά (SFField) ή πολλαπλών τιμών (MFField). Τα πεδία πολλαπλών τιμών αποτελούνται από ένα σύνολο μονών πεδίων, που σχηματίζουν μια λίστα ή ένα διάνυσμα. Η λίστα αποτελείται από τυχαίο αριθμό πεδίων τα οποία δεν είναι δυνατόν να αναφερθούν με βάση τη θέση τους σε αυτήν, ενώ το διάνυσμα αποτελείται από τόσα πεδία, όσο ο αριθμός `number_of_fields` και επιτρέπει την αναφορά σε καθένα από αυτά με βάση τη θέση του στο διάνυσμα.

```

<!-- Στοιχεία πεδίων -->
<!ELEMENT   Field
            (MFField | SFField)>

<!-- Σύνθετα Πεδία -->
<!ELEMENT   MFField (MFListDescription | MFVectorDescription)>

<!-- Λίστα πεδίων -->
<!ELEMENT   MFListDescription (ListSFField*)>

<!ELEMENT   ListSFField (SFField)>
<!ATTLIST   ListSFField
            next          %ref_type;  %required;>

<!-- Διάνυσμα πεδίων -->
<!ELEMENT   MFVectorDescription (VectorSFField*)>
<!ATTLIST   MFVectorDescription
            number_of_fields %integer; %implied;>

<!ELEMENT   VectorSFField (SFField)>
<!ATTLIST   VectorSFField
            index %integer;  %required;>

```

Σχήμα 3-38. Σύνθετα πεδία κόμβων

Τα μονά πεδία (SFField) ανήκουν σε έναν από τους βασικούς τύπους πεδίων: μονός κόμβος (SFNode), πεδίο δίτιμης λογικής (Bool), χρώμα (Color), αριθμός κινητής υποδιαστολής (Float), ακέραιος (Int), περιστροφή (Rotation), συμβολοσειρά (String), χρόνος (Time), ομοιόμορφη τοποθεσία πόρων (URL) και διάνυσμα δύο ή τριών αριθμών κινητής υποδιαστολής (Vector2f, Vector3f).

```

<!ELEMENT SFField
            (SFNode | Bool | Color | Float | Int |
             Rotation | String | Time | Url | Vector2f | Vector3f) >

<!ATTLIST   Bool          value %boolean; %implied;>

<!ATTLIST   Color

```

```

        red    %integer; %implied;
        green  %integer; %implied;
        blue   %integer; %implied;>

<!ATTLIST   Int    value %integer; %implied;>
<!ATTLIST   Float value %float; %implied; >
<!ATTLIST   String value %text; %implied;>
<!ATTLIST   Url    value %url; %implied;>

<!ATTLIST   Vector2f
        value1 %float; %implied;
        value2 %float; %implied; >

<!ATTLIST   Vector3f
        value1 %float; %implied;
        value2 %float; %implied;
        value3 %float; %implied; >

```

Σχήμα 3-39. Στοιχεία βασικών τύπων πεδίων

Οι κόμβοι των σκηνών του MPEG-4 αποτελούνται από τεσσάρων ειδών πεδία ως προς τη συμπεριφορά τους:

- πεδία που παίρνουν τιμή μόνο κατά την στιγμιοποίηση του κόμβου (τύπου field),
- πεδία που μπορούν να δεχτούν εισερχόμενα γεγονότα (eventIn),
- πεδία που εκπέμπουν γεγονότα (eventOut) και
- πεδία που θέτουν τιμές και ταυτόχρονα μπορούν να είναι δέκτες ή πομποί γεγονότων (exposedField).

```

<!-- Είδος της συμπεριφοράς των πεδίων των κόμβων -->

<!ENTITY %exposedField    "CDATA %fixed; &quot;exposedField&quot; ">
<!ENTITY %eventIn         "CDATA %fixed; &quot;eventIn&quot; ">
<!ENTITY %eventOut        "CDATA %fixed; &quot;eventOut&quot; ">

<!-- Πηγή δεδομένων URL για έναν κόμβο μέσω -->

<!ENTITY %url.attr        "url %url; &quot;&quot; %implied;">

<!-- Πηγή δεδομένων στοιχειώδους ροής που χρησιμοποιείται μέσω περιγραφέα
αντικειμένων -->
<!ENTITY % object_descriptor.ref  "od_ref IDREF #IMPLIED">

<!ELEMENT ObjectDescriptorRef EMPTY>
<!ATTLIST ObjectDescriptorRef
        %object_descriptor.ref; >

<!-- Χαρακτηριστικά για το χρόνο έναρξης και παύσης μιας χρονικής
διάρκειας -->

<!ENTITY %startTime.attr    "startTime %time; 0 %implied;">
<!ENTITY %stopTime.attr     "stopTime %time; 0 %implied;">

```

Στην περιγραφή των τυποποιημένων κόμβων, το είδος της συμπεριφοράς υπονοείται όταν αυτή είναι τύπου "field". Σε κάθε άλλη περίπτωση, δηλώνεται με ένα χαρακτηριστικό που έχει το όνομα "field_name.behaviour" και ακολουθείται από την παραμετρική οντότητα "%exposedField;", "%eventIn;" ή "%eventOut;", όπως φαίνεται στο παραπάνω σχήμα. Στο σχήμα αυτό φαίνεται και η δήλωση των χαρακτηριστικών τύπου URL που μπορεί να περιέχονται σε έναν κόμβο της σκηνής για να δηλώσουν τη θέση του αντικειμένου που περιέχει τα δεδομένα του αντικειμένου στο οποίο αντιστοιχεί ο κόμβος. Επίσης παρουσιάζεται και η δήλωση που αναφέρεται σε μια ταυτότητα περιγραφέα αντικειμένου, η οποία μπορεί να χρησιμοποιηθεί αντί για ένα URL για να δηλώσει ότι τα δεδομένα του αντικειμένου

Τέλος στο ίδιο σχήμα περιγράφονται οι δηλώσεις για τους τύπου χαρακτηριστικών startTime και stopTime που χρησιμοποιούνται από πολλά διαφορετικά είδη κόμβων για να περιγράψουν το χρόνο παρουσίασης των αντικειμένων.

3.9.3 Κατηγορίες των κόμβων της σκηνής

Οι κόμβοι των σκηνών του MPEG-4 κατηγοριοποιούνται με δύο τρόπους όπως φαίνεται στο παρακάτω σχήμα. Στην πρώτη κατηγοριοποίηση διακρίνονται σε πέντε κατηγορίες: κόμβοι κορυφής, κοινοί δυσδιάστατοι και τρισδιάστατοι κόμβοι, δυσδιάστατοι κόμβοι, τρισδιάστατοι κόμβοι και μικτοί κόμβοι. Η δεύτερη κατηγοριοποίηση διαχωρίζει τους κόμβους ανάλογα με τη λειτουργία τους, οπότε διακρίνονται οι εξής κατηγορίες: κόμβοι ομαδοποίησης, κόμβοι παρεμβολής, κόμβοι αισθητήρες, τρισδιάστατοι γεωμετρικοί κόμβοι, κόμβοι παιδιά, κόμβοι μέσων, κόμβοι κίνησης προσώπων και σωμάτων, κόμβοι διάφορων χαρακτηριστικών, κόμβοι κορυφής και κόμβοι μεταφοράς. Στο σχήμα αναλύονται οι κατηγορίες κόμβων ως προς τη λειτουργία τους σε ένα σύνολο τύπων η κάθε μία, ενώ η λεπτομερής περιγραφή του περιεχομένου των κόμβων γίνεται στις ενότητες 3.9.3.2 ως 3.9.3.5.

```
<!-- Τύποι κόμβων σκηνής ως προς τον αριθμό των διαστάσεων -->
<!ELEMENT TypedNodeDescription
  (TopNode | SharedNode | 2DNode | 3DNode | MixedNode )>

<!-- Τύποι κόμβων κατηγοριοποιημένοι ως προς τη λειτουργία τους -->
<!ELEMENT FunctionNodeDescription
  (GroupingNode | InterpolatorNode | SensorNode | 3DGeometryNode |
   BindableChildNode | ChildNode | MediaNode | FBANode | MiscAttrNode
   |TopNode | DeliveryNode )>

<!-- Κόμβοι ομαδοποίησης (κοινοί 2Δ & 3Δ, 2Δ, 3Δ και ήχου) -->
<!ELEMENT GroupingNode
  (Group | Inline | OrderedGroup | Switch |
```

```

    Form | Layer2D | Layout | Transform2D |
    Anchor | Billboard | Collision | Layer3D | LOD | Transform |
    AudioBuffer | AudioDelay | AudioFX | AudioMix | AudioSwitch) >

<!-- Κόμβοι παρεμβολής (κοινοί 2Δ & 3Δ, 2Δ και 3Δ) -->
<!ELEMENT InterpolatorNode
    (ColorInterpolator | ScalarInterpolator |
    PositionInterpolator2D | CoordinateInterpolator2D |
    PositionInterpolator | CoordinateInterpolator | NormalInterpolator
    | OrientationInterpolator)>

<!-- Κόμβοι αισθητήρες (κοινοί 2Δ & 3Δ, 2Δ και 3Δ) -->
<!ELEMENT SensorNode
    (Anchor | TimeSensor | TouchSensor |
    DiscSensor | PlaneSensor2D | ProximitySensor2D |
    Collision | CylinderSensor | PlaneSensor | ProximitySensor |
    SphereSensor | VisibilitySensor)>

<!-- Γεωμετρικοί κόμβοι (2Δ και 3Δ) -->
<!ELEMENT 3DgeometryNode
    (BitMap | Circle | Curve2D | IndexedFaceSet2D | IndexedLineSet2D |
    PointSet2D | Rectangle | Text |
    Box | Cone | Cylinder | ElevationGrid | Extrusion | IndexedFaceSet
    | IndexedLineSet | PointSet | Sphere)>

<!-- Προσαρτήσιμοι κόμβοι-παιδιά (2Δ και 3Δ) -->
<!ELEMENT BindableChildNode
    (Background2D |
    Background | Fog | ListeningPoint | NavigationInfo | Viewpoint)>

<!-- Κόμβοι παιδιά (κοινοί 2Δ & 3Δ, 2Δ και 3Δ) -->
<!ELEMENT ChildNode
    (AnimationStream | Conditional | Face | QuantizationParameter |
    Script | Shape | TermCap | Valuator | WordInfo |
    Sound2D |
    DirectionalLight | PointLight | Sound | SpotLight)>

<!-- Κόμβοι μέσων -->
<!ELEMENT MediaNode
    (Anchor | AnimationStream | AudioClip | AudioSource | Background |
    Background2D | ImageTexture | Inline | MovieTexture)>

<!-- Κόμβοι κίνησης προσώπων και σωμάτων -->
<!ELEMENT FBANode
    (Face | FaceDefMesh | FaceDefTables | FaceDefTransform | FDP | FIT
    | Viseme)>

<!-- Κόμβοι χαρακτηριστικά -->
<!ELEMENT MiscAttrNode
    (Appearance | Color | FontStyle | PixelTexture |
    Coordinate2D | Material2D |
    Coordinate | Material | Normal | TextureCoordinate |
    TextureTransform)>

<!-- Αρχικός κόμβος σκηνής -->

```



```
<!ELEMENT TopNode (Group2D | Layer2D | Group | Layer3D)>

<!-- Κόμβοι μεταφοράς -->

<!ELEMENT DeliveryNode
      (BIFSCCommand | BIFSAanim)>
```

Σχήμα 3-40. Κατηγορίες κόμβων σκηνής**3.9.3.1 Κόμβος-κορυφή της σκηνής**

Ο αρχικός κόμβος της σκηνής (TopNode) ανήκει σε έναν από τους παρακάτω κόμβους ομαδοποίησης, δηλαδή ομάδα στις δύο διαστάσεις, βαθμίδα στις δύο διαστάσεις και ομάδα ή βαθμίδα στις τρεις διαστάσεις. είναι κάποιος από τους κόμβους ομαδοποίησης, δηλαδή Group2D, Group, Layer2D, Layer3D ή OrderedGroup. Το πρώτο και δεύτερο είδος κόμβων κορυφής είναι σκηνές που ομαδοποιούν αντικείμενα στην ίδια διάσταση χωρίς την εφαρμογή κάποιου μετασχηματισμού. Οι κόμβοι ομαδοποίησης είναι χρήσιμοι για εφαρμογή κοινής συμπεριφοράς σε ένα σύνολο κόμβων. Για την τρισδιάστατη ομάδα κόμβων καθορίζεται ο τρόπος συνδυασμού των κόμβων της ομάδας όταν αυτοί είναι κόμβοι ήχου, ώστε να παραχθεί ένα σήμα εξόδου από τον κόμβο-ομάδα. Η δομή των δύο ειδών κόμβων ομαδοποίησης είναι ουσιαστικά η ίδια, με τη διαφορά ότι τα περιεχόμενα έχουν άλλη διάσταση. Οι βαθμίδες Layer2D και Layer3D συνδυάζουν αντικείμενα και των δύο διαστάσεων σε μία σκηνή.

Ο κόμβος Group2D περιέχει και μία λίστα παιδιών που θα προστεθούν και μία λίστα παιδιών που θα διαγραφούν. Αυτό στην ουσία χρειάζεται κατά τη διάρκεια του interaction με τη σκηνή. Τα παιδιά είναι 2-διάστατοι κόμβοι γι' αυτό γίνεται διάκριση στα στοιχεία MultipleFieldName και MultipleFieldName2D, τα οποία ουσιαστικά είναι ίδια, απλά στο Group2D υπάρχει ο περιορισμός τα παιδιά να είναι 2-διάστατοι κόμβοι.

```
<!-- Αρχικός κόμβος σκηνής -->
<!ELEMENT TopNode (Group2D | Layer2D | Group | Layer3D)>

<!-- Ομάδα δυσδιάστατων κόμβων -->
<!ELEMENT Group2D
      (Group2D.Children*, BBoxCenter, BBoxSize)>

<!ELEMENT Group2D.Children (2DNode | SharedNode)>

<!ELEMENT BBoxCenter (Vector2f)>
<!ELEMENT BBoxSize (Vector2f)>

<!-- Ομάδα τρισδιάστατων κόμβων -->
<!ELEMENT Group
      (Group.Children*, BBoxCenter, BBoxSize)>
```

Το στοιχείο Layer2D είναι ένα διαφανές ορθογώνιο ζωγραφίσματος της οθόνης, στο οποίο παρουσιάζεται μια 2D σκηνή. Είναι τμήμα της ιεραρχίας των βαθμίδων και μπορεί να

συντεθεί σε ένα 2D περιβάλλον με βάθος. Γενικά οι κόμβοι Layer2D και Layer3D επιτρέπουν τη σύνθεση πολλαπλών 2D και 3D σκηνών σε ένα 2D χώρο με βάθος.

Τα περιεχόμενα του τύπου στοιχείου Layer2D είναι ένα σύνολο από δυσδιάστατους κόμβους-παιδιά (στοιχείο "Layer2D.Children") τα οποία ορίζουν μια δυσδιάστατη σκηνή και των οποίων η σειρά μέσα στον κόμβο είναι και η σειρά ζωγραφίσματος. Επίσης αποτελείται από μία 2D βαθμίδα (στοιχείο "Layer2D.ChildrenLayer") που με τη σειρά της περιέχει μία 2D ή 3D βαθμίδα. Τα στοιχεία "Layer2D.Translation" και "Layer2D.Depth" είναι αυτά που καθορίζουν την τοποθέτηση των βαθμίδων. Ο κόμβος Layer2D είναι δυνατόν να δεχτεί και μηνύματα τροποποίησης του περιεχομένου του με τους κόμβους AddChildren, RemoveChildren, AddChildrenLayer και RemoveChildrenLayer. Το μέγεθος Size της βαθμίδας καθορίζεται ως κλάσμα του πλάτους και ύψους της βαθμίδας πατέρα του κόμβου ενώ στον κόμβο περιέχονται και τα στοιχεία του ορθογωνίου που περικλείει όλα τα παιδιά του κόμβου. Επίσης χρησιμοποιείται και ο μηχανισμός DEF, δηλαδή η ταυτότητα των κόμβων, ώστε να είναι δυνατή η διαμοίραση αντικειμένων ανάμεσα σε πολλές βαθμίδες.

```
<!-- 2D Βαθμίδα σκηνής -->

<!ELEMENT Layer2D
  (Layer2D.AddChildren?, Layer2D.RemoveChildren?,
   Layer2D.AddChildrenLayer?, Layer2D.RemoveChildrenLayer?,
   Layer2D.Children, Layer2D.ChildrenLayer,
   Layer.Size, Layer.Translation,
   Layer2D.BBoxCenter, Layer2D.BBoxSize)>

<!ATTLIST Layer2D
  %specializes;      "GroupingNode, TopNode, 2DNode, MixedNode"
  DEF %id_type;      %implied;
  depth %integer;    "0" %implied;>

<!ELEMENT Layer2D.AddChildren      (2DNode*)>
<!ELEMENT Layer2D.RemoveChildren   (2DNode*)>
<!ELEMENT Layer2D.AddChildrenLayer (2DNode*)>
<!ELEMENT Layer2D.RemoveChildrenLayer (2DNode*)>

<!ELEMENT Layer2D.Children      (2DNode*)>
<!ELEMENT Layer2D.ChildrenLayer (Layer2D | Layer3D)>

<!ELEMENT Layer.Size      (Vector2f)>
<!ELEMENT Layer.Translation (Vector2f)>
<!ELEMENT Layer2D.BBoxSize (Vector2f)>
<!ELEMENT Layer2D.BBoxCenter (Vector2f)>
```

Σχήμα 3-41. 2D βαθμίδα σκηνής

Μια τρισδιάστατη βαθμίδα έχει παρόμοια δομή με αυτήν της δυσδιάστατης, δηλαδή μέγεθος, κόμβους-παιδιά, μία βαθμίδα κόμβων-παιδιών, μετατόπιση της βαθμίδας παιδιών και μηνύματα ενημέρωσης των παιδιών της. Το σχήμα που περικλείει μια τρισδιάστατη βαθμίδα είναι και αυτό τρισδιάστατο, δηλαδή κύβος και επομένως το κέντρο και το

μέγεθός του είναι τρισδιάστατα μεγέθη. Επίσης η τρισδιάστατη βαθμίδα περιέχει τους κόμβους Background, Fog, NavigationInfo και Viewpoint που δηλώνουν επίσης γεγονότα εισόδου στον κόμβο-βαθμίδα. Οι κόμβοι αυτοί περιγράφουν χαρακτηριστικά της σκηνής που μπορεί να είναι ενεργά μόνο ένα κάθε χρονική στιγμή και σχηματίζουν μια στοίβα κατά τη διάρκεια της σκηνής.

```
<!-- 3D Βαθμίδα σκηνής -->

<!ELEMENT Layer3D
  (Layer3D.AddChildren?, Layer3D.RemoveChildren?,
   Layer3D.AddChildrenLayer?, Layer3D.RemoveChildrenLayer?,
   Layer3D.Children, Layer3D.ChildrenLayer,
   Layer.Size, Layer.Translation,
   Background, Fog, NavigationInfo, Viewpoint,
   Layer3D.BBoxCenter, Layer3D.BBoxSize)>

<!ATTLIST Layer3D
  %specializes;      "GroupingNode, TopNode, 3DNode, MixedNode"
  DEF %id_type;      %implied;
  depth %integer;    "0" %implied;>

<!ELEMENT Layer3D.AddChildren      (3DNode*)>
<!ELEMENT Layer3D.RemoveChildren   (3DNode*)>
<!ELEMENT Layer3D.AddChildrenLayer (3DNode*)>
<!ELEMENT Layer3D.RemoveChildrenLayer (3DNode*)>

<!ELEMENT Layer3D.Children      (3DNode*)>
<!ELEMENT Layer3D.ChildrenLayer (Layer2D | Layer3D)>

<!ELEMENT Layer3D.BBoxSize      (Vector3f)>
<!ELEMENT Layer3D.BBoxCenter    (Vector3f)>
```

Σχήμα 3-42. 3D βαθμίδα σκηνής

3.9.3.2 Κοινοί 2D και 3D κόμβοι

Οι κοινοί δυοδιάστατοι και τρισδιάστατοι κόμβοι φαίνονται στο παρακάτω σχήμα. Ο κόμβος τύπου "AnimationStream" χρησιμοποιείται για την περιγραφή των παραμέτρων ελέγχου των ροών τύπου "BIFSAnim". Τα χαρακτηριστικά `loop`, `StartTime`, `stopTime`, `isActive` ορίζονται στην VRML. Το χαρακτηριστικό `"speed"` ελέγχει την ταχύτητα αναπαραγωγής. Αν η ταχύτητα είναι θετική, τότε η αναπαραγωγή γίνεται με βάση τον τύπο *fmod(τρέχων_χρόνος-εκκίνηση, διάρκεια/ταχύτητα)*, ενώ όταν είναι αρνητική, η αναπαραγωγή γίνεται με βάση τον τύπο *διάρκεια+fmod(τρέχων_χρόνος-εκκίνηση, διάρκεια/ ταχύτητα)*. Αυτό σημαίνει ότι για θετικές τιμές της ταχύτητας αναπαραγωγής, η αναπαραγωγή γίνεται σε πολλαπλάσιο της αρχικής ταχύτητας. Για δεδομένα ροής (streaming media), οι αρνητικές τιμές της ταχύτητας αγνοούνται.

```
<!ELEMENT SharedNode
  ( AnimationStream |
    AudioDelay | AudioMix | AudioSource | AudioFX | AudioSwitch
    Conditional |
    MediaTimeSensor |
```

Comment [Γ.Κ.10]: να τα κοιτάξω ακριβώς

Comment [Γ.Κ.11]: τι συνάρτηση είναι αυτή;

```

    QuantizationParameter |
    TermCap |
    Valuator |
    SharedVRMLNode )

<!ATTLIST AnimationStream
    loop                %bool;          FALSE %implied;
    speed               %integer;       1      %implied;
    %startTime.attr;
    %stopTime.attr;
    %url.attr;
    isActive            %bool;          %implied;
    isActive.behavior   %eventOut; >

```

Σχήμα 3–43. Κοινοί 2Δ και 3Δ κόμβοι σκηνών

Ο κόμβος "AudioDelay" που περιγράφεται στο παρακάτω σχήμα, επιτρέπει τον έλεγχο του χρόνου εκκίνησης και ακινητοποίησης του ήχου. Ο ρόλος του κόμβου είναι να εισάγει καθυστέρηση στην αναπαραγωγή όλων των κόμβων-παιδιών του (στοιχεία που περιέχονται στο στοιχείο "Children"), σε σχέση με το χρόνο που καθορίζεται από τους κόμβους AudioSource. Το χαρακτηριστικό "delay" περιγράφει την χρονική καθυστέρηση που θα εφαρμοστεί σε καθένα από τα παιδιά, ενώ το χαρακτηριστικό "num_channels" περιγράφει τον αριθμό των καναλιών ήχου που εξάγονται από αυτόν τον κόμβο. Το στοιχείο "PhaseGroup" καθορίζει τις σχέσεις ανάμεσα στα διάφορα κανάλια εξόδου.

Το στοιχείο "AudioMix" χρησιμοποιείται για να περιγράψει τη μίξη διαφορετικών ηχητικών σημάτων με πολλαπλασιαστικό τρόπο. Το χαρακτηριστικό "num_inputs" που περιγράφει τον αριθμό των καναλιών πρέπει να έχει τιμή ίση με το άθροισμα των καναλιών των κόμβων-παιδιών του στοιχείου, ενώ το χαρακτηριστικό "num_channels" παριστάνει τον αριθμό των καναλιών εξόδου του κόμβου. Το στοιχείο "Matrix" περιγράφει τον πίνακα μίξης των καναλιών εισόδου με τα κανάλια εξόδου και έχει διαστάσεις "num_inputs x num_channels". Το στοιχείο "Phase" έχει την ίδια σημασιολογία με αυτή στον κόμβο "AudioDelay", δηλαδή περιγράφει τη σχέση φάση ανάμεσα στα κανάλια εξόδου.

```

<!ENTITY % num_channels.attr "num_channels %integer; 1 %implied;">
<!ENTITY % pitch.attr        "pitch %float; 1 %implied;">

<!-- Κόμβος εισαγωγής καθυστέρησης σε ομάδα ήχων -->
<!ELEMENT AudioDelay (Children, Phase*)>
<!ATTLIST AudioDelay
    delay                %time;          "0" %implied;
    %num_channels.attr;>

<!ATTLIST Phase
    phase                %integer;       %implied;>

<!-- Κόμβος μίξης ήχων -->
<!ELEMENT AudioMix (Children, Matrix, Phase*)>
<!ATTLIST AudioMix
    num_inputs           %integer;       "1" %implied;
    %num_channels.attr;>

<!-- Πηγή ήχου -->

```

Comment [Γ.Κ.12]: Τι είδους κόμβοι είναι τα παιδιά;

Comment [Γ.Κ.13]: Τι τύπου είναι;

χαρακτηριστικού "activate" ή το γεγονός FALSE στην τιμή του "reverseActivate", τότε ερμηνεύει τα περιεχόμενα του απομονωτή ως λειτουργίες ενημέρωσης της σκηνής.

Ο κόμβος "MediaTimeSensor" παρέχει ένα μηχανισμό προσκόλλησης μιας ψηφιοσειράς μέσω σε ένα συγκεκριμένο κόμβο τύπου "TimeSensor", κάνοντας τον μετρητή του χρόνου να ακολουθεί τη χρονική βάση της ψηφιοσειράς. Τα στοιχεία "Media" και "Timer" του κόμβου είναι και τα ίδια κόμβοι.

Το στοιχείο "QuantizationParameter" χρησιμοποιείται για να περιγράψει τις τιμές κβαντοποίησης που θα εφαρμοστούν σε μονά πεδία αριθμητικού τύπου, δηλαδή για κάθε κατηγορία πεδίου περιέχει τη μέγιστη και την ελάχιστη τιμή καθώς και των αριθμό των δυαδικών ψηφίων που χρησιμοποιούνται για την αναπαράσταση του πεδίου. Ο κόμβος αυτός μπορεί να είναι παιδί μόνο κάποιου κόμβου ομαδοποίησης.

Ο κόμβος "TermCap" περιγράφει τις ικανότητες του τερματικού ως προς τα εξής χαρακτηριστικά: ρυθμός εικόνων, βάθος χρώματος, μέγεθος οθόνης, υλικό για γραφικά, μορφή εξόδου του ήχου, μέγιστος ρυθμός δειγματοληψίας του ήχου, χωρική δυνατότητα ήχου καθώς και φόρτος του επεξεργαστή και της μνήμης. Στο πρότυπο MPEG-4, ο κόμβος περιέχει ένα πεδίο που ονομάζεται "evaluate", ένα πεδίο που ονομάζεται "capability" και το πεδίο "value". Το πεδίο "evaluate" χρησιμοποιείται ως γεγονός που θα πυροδοτήσει τον υπολογισμό της τιμής που αντιπροσωπεύει ο κωδικός "capability" και θα θέσει την τιμή που υπολογίστηκε, στο πεδίο "value" του κόμβου. Επομένως ο κόμβος αυτός έχει δυναμική συμπεριφορά, δηλαδή δέχεται και εκπέμπει γεγονότα. Παρακάτω περιγράφεται η δομή του κόμβου ως ένα στατικό στοιχείο που περιέχει την τιμή για κάποια δυνατότητα του τερματικού, χωρίς να καθορίζεται η διαδικασία δυναμικής συμπεριφοράς. Το στοιχείο "FrameRate" που περιγράφει το ρυθμό των εικόνων, παίρνει τιμές κωδικούς που αντιστοιχούν σε ζώνες ρυθμός όπως φαίνεται στο σχήμα. Το χρώμα περιγράφεται ως προς τον αριθμό των δυαδικών ψηφίων ανά εικονοστοιχείο, το μέγεθος της εικόνας περιγράφεται με γραμμές, ενώ για τα γραφικά περιγράφεται το είδος της επιτάχυνσης που χρησιμοποιείται. Η έξοδος του ήχου μπορεί να είναι μονοφωνική, σε στερεοφωνικά ηχεία ή ακουστικά, surround πέντε καναλιών ή περισσότερα από πέντε κανάλια. Ο φόρτος του επεξεργαστή καθορίζεται ως ποσοστό των μέγιστων δυνατοτήτων, ενώ ο φόρτος της μνήμης ως απόλυτο ποσό διαθέσιμης μνήμης. Τα υπόλοιπα στοιχεία δυνατοτήτων παρατίθενται στο Παράρτημα Α.

```
<!-- Ενεργοποίηση εντολών τροποποίησης της σκηνής -->
<!ATTLIST Conditional
    activate          %bool; "FALSE" %implied;
    activate.behavior %eventIn;
    reverseActivate   %bool; "FALSE" %implied;
    reverseActivate.behavior %eventIn;
```

```

        buffer          %text;
        isActive        %bool; %implied;
        isActive.behavior %eventOut;>

<!ELEMENT MediaTimeSensor (Media, Timer)>

<!-- Δυνατότητες του τερματικού -->

<!ELEMENT TermCap (Capability)>
<!ELEMENT Capability
    (FrameRate | ColorDepth | ScreenSize | GraphicsHardware |
     AudioOutputFormat | MaxAudioSamplRate | SpatialAudioCap |
     CPUload | MemLoad)>

<!ATTLIST FrameRate
    value (unknown | less_than_5fps | 5_to_10_fps | 10_to_20fps |
          20_to_40fps | more_than_40fps) %implied; >

```

Το στοιχείο τύπου "Valuator" λαμβάνει γεγονότα οποιουδήποτε τύπου και πυροδοτεί γεγονότα διαφόρων τύπων με βάση τις δεδομένες τιμές και επομένως λειτουργεί ως προσαρμογέας γεγονότων. Μία χρήση αυτού του κόμβου είναι η τροποποίηση της επιλογής από έναν κόμβο τύπου "Switch" ως αποτέλεσμα κάποιου γεγονότος. Ο κόμβος δηλαδή δέχεται ένα γεγονός που πρέπει να μεταβάλλει την επιλογή και κατευθύνει το γεγονός μεταβολής του πεδίου "whichChoice" του κόμβου επιλογής.

```

<!ELEMENT SharedVRMLNode
    (Appearance |
     AudioClip |
     Color | ColorInterpolator |
     FontStyle |
     ImageTexture | MovieTexture |
     ScalarInterpolator |
     Shape |
     Switch |
     Text |
     TextureCoordinate | TextureTransform |
     TimeSensor |
     TouchSensor |
     WorldInfo) >

<!ELEMENT Appearance (Material, Texture, TextureTransform)>

```

Στους κοινούς δυσδιάστατους και τρισδιάστατους κόμβους ανήκουν και κόμβοι που είναι κοινοί με τη γλώσσα VRML με κάποιες επεκτάσεις για το MPEG-4, όπως φαίνεται στο επόμενο σχήμα.

Ο κόμβος "AudioClip"+ χρησιμοποιείται για να μπορέσουν μικρά τμήματα ήχου, όπως ο ήχος που προκαλείται από ενέργειες με το ποντίκι, να χρησιμοποιηθούν σε μια αλληλεπιδραστική σκηνή. Στο BIFS ο κόμβος αυτός συσσωρεύει τα δεδομένα σε έναν απομονωτή, ώστε να μπορούν να ξαναπαιχτούν μετά από αλληλεπίδραση του χρήστη. Στην ουσία ο κόμβος αποτελεί μηχανισμό μετατροπής των ρών ήχου σε στατική μορφή, ώστε να μπορεί να χρησιμοποιηθεί αλληλεπιδραστικά.

Ο κόμβος MovieTexture παρουσιάζει ένα καρτέ ή επίπεδο αντικειμένου βίντεο. Όταν η ταχύτητα παρουσίασης είναι 0, παρουσιάζεται μόνο το πρώτο καρτέ ή επίπεδο αντικειμένου βίντεο. Για θετικές τιμές της ταχύτητας, υπολογίζεται σε κάθε χρονική στιγμή ποια μονάδα του βίντεο (καρτέ ή VOP) θα παρουσιαστεί, με βάση το χρόνο που έχει περάσει από το χρόνο έναρξης, τη διάρκεια και την ταχύτητα. Για αρνητικές τιμές της ταχύτητας, η παρουσίαση αρχίζει από τις μονάδες που από το τέλος του αντικειμένου.

```
<!ELEMENT AudioClip (Loop, Pitch, StartTime, StopTime, URL,
    DurationChnaged, IsActive)>

<!ELEMENT ColorInterpolator (SetFraction, Key, KeyValue, ValueChanged)>

<!ELEMENT FontStyle (Family, Horizontal, Justify, Language, LeftToRight,
    Size, Spacing, Style, TopToBottom)>

<!ELEMENT ImageTexture (URL, RepeatS, RepeatT)>
<!ELEMENT MovieTexture (Loop, Speed, StartTime, StopTime, URL, ReapeatS,
    RepeatT, DurationChanged, IsActive)>
<!ELEMENT ScalarInterpolator (SetFraction, Key, KeyValue, ValueChanged)>

<!ELEMENT Shape (Appearance, Geometry)>

<!ELEMENT Switch (Choice, WhichChoice)>

<!ELEMENT Text (String, Length, FontStyle, MaxExtent)>

<!ELEMENT TextureCoordinate (Point)>

<!ELEMENT TextureTransform (Center, Rotation, Scale, Translation)>

<!ELEMENT TimeSensor (CycleInterval, Enabled, Loop, StartTime, StopTime,
    CycleTime, FractionChanged, IsActive, Time)>

<!ELEMENT TouchSensor (Enabled, HitNormalChanged, HitPointChanged,
    HitTextCoordChanged, IsActive, IsOver, TouchTime)>

<!ELEMENT WorldInfo (Info, Title)>
```

3.9.3.3 Κόμβοι δύο διαστάσεων

Οι κόμβοι των δύο διαστάσεων είναι αυτοί που μπορούν να χρησιμοποιηθούν σε σκηνές δύο διαστάσεων και σε κόμβους που επιτρέπουν τη χρήση δυσδιάστατων στοιχείων σε τρισδιάστατες σκηνές.

```
<!ELEMENT 2DNode (Background2D | Circle | Coordinate2D | Curve2D |
    DiscSensor | Form |
    Group2D |
    Image2D |
    IndexedFaceSet2D | IndexedLineSet2D |
    Inline2D |
    Layout |
    LineProperties | Material2D | PlaneSensor2D |
    PointSet2D |
    Position2DInterpolator | Proximity2DSensor |
    Rectangle |
    Sound2D |
    Switch2D |
```



```
Transform2D |
VideoObject2D) >
```

Κατά την παρουσίαση μιας σκηνής, υπάρχει μια στοίβα από αντικείμενα δυσδιάστατου φόντου που μπορούν να παρουσιαστούν μέσω του κόμβου Background2D. Ο κόμβος αυτός δεν είναι γεωμετρικός και δεν υπόκειται σε μετασχηματισμούς κλιμάκωσης ή περιστροφής. Αποτελείται από τρία πεδία: το πεδίο SetBind που δηλώνει ένα εισερχόμενο γεγονός (eventIn) και όταν πάρει την τιμή TRUE μετακινείται στην κορυφή της στοίβας των δυσδιάστατων φόντων. Αντίστοιχα το πεδίο IsBound είναι ένα εξερχόμενο πεδίο που λαμβάνει την τιμή TRUE όταν μόλις το συγκεκριμένο φόντο βρεθεί στην κορυφή της στοίβας. Το στοιχείο URL χρησιμοποιείται όταν υπάρχει δείχνει την πηγή των δεδομένων για το δυσδιάστατο φόντο σε κάποια τοποθεσία που έχει διεύθυνση στον Παγκόσμιο Ιστό, εναλλακτικά με το στοιχείο ObjectDescriptorRef που χρησιμοποιείται για να δηλώσει ότι τα δεδομένα του κόμβου βρίσκονται σε μια στοιχειώδη ροή, η οποία ταυτοποιείται μέσω ενός περιγραφέα αντικειμένων.

Το στοιχείο DiscSensor περιγράφει την περιστροφή ενός αντικειμένου σε ένα επίπεδο δύο διαστάσεων, γύρω από έναν άξονα στο τοπικό σύστημα συντεταγμένων.

```
<!ELEMENT Background2D (URL | ObjectDescriptorRef)>
<!ATTLIST Background2D
    set_bind                %boolean;
    set_bind.behaviour      %eventIn;
    is_bound                %boolean;
    set_bind.behaviour      %eventOut; >

<!ELEMENT Circle (Radius)>

<!ELEMENT Coordinate2D (Point)>

<!ELEMENT Curve2D (Points, Fineness)>
<!ELEMENT DiscSensor (AutoOffset, Center, Enabled, MaxAngle, MinAngle,
    Offset, IsActive, RotationChanged, TrackPointChanged)>

<!ELEMENT Form (Children, Size, Groups, Constraint) >
<!ELEMENT Group2D (AddChildren, RemoveChildren, Children, BBoxCenter,
    BBoxSize)>
```

Σχήμα 3-45. Δυσδιάστατοι κόμβοι σκηνών

Το στοιχείο Form καθορίζει την τοποθέτηση των κόμβων-παιδιών του σύμφωνα με μια σχετική ευθυγράμμιση και μία κατανομή μεταξύ τους. Η διάταξη αυτή γίνεται σε ένα καρέ και τα αντικείμενα μπορούν να οργανωθούν σε ομάδες για να εφαρμοστούν σε αυτά οι περιορισμοί της διάταξης.

Μια ομάδα δύο διαστάσεων αποτελεί στην ουσία από μόνη της μια δυσδιάστατη σκηνή, η οποία δεν περιέχει στη ρίζα της κάποιο μετασχηματισμό. Περιέχει κόμβους που

περιγράφουν τη δυναμική προσθήκη ή αφαίρεση παιδιών της ομάδας, τα ίδια τα παιδιά της, καθώς και το κέντρο και το μέγεθος του περικλειόντος ορθογωνίου.

Το στοιχείο Image2D είναι κόμβος που επιτρέπει την τοποθέτηση μιας εικόνας στο φυσικό της μέγεθος μέσα σε μια δυοδιάστατη σκηνή. Η εικόνα αυτή δεν μετασχηματίζεται για να παρουσιαστεί, ούτε αντιστοιχίζεται ως υφή.

```
<!ELEMENT Image2D (URL)>

<!ELEMENT IndexedFaceSet2D
  (SetColorIndex, SetCoordIndex, SetTexCoordIndex, Color, Coord,
   TexCoord, ColorIndex, ColorPerVertex, Convex, CoordIndex,
   TexCoordIndex)>

<!ELEMENT IndexedLineSet2D (SetColorIndex, SetCoordIndex, Color, Coord,
  ColorIndex, ColorPerVertex, CoordIndex)>

<!ELEMENT Inline2D (URL, BBoxCenter, BBoxSize)>

<!ELEMENT Layout (Children, Wrap, Size, Horizontal, Justify, LeftToRight,
  TopToBottom, Spacing, SmoothScroll, Loop, ScrollVertical,
  ScrollRate, AddChildren, RemoveChildren)>

<!ELEMENT LineProperties (LineColor, LineStyle, Width)>

<!ELEMENT Material2D (EmissiveColor, Filled, LineProps, Transparency)>

<!ELEMENT PlaneSensor2D (AutoOffset, Enabled, MaxPosition, MinPosition,
  Offset, TrackPointChanged)>

<!ELEMENT PointSet2D (Color, Coord)>

<!ELEMENT Position2DInterpolator (SetFraction, Key, KeyValue,
  ValueChanged)>

<!ELEMENT Proximity2DSensor (Center, Size, Enabled, IsActive,
  PositionChanged, OrientationChanged, EnterTime, ExitTime)>

<!ELEMENT Rectangle (Size)>

<!ELEMENT Sound2D (Intensity, Location, Source, Spatialize)>

<!ELEMENT Switch2D (Choice, WhichChoice)>

<!ELEMENT Transform2D (AddChildren, RemoveChildren, Center, Children,
  RotationAngle, Scale, ScaleOrientation, DrawingOrder,
  Translation, BBoxCenter, BBoxSize)>

<!ELEMENT VideoObject2D (Loop, Speed, StartTime, StopTime, URL,
  DurationChanged, IsActive)>
```

3.9.3.4 Κόμβοι τριών διαστάσεων

Στο επόμενο σχήμα παρουσιάζεται η περιγραφή των τρισδιάστατων κόμβων. Το στοιχείο Sound χρησιμοποιείται για την προσθήκη ήχου στη σκηνή, δίνοντάς του χωρική διάσταση και συσχετίζοντας τον με το οπτικό περιεχόμενο. Ο κόμβος αυτός συσχετίζει ένα ακουστικό υποδέντρο με την υπόλοιπη σκηνή.

```

<!ELEMENT 3DNode
  (Group | Collision | Billboard | Transform | LOD |
   Background | ListeningPoint | Viewpoint |
   Sound | SpotLight |
   ProximitySensor |
   NormalInterpolator | OrientationInterpolator |
   CoordinateInterpolator | PositionInterpolator |
   Box | Cone | Coordinate | Sphere | Cylinder | ElevationGrid |
   Extrusion | IndexedFaceSet | IndexedLineSet | PointSet |
   DirectionalLight | PointLight |
   Inline |
   Face |
   Material | Normal)>

<!ELEMENT ListeningPoint (SetBind, Jump, Orientation, Position,
  Description, BindTime, IsBound)>

<!ELEMENT Face (FIT, FDP, FAP, URL, RenderedFace)>

<!ELEMENT FAP (Viseme, Expression, OpenJaw, LowerTMidlip, RaiseBMidlip,
  StretchLCorner, StretchRCorner, LowerTLipLM)>

<!ELEMENT Sound (Direction, Intensity, Location, MaxBack, MaxFront,
  MinBack, MinFront, Priority, Source, Spatialize)>

```

3.9.3.5 Μικτοί κόμβοι

Οι μικτοί κόμβοι είναι κατάλληλοι για τη δημιουργία σκηνών με μικτό δυσδιάστατο και τρισδιάστατο περιεχόμενο. Οι μικτές σκηνές περιέχουν διάφορες ζωγραφισμένες 2Δ ή 3Δ σκηνές, χρησιμοποιούν μια ζωγραφισμένη 2Δ ή 3Δ σκηνή ως χάρτη υψής ή ζωγραφίζουν μια 2Δ σκηνή σε ένα 3Δ σύστημα συντεταγμένων. Οι κόμβοι βαθμίδες περιγράφονται στην ενότητα 3.9.3.1 γιατί αποτελούν και κόμβους-κορυφή της σκηνής.

```

<!-- Μικτοί 2Δ και 3Δ κόμβοι -->

<!ELEMENT MixedNode
  (Layer2D | Layer3D |
   Composite2DTexture | Composite3DTexture | CompositeMap)>

```

Ένας άλλος τρόπος συνδυασμού δυσδιάστατων και τρισδιάστατων στοιχείων σε μια σκηνή είναι τα στοιχεία τύπου Composite2Dtexture, Composite3Dtexture, CompositeMap. Το πρώτο παριστάνει μια υφή που συντίθεται από μία δυσδιάστατη σκηνή και η οποία μπορεί να αντιστοιχιστεί σε ένα 2Δ αντικείμενο. Αποτελείται από ένα σύνολο 2Δ κόμβων-παιδιών που ορίζουν τη σκηνή που θα χρησιμοποιηθεί ως χάρτης υψής και από το ιδανικό πλάτος και ύψος της υψής σε μονάδες εικονοστοιχείων. Το στοιχείου τύπου Composite3Dtexture μπορεί να σχηματίσει έναν χάρτη τρισδιάστατης υψής από μία 3Δ σκηνή. Η υφή αυτή μπορεί να χρησιμοποιηθεί σε ένα τρισδιάστατο αντικείμενο. Οι κόμβοι Background, Fog, NavigationInfo και Viewpoint δηλώνουν γεγονότα εισόδου που αναπαριστούν τις τρέχουσες τιμές των φύλλων της σκηνής. Ο κόμβος τύπου CompositeMap επιτρέπει σε μία δυσδιάστατη σκηνή να εμφανιστεί σε ένα επίπεδο μιας τρισδιάστατης σκηνής. Η

λειτουργικότητά του είναι παρόμοια με του στοιχείου Composite2Dtexture με τη διαφορά ότι εδώ δεν είναι απαραίτητο να ζωγραφιστεί η 2D σκηνή στο τοπικά ΧΥ επίπεδο.

3.9.4 ROUTE

Το στοιχείο τύπου ROUTE περιέχει τον κόμβο αποστολέα του γεγονότος, το πεδίο του κόμβου αποστολέα που στέλνει το γεγονός (τύπου eventOut ή exposedField), καθώς και τον κόμβο δέκτη και το πεδίο του που δέχεται το γεγονός (τύπου eventIn ή exposedField). Η αναφορά στους κόμβους αποστολής και λήψης του γεγονότος υλοποιείται στην περιγραφή μέσω των ταυτοτήτων των κόμβων και των ονομάτων των χαρακτηριστικών τους, όπως φαίνεται στο παρακάτω σχήμα.

```
<!-- ROUTE -->
<ELEMENT ROUTE (UpdatableROUTE?)>
<ATTLIST ROUTE
    out_node_ID      %ref_type;      %required;
    out_field_ref    CDATA          %required;
    in_node_ID       %ref_type;      %required;
    in_field_ref     CDATA          %required;>
<ATTLIST UpdatableROUTE
    route_ID        %id_type;      %required;>
<!-- Στοιχεία πολλών ROUTEs -->
<ELEMENT ROUTEs (ListROUTEs | VectorROUTEs)>
<ELEMENT ListROUTEs (ROUTE*)>
<ELEMENT VectorROUTEs (ROUTE*)>
<ATTLIST VectorROUTEs
    length %integer; >
```

Σχήμα 3-46. ROUTEs

Επίσης αν το στοιχείο τύπου ROUTE έχει τη δυνατότητα να γίνει αντικείμενο εντολών ενημέρωσης, τότε περιέχει και την ταυτότητά του (UpdatableROUTE.route_ID).

3.10 Ανακεφαλαίωση

Στο κεφάλαιο αυτό παρουσιάστηκε η αναπαράσταση οπτικοακουστικού περιεχομένου που είναι κωδικοποιημένο με το πρότυπο MPEG-4. Η αναπαράσταση αυτή αποτελεί βασικό τμήμα της γενικής υποδομής υποστήριξης υπηρεσιών διαχείρισης οπτικοακουστικού περιεχομένου και αποτελείται από τη Γλώσσα Περιγραφής Οπτικοακουστικού Περιεχομένου AVDL. Η γλώσσα AVDL περιγράφει τη δομή και τη σημασιολογία του περιεχομένου με απλό και ευέλικτο τρόπο, ο οποίος επειδή στηρίζεται στη μεταγλώσσα αναπαράστασης εγγράφων XML είναι κατάλληλος και για εφαρμογές του Παγκόσμιου Ιστού. Η AVDL περιγράφει τα οπτικά αντικείμενα του MPEG-4, τους περιγραφείς

αντικειμένων οι οποίοι επιτρέπουν τη συσχέτιση μεταξύ τους, αλλά και τη σύνθεση των αντικειμένων σε οπτικοακουστικές σκηνές. Στο επόμενο κεφάλαιο η γλώσσα AVDL επεκτείνεται ώστε να περιγράψει το περιεχόμενο του προτύπου MPEG-2, το οποίο αποτελεί σημαντικό πρότυπο για τις οπτικοακουστικές εφαρμογές.

Η αναπαράσταση ή περιγραφή του περιεχομένου των MPEG-4 και MPEG-2 με την γλώσσα AVDL, χρησιμοποιείται στον πυρήνα του πληροφοριακού συστήματος εξυπηρέτησης οπτικοακουστικών εφαρμογών μέσω τριών βασικών λειτουργιών: έλεγχος της παρουσίασης του υλικού, ανάκτηση βάσει περιεχομένου, διαβαθμισμένη μεταβίβαση με βάση προφίλ χρηστών και συσκευών και ανταλλαγή περιεχομένου. Η λειτουργικότητα αυτή υποστηρίζεται από τη γλώσσα AVDL σε συνδυασμό με την κατασκευή βάσεων δεδομένων οπτικοακουστικού περιεχομένου οι οποίες περιγράφονται στο Κεφάλαιο 5.

Κεφάλαιο 4

Διαχείριση MPEG περιεχομένου σε εξυπηρετητή πολυμέσων

Στο κεφάλαιο αυτό περιγράφεται η διαχείριση οπτικοακουστικού περιεχομένου που υλοποιήθηκε στο σύστημα ΚΥΔΩΝΙΑ στα πλαίσια του προγράμματος SICMA. Επίσης παρουσιάζεται η επέκταση και τροποποίηση της εν λόγω υπηρεσίας για την υποστήριξη της γενικής υποδομής διαχείρισης οπτικοακουστικών μέσων στον πυρήνα του πληροφοριακού συστήματος ΚΥΔΩΝΙΑ. Η διαχείριση των οπτικοακουστικών μέσων περιλαμβάνει τη μοντελοποίηση του περιεχομένου τους στην αντικειμενοστρεφή βάση δεδομένων του συστήματος ΚΥΔΩΝΙΑ, τη δεικτοδότηση MPEG-2 βίντεο και MPEG-1 συγχρονισμένου βίντεο και ήχου για την παροχή υπηρεσιών περιήγησης και την υποστήριξη της Υπηρεσίας Περιεχομένου σύμφωνα με το πρότυπο DAVIC. Για την υποστήριξη της γενικής υποδομής διαχείρισης οπτικοακουστικής πληροφορίας παρουσιάζεται η κατασκευή της γλώσσας αναπαράστασης των δεδομένων του MPEG-2 η οποία συμπληρώνει την αναπαράσταση του περιεχομένου με ενιαίο τρόπο για τα δύο πρότυπα κωδικοποίησης. Επίσης παρουσιάζεται ο σχεδιασμός κατάλληλων δομών δεικτοδότησης για την πληροφορία του MPEG-4, οι οποίες χρησιμοποιούνται για την υποστήριξη υπηρεσιών περιήγησης και ελεγχόμενης παρουσίασης σε οπτικά αντικείμενα και σκηνές του προτύπου αυτού.

4.1 Εισαγωγή

Οι εξυπηρετητές πολυμέσων παρέχουν ένα σύνολο από υπηρεσίες διαχείρισης της πληροφορίας, οι οποίες αφορούν στην αποθήκευσή της με τεχνικές που επιτρέπουν τη μετάδοσή της πάνω από δίκτυα σε πραγματικό χρόνο, στη διαβαθμισμένη μεταβίβασή της σύμφωνα με τα χαρακτηριστικά του δικτύου και της εφαρμογής και στην αλληλεπίδραση

του χρήστη με το περιεχόμενο για τον έλεγχο της παρουσίας του. Το σύστημα ΚΥΔΩΝΙΑ που παρουσιάστηκε στο κεφάλαιο 2, παρέχει το σύνολο αυτών των υπηρεσιών σύμφωνα με το πρότυπο DAVIC. Οι απαιτήσεις του προτύπου αυτού για τη διαχείριση του περιεχομένου των εφαρμογών, δηλαδή της οπτικοακουστικής πληροφορίας, περιλαμβάνουν την αποθήκευση και διαχείριση ροών οπτικοακουστικών δεδομένων που είναι κωδικοποιημένες με το πρότυπο MPEG-2, την αποθήκευση του περιεχομένου σύμφωνα με τις ανάγκες του ιδιοκτήτη του, τη μετάδοσή του μέσω διαβαθμισμένων υπηρεσιών και την παροχή υπηρεσιών περιήγησης στο οπτικοακουστικό υλικό. Στα πλαίσια της διατριβής και για τις ανάγκες του ερευνητικού προγράμματος SICMA, υποστηρίχθηκαν οι υπηρεσίες διαχείρισης MPEG-2 περιεχομένου και η υπηρεσία εγκατάστασης περιεχομένου στον εξυπηρετητή. Οι υπηρεσίες διαχείρισης περιεχομένου αποτελούνται από την μοντελοποίηση των δεδομένων του MPEG-2, την αποθήκευση μεταπληροφορίας γι' αυτά στην αντικειμενοστρεφή βάση δεδομένων του συστήματος και τη δεικτοδότηση των ροών του MPEG-2 για την υποστήριξη των υπηρεσιών αλληλεπιδραστικής παρουσίασης. Η υπηρεσία εγκατάστασης περιεχομένου σύμφωνα με το DAVIC πρότυπο αποτελείται από την ταυτοποίηση και κατηγοριοποίηση του εισερχόμενου περιεχομένου και των μεταδεδομένων του και τη διαχείρισή του από τον ιδιοκτήτη του.

Οι υπηρεσίες αυτές επεκτάθηκαν και ενοποιήθηκαν στα πλαίσια της διατριβής με στόχο να παρέχουν έναν πυρήνα λειτουργιών χρήσιμο για την υποστήριξη οποιασδήποτε εφαρμογής οπτικοακουστικού περιεχομένου, ακόμα και εφαρμογών μη συμβατών με το πρότυπο DAVIC. Το σύνολο αυτών των λειτουργιών στηρίζεται στην αναπαράσταση των δεδομένων του MPEG-2 με τη Γλώσσα Περιγραφής Οπτικοακουστικού Περιεχομένου AVDL, που χρησιμοποιήθηκε στο προηγούμενο κεφάλαιο για την περιγραφή των δεδομένων του MPEG-4. Η περιγραφή του MPEG-2 χρησιμοποιείται επίσης για την υλοποίηση των υπηρεσιών δεικτοδότησης του περιεχομένου που παρέχουν την αλληλεπιδραστική παρουσίασή του.

Στην ενότητα 4.2 περιγράφεται η διαχείριση των πολυμέσων στο ΚΥΔΩΝΙΑ και η μοντελοποίησή τους στην αντικειμενοστρεφή βάση δεδομένων του συστήματος. Στην επόμενη ενότητα παρουσιάζεται η δεικτοδότηση MPEG-2 βίντεο για την υποστήριξη υπηρεσιών περιήγησης στο περιεχόμενο. Η ενότητα 4.4 περιγράφει την υποστήριξη της Υπηρεσίας Περιεχομένου στο σύστημα ΚΥΔΩΝΙΑ σύμφωνα με το πρότυπο DAVIC.

Comment [Γ.Κ.14]: εδώ πρέπει να μπει για την επόμενη παράγραφο, η οποία θα μιλάει για την αναπαράσταση του MPEG-2

4.2 Διαχείριση πολυμέσων στο σύστημα ΚΥΔΩΝΙΑ

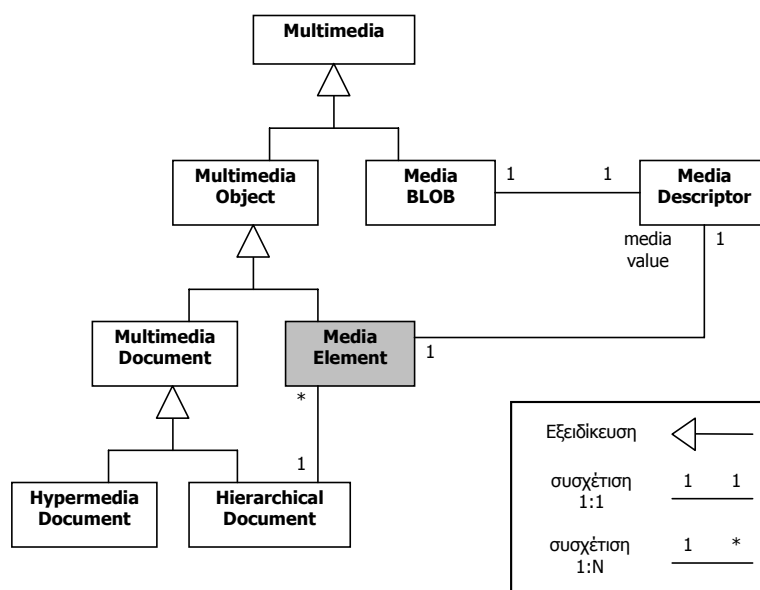
Για οπτικοακουστικές εφαρμογές όπως βίντεο κατ' απαίτηση, είναι απαραίτητο να παρέχεται λειτουργικότητα περιήγησης στο περιεχόμενο των οπτικοακουστικών μέσων. Η λειτουργικότητα αυτή επιτυγχάνεται με μεθόδους μοντελοποίησης και δεικτοδότησης του περιεχομένου, οι οποίες υποστηρίζουν τη διαχείρισή του.

Γενικά οι τιμές των πολυμέσων αφού μετατραπούν από μία αναλογική πηγή σε ψηφιακή μορφή αναπαρίστανται ως μεγάλες μη δομημένες σειρές πληροφορίας με συγκεκριμένες απαιτήσεις παρουσίασης. Όμως οι εφαρμογές που χρησιμοποιούν τα πολυμέσα έχουν ανάγκη από ένα μοντέλο διαχείρισης που να προσφέρει περισσότερες δυνατότητες και να είναι πιο ευέλικτο. Για παράδειγμα σε μια εφαρμογή ταινιών κατ' απαίτηση, η έννοια της ταινίας είναι η κυρίαρχη και η εφαρμογή πρέπει να παρέχει τη λειτουργικότητά της με βάση αυτήν και με βάση τις ιδιότητές της που αντιλαμβάνεται ο χρήστης. Το σύστημα ΚΥΔΩΝΙΑ για να εξυπηρετήσει τις ανάγκες των εφαρμογών πολυμέσων υποστηρίζει ένα αντικειμενοστρεφές μοντέλο κατάλληλο για την δομημένη αναπαράσταση των πολυμέσων, με τα εξής γενικά χαρακτηριστικά:

- Επιτρέπει τη δόμηση των δεδομένων πολυμέσων σε εννοιολογικές οντότητες των εφαρμογών. Για το λόγο αυτό παρέχει τον πυρήνα ή βάση για την αναπαράσταση των πολυμέσων αλλά και άλλων εννοιολογικών οντοτήτων.
- Υποστηρίζει μεθόδους προσπέλασης και πλοήγησης στο περιεχόμενο των πολυμέσων.
- Είναι επεκτάσιμο ώστε να μπορεί να καλύψει τις ανάγκες διαφορετικών εφαρμογών.

Στο σύστημα ΚΥΔΩΝΙΑ η μοντελοποίηση των πολυμέσων έγινε σε τρία επίπεδα διαχείρισης: αποθήκευση, ανάκτηση και παρουσίαση. Αυτά τα επίπεδα διαχείρισης υποστηρίζονται από το φυσικό, το λογικό, και το εννοιολογικό μοντέλο πληροφορίας. Το φυσικό μοντέλο καθορίζει τον τρόπο αποθήκευσης, ενώ το λογικό μοντέλο εξυπηρετεί την ανάγκη δόμησης του περιεχομένου των πολυμέσων. Το εννοιολογικό μοντέλο επιτρέπει τη συσχέτιση του περιεχομένου με έννοιες από τον πραγματικό κόσμο, όπως καθορίζεται από τις εφαρμογές. Τα μη δομημένα οπτικοακουστικά δεδομένα δηλαδή κατά την εισαγωγή τους στο σύστημα ΚΥΔΩΝΙΑ αποθηκεύονται με τρόπο που να διευκολύνει τις απαιτήσεις ανάκτησής τους, κατόπιν αποκτούν μια εσωτερική δομή κατάλληλη για τη διαχείρισή τους και στο τελευταίο στάδιο συσχετίζονται με άλλα αντικείμενα της βάσης που μοντελοποιούν έννοιες από τον πραγματικό κόσμο.

Στο παρακάτω σχήμα, που είναι ένα UML διάγραμμα κλάσεων, φαίνεται το βασικό λογικό μοντέλο αναπαράστασης των αντικειμένων πολυμέσων όπως υλοποιήθηκε στο σύστημα ΚΥΔΩΝΙΑ. Με διαφορετικό χρώμα διακρίνονται οι κλάσεις που υλοποιήθηκαν στα πλαίσια της διατριβής. Το μοντέλο που απεικονίζεται αποτελείται από κλάσεις αντικειμένων που οργανώνονται σε μια ιεραρχία εξειδίκευσης. Τα μεγάλα τμήματα του μη δομημένου περιεχομένου μοντελοποιούνται από την κλάση MediaBLOB και αποθηκεύονται στο σύστημα ως αντικείμενα που ανήκουν σε κάποιον από τους βασικούς τύπους πολυμέσων, δηλαδή κείμενο, εικόνες βίντεο, ή ήχο. Η μορφή και η κωδικοποίηση του περιεχομένου τους εκφράζονται από εξειδικευμένες κλάσεις του συστήματος. Η αποθήκευση και διαχείριση του περιεχομένου των μεγάλων αντικειμένων γίνεται από το Διαχειριστή Αποθήκευσης με μεθόδους όπως παράλληλη αποθήκευση.



Σχήμα 4-1. Μοντελοποίηση αντικειμένων πολυμέσων στο σύστημα ΚΥΔΩΝΙΑ

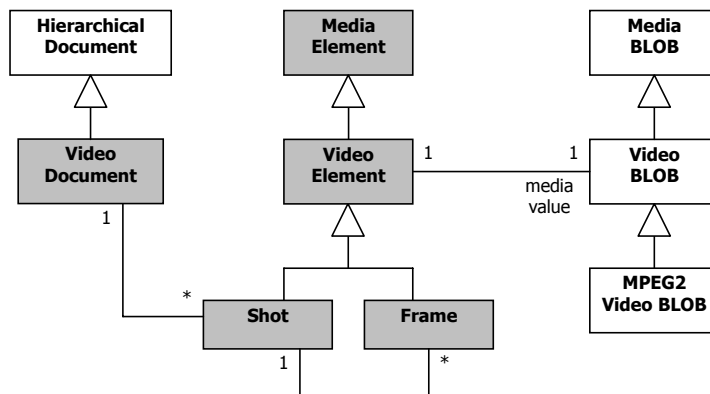
Τα αντικείμενα πολυμέσων ανήκουν στη γενικευμένη κλάση Multimedia και διακρίνονται σε μεγάλα μη δομημένα αντικείμενα της κλάσης MediaBLOB και δομημένα αντικείμενα της κλάσης MultimediaObject. Η κλάση MediaBLOB συσχετίζεται με την κλάση MediaDescriptor που επιτρέπει τη διαχείριση της μη δομημένης πληροφορίας. Η κλάση MultimediaObject εξειδικεύεται περαιτέρω σε δύο υποκλάσεις. Η κλάση MultimediaDocument μοντελοποιεί σύνθετα έγγραφα πολυμέσων, που εξειδικεύονται σε HypermediaDocuments και HierarchicalDocuments. Η κλάση MediaElement μοντελοποιεί τα στοιχεία της δομής των πολυμέσων. Ένα ιεραρχικό έγγραφο αποτελείται από ένα σύνολο αντικειμένων της κλάσης MediaElement.

Για κάθε ένα από αυτά τα μεγάλα αντικείμενα, δημιουργείται και αποθηκεύεται ένα μικρό αντικείμενο στη βάση, το οποίο αποτελεί μια μικρή περιγραφή κάποιων χαρακτηριστικών

του, που είναι απαραίτητα για την εσωτερική διαχείριση και ανάκτησή τους. Αυτά τα μικρά αντικείμενα που ονομάζονται *περιγραφείς αντικειμένων* και ανήκουν στην κλάση MediaDescriptor, αποτελούν για το σύστημα το βασικότερο επίπεδο διαχείρισης των μεγάλων μη δομημένων αντικειμένων και χρησιμοποιούνται ως υποκατάστατά τους επειδή έχουν πολύ μικρό μέγεθος. Τα μεγάλα αντικείμενα μέσων, πλέον ανακτώνται μόνο όταν γίνει αίτηση αποστολής του περιεχομένου τους για αναπαραγωγή και δεν ανακτώνται εσωτερικά στον εξυπηρετητή. Η δόμηση των μεγάλων αντικειμένων πληροφορίας πολυμέσων σε μικρότερα αντικείμενα, εξασφαλίζεται από την κλάση MediaElement και τη συσχέτιση ανάμεσα σε MultimediaBLOB's και MediaElement's. Τα αντικείμενα της κλάσης MultimediaObject εξειδικεύονται σε έγγραφα πολυμέσων και μπορεί να έχουν ιεραρχική δομή ή δομή υπερκειμένου.

Comment [Γ.Κ.15]: ίσως η συσχέτιση είναι μέσω του descriptor και όχι απευθείας με το BLOB.

Για κάθε τύπο μέσων, το παραπάνω μοντέλο επεκτείνεται με πρόσθετους τύπους. Για παράδειγμα τα στοιχειώδη αντικείμενα βίντεο διακρίνονται σε καρτέ (frames) και πλάνα (shots). Τα πλάνα ορίζονται ως σειρές εικόνων που είναι αποτέλεσμα μιας συνεχόμενης λήψης και είναι δυνατόν να εξαχθούν από ολόκληρη τη σειρά του βίντεο με ειδικές μεθόδους που ανιχνεύουν αλλαγές στα χαρακτηριστικά της οπτικής πληροφορίας. Τα αντικείμενα βίντεο μπορούν να είναι απλά ή σύνθετα έγγραφα βίντεο που αποτελούνται από στοιχειώδη αντικείμενα βίντεο. Η ιεραρχία αυτή φαίνεται στο παρακάτω σχήμα.



Σχήμα 4-2. Μοντελοποίηση αντικειμένων βίντεο στο σύστημα ΚΥΔΩΝΙΑ

Τα μεγάλα μη δομημένα αντικείμενα βίντεο μοντελοποιούνται από την κλάση VideoBLOB και την εξειδικευμένη κλάση MPEG2VideoBLOB. Τα στοιχειώδη αντικείμενα βίντεο (κλάση VideoElement) εξειδικεύονται σε πλάνα (Shot) και καρτέ (Frame), τα οποία αποκτούν περιεχόμενο μέσω της συσχέτισής τους με αντικείμενα της κλάσης VideoBLOB. Τα ιεραρχικά έγγραφα βίντεο της κλάσης VideoDocument, αποτελούνται από ένα σύνολο πλάνων λήψης.

Το παραπάνω μοντέλο πολυμέσων καλύπτει τις βασικές ανάγκες αναπαράστασης του περιεχομένου των οπτικοακουστικών μέσων στην αντικειμενοστρεφή βάση δεδομένων του συστήματος ΚΥΔΩΝΙΑ. Μοντελοποιεί τις βασικές κλάσεις αντικειμένων, δομημένων και μη, επιτρέπει τη συσχέτιση των δομημένων αντικειμένων με περιεχόμενο από τη μη δομημένη πληροφορία και τέλος παρέχει ένα εσωτερικό επίπεδο διαχείρισης και συσχέτισης μέσω των περιγραφών αντικειμένων.

4.3 Δεικτοδότηση MPEG περιεχομένου στο σύστημα ΚΥΔΩΝΙΑ

Στα πλαίσια της διατριβής υποστηρίχθηκαν στο σύστημα ΚΥΔΩΝΙΑ υπηρεσίες περιήγησης στο περιεχόμενο αντικειμένων βίντεο με κωδικοποίηση MPEG. Η περιήγηση και ο έλεγχος της παρουσίασης του περιεχομένου ήταν αναγκαίο να υποστηρίζει τις βασικές λειτουργίες αναπαραγωγής του περιεχομένου από τυχαίο χρονικό σημείο, παρουσίασης του βίντεο με τυχαίο και μεταβαλλόμενο ρυθμό (γρήγορα, πολύ γρήγορα), αναπαραγωγής προς την αντίθετη χρονική κατεύθυνση και παύση της παρουσίασης και συνέχισή της από το ίδιο σημείο. Η λειτουργικότητα αυτή μπορεί να περιγραφεί ως εξής:

- Προσωρινή παύση της παρουσίασης σε οποιοδήποτε σημείο του βίντεο.
- Αναπαραγωγή από τυχαίο σημείο του βίντεο. Η λειτουργία αυτή είναι χρήσιμη όταν έχει προηγηθεί προσωρινή παύση ή όταν στην εφαρμογή υπάρχει κινούμενη μπάρα που επιτρέπει τη μετάβαση σε τυχαίο σημείο του βίντεο.
- Γρήγορη αναπαραγωγή προς τα μπροστά ή προς τα πίσω, από τυχαίο σημείο μέσα στη ροή.
- Εναλλαγή λειτουργιών κανονικής, γρήγορης και γρήγορης προς τα πίσω αναπαραγωγής από τυχαία σημεία της παρουσίασης.

Για να υποστηριχθεί μεγάλο εύρος δυνατοτήτων των τερματικών, η υλοποίηση των παραπάνω λειτουργιών ήταν αναγκαίο να υποστηριχθεί στον πυρήνα του εξυπηρετητή με ενιαίο τρόπο για όλα τα προφίλ δυνατοτήτων. Επίσης οι λειτουργίες αυτές ήταν αναγκαίο να υποστηριχθούν για μεγάλο αριθμό πελατών και αιτήσεων εξυπηρέτησης λαμβάνοντας υπόψη τους περιορισμούς του συγκεκριμένου εξυπηρετητή. Οι περιορισμοί αυτοί αφορούσαν κυρίως στην επιβάρυνση του συστήματος ανάκτησης των ροών πραγματικού χρόνου, στη χρήση κύριας μνήμης στον εξυπηρετητή και σε επιπλέον αποθηκευτικό χώρο. Επιπλέον η υποστήριξη των παραπάνω λειτουργιών περιήγησης ήταν αναγκαίο να καλύπτει

διαφορετικές μορφές κωδικοποιημένης πληροφορίας MPEG όπως βίντεο αλλά και συγχρονισμένο βίντεο και ήχο.

Για να υποστηριχθεί η παρουσίαση τυχαίων τμημάτων μιας σειράς βίντεο είναι απαραίτητη η δεικτοδότηση των δεδομένων της σειράς, δηλαδή η κατασκευή μιας δομής η οποία θα παρέχει πληροφορία για τυχαία τμήματα μέσα στην ψηφιοσειρά του βίντεο. Επίσης είναι αναγκαίο να κατασκευαστεί μέθοδος προσπέλασης της δομής δεικτοδότησης, ώστε να είναι δυνατόν να ταυτοποιηθεί μέσα στη σειρά το τυχαίο τμήμα που πρέπει να αναπαραχθεί. Το πλεονέκτημα της κωδικοποίησης του MPEG είναι ότι παρέχει ανεξάρτητα σημεία προσπέλασης μέσα στην ψηφιοσειρά των κωδικοποιημένων δεδομένων επιτρέποντας έτσι την υποστήριξη των λειτουργιών περιήγησης απευθείας στα κωδικοποιημένα δεδομένα, δηλαδή στον πυρήνα του εξυπηρετητή πολυμέσων όπου είναι αποθηκευμένα.

Οι λειτουργίες αναπαραγωγής του βίντεο σε ρυθμούς διαφορετικούς από το ρυθμό εγγραφής και κωδικοποίησής του, δηλαδή γρήγορα, πολύ γρήγορα και προς τα πίσω, υποστηρίζονται με παρουσίαση ενός υποσυνόλου των καρτέ της σειράς. Οι λειτουργίες αυτές επίσης απαιτούν την κατασκευή ειδικών τεχνικών οι οποίες να μπορούν να παρέχουν πρόσβαση σε καρτέ που αντιστοιχούν σε ένα κλάσμα του συνολικού αριθμού των καρτέ. Ο ρυθμός αναπαραγωγής δηλαδή καθορίζει τον αριθμό των συνεχόμενων καρτέ που θα παραλείπονται. Το τυπικό μέγεθος μιας ταινίας 2 ωρών η οποία αποτελείται από 30 καρτέ για κάθε δευτερόλεπτο, είναι 216000 καρτέ καθένα από τα οποία αντιστοιχεί σε διακριτά χρονικά σημεία μέσα στην ταινία. Η γρήγορη αναπαραγωγή της ταινίας με ρυθμό 10 φορές του κανονικού σημαίνει ότι για κάθε 10 συνεχόμενα καρτέ μόνο το ένα θα πρέπει να αναπαράγεται, οπότε η σειρά σε δεκαπλάσιο ρυθμό θα αποτελείται από το ένα δέκατο των καρτέ του αρχικού βίντεο. Στα πλαίσια της διατριβής, αναπτύχθηκε στο σύστημα ΚΥΔΩΝΙΑ ένα σύνολο τεχνικών με τρόπο που να καλύπτουν τις ανάγκες των εφαρμογών του προγράμματος SICMA. Αρχικά αναπτύχθηκε μια μέθοδος δεικτοδότησης βίντεο, η οποία σε δεύτερη φάση χρησιμοποιήθηκε για την δημιουργία νέων ροών πληροφορίας με μεγάλες ταχύτητες αναπαραγωγής. Αργότερα η δομή δεικτοδότησης επεκτάθηκε για την υποστήριξη περιήγησης σε συγχρονισμένο βίντεο και ήχο και τέλος για την υποστήριξη εναλλαγής των λειτουργιών περιήγησης.

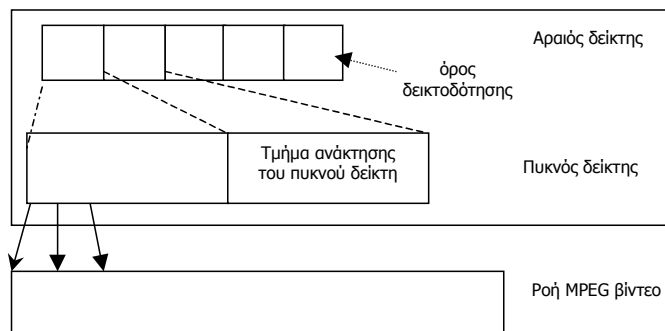
Comment [Γ.Κ.16]: Λίγο καλύτερα;

4.3.1 Δομή δεικτοδότησης MPEG βίντεο

Αρχικά αναπτύχθηκε μια μέθοδος δεικτοδότησης του βίντεο με τρόπο έτσι ώστε να ελαχιστοποιείται ο επιπλέον χώρος που απαιτείται για τα δεδομένα του δείκτη. Η μέθοδος αυτή στηρίζεται σε συντακτική ανάλυση της σειράς του MPEG βίντεο έτσι ώστε να εξαχθεί

πληροφορία για τη θέση των καρτέ μέσα στη σειρά. Η θέση των καρτέ καθορίζεται από την απόστασή τους από την αρχή της σειράς. Οι όροι δεικτοδότησης (indexing terms) αποτελούνται εκτός από τη θέση των καρτέ μέσα στην ψηφιοσειρά, από τη σειρά παρουσίασής τους, το είδος τους και το μέγεθός τους. Η σειρά παρουσίασης είναι σε σχέση με το συνολικό αριθμό των εικόνων του βίντεο και αποτελεί το κύριο κλειδί του δείκτη. Το είδος του καρτέ αφορά στην κωδικοποίηση MPEG όπου υπάρχουν τρία διαφορετικά είδη εικόνων: I (intracoded) ή ενδοκωδικοποιημένα, P (predictive) ή προφητικά και B (bidirectionally predictive) ή προφητικά προς τις δύο κατευθύνσεις. Το είδος των εικόνων αφορά στον τρόπο αποκωδικοποίησής τους. Δηλαδή τα καρτέ τύπου I μπορούν να αποκωδικοποιηθούν αυτόνομα, ενώ αυτά του τύπου P χρησιμοποιούν πρόβλεψη από προηγούμενα καρτέ I ή P. Τα καρτέ τύπου B τέλος χρησιμοποιούν προηγούμενα και επόμενα καρτέ τύπου I ή P.

Οι όροι του δείκτη που είναι της μορφής (*σειρά παρουσίασης καρτέ, είδος καρτέ, απόσταση καρτέ από την αρχή του βίντεο, μέγεθος καρτέ*) είναι ταξινομημένοι με αύξουσα σειρά ως προς το κύριο κλειδί. Για να γίνεται γρήγορη προσπέλαση του δείκτη δημιουργείται ένα ακόμα επίπεδο δεικτοδότησης με λιγότερη πληροφορία, το οποίο αποτελεί έναν αραιό δείκτη που περιέχει μόνο μερικούς από τους όρους του αρχικού πυκνού δείκτη. Η επιλογή των όρων για τον αραιό δείκτη έχει να κάνει με το μέγεθος του τμήματος ανάκτησης από τη δευτερεύουσα μνήμη. Δηλαδή ο πυκνός δείκτης διαιρείται σε τμήματα ίσα με το μέγεθος του τμήματος ανάκτησης και για κάθε ένα από αυτά επιλέγεται η πρώτη τριάδα (αριθμός καρτέ, απόσταση καρτέ, μέγεθος καρτέ) ως αντιπροσωπευτικό του αραιού δείκτη. Η δομή του δείκτη φαίνεται στο παρακάτω σχήμα.



Σχήμα 4-3. Δεικτοδότηση ροών MPEG βίντεο

Η δομή δεικτοδότησης του MPEG βίντεο αποτελείται από 2 επίπεδα, ένα αραιό και ένα πυκνό. Το πυκνό επίπεδο αποτελείται από όρους δεικτοδότησης για όλα τα καρτέ του κωδικοποιημένου βίντεο. Το αραιό επίπεδο παρέχει πρόσβαση σε ομάδες (blocks) ανάκτησης του πυκνού δείκτη.

Για ένα βίντεο 2 ωρών όπως ήδη αναφέρθηκε, το βίντεο αποτελείται από 216000 καρτέ. Ο δείκτης περιλαμβάνει 21600 όρους για κάθε ένα από τους οποίους χρειάζονται 12 bytes πληροφορίας. Το μέγεθος του πυκνού δείκτη και επομένως το επιπλέον κόστος δευτερεύουσας αποθήκευσης υπολογίζεται ότι είναι $12 \times 216000 \approx 2,5$ MBytes. Το αντίστοιχο μέγεθος της ταινίας είναι για κωδικοποίηση με ρυθμό 1,5Mbps (σχετικά χαμηλό ρυθμό) είναι περίπου 2,6 GBytes, δηλαδή μία τάξη μεγέθους μεγαλύτερο. Για το περιεχόμενο των εφαρμογών του προγράμματος SICMA όπου το μέγεθος κάθε βίντεο ήταν της τάξης των 5 λεπτών το μέγεθος του δείκτη άρα και το επιπλέον αποθηκευτικό κόστος ήταν πολύ μικρότερο (περίπου το 1/20 των 2,5 Mbytes).

Η μέθοδος προσπέλασης της δομής δεικτοδότησης στηρίζεται στην αναζήτηση των κλειδιών, δηλαδή γίνεται με βάση τη σειρά παρουσίασης των καρτέ. Η αναζήτηση αυτή επιστρέφει τη θέση του καρτέ μέσα στην αποθηκευμένη σειρά του βίντεο καθώς και το μέγεθός του. Ο Διαχειριστής Αποθήκευσης έχοντας την πληροφορία για τη θέση και το μέγεθος κάποιου καρτέ, μπορεί να το ανακτήσει χωρίς να προσπελάσει ή να αποκωδικοποιήσει το ίδιο το βίντεο. Με κατάλληλη διαχείριση της κύριας μνήμης, η προσπέλαση του δείκτη περιλαμβάνει το πολύ ένα διάβασμα από τη δευτερεύουσα αποθήκευση. Ο αραιός δείκτης που χρησιμοποιείται για να δώσει τη θέση του τμήματος του πυκνού δείκτη που μας ενδιαφέρει, έχει μικρό μέγεθος και βρίσκεται συνεχώς στην κύρια μνήμη του εξυπηρετητή. Ο αραιός δείκτης προσπελαύνεται για να βρεθεί το τμήμα ανάκτησης του πυκνού δείκτη που περιέχει το κλειδί που μας ενδιαφέρει. Το τμήμα αυτό καθορίζεται από την απόστασή του από την αρχή του πυκνού δείκτη. Όταν το τμήμα αυτό ανακτηθεί στην κύρια μνήμη, γίνεται αναζήτηση με βάση το κλειδί για να βρεθεί η εγγραφή του δείκτη που το περιέχει και στη συνέχεια να είναι δυνατόν να ανακτηθεί το ίδιο το καρτέ από τη σειρά του βίντεο. Η λειτουργικότητα αυτή είναι ιδιαίτερα χρήσιμη στην αναπαραγωγή του βίντεο από ένα τυχαίο σημείο μέχρι κάποιο άλλο. Αν το δεύτερο σημείο είναι το τέλος του βίντεο χρειάζεται το πολύ μία προσπέλαση του δείκτη, ενώ αν το δεύτερο σημείο δεν είναι το τέλος του βίντεο χρειάζονται το πολύ δύο προσπελάσεις του δείκτη για ν' αρχίσει η αναπαραγωγή του κομματιού του βίντεο. Ο δείκτης επομένως παρέχει γρήγορη προσπέλαση σε «τυχαία» καρτέ μέσα στη σειρά των εικόνων.

4.3.2 Γρήγορη επισκόπηση βίντεο

Για την υλοποίηση της αναπαραγωγής του βίντεο σε γρήγορη κίνηση, είναι αναγκαίο να καθοριστεί ο ρυθμός αναπαραγωγής ή ισοδύναμα το κλάσμα παράλειψης καρτέ (skipping factor) από το αρχικό βίντεο. Από το κλάσμα παράλειψης καρτέ υπολογίζεται ο συνολικός αριθμός των καρτέ που πρέπει να παρουσιαστούν. Για παράδειγμα αν ο ζητούμενος ρυθμός γρήγορης επισκόπησης είναι δεκαπλάσιος από τον κανονικό, το κλάσμα παράλειψης καρτέ

είναι $1/10$. Δηλαδή την γρήγορη παρουσίαση του βίντεο πρέπει να μεταδοθεί το $1/10$ των συνολικών καρτέ του βίντεο και έτσι ώστε να μεταδίδεται ένα καρτέ, να παραλείπονται 10, να μεταδίδεται ένα, κλπ.

Η δομή δεικτοδότησης που περιγράφεται στην προηγούμενη ενότητα χρησιμοποιείται για να υπολογιστεί η θέση των καρτέ που θα χρησιμοποιηθούν για την γρήγορη επισκόπηση μέσα στην κωδικοποιημένη σειρά του βίντεο. Με βάση τη θέση αυτών των καρτέ είναι δυνατόν να γίνει η ανάκτησή τους από τη δευτερεύουσα αποθήκευση. Ανάλογα με το ρυθμό αναπαραγωγής που ζητείται από την εφαρμογή ακολουθούνται διαφορετικές πολιτικές για τον καθορισμό των καρτέ που θα ανακτηθούν τελικά και θα αποτελέσουν τη γρήγορη παρουσίαση του περιεχομένου. Γενικά υποθέτουμε ότι ολόκληρο το βίντεο αποτελείται από N καρτέ και S είναι ταχύτητα γρήγορης αναπαραγωγής, όπου $S > 1$. Η σειρά των καρτέ που αποτελούν τη γρήγορη επισκόπηση με αυτόν το ρυθμό είναι $S, 2*S, 3*S$, κλπ. Αν όμως κάποιο από αυτά τα καρτέ είναι τύπου P ή B , η παρουσίασή του από μόνο του δεν είναι δυνατή, γιατί εξαρτάται από άλλα καρτέ και θα παρουσιαστούν σφάλματα στον αποκωδικοποιητή. Επομένως η ανάκτηση της σειράς εικόνων $S, 2S, 3S$, κλπ δεν έχει πάντα σωστό αποτέλεσμα.

Έτσι εφαρμόζεται μέθοδος που επιλέγει κατά προσέγγιση τα καρτέ που θα αποτελέσουν τη γρήγορη επισκόπηση του βίντεο με βήμα S . Αρχικά γίνεται με ομαδοποίηση των καρτέ του βίντεο σε N/S ομάδες των S καρτέ η κάθε μία. Για ένα καρτέ που θα περιέχεται στη σειρά γρήγορης επισκόπησης, μας ενδιαφέρει να υπολογίζουμε σε ποια ομάδα ανήκει. Στο σημείο αυτό ακολουθήθηκαν δύο διαφορετικές προσεγγίσεις με διαφορετικό αποτέλεσμα στην απόκριση του συστήματος. Η πρώτη προσέγγιση εξασφαλίζει μεγαλύτερη ακρίβεια στην παρουσίαση της γρήγορης επισκόπησης, ενώ η δεύτερη διατηρεί χαμηλό ρυθμό μετάδοσης για το σύστημα του εξυπηρετητή.

Στην πρώτη προσέγγιση, αν κάποιο από τα καρτέ που ανήκουν στη σειρά γρήγορης αναπαραγωγής είναι τύπου I , τότε το καρτέ αυτό θα μεταδοθεί μαζί με τα υπόλοιπα της σειράς. Αν δεν είναι τύπου I και δεν είναι δυνατόν να αποκωδικοποιηθεί αυτόνομα, τότε διαλέγουμε από την ομάδα που ανήκει το I πλησιέστερο καρτέ της ομάδας για να το αντικαταστήσουμε. Στη μέθοδο αυτή είναι δυνατόν να εφαρμοστεί και όριο στην επιτρεπτή ακρίβεια, δηλαδή στην απόσταση του καρτέ από το I καρτέ που θα το αντικαταστήσει.

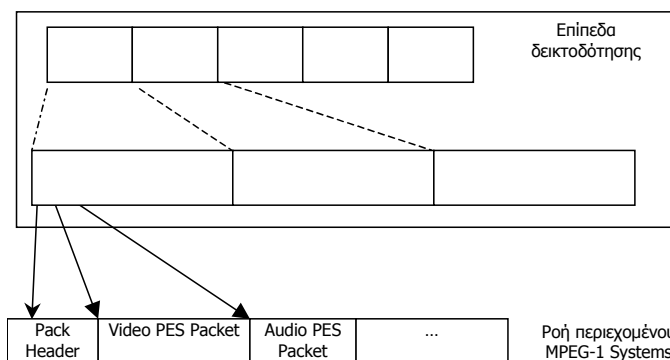
Στη δεύτερη μέθοδο, που εξασφαλίζει μικρότερο ρυθμό ανάκτησης και μετάδοσης, υπολογίζουμε το ποσοστό των καρτέ που πρέπει να ανακτηθούν από το σύνολο της σειράς και επιλέγουμε περισσότερα από ένα καρτέ από κάθε ομάδα. Ο στόχος είναι να μειώσουμε το μέγεθος της σειράς διατηρώντας έναν επιτρεπτό ρυθμό μετάδοσης. Έτσι

περιλαμβάνουμε σε αυτήν και καρτέ που περιέχουν πρόβλεψη κίνησης (P ή B), τα οποία έχουν πολύ μικρότερο μέγεθος. Η μέθοδος αυτή παρέχει μικρότερη ακρίβεια, αλλά δε θυσιάζει την απόδοση του εξυπηρετητή.

Και στις δύο παραπάνω περιπτώσεις, το μειονέκτημα είναι ότι ο ρυθμός ανάκτησης των καρτέ για την επιλογή αυτών που θα μεταδοθούν, παραμένει πολλαπλάσιος του ρυθμού ανάκτησης του βίντεο, γεγονός που επηρεάζει αρνητικά την απόδοση του εξυπηρετητή. Εξαιτίας του προβλήματος αυτού, ο δείκτης χρησιμοποιήθηκε για την κατασκευή και αποθήκευση των καρτέ γρήγορης αναπαραγωγής ως ξεχωριστής σειράς βίντεο. Οι νέες σειρές κατασκευάζονται σε μη πραγματικό χρόνο και αποθηκεύονται στο σύστημα ως μεγάλα αντικείμενα βίντεο. Έτσι όταν η λειτουργία είναι γρήγορη αναπαραγωγή του βίντεο, το υποσύστημα διαχείρισης των αντικειμένων του βίντεο στον εξυπηρετητή, αποφασίζει ποια ροή πρέπει να ανακτηθεί.

4.3.3 Δεικτοδότηση συγχρονισμένου βίντεο και ήχου

Ένας από τους στόχους της δεύτερης φάσης του προγράμματος SICMA ήταν να παρέχει πλήρη λειτουργικότητα περιήγησης σε συγχρονισμένες ροές βίντεο και ήχου, κωδικοποιημένες σύμφωνα με το πρότυπο MPEG-1. Η προσέγγιση που ακολουθήθηκε ήταν η δημιουργία δύο νέων ροών για κάθε βίντεο, μία για γρήγορη αναπαραγωγή προς τα μπροστά και μία για γρήγορη αναπαραγωγή προς τα πίσω. Οι δύο σειρές ψηφιοποιήθηκαν από την αρχή για να επιτευχθεί ο μέγιστος ρυθμός μετάδοσης των 1.5 Mbps που ήταν πραγματικός περιορισμός στο περιβάλλον της εφαρμογής απομακρυσμένης διδασκαλίας. Η υλοποίηση των λειτουργιών γρήγορης αναπαραγωγής έγινε επομένως με στατική ταχύτητα.



Σχήμα 4-4. Δεικτοδότηση συγχρονισμένων ροών MPEG-1 Systems

Η δομή δεικτοδότησης MPEG περιεχομένου χρησιμοποιείται για συγχρονισμένες ροές βίντεο και ήχου. Οι όροι δεικτοδότησης αποτελούνται από πληροφορία για τα πακέτα

βίντεο και ήχου που σχηματίζουν τη συγχρονισμένη ροή. Σε λειτουργίες τυχαίας προσπέλασης ανακτάται ο συγχρονισμένος συνδυασμός πακέτων βίντεο και ήχου. Σε λειτουργίες γρήγορης αναπαραγωγής χρησιμοποιείται μόνο το βίντεο της σειράς.

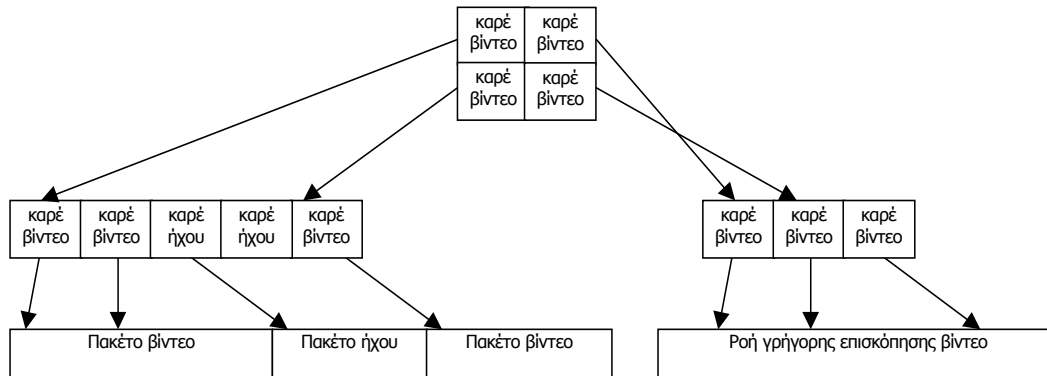
Για την υποστήριξη τυχαίας προσπέλασης στο κωδικοποιημένο περιεχόμενο, κατασκευάζονται δομές δεικτοδότησης παρόμοιες με αυτές για το βίντεο. Στο παραπάνω σχήμα απεικονίζεται η δομή δεικτοδότησης μιας συγχρονισμένης ροής βίντεο και ήχου σύμφωνα με το τμήμα Systems του προτύπου MPEG-1. Η διαφορά της δομής αυτής είναι ότι πρέπει σε λειτουργίες τυχαίας προσπέλασης να διατηρεί το συγχρονισμό βίντεο και ήχου. Το βίντεο και ο ήχος πολυπλέκονται μέσα στη ροή δεδομένων και πρέπει να ταυτοποιείται η θέση τους μέσα στη ροή για να μπορεί να γίνει συσχετισμός των σύγχρονων τμημάτων. Επιπλέον στη γρήγορη αναπαραγωγή ο ήχος δεν μεταδίδεται στο τερματικό, επομένως είναι αναγκαίο να διαχωρίζεται μέσω της δομής δεικτοδότησης από το περιεχόμενο βίντεο.

4.3.4 Εναλλαγή λειτουργιών περιήγησης

Οι τεχνικές που παρουσιάστηκαν στις δύο προηγούμενες υποενότητες δεν αρκούν για την υποστήριξη ελεύθερης περιήγησης στο περιεχόμενο του βίντεο. Εξασφαλίζουν μόνο την αποδοτική υποστήριξη τυχαίας προσπέλασης ή την γρήγορη επισκόπηση του περιεχομένου του βίντεο. Για πιο ευέλικτη λειτουργικότητα περιήγησης είναι επιθυμητό να χρησιμοποιούνται αυτές οι δύο λειτουργίες μαζί και με την κανονική παρουσίαση του βίντεο, με τυχαία σειρά και από τυχαία σημεία στο περιεχόμενο. Για παράδειγμα, ένας φοιτητής που παρακολουθεί το βίντεο ενός μαθήματος από απόσταση, μπορεί να κάνει γρήγορη επισκόπηση μέχρι το σημείο που τον ενδιαφέρει και κατόπιν να βλέπει το βίντεο κανονικά. Η αποκωδικοποίηση MPEG βίντεο (MPEG-1 ή MPEG-2, MPEG-1 Systems ή MPEG-2 Transport Stream) δίνουν τη δυνατότητα να γνωρίζουμε τη χρονική εξέλιξη της αποκωδικοποίησης, δηλαδή της παρουσίασης. Η πληροφορία αυτή αλλά και το είδος της λειτουργίας περιήγησης που επιλέγεται από το χρήστη, χρησιμοποιείται από την Υπηρεσία Ροής του συστήματος ΚΥΔΩΝΙΑ-D και μεταφέρεται στο υποσύστημα υποστήριξης των υπηρεσιών περιήγησης. Το υποσύστημα αυτό αναλαμβάνει να καθοδηγήσει την ανάκτηση της κατάλληλης πληροφορίας κάνοντας χρήση των δομών δεικτοδότησης.

Στην εφαρμογή εκπαίδευσης από απόσταση του προγράμματος SICMA, η συνολική αντιμετώπιση για την υποστήριξη περιήγησης, οδήγησε στην κατασκευή μιας μεθόδου αντιστοίχισης ανάμεσα στις κανονικές ροές βίντεο και στις ροές περιήγησης (γρήγορης αναπαραγωγής). Οι τελευταίες είχαν ψηφιοποιηθεί σε μη πραγματικό χρόνο από τα αναλογικά δεδομένα. Για να υποστηριχθεί εναλλαγή των λειτουργιών περιήγησης καταρχήν κατασκευάζονται και αποθηκεύονται στο σύστημα οι δείκτες που παρέχουν

τυχαία προσπέλασης. Οι δείκτες αυτοί κατασκευάζονται για την κανονική σειρά βίντεο αλλά και για τις γρήγορες σειρές. Κατόπιν δημιουργείται ένας επιπλέον δείκτης ο οποίος διατηρεί την αντιστοιχία ανάμεσα στο κανονικό βίντεο και τις δύο εκδόσεις γρήγορης αναπαραγωγής.



Σχήμα 4-5. Δομή των δεικτών ροών MPEG-1 και ροών γρήγορης επισκόπησης

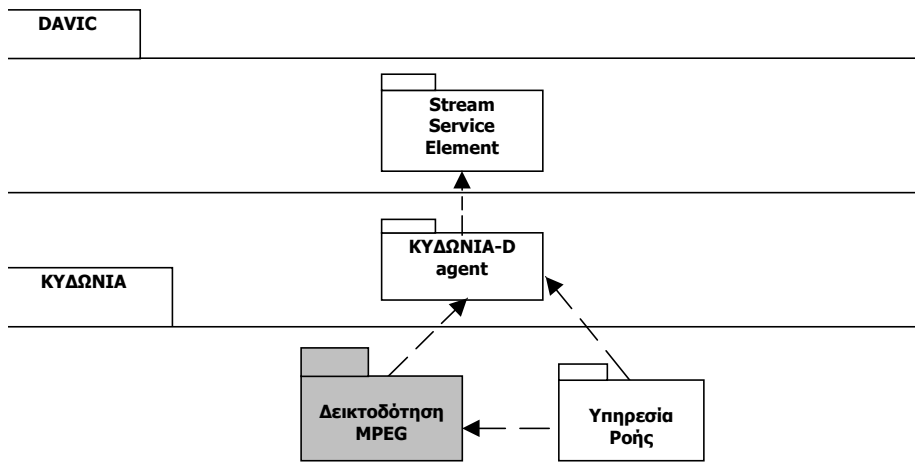
Για τις ροές συγχρονισμένου βίντεο και ήχου και για τις ροές γρήγορης επισκόπησης, κατασκευάζονται δύο ξεχωριστές δομές δεικτοδότησης. Οι δομές αυτές επιτρέπουν την τυχαία προσπέλαση μέσα στις δύο ροές πληροφορίας και επομένως την υποστήριξη υπηρεσιών περιήγησης. Σε συνδυασμό με μία τρίτη δομή που παρέχει αντιστοίχιση ανάμεσα στα καρέ βίντεο που παρουσιάζονται και στις δύο ροές, είναι δυνατή η ευέλικτη εναλλαγή των λειτουργιών περιήγησης με αποδοτικό τρόπο.

Στο παραπάνω σχήμα απεικονίζεται η υποστήριξη της εναλλαγής των λειτουργιών περιήγησης με δύο διαφορετικές δομές δεικτοδότησης για τη ροή βίντεο και ήχου και για τη ροή γρήγορης επισκόπησης που αποτελείται μόνο από καρέ βίντεο. Οι δομές συνδέονται με μία δομή αντιστοίχισης καρέ βίντεο που εμφανίζονται στην αρχική ροή αλλά και στη ροή γρήγορης επισκόπησης.

4.3.5 Αρχιτεκτονική του συστήματος υποστήριξης λειτουργιών περιήγησης

Η υποστήριξη των λειτουργιών περιήγησης σε περιεχόμενο MPEG έγινε στο σύστημα ΚΥΔΩΝΙΑ-D σύμφωνα με το παράδειγμα του αντικειμενοστρεφούς προγραμματισμού. Στο παρακάτω διάγραμμα απεικονίζεται η αρχιτεκτονική της υποδομής υποστήριξης ως προς πακέτα ή σύνολα κλάσεων του συστήματος. Τα πακέτα κλάσεων συγκεντρώνουν διαφορετικά τμήματα της λειτουργικότητας του συστήματος σε ομάδες κλάσεων με κοινό σκοπό. Η συνεισφορά της διατριβής εστιάζεται στο πακέτο δεικτοδότησης, το οποίο

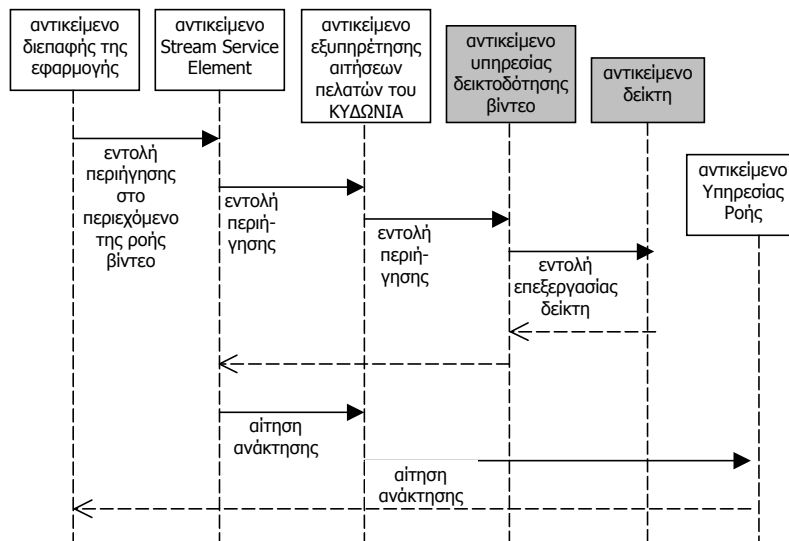
καθορίζει τη συμπεριφορά της Υπηρεσίας Ροής η οποία αναλαμβάνει την ανάκτηση και μεταβίβαση των ροών πληροφορίας προς τα τερματικά.



Σχήμα 4-6. Διάγραμμα πακέτων της αρχιτεκτονικής υποστήριξης περιήγησης

Το σύστημα KYΔΩΝΙΑ-D αποτελείται από ένα σύνολο κλάσεων που ομαδοποιούνται σε πακέτα με κοινό στόχο. Το πακέτο Δεικτοδότησης MPEG που αναπτύχθηκε στα πλαίσια της διατριβής, περιέχει τη λειτουργικότητα των κλάσεων δεικτοδότησης και προσπέλασης του περιεχομένου MPEG. Επίσης υπάρχει εξάρτηση ανάμεσα σε αυτό το πακέτο και το πακέτο που υλοποιεί την Υπηρεσία Ροής του εξυπηρετητή KYΔΩΝΙΑ. Τα δύο αυτά πακέτα αποτελούν μέρος του KYΔΩΝΙΑ. Για την επικοινωνία με τον εξωτερικό κόσμο και το φλοιό του DAVIC, χρησιμοποιείται το πακέτο KYΔΩΝΙΑ-D agent. Ο φλοιός του DAVIC είναι πακέτο κλάσεων που παρέχουν τη διεπικοινωνία που καθορίζει το DAVIC για να πετύχει διαλειτουργικότητα. Μία από αυτές τις κλάσεις αφορά στην εξυπηρέτηση ροών (Stream Service Element).

Στο παρακάτω σχήμα φαίνεται η ροή επεξεργασίας των αιτήσεων περιήγησης στον εξυπηρετητή KYΔΩΝΙΑ-D. Τα ορθογώνια παριστάνουν αντικείμενα του συστήματος, τα οποία αποτελούν στιγμιότυπα των κλάσεων του προηγούμενου διαγράμματος. Εκτός από τα αντικείμενα απεικονίζεται η διαδοχή των μηνυμάτων που ανταλλάσσουν μεταξύ τους. Υποθέτουμε ότι η πυροδότηση των ενεργειών περιήγησης έχει αφετηρία το χρήστη και μέσω της εφαρμογής (αντικείμενο διεπαφής) κατευθύνεται σε ένα αντικείμενο εξυπηρέτησης ροών DAVIC που υπάρχει στον εξυπηρετητή.



Σχήμα 4-7. Διάγραμμα σειράς για την υποστήριξη των αιτήσεων περιήγησης στο σύστημα ΚΥΔΩΝΙΑ

Μία εντολή περιήγησης στέλνεται από την εφαρμογή στο σύστημα ΚΥΔΩΝΙΑ-D μέσω των αντικειμένων Stream Service Element του DAVIC και εξυπηρέτησης αιτήσεων πελατών του ΚΥΔΩΝΙΑ. Η αίτηση αυτή μεταφέρεται στα αντικείμενα δεικτοδότησης και το αποτέλεσμα της επεξεργασίας των δεικτών είναι η ανάκτηση δεδομένων ροής από την Υπηρεσία Ροής του ΚΥΔΩΝΙΑ.

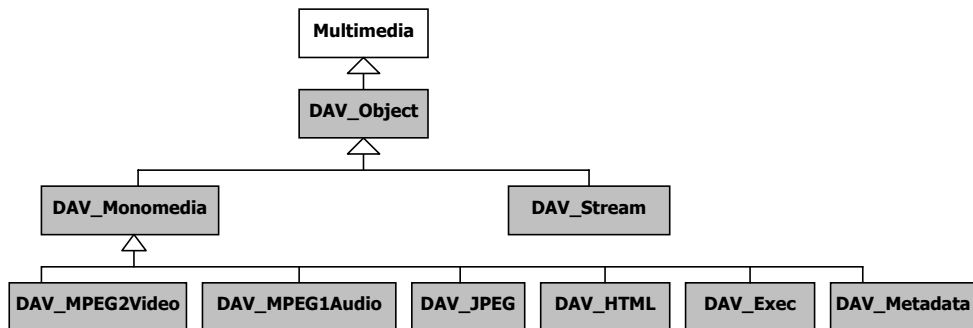
Το μήνυμα κατόπιν μεταφορά σε ένα αντικείμενο εξυπηρέτησης αιτήσεων και από εκεί κατευθύνεται στο αντικείμενο υπηρεσιών δεικτοδότησης και στο αντικείμενο του δείκτη. Η ανάκτηση των ροών προκαλείται από ένα αντικείμενο Stream Service Element και κατευθύνεται προς την Υπηρεσία Ροής, η οποία τελικά αποστέλλει τα δεδομένα στον πελάτη.

4.4 Υποστήριξη της Υπηρεσίας Περιεχομένου στο σύστημα ΚΥΔΩΝΙΑ-D

Η διαχείριση του περιεχομένου στο σύστημα ΚΥΔΩΝΙΑ-D κάλυψε τις ανάγκες των εφαρμογών του προγράμματος SICMA. Στην εφαρμογή ιδεατής πραγματικότητας του Μουσείου Φυσικής Ιστορίας του Λονδίνου ο παροχέας του περιεχομένου ήταν το ίδιο το μουσείο. Το είδος του περιεχομένου περιείχε σειρές βίντεο και ήχου συγχρονισμένες μεταξύ τους και εικόνες. Η μορφή των συγχρονισμένων σειρών βίντεο και ήχου ήταν Ροές Μετάδοσης του MPEG-2 (MPEG-2 Transport Streams). Οι απαιτήσεις από τον παροχέα του περιεχομένου ήταν η δυνατότητα εγκατάστασης και προσπέλασης του περιεχομένου στο

σύστημα εξυπηρέτησης σε μη πραγματικό χρόνο. Η εγκατάσταση, διαχείριση και προσπέλαση του περιεχομένου υποστηρίζεται από το υποσύστημα Εγκατάστασης Περιεχομένου, το τμήμα Ονομασίας Περιεχομένου και το τμήμα Διαχείρισης Αντικειμένων του ΚΥΔΩΝΙΑ-D.

Τα αντικείμενα αποθηκεύονται από το Διαχειριστή Αποθήκευσης και εξάγονται στις εφαρμογές από το Διαχειριστή Αντικειμένων και Συνόδων του ΚΥΔΩΝΙΑ και δεν αποτελούν συνεισφορά της διατριβής. Το μοντέλο αναπαράστασης της πληροφορίας του DAVIC αντιστοιχίζεται σε ένα μοντέλο αντικειμένων που υποστηρίζεται στο ΚΥΔΩΝΙΑ με βάση το βασικό μοντέλο τύπων του αντικειμενοστρεφούς πυρήνα του συστήματος. Οι τύποι μονομέσων που μοντελοποιούνται στο ΚΥΔΩΝΙΑ είναι MPEG-2 βίντεο, MPEG-1 ήχος, HTML κείμενο, εκτελέσιμα εικόνες της μορφής JPEG και RGB και ροές μετάδοσης. Οι εικόνες, το κείμενο και τα εκτελέσιμα αντικείμενα μοντελοποιούνται ως αντικείμενα που μεταδίδονται χωρίς απαιτήσεις πραγματικού χρόνου. Τα αντικείμενα βίντεο και ήχου μεταδίδονται πακεταρισμένα σε συγχρονισμένες ροές μετάδοσης MPEG-2 πάνω από ATM. Το πλεονέκτημα της προσέγγισης που ακολουθείται στο σύστημα ΚΥΔΩΝΙΑ για την εξυπηρέτηση εφαρμογών συμβατών με το πρότυπο DAVIC, είναι ότι το μοντέλο αντικειμένων παρέχει όλα τα πλεονεκτήματα των αντικειμενοστραφών βάσεων δεδομένων, όπως συνδιαλλαγές, δεικτοδότηση, ερωτήσεις, ενθυλάκωση, κληρονομικότητα και επεκτασιμότητα. Το μοντέλο αυτό φαίνεται στο ακόλουθο σχήμα. Η ιεραρχία των κλάσεων του συστήματος έχει επεκταθεί κάτω από τη βασική κλάση για τα πολυμέσα, με την κλάση DAV_Object, που μοντελοποιεί όλα τα αντικείμενα πολυμέσων του DAVIC.



Σχήμα 4-8. Μοντελοποίηση αντικειμένων του DAVIC στο σύστημα ΚΥΔΩΝΙΑ

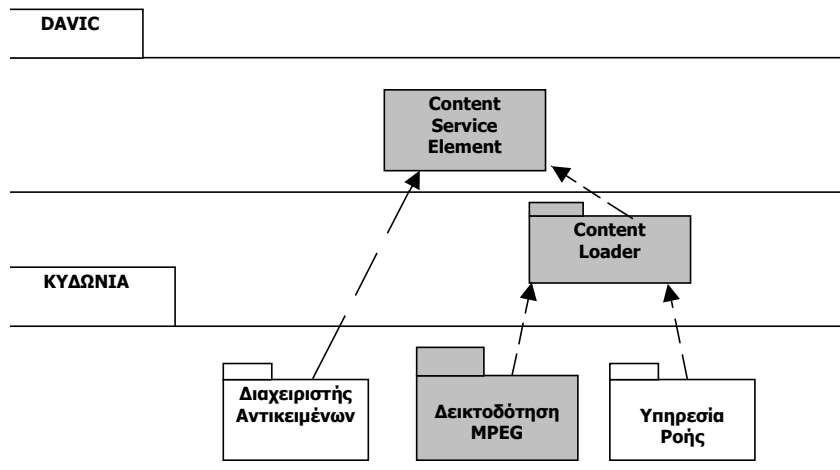
Το περιεχόμενο των εφαρμογών του DAVIC μοντελοποιείται από ένα σύνολο κλάσεων του συστήματος ΚΥΔΩΝΙΑ, οι οποίες εξειδικεύουν την κλάση Multimedia και διακρίνονται σε "μονομέσα" και σύνθετες ροές μετάδοσης. Τα μονομέσα περιλαμβάνουν MPEG βίντεο και ήχο, εικόνες, κείμενο και εκτελέσιμα.

Τα μεταδεδομένα για το περιεχόμενο μοντελοποιήθηκαν επίσης στο αντικειμενοστρεφές μοντέλο του ΚΥΔΩΝΙΑ ως αντικείμενα που σχετίζονται με τα αντικείμενα περιεχομένου. Η διαδικασία δημιουργίας των στιγμιότυπων των κλάσεων, δηλαδή των αντικειμένων μπορεί να γίνει είτε εσωτερικά στον εξυπηρετητή, είτε από απομακρυσμένες εφαρμογές. Για την υποστήριξη της Υπηρεσίας Περιεχομένου του προτύπου DAVIC, η δημιουργία των στιγμιότυπων γίνεται τοπικά στον εξυπηρετητή, εφόσον έχει προηγηθεί η κλήση κάποιας από τις λειτουργίες που παρέχει η συγκεκριμένη υπηρεσία.

Όσον αφορά στην αρχιτεκτονική της Υπηρεσίας Περιεχομένου, το σημείο εισόδου είναι η Υπηρεσία Εισόδου στο σύστημα με «κλειδί» το όνομα μιας συγκεκριμένης υπηρεσίας περιεχομένου. Οι υπηρεσίες περιεχομένου αντιστοιχούν σε διαφορετικές «περιοχές» περιεχομένου που μπορεί να ανήκουν σε διαφορετικούς παροχείς περιεχομένου και να εξυπηρετούν διαφορετικές εφαρμογές η κάθε μία. Η Υπηρεσία Εισόδου ελέγχει την ύπαρξη του συγκεκριμένου στιγμιότυπου της Υπηρεσίας Περιεχομένου και αν υπάρχει εγκατεστημένο και αποθηκευμένο στο σύστημα, δημιουργεί το στιγμιότυπό του και στην κύρια μνήμη του εξυπηρετητή, ώστε να είναι έτοιμο να δεχτεί αιτήσεις εξυπηρέτησης. Τα ονόματα των αντικειμένων και υπηρεσιών του DAVIC που αντιστοιχούν σε μόνιμα αντικείμενα στο σύστημα ΚΥΔΩΝΙΑ, αντιστοιχίζονται σε ταυτότητες αντικειμένων. Το υποσύστημα διαχείρισης αντικειμένων και συνόδων του ΚΥΔΩΝΙΑ διατηρεί έναν αποθηκευμένο χώρο ονομάτων, όπου γίνεται αντιστοίχιση από όνομα σε ταυτότητα αντικείμενου. Κάθε αντικείμενο που εγκαθίσταται στο σύστημα συσχετίζεται με ένα όνομα όπως ορίζεται από το πρότυπο CORBA και μέσω αυτού του ονόματος προσπελάζεται από τις συναρτήσεις του εξωτερικού φλοιού. Οι εξωτερικές λειτουργίες στο περιεχόμενο καθορίζονται ως προς το μοναδικό όνομα, ενώ εσωτερικά στο σύστημα εξυπηρέτησης η διαχείριση γίνεται μέσω της επίσης μοναδικής ταυτότητας. Στο παρακάτω σχήμα φαίνεται η αρχιτεκτονική της Υπηρεσίας Περιεχομένου όπως υλοποιήθηκε στο Σύστημα Εξυπηρέτησης.

Οι λειτουργίες της Υπηρεσίας Περιεχομένου είναι εγκατάσταση, αντιγραφή, τροποποίηση και μετακίνηση. Η εγκατάσταση περιεχομένου επιτρέπει την ατομική εγκατάσταση αντικειμένων κάτω από μία ιεραρχία ονομάτων στο χώρο ονομάτων του συστήματος εξυπηρέτησης. Η επιλογή της ιεραρχίας των ονομάτων γίνεται από τον παροχέα του περιεχομένου. Εκτός από το όνομα του αντικείμενου που εγκαθίσταται στο σύστημα είναι δυνατόν να καθοριστεί και λίστα από προκαθορισμένα χαρακτηριστικά του αντικείμενου μαζί με τις τιμές τους. Η λίστα αυτή κωδικοποιείται και αποκωδικοποιείται όπως ορίζεται από το πρότυπο CORBA. Ένα από τα χαρακτηριστικά που είναι δυνατόν να καθοριστούν είναι η πηγή του περιεχομένου ή της τιμής κάποιου χαρακτηριστικού του αντικείμενου. Η πηγή μπορεί να είναι ένα τοπικό αρχείο το οποίο θα χρησιμοποιηθεί για εγκατάσταση στον

εξυπηρετητή. Η λειτουργία αυτή εσωτερικά αρχικοποιεί μια συνδιαλλαγή στη βάση δεδομένων για τη δημιουργία του αντικειμένου που θα αποθηκευτεί. Είναι επίσης υπεύθυνη για την ανίχνευση του τρόπου αποθήκευσης του αντικειμένου, δηλαδή σαν αντικείμενο της βάσης δεδομένων ή σαν ροή στον εξυπηρετητή ροών πραγματικού χρόνου.



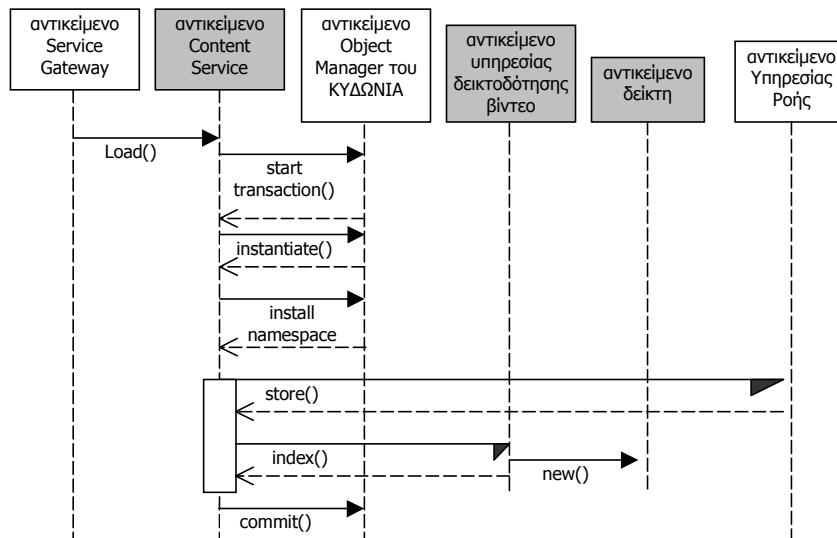
Σχήμα 4-9. Η Υπηρεσία Περιεχομένου στον εξυπηρετητή ΚΥΔΩΝΙΑ-D

Η Υπηρεσία Περιεχομένου του DAVIC υποστηρίζεται στον εξυπηρετητή ΚΥΔΩΝΙΑ-D από τρία βασικά τμήματα: την κλάση Content Service Element, το πακέτο κλάσεων Content Loader και το πακέτο δεικτοδότησης MPEG. Η κλάση Content Service Element εξάγει τις λειτουργίες διαχείρισης του περιεχομένου από τον έξω κόσμο και αποτελεί το front-end του συστήματος ΚΥΔΩΝΙΑ γι' αυτό το σκοπό. Το πακέτο Content Loader υλοποιεί την απομακρυσμένη αποθήκευση του περιεχομένου. Το τελευταίο είναι επίσης απαραίτητο να δεικτοδοτηθεί στο σύστημα ΚΥΔΩΝΙΑ για να υποστηριχθούν οι υπηρεσίες περιήγησης στο περιεχόμενο.

Η αντιγραφή περιεχομένου γίνεται με βάση το όνομα κάποιου αντικειμένου που έχει ήδη εγκατασταθεί στο σύστημα. Έτσι είναι δυνατόν να δημιουργηθεί ένα αντίγραφο του με άλλο όνομα ή σε άλλο χώρο ονομάτων, εφόσον τα ονόματα σχηματίζουν μια ιεραρχία. Η λειτουργία αυτή ελέγχει αν το αντικείμενο το όνομα του οποίου καθορίζεται, υπάρχει στο σύστημα. Ανάλογα με τον τύπο του αντικειμένου, δημιουργείται ένα στιγμιότυπο του ίδιου τύπου, αντίγραφο του αρχικού. Η τροποποίηση του περιεχομένου επιτρέπει την ενημέρωση κάποιας(ων) από τις ιδιότητες του αποθηκευμένου αντικειμένου, ενώ η μετακίνηση περιεχομένου αφορά στην αλλαγή του ονόματος ή του χώρου ονομάτων όπου ανήκει ένα ήδη εγκατεστημένο αντικείμενο. Η ανάκτηση και ενημέρωση των ιδιοτήτων αντικειμένων. Οι λειτουργίες αυτές προβλέπονται από το DSM-CC. Όλες οι λειτουργίες

διαχείρισης περιεχομένου εξασφαλίζουν την εκτέλεσή τους ως συνδιαλλαγές του συστήματος διαχείρισης της βάσης δεδομένων.

Στο σχήμα που ακολουθεί παρουσιάζεται η διαδοχή των μηνυμάτων εγκατάστασης περιεχομένου σύμφωνα με το πρότυπο DAVIC και όπως αυτή υποστηρίζεται από τα αντικείμενα του συστήματος.



Σχήμα 4-10. Εγκατάσταση MPEG περιεχομένου στον εξυπηρετητή

Η εγκατάσταση περιεχομένου είναι απομακρυσμένη και έχει ως αφηγηρία το αντικείμενο Service Gateway που αποτελεί την είσοδο στο σύστημα του εξυπηρετητή. Το αντικείμενο της υπηρεσίας περιεχομένου αναλαμβάνει την επικοινωνία με το σύστημα ΚΥΔΩΝΙΑ για την ενεργοποίηση των απαραίτητων συνδιαλλαγών για την δημιουργία και αποθήκευση του στιγμιότυπου στη βάση δεδομένων αλλά και στον εξυπηρετητή αποθήκευσης της Υπηρεσίας Ροής όταν πρόκειται για δεδομένα ευαίσθητα σε καθυστερήσεις.

Η εγκατάσταση του περιεχομένου γίνεται με την αποστολή του μηνύματος load() σε ένα αντικείμενο της Υπηρεσίας Περιεχομένου. Το αντικείμενο αυτό επικοινωνεί με το Διαχειριστή Αντικειμένων του ΚΥΔΩΝΙΑ ενεργοποιώντας μια συνδιαλλαγή στη βάση δεδομένων. Η συνδιαλλαγή περιλαμβάνει την δημιουργία ενός στιγμιότυπου για το περιεχόμενο. Επίσης γίνεται διαχείριση των χώρων ονομάτων με την εγκατάσταση νέων. Κατόπιν το αντικείμενο της Υπηρεσίας Περιεχομένου ενεργοποιεί ταυτόχρονα δύο λειτουργίες: την αποθήκευση του περιεχομένου και την δημιουργία των δομών δεικτοδότησης, η αποθήκευση του περιεχομένου στην περίπτωση που αυτό έχει απαιτήσεις ανάκτησης σε πραγματικό χρόνο γίνεται με την αποστολή των δεδομένων στη

Υπηρεσία Ροής του συστήματος. Η δημιουργία των δεικτών γίνεται από ένα αντικείμενο της υπηρεσίας δεικτοδότησης.

4.5 Αναπαράσταση MPEG-2 βίντεο με την γλώσσα AVDL

Η διαχείριση MPEG περιεχομένου στον εξυπηρετητή ΚΥΔΩΝΙΑ-D, όπως παρουσιάστηκε στις προηγούμενες ενότητες, έγινε με στόχο την κάλυψη των αναγκών εφαρμογών βίντεο κατ' απαίτηση και εκπαίδευσης από απόσταση με πολυμέσα. Η προσέγγιση αυτή επεκτείνεται στα πλαίσια της διατριβής για την ανάπτυξη μιας γενικής υποδομής διαχείρισης MPEG περιεχομένου σε οποιοδήποτε περιβάλλον ή σύστημα υποστήριξης οπτικοακουστικών εφαρμογών. Για την υποστήριξη αυτής της γενικής υποδομής στον πυρήνα των πληροφοριακών συστημάτων παρουσιάστηκε στο Κεφάλαιο 3 η γλώσσα αναπαράστασης οπτικοακουστικού περιεχομένου AVDL, η οποία παρέχει έναν τρόπο περιγραφής του περιεχομένου που είναι απλός, ευέλικτος, στηρίζεται σε ευρέως αποδεκτά πρότυπα και είναι κατάλληλος και για εφαρμογές του Παγκόσμιου Ιστού. Επιπλέον παρουσιάστηκε αναλυτικά το τμήμα της AVDL που περιγράφει το κωδικοποιημένο περιεχόμενο του προτύπου MPEG-4. Αν και το πρότυπο MPEG-4 είναι πρωτοποριακό και πρόκειται να διαδραματίσει σημαντικό ρόλο στην εξάπλωση των οπτικοακουστικών εφαρμογών, ιδιαίτερα σε περιβάλλοντα δικτύων με χαμηλούς ρυθμούς μετάδοσης, το πρότυπο MPEG-2 είναι αυτό που έχει ήδη καθιερωθεί στο πεδίο των εφαρμογών φυσικού βίντεο και τα ελκυστικά χαρακτηριστικά του αναμένεται να το διατηρήσουν σε αυτή τη θέση για αρκετά μεγάλο διάστημα. Επομένως αποτελεί αναγκαιότητα των συστημάτων υποστήριξης υπηρεσιών διαχείρισης οπτικοακουστικού περιεχομένου να παρέχουν υπηρεσίες διαχείρισης MPEG-2 περιεχομένου στον πυρήνα της λειτουργικότητάς τους.

Στην παρούσα ενότητα η γλώσσα περιγραφής οπτικοακουστικού περιεχομένου AVDL επεκτείνεται ώστε να περιγράψει τα δεδομένα του προτύπου MPEG-2 με στόχο η περιγραφή που θα προκύψει να χρησιμοποιηθεί για να υποστηριχθούν οι λειτουργίες παρουσίασης και περιήγησης, ανάκτησης βάσει περιεχομένου, διαβαθμισμένης μεταβίβασης και ανταλλαγής.

Η γλώσσα AVDL όπως παρουσιάστηκε στο προηγούμενο κεφάλαιο ορίζει την αναπαράσταση οπτικοακουστικών αντικειμένων. Τα αντικείμενα αυτά αναπαρίστανται με στοιχεία του τύπου `mpeg4` ή `mpeg2`, όπως φαίνεται στο Σχήμα 3-7. Δήλωση τύπου εγγράφου και στοιχείου - ρίζας. Η αναπαράσταση των στοιχείων του MPEG-2 αποτελεί διαφορετικό προφίλ της περιγραφής του περιεχομένου, οπότε περιέχεται σε τμήμα υπό

συνθήκη που επιτρέπει την αγνόησή του σε ένα έγγραφο της AVDL που δεν περιγράφει περιεχόμενο του MPEG-2.

Τα στοιχεία του τύπου mpeg2 διακρίνονται σε στοιχεία σειρών βίντεο (VideoSequence), τα οποία περιγράφουν τη δομή των ψηφιοσειρών βίντεο, όπως φαίνεται στο παρακάτω σχήμα. Οι σειρές αυτές αποτελούνται από μια επικεφαλίδα (SequenceHeader) και το στοιχείο Extension που με τη σειρά του αποτελείται από το σύνολο των εικόνων της σειράς.

```
<!-- Δηλώσεις για τη χρήση του προφίλ περιγραφής MPEG-2 βίντεο -->
<!ENTITY % MPEG2.Description "IGNORE">

<![ %MPEG2.Description ; [
    <!-- Στοιχείο ρίζα της αναπαράστασης MPEG-2 βίντεο -->
    <!ELEMENT mpeg2 (MPEG2Video) >

    <!ENTITY % MPEG2Video.Description "INGORE">

    <![ %MPEG2Video.Description ; [

        <!ELEMENT MPEG2Video (VideoSequence) >

        <!-- Σειρά MPEG-2 βίντεο -->
        <!ELEMENT VideoSequence (SequenceHeader, Extension) >
        <!ATTLIST VideoSequence
            id      %id_type; %implied; >

        <!ELEMENT Extension
            (SequenceExtension,
             ((SequenceDisplayExtension | SequenceScalableExtension),
              UserData), Picture*,
             (SequenceHeader, SequenceExtension) )* >

        <!-- Εικόνα βίντεο -->
        <!ELEMENT Picture
            (GOPHeader?, PictureHeader, PictureCodingExtension,
             (QuantMatrixExtension | PictureDisplayExtension |
              PictureSpatialScalableExtension |
              PictureTemporalScalableExtension),
             UserData, PictureData) >

        <!ATTLIST Picture
            id      %id_type; %implied; >
```

Σχήμα 4-11. Αναπαράσταση σειρών MPEG-2 βίντεο με τη γλώσσα AVDL

Στο επόμενο σχήμα περιγράφεται η δομή και η σημασιολογία της επικεφαλίδας μιας σειράς MPEG-2 βίντεο. Το οριζόντιο και κατακόρυφο μέγεθος σχηματίζεται από τα αντίστοιχα χαρακτηριστικά του SequenceHeader και από αυτά του SequenceExtension. Αποτελούν το πλάτος και το ύψος των ορατών τμημάτων της φωτεινότητας των εικόνων και ορίζονται ως προς δείγματα. Ο λόγος των διαστάσεων των δειγμάτων (AspectRatio), είτε είναι 1 που υπονοεί τετράγωνα δείγματα, είτε υπολογίζεται από τον τύπο

$\text{display_aspect_ratio} \times \frac{\text{horizontal_size}}{\text{vertical_size}}$. Ο ρυθμός των εικόνων είναι πολλαπλάσιο των 400

bits/sec. Το συνολικό μέγεθος του απομονωτή αποκωδικοποίησης είναι $B = 16 \times 1024 \times \text{vbn_buffer_size}$. Ο ρυθμός των εικόνων ξεκινάει από 23,976 εικόνες το δευτερόλεπτο και φτάνει μέχρι 60 fps.

```
<!-- Επικεφαλίδα της σειράς βίντεο -->
<!ELEMENT SequenceHeader
  (AspectRatio, IntraQuantiserMatrix | NonIntraQuantiserMatrix) >

<!-- ATTLIST SequenceHeader
  horizontal_size          %integer;
  vertical_size            %integer;
  frame_rate               ( 23_976 | 24 | 25 | 29_97 | 30 | 50
  | 59_94 | 60 ) %implied; >
  bit_rate                 %integer;
  vbn_buffer_size          %integer;
  constrained_parameters   %fixed; 0
  load_intra_quantiser_matrix %boolean; >

<!ELEMENT AspectRatio ( DisplayAspectRatio? )>
<!-- ATTLIST AspectRatio
  sample_aspect_ratio ( square | display_x_size_ratio ) %implied; >

<!-- ATTLIST DisplayAspectRatio
  display_aspect_ratio ( 3_4 | 9_16 | 1_2_21 ) %implied; >
```

Η σειρά του βίντεο εκτός από την επικεφαλίδα της αποτελείται από ένα σύνολο δεδομένων που αφορούν στην παρουσίαση και κωδικοποίηση των εικόνων που περιέχει. Τα δεδομένα κωδικοποίησης αυτά συγκεκριμένα αφορούν στην κλιμάκωση της σειράς και στην χρήση πινάκων κβαντοποίησης όπως φαίνεται στο ακόλουθο σχήμα. Οι πίνακες κβαντοποίησης που αποτελούν μέρος της δομής των αντικειμένων βίντεο, όπως φαίνεται από το στοιχείο `ExtensionData`, έχουν διαστάσεις 8x8 και αποτελούνται από προκαθορισμένες τιμές για τα διαφορετικά είδη εικόνων.

```
<!-- ELEMENT ExtensionData
  ((SequenceDisplayExtension | SequenceScalableExtension) |
  (QuantMatrixExtension |
  PictureDisplayExtension |
  PictureSpatialScalableExtension |
  PictureTemporalScalableExtension)
  )>
```

Αν η κωδικοποίηση του βίντεο είναι προοδευτική, τότε αυτό περιέχει εικόνες με κωδικοποίηση καρτέ (frame-based), ενώ αν δεν είναι προοδευτική περιέχει εικόνες κωδικοποίησης πεδίου (field-based) αλλά και κωδικοποίησης καρτέ οι οποίες μπορεί να είναι πεπλεγμένες. Η μορφή του χρώματος του βίντεο έχει τρεις δυνατές τιμές, δηλαδή 4:2:0, 4:2:2 ή 4:4:4. Αν το βίντεο περιέχει επίπεδα επέκτασης, τότε χαρακτηριστικά όπως το οριζόντιο και κατακόρυφο μέγεθος, το μέγεθος του απομονωτή αποκωδικοποίησης και ο ρυθμός των εικόνων μεταβάλλονται για να λάβουν υπόψη τους την επέκταση. Το προφίλ

των αντικειμένων MPEG-2 βίντεο είναι πέντε ειδών: Απλό (Simple), Βασικό (Main), SNR, Χωρικό (Spatial) και Υψηλό (High). Τα επίπεδα διακρίνονται σε Χαμηλό, Βασικό, Υψηλό-1440 και Υψηλό.

```
<!-- ELEMENT SequenceExtension (MPEG2ProfileAndLevel) -->
<!-- ATTLIST SequenceExtension
    progressive_sequence          %boolean;
    chroma_format                 ( 4_2_0 | 4_2_2 | 4_4_4 ) %implied;
    horizontal_size_extension     %byte;
    vertical_size_extension       %byte;
    bit_rate_extension            %integer;
    vbv_buffer_size_extension     %byte;
    low_delay                     %boolean;
    frame_rate_extension_n        %byte;
    frame_rate_extension_d        %byte; >

<!-- ATTLIST MPEG2ProfileAndLevel
    profile ( simple | main | SNR | spatial | high ) %implied;
    level   ( low | main | high-1440 | high ) %implied; >
```

Σχήμα 4-12. Προφίλ και επίπεδα σειρών MPEG-2 βίντεο στη γλώσσα AVDL

Η μορφή του βίντεο πριν κωδικοποιηθεί αναπαρίσταται από το χαρακτηριστικό video_format το οποίο έχει τις ίδιες τιμές για το βίντεο MPEG-2 και MPEG-4. Η περιγραφή των χρωμάτων (στοιχεία ColorPrimaries, TransferCharacteristics και MatrixCoefficients) έχει επίσης κοινά στοιχεία με του MPEG-4. Τα κοινά αυτά στοιχεία αποτελούν μέρος των κοινών δηλώσεων της AVDL. Το οριζόντιο και κατακόρυφο μέγεθος του display ορίζουν ένα ορθογώνιο που μπορεί να θεωρηθεί η ενεργός περιοχή παρουσίασης.

```
<!-- Χαρακτηριστικά της μορφής του βίντεο -->
<!-- ELEMENT SequenceDisplayExtension (MPEG2ColourDescription?)
<!-- ATTLIST SequenceDisplayExtension
    %video_format;
    display_horizontal_size %integer;
    display_vertical_size   %integer; >

<!-- ATTLIST MPEG2ColourDescription
    (MPEG2ColorPrimaries, TransferCharacteristics, MatrixCoefficients)>

<!-- ELEMENT MPEG2ColorPrimaries
    ( Unspecified | ITU-R.709 | ITU-R.624-4_System_M | ITU-R.624-
4SystemBG | SMPTE_170M | SMPTE_240M ) > <!-- ATTLIST ITU-R.709
    %common. ITU-R.709;
    %whiteD65; "0,3127 0,3290">
```

Η επέκταση των ψηφιοσειρών του MPEG-2 βίντεο είναι δυνατόν να περιέχει κλιμάκωση δύο ειδών, χωρική ή χρονική. Το χαρακτηριστικό scalable_mode δείχνει το είδος της κλιμάκωσης. Η χωρική κλιμάκωση καθορίζει τις παραμέτρους υποδειγματοληψίας της βασικής βαθμίδας. Στη χρονική κλιμάκωση τα χαρακτηριστικά picture_max_order και picture_max_factor δείχνουν τον αριθμό των βελτιωτικών εικόνων που παρεμβάλλονται ανάμεσα στις εικόνες της βασικής βαθμίδας.

```
<!-- Βαθμίδα κλιμάκωσης βίντεο -->
<!-- ELEMENT SequenceScalableExtension
```

```

        (SpatialScalability | TemporalScalability )>
<!-- ATTLIST SequenceScalableExtension
        scalable_mode (data_partitioning | spatial | SNR | temporal)
%implied;
        layer_id          %ref_type; %required;>

<!-- Χωρική κλιμάκωση MPEG-2 βίντεο -->
<!-- ATTLIST SpatialScalability
        lower_layer_prediction_horizontal_size    %integer;
        lower_layer_prediction_vertical_size      %integer;
        horizontal_subsampling_factor_m           %integer;
        horizontal_subsampling_factor_n           %integer;
        vertical_subsampling_factor_m             %integer;
        vertical_subsampling_factor_n             %integer; >

<!-- Χρονική κλιμάκωση MPEG-2 βίντεο -->
<!-- ELEMENT TemporalScalability (MuxToProgressiveSequence?)
<!-- ATTLIST TemporalScalability
        picture_mux_enable    %boolean;
        picture_mux_order     %byte;
        picture_mux_factor    %byte; >

<!-- ATTLIST MuxToProgressiveSequence
        mux_to_progressive_sequence    %boolean; >

```

Η επικεφαλίδα μιας ομάδας εικόνων (GOP) αποτελείται από χρόνο παρουσίασης (TimeCode) της πρώτης ομάδας, όπως καθορίζεται και για το MPEG-4. Ο τύπος των κωδικοποιημένων εικόνων είναι I, P ή B έχοντας την ίδια σημασία όπως των Επιπέδων Εικόνων Βίντεο (VOP) του MPEG-4.

```

<!-- Επικεφαλίδα ομάδας εικόνων -->
<!-- ELEMENT GOPHeader (TimeCode)>

<!-- ATTLIST GOPHeader
        drop_frame_flag    %boolean;
        time_code_pictures  %minutes;
        closed_gop         %boolean;
        broken_link         %boolean;>

<!-- Επικεφαλίδα εικόνων -->
<!-- ELEMENT PictureHeader
        ( PBHeader? , BHeader?, ExtraInfoPicture* )>

<!-- ATTLIST PictureHeader
        temporal_reference    %integer;
        picture_coding_type   ( I | P | B ) %implied;
        vbv_delay             %integer; >

<!-- ATTLIST PBHeader
        full_pixel_forward_vector    %fixed; "0"
        forward_code                  %fixed; "7">

<!-- ATTLIST BHeader
        full_pixel_backward_vector    %fixed; "0"
        backward_code                  %fixed; "7">

<!-- ATTLIST ExtraInfoPicture
        extra_info_picture            %byte; >

```

Τα χαρακτηριστικά του στοιχείου PictureCodingExtension Το στοιχείο CompositeDisplay ορίζεται αν η μορφή του βίντεο πριν την κωδικοποίηση ήταν Composite και η πληροφορία αυτή αφορά σε μεμονωμένες εικόνες της σειράς.

```
<!ELEMENT PictureCodingExtension (CompositeDisplay?)>

<!--ATTLIST PictureCodingExtension
    f_code_forward_horizontal      (1|2|3|4|5|6|7|8|9 | 15) %implied;
    f_code_forward_vertical        (1|2|3|4|5|6|7|8|9 | 15) %implied;
    f_code_backward_horizontal     (1|2|3|4|5|6|7|8|9 | 15) %implied;
    f_code_backward_vertical       (1|2|3|4|5|6|7|8|9 | 15) %implied;
    intra_dc_precision             ( 8 | 9 | 10 | 11 ) %implied;
    picture_structure               ( top_field | bottom_field |
frame_picture ) %implied;
    top_field_first                 %boolean;
    frame_pred_frame_dct            %boolean;
    concealment_motion_vectors     %boolean;
    q_scale_type                   %boolean;
    intra_vlc_format               %boolean;
    alternate_scan                  %boolean;
    repeat_field_first              %boolean;
    progressive_frame               %boolean;
    composite_display_flag          %boolean;>

<!--ATTLIST CompositeDisplay
    v_axis                          %boolean;
    field_sequence                  ( 1 | 2 | 3 ) %implied;
    sub_carrier_phase               %integer;
    burst_amplitude                 %byte;
    sub_carrier_phase               %byte;>
```

Το στοιχείο PictureDisplayExtension επιτρέπει την τοποθέτηση του ορθογωνίου παρουσίασης που ορίζεται στο στοιχείο SequenceDisplayExtension να μεταβάλλεται από εικόνα σε εικόνα του βίντεο. Τα χαρακτηριστικά (frame_centre_horizontal, frame_centre_vertical_offset) καθορίζουν την απόσταση του κέντρου της εικόνας από το κέντρο του ορθογωνίου παρουσίασης. Οι αποστάσεις αυτές καθορίζονται σε κλάσματα των δειγμάτων των εικόνων.

```
<!ELEMENT PictureDisplayExtension (FrameCentreOffsets*)>
<!--ATTLIST PictureDisplayExtension
    number_of_frame_centre_offsets %integer;>

<!--ATTLIST FrameCentreOffsets
    frame_centre_horizontal_offset %integer;
    frame_centre_vertical_offset   %integer;>
```

Η πληροφορία χρονικής κλιμάκωσης περιέχει τη χρονική ταυτότητα των εικόνων που θα χρησιμοποιηθούν για χρονική πρόβλεψη (εμπρόσθια ή προς τα πίσω). Με παρόμοιο τρόπο η χωρική κλιμάκωση καθορίζει τη χρονική ταυτότητα των εικόνων που θα χρησιμοποιηθούν για χωρική κλιμάκωση καθώς και τη σχετική θέση των δύο εικόνων (χαμηλού επιπέδου και βελτιωτικής).

```
<!--ATTLIST PictureTemporalScalableExtension
```

```

reference_select_code      %byte;
forward_temporal_reference %integer;
backward_temporal_reference %integer; >

<!ATTLIST PictureSpatialScalableExtension
  lower_layer_temporal_reference %integer;
  lower_layer_horizontal_offset %integer;
  lower_layer_vertical_offset %integer;
  spatial_temporal_weight_code_table_index
  lower_layer_progressive_frame %boolean;
  lower_layer_deinterlaced_field_select %boolean; >

```

Οι εικόνες του βίντεο αποτελούνται από "φέτες" (slices), δηλαδή λωρίδες αποτελούμενες από μικρότερα τμήματα που ονομάζονται macroblocks. Οι λωρίδες αυτές επιτρέπουν την κωδικοποίηση τυχαίων τμημάτων μιας εικόνας με ορθογώνιες διαστάσεις. Τα περισσότερα επίπεδα (levels) που καθορίζει το πρότυπο απαιτούν τις λωρίδες να καλύπτουν όλη την περιοχή της εικόνας. Τα macroblocks αποτελούνται από τα κωδικοποιημένα τμήματα φωτεινότητας και χρωματότητας της εικόνας. Η αναπαράσταση των τμημάτων αυτών και του περιεχομένου τους δεν αναλύεται περαιτέρω, εφόσον είναι χρήσιμη μόνο στον αποκωδικοποιητή των δεδομένων.

```

<!ELEMENT PictureData (Slice*)>

<!ELEMENT Slice
  (DataPartitioningScalableExtensionSlice?, IntraSlice?,
   ExtraInfoSlice*, Macroblock*)>

<!ATTLIST IntraSlice
  intra_slice %boolean; >

<!ATTLIST ExtraInfoSlice
  extra_info_slice %byte; >

<!ELEMENT Macroblock
  (MacroblockEscape*, MacroblockModes, QuantiserScaleCode?,
   MotionVectors0?, MotionVectors1?,
   CodedBlockPattern?, Block*) >

```

Το τμήμα της γλώσσας AVDL που παρουσιάστηκε στην ενότητα αυτή για την αναπαράσταση του περιεχομένου MPEG-2 βίντεο, επιτρέπει την περιγραφή όλων εκείνων των στοιχείων που μπορούν να χρησιμοποιηθούν για την υποστήριξη των βασικών λειτουργιών διαχείρισης οπτικοακουστικού περιεχομένου, δηλαδή ελέγχου της περιήγησης και παρουσίασης, ανάκτησης βάσει περιεχομένου, διαβαθμισμένης μεταβίβασης σύμφωνα με το προφίλ χρηστών και εφαρμογών και ανταλλαγής εγγράφων. Η χρησιμότητα της αναπαράστασης παρουσιάζεται αναλυτικά στο Κεφάλαιο 6 μέσα από την περιγραφή της υποστήριξης αυτών των λειτουργιών.

4.6 Επέκταση των δομών δεικτοδότησης για τα δεδομένα του MPEG-4

Στην ενότητα αυτή προτείνεται η επέκταση των δομών δεικτοδότησης που αναπτύχθηκαν για το περιεχόμενο MPEG-1 και MPEG-2 και στο περιεχόμενο του MPEG-4. Η δεικτοδότηση του περιεχομένου αποτελεί βασικό στοιχείο του πληροφοριακού συστήματος για την υποστήριξη του πυρήνα λειτουργικότητας των οπτικοακουστικών εφαρμογών. Η δεικτοδότηση του MPEG-2 και MPEG-1 περιεχομένου που παρουσιάστηκε στο κεφάλαιο αυτό, παρέχουν τυχαία προσπέλαση στο περιεχόμενο των ροών και υποστηρίζουν τη λειτουργικότητα περιήγησης σε αυτό. Οι δομές δεικτοδότησης που προτείνονται σε αυτήν την ενότητα στηρίζονται στη γενική ιδέα των δομών για το MPEG-2 και MPEG-1. Δηλαδή προτείνεται να παρέχουν προσπέλαση σε ανεξάρτητα τμήματα του περιεχομένου.

Οι δομές δεικτοδότησης για το MPEG-4 πρέπει να περιλαμβάνουν δεικτοδότηση των ροών οπτικών αντικειμένων και της σκηνής. Τα οπτικά αντικείμενα ροών πραγματικού χρόνου διακρίνονται σε αντικείμενα βίντεο, πλέγματα και πρόσωπα. Τα ανεξάρτητα σημεία προσπέλασης μέσα στο περιεχόμενο αυτών των ροών είναι τα Επίπεδα Αντικειμένων Βίντεο (Video Object Planes), Αντικειμένων Πλεγμάτων (Mesh Object Planes) και Αντικειμένων Προσώπων (Face Object Planes) αντίστοιχα. Επομένως προτείνεται οι δομές δεικτοδότησης των αντικειμένων βίντεο, πλεγμάτων και προσώπων να περιέχουν όρους για τα ανεξάρτητα σημεία προσπέλασης Επιπέδων βίντεο, πλεγμάτων και προσώπων αντίστοιχα.

Στα αντικείμενα βίντεο η δομή των κωδικοποιημένων ροών είναι δυνατόν να περιέχει βαθμίδες κλιμάκωσης, οι οποίες έπονται της βαθμίδας βάσης. Η τυχαία προσπέλαση των βαθμίδων είναι εξίσου σημαντική με την τυχαία προσπέλαση των Επιπέδων, επομένως είναι σημαντικό η δομή δεικτοδότησης να περιέχει όρους και για την διάκριση βαθμίδων βάσης και κλιμάκωσης. Για παράδειγμα, κατά την τυχαία προσπέλαση σε ένα χρονικό σημείο μιας σειράς βίντεο που περιέχει κλιμάκωση, είναι απαραίτητο να μεταδοθεί το Επίπεδο από τη βαθμίδα βάσης μαζί με την επιπλέον πληροφορία του Επιπέδου στη βαθμίδα κλιμάκωσης. Η δομή των όρων δεικτοδότησης MPEG-4 βίντεο φαίνεται στο ακόλουθο σχήμα, όπου οι όροι δεικτοδότησης είναι τριών ειδών: βαθμίδας βάσης, βαθμίδων κλιμάκωσης και Επιπέδων Αντικειμένου Βίντεο.

Όροι δεικτοδότησης
βαθμίδων MPEG-4 βίντεο

Πρώτο Επίπεδο βαθμίδας βάσης	Μέγεθος βαθμίδας βάσης	Πρώτο Επίπεδο βαθμίδας	Μέγεθος πρώτης βαθμίδας κλιμάκωσης	Είδος πρώτης βαθμίδας κλιμάκωσης	...
---------------------------------	---------------------------	---------------------------	---------------------------------------	-------------------------------------	-----

Όρος δεικτοδότησης
MPEG-4 βίντεο

Σειρά παρουσίασης Επιπέδου Αντικειμένου Βίντεο	Μέγεθος Επιπέδου Αντικειμένου Βίντεο	Είδος Επιπέδου Αντικειμένου Βίντεο	Κλιμάκωση	...
--	---	---------------------------------------	-----------	-----

Σχήμα 4-13. Όροι δεικτοδότησης ροών MPEG-4 βίντεο

Οι όροι των δομών δεικτοδότησης για το βίντεο του MPEG-4 αφορούν στη δομή των βαθμίδων και στα πραγματικά σημεία προσπέλασης που είναι τα Επίπεδα Αντικειμένων Βίντεο. Οι βαθμίδες δεικτοδοτούνται για τη διευκόλυνση της τυχαιάς προσπέλασης Επιπέδων Βίντεο που περιέχουν κλιμάκωση. Οι όροι δεικτοδότησης των Επιπέδων αποτελούνται από τη σειρά χρονικής παρουσίασης, το μέγεθος του επιπέδου, το είδος του (I, P, B ή Sprite) και αν είναι επίπεδα κλιμάκωσης ή όχι.

Για την οργάνωση των όρων του δείκτη προτείνεται η δομή που κατασκευάστηκε και για τη δεικτοδότηση των δεδομένων του MPEG-2 και του MPEG-1, όπου ο δείκτης αποτελείται από ένα πυκνό επίπεδο και ένα ή περισσότερα αραιά επίπεδα που διευκολύνουν την ταχύτερη προσπέλαση του δείκτη στη δευτερεύουσα μνήμη και μειώνουν το επιπλέον κόστος προσπέλασης.

Τα Επίπεδα Αντικειμένων Πλεγμάτων έχουν ένα είδος και οι σειρές δεν περιέχουν κλιμάκωση. Οι όροι δεικτοδότησης επομένως περιέχουν τη σειρά παρουσίασης, το μέγεθος του Επιπέδου Πλέγματος και το είδος του Επιπέδου. Τα είδη των Επιπέδων Πλεγμάτων είναι δύο: ενδοκωδικοποιημένα και μη. Τα ενδοκωδικοποιημένα Επίπεδα είναι ανεξάρτητα από οποιοδήποτε άλλο Επίπεδο της σειράς και αποτελούν Επίπεδα αναφοράς για τα μη ενδοκωδικοποιημένα. Στο ακόλουθο σχήμα παριστάνεται η δομή των όρων δεικτοδότησης.

Όρος δεικτοδότησης
MPEG-4 πλέγματος

Σειρά παρουσίασης Επιπέδου Αντικειμένου Πλέγματος	Μέγεθος Επιπέδου Αντικειμένου Πλέγματος	Είδος Επιπέδου Αντικειμένου Πλέγματος	...
---	---	---	-----

Σχήμα 4-14. Όροι δεικτοδότησης ροών MPEG-4 πλεγμάτων

Οι όροι των δομών δεικτοδότησης για τα πλέγματα του MPEG-4 αποτελούνται από τη σειρά χρονικής παρουσίασης των Επιπέδων, το μέγεθός τους και το είδος τους (αν είναι ενδοκωδικοποιημένα ή μη).

Οι όροι της δομής δεικτοδότησης για τα Επίπεδα Αντικειμένων Προσώπων είναι παρόμοιοι με αυτούς των αντικειμένων πλεγμάτων. Αποτελούνται δηλαδή από τη σειρά παρουσίασης,

το μέγεθος και το είδος του Επιπέδου. Και στα κωδικοποιημένα αντικείμενα προσώπων το είδος των Επιπέδων είναι ενδοκωδικοποιημένο και μη ενδοκωδικοποιημένα, δηλαδή περιέχουν πρόβλεψη από άλλα Επίπεδα.

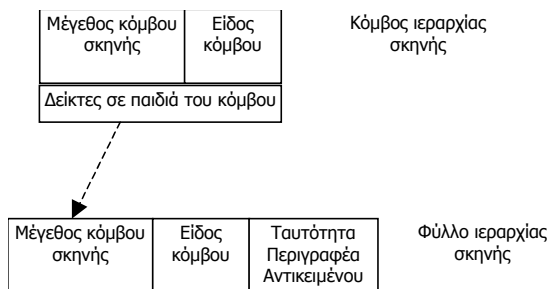
Όροι δεικτοδότησης
MPEG-4 προσώπου

Σειρά παρουσίασης Επιπέδου Αντικειμένου Προσώπου	Μέγεθος Επιπέδου Αντικειμένου Προσώπου	Είδος Επιπέδου Αντικειμένου Προσώπου	...
--	--	--	-----

Σχήμα 4-15. Όροι δεικτοδότησης ροών MPEG-4 προσώπων

Οι όροι των δομών δεικτοδότησης για τα πρόσωπα του MPEG-4 αποτελούνται από τη σειρά χρονικής παρουσίασης των Επιπέδων Προσώπων, το μέγεθός τους και το είδος τους (αν είναι ενδοκωδικοποιημένα ή μη).

Η λογική δομή της σκηνής του MPEG-4 είναι μια ιεραρχία κόμβων που αντιστοιχούν σε οπτικοακουστικά αντικείμενα, γραφικά αντικείμενα παρουσίασης και ομαδοποίησης και αντικείμενα αλληλεπίδρασης με τη σκηνή.



Σχήμα 4-16. Όροι δεικτοδότησης σκηνών του MPEG-4

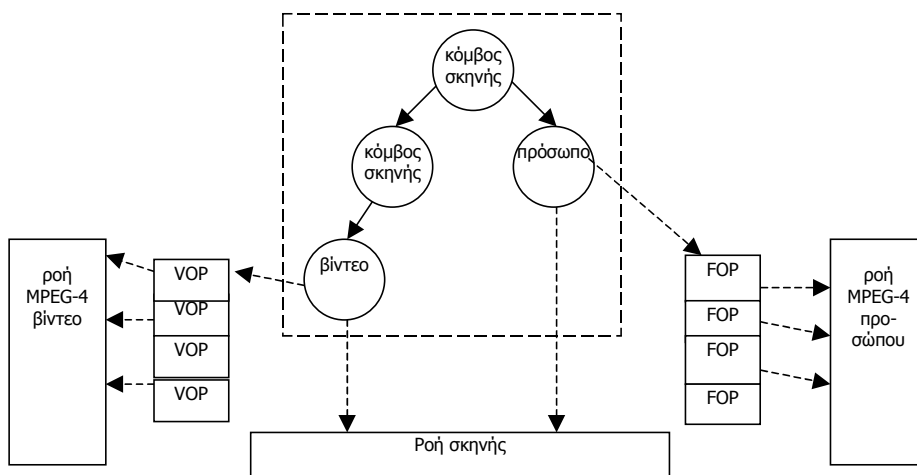
Η δομή δεικτοδότησης των σκηνών του MPEG-4 σχηματίζουν μια ιεραρχία (δενδρική δομή) όρων που αντιστοιχούν στους κόμβους της σκηνής. Οι όροι δεικτοδότησης των κόμβων είναι δύο ειδών: εσωτερικοί κόμβοι και φύλλα. Οι εσωτερικοί κόμβοι περιέχουν το είδος τους καθώς και δείκτες στα επόμενα επίπεδα της ιεραρχίας, ενώ οι κόμβοι φύλλα αντιστοιχούν σε αντικείμενα μέσω και περιέχουν την ταυτότητα του περιγραφέα που αναφέρεται σε αυτά.

Η δομή δεικτοδότησης που προτείνεται και απεικονίζεται στο παραπάνω σχήμα αποτελεί επίσης μια δενδρική δομή όπου οι όροι δεικτοδότησης αντιστοιχούν στους κόμβους της σκηνής και περιέχουν πληροφορία για τη θέση τους μέσα σε μια σειρά κωδικοποιημένη με τη μορφή BIFS. Οι όροι αυτοί είναι δύο ειδών, για τους εσωτερικούς και για τα φύλλα της ιεραρχίας. Οι εσωτερικοί κόμβοι περιέχουν το είδος και το μέγεθός τους, ενώ οι κόμβοι-φύλλα αντιστοιχούν σε αντικείμενα μέσω που μεταφέρονται σε ξεχωριστές

κωδικοποιημένες ροές. Η αντιστοίχιση των φύλλων με τα πραγματικά δεδομένα γίνεται μέσω των περιγραφών αντικειμένων όπως και στα Συστήματα του MPEG-4.

Comment [Γ.Κ.17]: αυτό δεν είναι πολύ σαφές. Τουλάχιστον πρέπει να αναφέρω στο κομμάτι της γλώσσας για τους περιγραφείς ότι χρησιμοποιούνται ως σύνδεσμοι με την πηγή πληροφορίας, αν και δεν περιέχουν τη θέση των ροών σε ένα πληροφοριακό σύστημα, π.χ. ένα stream_object_id ή ένα stream_file_name, κλπ

Στο επόμενο σχήμα φαίνεται η συνολική δομή των δεικτών που αφορούν στη σκηνή και στα αντικείμενα του MPEG-4 και η οποία αποτελεί τη σύνδεση μεταξύ τους για την υποστήριξη των λειτουργιών περιήγησης σε ένα ολοκληρωμένο περιβάλλον εφαρμογής. Στην περίπτωση αυτή οι δείκτες χρησιμοποιούνται μέσω της δομής της σκηνής, η οποία αποτελεί και την οντότητα αλληλεπίδρασης με το χρήστη.



Σχήμα 4-17. Δομές δεικτοδότησης σκηνών και αντικειμένων του MPEG-4

Η δεικτοδότηση των σκηνών και των άλλων αντικειμένων του MPEG-4 υποστηρίζει την τυχαία προσπέλαση στο υλικό με αφετηρία τη σκηνή, η οποία αποτελεί την σύνθετη οντότητα με την οποία αλληλεπιδρά ο χρήστης. Τα φύλλα του δέντρου δεικτοδότησης της σκηνής αναφέρονται σε αντικείμενα μέσω των οποίων έχουν κατασκευαστεί ειδικές δομές δεικτοδότησης.

Η αρχιτεκτονική και οι λεπτομέρειες της λειτουργικότητας περιήγησης που στηρίζεται στις δομές δεικτοδότησης παρουσιάζεται στο κεφάλαιο 6 της διατριβής.

4.7 Ανακεφαλαίωση

Στο κεφάλαιο αυτό παρουσιάστηκε η διαχείριση του MPEG-2 περιεχομένου σε εξυπηρετητή πολυμέσων. Η διαχείριση αυτή αποτελείται από τη μοντελοποίηση του περιεχομένου στην αντικειμενοστρεφή βάση δεδομένων του συστήματος ΚΥΔΩΝΙΑ, τη δεικτοδότηση του περιεχομένου και την υποστήριξη της Υπηρεσίας Περιεχομένου του DAVIC. Η μοντελοποίηση του περιεχομένου παρέχει έναν τρόπο δόμησης του

περιεχομένου και διάκρισης των αντικειμένων που είναι σημαντικά για την υποστήριξη των υπηρεσιών του εξυπηρετητή ΚΥΔΩΝΙΑ-D. Η αποθήκευσή του στη βάση δεδομένων του συστήματος αντιμετωπίζεται με ενιαίο τρόπο με όλα τα αντικείμενα της βάσης και γίνεται μέσα από συνδιαλλαγές που εξασφαλίζουν τη συνέπεια των δεδομένων. Η δεικτοδότηση του περιεχομένου εξασφαλίζει την αποδοτική υποστήριξη των υπηρεσιών περιήγησης στο περιεχόμενο του MPEG-2 βίντεο και του MPEG-1 συγχρονισμένου βίντεο και ήχου. Η Υπηρεσία Περιεχομένου του DAVIC παρέχει στους ιδιοκτήτες του περιεχομένου έναν τρόπο εγκατάστασής του στον απομακρυσμένο εξυπηρετητή για να χρησιμοποιηθεί από τις εφαρμογές, αλλά και διαχείρισής του. Η συνεισφορά των παραπάνω είναι ότι παρέχουν ένα μέρος της γενικής υποδομής υποστήριξης υπηρεσιών διαχείρισης οπτικοακουστικού περιεχομένου σε εξυπηρετητή πολυμέσων και αντικειμενοστρεφή βάση δεδομένων. Επίσης αποτέλεσαν σημείο αναφοράς για την ανάπτυξη των υπόλοιπων μεθόδων που παρουσιάζονται στην παρούσα διατριβή και αφορούν στην κατασκευή γλώσσας αναπαράστασης περιεχομένου, στο σχεδιασμό δομών δεικτοδότησης για την πληροφορία και στην κατασκευή βάσεων δεδομένων για την αποθήκευση και ανάκτηση του περιεχομένου.

Έτσι παρουσιάστηκε περιγραφή των δεδομένων του MPEG-2 με τη γλώσσα Περιγραφής Οπτικοακουστικού Περιεχομένου AVDL, μέρος της οποίας παρουσιάστηκε στο προηγούμενο κεφάλαιο για την περιγραφή των δεδομένων του MPEG-4. Το κεφάλαιο αυτό επομένως συμπληρώνει την αναπαράσταση του οπτικοακουστικού περιεχομένου ούτως ώστε στο Κεφάλαιο 5 να αναπτυχθεί το μοντέλο βάσεων δεδομένων για το οπτικοακουστικό περιεχόμενο.

Επιπλέον παρουσιάστηκε ο σχεδιασμός των δομών δεικτοδότησης για τα οπτικά αντικείμενα και τις σκηνές του MPEG-4. Οι δομές αυτές παρέχουν τυχαία προσπέλαση σε τμήματα της πληροφορίας που είναι κωδικοποιημένη σύμφωνα με το πρότυπο MPEG-4 με παρόμοιο τρόπο με τις δομές που έχουν αναπτυχθεί για τη δεικτοδότηση του MPEG-2 και MPEG-1 περιεχομένου και συμπληρώνουν το τμήμα του πυρήνα του πληροφοριακού συστήματος που αφορά στη δεικτοδότηση της πληροφορίας.

Οι μηχανισμοί και τα εργαλεία που παρουσιάστηκαν στο κεφάλαιο αυτό είναι επεκτάσιμα και ως προς την αναπαράσταση πληροφορίας αλλά και ως προς τη δεικτοδότηση του περιεχομένου. Η γλώσσα AVDL μπορεί να επεκταθεί με κατάλληλες δηλώσεις τύπων ώστε να αναπαραστήσει επιπλέον πληροφορία για το περιεχόμενο ανάλογα με τις ανάγκες των εφαρμογών. Η μεταγλώσσα XML επιτρέπει σε κάθε έγγραφο, δηλαδή σε κάθε στιγμιότυπο του τύπου εγγράφου να περιέχει επιπλέον δηλώσεις τύπων στοιχείων που μπορούν να χρησιμοποιηθούν για την αναπαράσταση πληροφορίας.

Επίσης οι δομές δεικτοδότησης MPEG πληροφορίας που παρουσιάστηκαν είναι επεκτάσιμες, με την έννοια ότι μπορούν να χρησιμοποιηθούν για την κατασκευή δεικτών με βάση χαρακτηριστικά (features) του περιεχομένου. Οι νέοι δείκτες στη συγκεκριμένη περίπτωση είναι δυνατόν να περιέχουν όρους δεικτοδότησης που σε συνδυασμό με τις δομές που παρουσιάστηκαν στο παρόν κεφάλαιο θα επιτρέπουν την τυχαία ανάκτηση της πληροφορίας με βάση χαρακτηριστικά που δεν περιορίζονται στη χρονική σειρά παρουσίασης.

Κεφάλαιο 5

Βάση δεδομένων οπτικοακουστικού περιεχομένου

Το κεφάλαιο αυτό παρουσιάζει την κατασκευή βάσεων δεδομένων για την αποθήκευση και διαχείριση οπτικοακουστικού περιεχομένου. Οι βάσεις δεδομένων οπτικοακουστικού περιεχομένου αποτελούν μέρος της υποδομής υποστήριξης των οπτικοακουστικών εφαρμογών και παρέχουν μεθόδους ανάκτησης της οπτικοακουστικής πληροφορίας. Για την κατασκευή τους αναπτύσσεται ένα λογικό μοντέλο δεδομένων που στηρίζεται στο μοντέλο οντοτήτων-συσχετίσεων. Το μοντέλο αυτό υποστηρίζει την αποθήκευση της αναπαράστασης του περιεχομένου από τη γλώσσα AVDL στη βάση δεδομένων, αλλά και τη δημιουργία περιγραφών με ανάκτησή τους από τη βάση. Τέλος παρουσιάζεται παράδειγμα χρήσης της βάσης δεδομένων για την αποθήκευση της περιγραφής του περιεχομένου.

5.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο παρουσιάσαμε τη γλώσσα περιγραφής του οπτικοακουστικού περιεχομένου καθώς και το γενικό πλαίσιο χρήσης της περιγραφής σε συστήματα υποστήριξης εφαρμογών. Τα συστήματα αυτά μπορεί να αποτελούν εξυπηρετητές μετάδοσης πολυμέσων σε πραγματικό χρόνο, συστήματα εκπομπής (ή ευρείας μετάδοσης), ψηφιακές βιβλιοθήκες αρχειακού υλικού, συστήματα ανάκτησης πληροφορίας (information retrieval systems) ή βάσεις δεδομένων πολυμέσων. Τα περισσότερα από αυτά τα συστήματα επεξεργάζονται μεγάλο όγκο δεδομένων και εξυπηρετούν πολλούς πελάτες, ιδιαίτερα σε εφαρμογές του Παγκόσμιου Ιστού. Συγκεκριμένα, όσον αφορά στην

περιγραφή του οπτικοακουστικού υλικού, η οποία χρησιμοποιείται για την υποστήριξη αυξημένης λειτουργικότητας των εφαρμογών, διακρίνουμε τα εξής στοιχεία:

- Το οπτικοακουστικό υλικό έχει το χαρακτηριστικό ότι είναι μεγάλο σε όγκο, επομένως είναι δυνατόν να υπάρχει μεγάλο πλήθος περιγραφών ακόμα και για το υλικό μίας εφαρμογής. Ο μεγάλος όγκος των περιγραφών είναι απαραίτητο να επεξεργάζεται από το σύστημα με τον καλύτερο δυνατό τρόπο, ώστε να εξασφαλίζεται η αποδοτική λειτουργία των εφαρμογών.
- Η πλειονότητα των εφαρμογών οπτικοακουστικού περιεχομένου που ανήκουν στην τεχνολογία αιχμή είναι αρκετά πολύπλοκες. Η πολυπλοκότητα των εφαρμογών επιδρά άμεσα στην πολυπλοκότητα των περιγραφών του περιεχομένου, εφόσον οι τελευταίες από τη μια αναπαριστούν πολύπλοκο περιεχόμενο και από την άλλη έχουν στόχο να διευκολύνουν την υποστήριξη πολύπλοκων λειτουργιών.
- Η επεξεργασία και χρήση των περιγραφών είναι αναγκαίο να γίνεται με μεγάλη ταχύτητα, έτσι ώστε να μην επηρεάζεται αρνητικά η απόκριση του συστήματος, ειδικά στις εφαρμογές μετάδοσης σε πραγματικά χρόνο.
- Υπάρχει ανάγκη επεξεργασίας των περιγραφών και συγκεκριμένα ανάγκη πραγματοποίησης πολύπλοκων ερωτήσεων για το περιεχόμενό τους. Οι ερωτήσεις είναι αναγκαίες για την υλοποίηση και των τριών στόχων της λειτουργικότητας των εφαρμογών, δηλαδή αλληλεπίδραση, εξατομίκευση και επιλογή.
- Υπάρχει ανάγκη υποστήριξης πολλαπλών εφαρμογών που είναι δυνατόν να χρησιμοποιούν το ίδιο περιεχόμενο. Στην συγκεκριμένη περίπτωση, η περιγραφή του περιεχομένου πρέπει να είναι διαθέσιμη σε όλες τις εφαρμογές με τρόπο που να εξασφαλίζει την ανεξάρτητη υλοποίησή τους, αλλά και την προσαρμογή του περιεχομένου στις ανάγκες κάθε εφαρμογής.

Οι παραπάνω ανάγκες καλύπτονται γενικά από συστήματα διαχείρισης βάσεων δεδομένων, τα οποία παρέχουν αποδοτική αποθήκευση μεγάλου όγκου πληροφορίας και γρήγορη προσπέλαση σε αυτήν και υποστηρίζουν την υλοποίηση ερωτήσεων και την προσπέλαση από πολλούς χρήστες ταυτόχρονα. Επομένως για την αποδοτική χρήση των περιγραφών του οπτικοακουστικού περιεχομένου είναι αναγκαίο να κατασκευαστεί σύστημα διαχείρισης βάσεων δεδομένων που θα αντιμετωπίζει τα παραπάνω προβλήματα. Για την κατασκευή οποιουδήποτε συστήματος βάσης δεδομένων είναι απαραίτητο να καθοριστεί καταρχήν ένα λογικό μοντέλο για τα δεδομένα. Το μοντέλο αυτό αποτελεί τον τρόπο με τον οποίο θα οργανωθεί η πληροφορία στη βάση, δηλαδή τη δομή της, αλλά και τον τρόπο με τον οποίο θα προσπελούνται τα δεδομένα μετά την αποθήκευσή τους. Στην

παρούσα διατριβή αναπτύχθηκε μοντέλο βάσης δεδομένων για την αποθήκευση των περιγραφών του οπτικοακουστικού περιεχομένου. Οι στόχοι του μοντέλου της βάσης δεδομένων είναι:

- Να αποτελέσει μοντέλο αποθήκευσης των περιγραφών του περιεχομένου, δηλαδή να διευκολύνει την αποθήκευση των περιγραφών που αναπαρίστανται με τη Γλώσσα Περιγραφής του MPEG-4 (AVDL) σε ένα σύστημα διαχείρισης βάσεων δεδομένων.
- Να υποστηρίζει όλη τη σημασιολογία και τις ιδιότητες της περιγραφής του οπτικοακουστικού περιεχομένου, δηλαδή τη δομή των στοιχείων περιγραφής, τα χαρακτηριστικά τους, αλλά και τις σχέσεις μεταξύ τους, ώστε αν είναι ισοδύναμο με την ίδια την περιγραφή.
- Να υποστηρίζει την αντίστροφη διαδικασία της αποθήκευσης, δηλαδή την ανακατασκευή της περιγραφής από δεδομένα που είναι αποθηκευμένα στη βάση. Αυτός ο στόχος εξασφαλίζει την ανεξαρτησία του συστήματος από τη βάση δεδομένων κάνοντας διαφανή την αποθήκευση των περιγραφών. Έτσι η λειτουργικότητα που υποστηρίζει τις εφαρμογές υλοποιείται με βάση την περιγραφή και όχι τη μορφή αποθήκευσής της σε οποιοδήποτε σύστημα και με οποιοδήποτε τρόπο.
- Να υποστηρίζει ερωτήσεις ανάκτησης των δεδομένων της περιγραφής αλλά και του ίδιου του περιεχομένου. Ο στόχος αυτός απαιτεί προσεκτικό σχεδιασμό του μοντέλου της βάσης με γνώμονα την υποστήριξη των σημαντικών ερωτήσεων αναζήτησης και ανάκτησης.
- Να είναι κατάλληλο για υλοποίηση σε διαφορετικές βάσεις δεδομένων, δηλαδή σχεσιακές αλλά και αντικειμενοστρεφείς και να παρέχει ευελιξία στην επιλογή του συστήματος υλοποίησης.
- Να μπορεί να αποτελέσει βάση δεδομένων οπτικοακουστικού περιεχομένου, δηλαδή να μπορεί να αποθηκεύσει περιεχόμενο και όχι μόνο περιγραφές, υποστηρίζοντας τη συσχέτισή του με τα αποθηκευμένα δεδομένα της περιγραφής.
- Να είναι ευέλικτο και επεκτάσιμο ώστε να μπορεί να υποστηρίξει εμπλουτισμό του περιεχομένου, δηλαδή των περιγραφών.

Ο σχεδιασμός και η παρουσίαση του μοντέλου δεδομένων για την περιγραφή του οπτικοακουστικού περιεχομένου στηρίζεται στο μοντέλο οντοτήτων-συσχετίσεων, το οποίο είναι πολύ διαδεδομένο στο πεδίο των σχεσιακών βάσεων δεδομένων, αλλά ταυτόχρονα υποστηρίζει όλες τις έννοιες του αντικειμενοστρεφούς μοντέλου σχετικά με τη λογική αναπαράσταση δεδομένων. Επομένως, το μοντέλο οντοτήτων-συσχετίσεων που

αναπτύχθηκε στα πλαίσια της διατριβής μπορεί να χρησιμοποιηθεί σε σχεσιακά, αντικειμενοστρεφή αλλά και στα ενδιάμεσα αντικειμενο-σχεσιακά (object-relational) συστήματα διαχείρισης βάσεων δεδομένων.

Στο κεφάλαιο αυτό παρουσιάζεται αναλυτικά το μοντέλο οντοτήτων-συσχετίσεων για τη βάση δεδομένων των περιγραφών. Στην επόμενη ενότητα γίνεται συνοπτική παρουσίαση του μοντέλου οντοτήτων-συσχετίσεων. Στην ενότητα 5.2.1 προτείνονται κάποιες επεκτάσεις στην παρουσίαση του μοντέλου οντοτήτων-συσχετίσεων. Στις ενότητες 5.3 και 5.4 παρουσιάζεται αναλυτικά η χρήση της βάσης δεδομένων στο σύστημα, ενώ στην ενότητα 5.5 γίνεται λεπτομερής παρουσίαση του μοντέλου δεδομένων για τη βάση.

5.2 Το μοντέλο οντοτήτων-συσχετίσεων

Τα βασικά στοιχεία του μοντέλου οντοτήτων-συσχετίσεων είναι τα *σύνολα οντοτήτων* (entity sets) και τα *σύνολα συσχετίσεων* (relationship sets). Τα σύνολα οντοτήτων παριστάνουν οντότητες του ίδιου τύπου οι οποίες έχουν τις ίδιες ιδιότητες. Οι ιδιότητες αυτές είναι είτε συσχετίσεις με άλλες οντότητες, είτε *χαρακτηριστικά* (attributes). Τα χαρακτηριστικά των οντοτήτων ανήκουν στους παρακάτω τύπους:

- *Απλά ή σύνθετα*, όπου τα σύνθετα αποτελούνται από άλλα χαρακτηριστικά.
- *Μονής τιμής ή πολλαπλών τιμών* (multi-valued attributes). Τα χαρακτηριστικά πολλαπλών τιμών επιτρέπουν την ύπαρξη πολλών διαφορετικών τιμών για το ίδιο χαρακτηριστικό μιας οντότητας. Τα όρια του πλήθους των διαφορετικών τιμών είναι δυνατόν να καθοριστούν από το μοντέλο.
- *Εξαγόμενα χαρακτηριστικά* (extracted attributes) από άλλα της ίδιας οντότητας ή άλλων οντοτήτων.

Οι συσχετίσεις συνδέουν οντότητες και ένα σύνολο συσχετίσεων αποτελείται από συσχετίσεις του ίδιου τύπου. Μερικές ιδιότητες των συσχετίσεων είναι οι εξής:

- Οι συσχετίσεις είναι δυνατόν να αποτελούνται από χαρακτηριστικά.
- Ένα σύνολο οντοτήτων είναι δυνατόν να συμμετέχει σε ένα σύνολο συσχετίσεων περισσότερες από μία φορές, δημιουργώντας μια *αναδρομική συσχέτιση* (recursive relationship) στην οποία πρέπει να καθοριστεί ο ρόλος των οντοτήτων σε κάθε περίπτωση.

- Οι συσχετίσεις είναι *δυναδικές* (binary), τριαδικές, ή *ν-αδικές* (n-ary), δηλαδή ανάμεσα σε δύο, τρία ή ν σύνολα οντοτήτων. Οι ν-αδικές συσχετίσεις ισοδυναμούν με ένα νέο σύνολο οντοτήτων που συσχετίζεται με ξεχωριστές δυναδικές συσχετίσεις με κάθε ένα από τα αρχικά σύνολα οντοτήτων.
- Οι συσχετίσεις μεταξύ οντοτήτων είναι δυνατόν να περιέχουν και τον *πληθάρημο αντιστοίχισης* (cardinality) ανάμεσα στα σύνολα οντοτήτων. Για μια δυναδική συσχέτιση ο πληθάρημος αντιστοίχισης των οντοτήτων είναι δυνατόν να είναι μία προς μία (συμβολίζεται 1:1), μία προς πολλές (1:N), πολλές προς μία (N:1) και πολλές με πολλές (M:N).
- Η συμμετοχή μιας οντότητας σε ένα σύνολο συσχετίσεων μπορεί να είναι *μερική* (partial) ή *ολική* (total participation), δηλαδή να συμμετέχουν στη συσχέτιση μερικές ή όλες οι οντότητες του συνόλου, αντίστοιχα. Στη δεύτερη περίπτωση είναι δυνατόν να υπάρχει *εξάρτηση ύπαρξης* (existence dependency) ανάμεσα στην οντότητα που συμμετέχει ολικά στη συσχέτιση και στην άλλη οντότητα. Η εξάρτηση ύπαρξης είναι δυνατή όταν το ένα σύνολο οντοτήτων είναι *αδύναμο* (weak entity set), δηλαδή οι οντότητές του δεν είναι δυνατόν να υπάρχουν ανεξάρτητα από μία οντότητα του *ισχυρού* (strong) συνόλου οντοτήτων. Επομένως, η εξάρτηση ύπαρξης προϋποθέτει αδυναμία αυτόνομης ύπαρξης (ή διάκρισης) ενός συνόλου οντοτήτων και συσχέτιση 1:N ανάμεσα στο ισχυρό και το αδύναμο σύνολο.

Τα σύνολα οντοτήτων αλλά και τα σύνολα συσχετίσεων είναι δυνατόν να περιέχουν *κλειδιά* (keys), τα οποία διακρίνουν τις οντότητες ή τις συσχετίσεις μεταξύ τους και τις ταυτοποιούν με μοναδικό τρόπο. Τα κλειδιά είναι σύνολα χαρακτηριστικών και στη βάση δεδομένων επιλέγεται ένα κλειδί το οποίο ονομάζεται *κύριο κλειδί* (primary key).

Εκτός από τα παραπάνω στοιχεία του μοντέλου χρησιμοποιούνται στην παρούσα και δημοφιλείς επεκτάσεις του όπως οι συσχετίσεις εξειδίκευσης-γενίκευσης. Η *εξειδίκευση* (specialisation) χρησιμοποιείται όταν ένα υποσύνολο κάποιου συνόλου οντοτήτων έχει ένα κοινό σύνολο χαρακτηριστικών που δεν έχουν οι υπόλοιπες οντότητες, ή συμμετέχει σε συσχετίσεις που δεν συμμετέχουν οι υπόλοιπες οντότητες του συνόλου. Η *γενίκευση* (generalisation) είναι το αντίστροφο της εξειδίκευσης, όπου γίνεται σύνθεση κοινών χαρακτηριστικών από δύο ή περισσότερα σύνολα οντοτήτων σε ένα νέο γενικευμένο σύνολο. Η έννοια της εξειδίκευσης είναι ότι τα χαρακτηριστικά και οι συσχετίσεις του γενικού συνόλου οντοτήτων κληρονομούνται στις οντότητες των εξειδικευμένων συνόλων. Η γενίκευση μπορεί να προέρχεται από περιορισμούς ή συνθήκες που αφορούν στην τιμή κάποιου χαρακτηριστικού (οντότητες με την ίδια τιμή σε κάποιο χαρακτηριστικό ανήκουν στο ίδιο σύνολο οντοτήτων) ή να καθορίζονται ελεύθερα από το σχεδιαστή της βάσης.

Ένα άλλο χαρακτηριστικό της εξειδίκευσης είναι αν επιτρέπει σε μία οντότητα να συμμετέχει ή όχι σε περισσότερα από ένα εξειδικευμένα σύνολα οντοτήτων, οπότε τα σύνολα αυτά επικαλύπτονται ή είναι ξένα μεταξύ τους. Η γενίκευση χαρακτηρίζεται επίσης από την ιδιότητα της πληρότητας, όπου υπάρχουν δύο διαφορετικές λύσεις σχεδιασμού, δηλαδή οι γενικευμένες οντότητες να ανήκουν οπωσδήποτε και σε κάποιο εξειδικευμένο σύνολο ή όχι.

Μία επίσης επέκταση του μοντέλου οντοτήτων-συσχετίσεων είναι η *συσσώρευση* (aggregation) ενός υποσυνόλου του μοντέλου σε ένα σύνολο οντοτήτων το οποίο επιτρέπει τη συνολική συσχέτισή του με άλλα σύνολα οντοτήτων.

Ο πιο δημοφιλής τρόπος παρουσίασης του μοντέλου οντοτήτων-συσχετίσεων, ο οποίος χρησιμοποιείται και στην παρούσα αναφορά είναι τα *διαγράμματα οντοτήτων-συσχετίσεων* (entity-relationship diagram):

- Τα σύνολα οντοτήτων παριστάνονται ως ορθογώνια, τα οποία συνδέονται με χαρακτηριστικά που παριστάνονται με ελλείψεις.
- Τα χαρακτηριστικά πολλαπλών τιμών παριστάνονται από ελλείψεις με διπλό περίγραμμα, ενώ τα εξαγόμενα χαρακτηριστικά από ελλείψεις με διακεκομμένο περίγραμμα. Επίσης τα χαρακτηριστικά που αποτελούν το κύριο κλειδί ενός συνόλου υπογραμμίζονται.
- Τα σύνολα συσχετίσεων αναπαρίστανται με ρόμβους που συνδέονται με ευθείες γραμμές με ένα ή περισσότερα σύνολα οντοτήτων και είναι δυνατόν να έχουν χαρακτηριστικά που παριστάνονται με τον ίδιο τρόπο με τα χαρακτηριστικά των οντοτήτων.
- Ο πληθάριθμος αντιστοίχισης αναγράφεται πάνω στις ευθείες σύνδεσης των οντοτήτων με τη συσχέτιση και κοντά στα ορθογώνια που παριστάνουν τα σύνολα οντοτήτων.
- Η κατεύθυνση μιας συσχέτισης υπονοείται από τον πληθάριθμό της. Συγκεκριμένα, μια συσχέτιση 1:N σημαίνει ότι η κατεύθυνση προσπέλασής της είναι από το σύνολο οντοτήτων που συμμετέχει με πληθάριθμο αντιστοίχισης N προς το σύνολο οντοτήτων που συμμετέχει με πληθάριθμο 1. Στα διαγράμματα οντοτήτων-συσχετίσεων η κατεύθυνση δηλώνεται με ένα βέλος στο άκρο της αντιστοίχισης με πληθάριθμο 1.
- Οι συσχετίσεις πλήρους συμμετοχής συμβολίζονται με διπλή γραμμή σύνδεσης ανάμεσα στη συσχέτιση και την οντότητα που συμμετέχει πλήρως.

- Η εξάρτηση ύπαρξης συμβολίζεται με διπλό περίγραμμα γύρω από το αδύναμο σύνολο οντοτήτων και γύρω από τη συσχέτιση που καθορίζει την εξάρτηση, επιπλέον του συμβολισμού ολικής συμμετοχής του αδύναμου συνόλου. Επίσης η ύπαρξη εξάρτησης σημαίνει ότι το αδύναμο σύνολο οντοτήτων δεν έχει δικό του κύριο κλειδί, αλλά είναι δυνατόν να έχει ένα *διακριτικό* χαρακτηριστικό, το οποίο υπογραμμίζεται με διακεκομμένη γραμμή.
- Η γενίκευση και η εξειδίκευση συμβολίζονται με ένα τρίγωνο που περιέχει τη λέξη "isa" (είναι) και συνδέει τα γενικευμένα με τα εξειδικευμένα σύνολα οντοτήτων.
- Η συσσώρευση τμήματος του μοντέλου σε ένα σύνολο οντοτήτων συμβολίζεται διαγραμματικά με ένα ορθογώνιο γύρω από το αντίστοιχο τμήμα του διαγράμματος. Το ορθογώνιο αυτό μπορεί να συμμετέχει σε συσχετίσεις με άλλα σύνολα.

5.2.1 Επεκτάσεις του μοντέλου οντοτήτων-συσχετίσεων

Στην παρούσα ενότητα παρουσιάζονται κάποιες τροποποιήσεις ή επεκτάσεις στη διαγραμματική αναπαράσταση του μοντέλου οντοτήτων συσχετίσεων, οι οποίες αποσαφηνίζουν τις μοντελοποιημένες έννοιες και προσθέτουν στο μοντέλο επιπλέον σημασιολογία.

Όταν ο πληθάριθμος μιας συσχέτισης υπόκειται σε συγκεκριμένους περιορισμούς, τότε οι περιορισμοί αυτοί αναγράφονται δίπλα στα άκρα της συσχέτισης. Συγκεκριμένα, ένας ακέραιος k στη θέση του πληθάρθμου σημαίνει ότι οι οντότητες του άλλου συνόλου (δηλαδή στο άλλο άκρο) της συσχέτισης συσχετίζονται με ακριβώς k οντότητες αυτού του συνόλου. Κεφαλαία γράμματα της αλφαβήτας όπως M , N δηλώνουν τυχαιούς ακέραιους. Επίσης είναι δυνατόν να καθορίζεται διάστημα τιμών για τον πληθάρθμο, π.χ. $N \geq 5$ όταν μία οντότητα συσχετίζεται με τουλάχιστον 5 οντότητες του άλλου συνόλου. Ο πληθάρθμος μιας συσχέτισης δεν είναι δυνατόν να έχει την τιμή 0 ακριβώς, γιατί σύμφωνα με τη σημασιολογία του μοντέλου, θα ίσχυε ότι οι οντότητες που συμμετέχουν στη συσχέτιση συσχετίζονται με 0 οντότητες του άλλου συνόλου, δηλαδή στην ουσία δεν υπάρχει συσχέτιση. Επιτρέπεται όμως το διάστημα του πληθάρθμου να είναι κλειστό στο 0. Για παράδειγμα, πληθάρθμος $N \geq 0$ σημαίνει ότι η συσχέτιση είναι μερική από το άλλο σύνολο που συμμετέχει στη συσχέτιση.

Για την κατεύθυνση της συσχέτισης χρησιμοποιείται και το βέλος, ανεξάρτητα από τον πληθάρθμο της συσχέτισης. Μια παρατήρηση εδώ είναι ότι το βέλος της κατεύθυνσης στη βιβλιογραφία δείχνει τη φορά προσπέλασης της συσχέτισης όπως αυτή μετατρέπεται στο

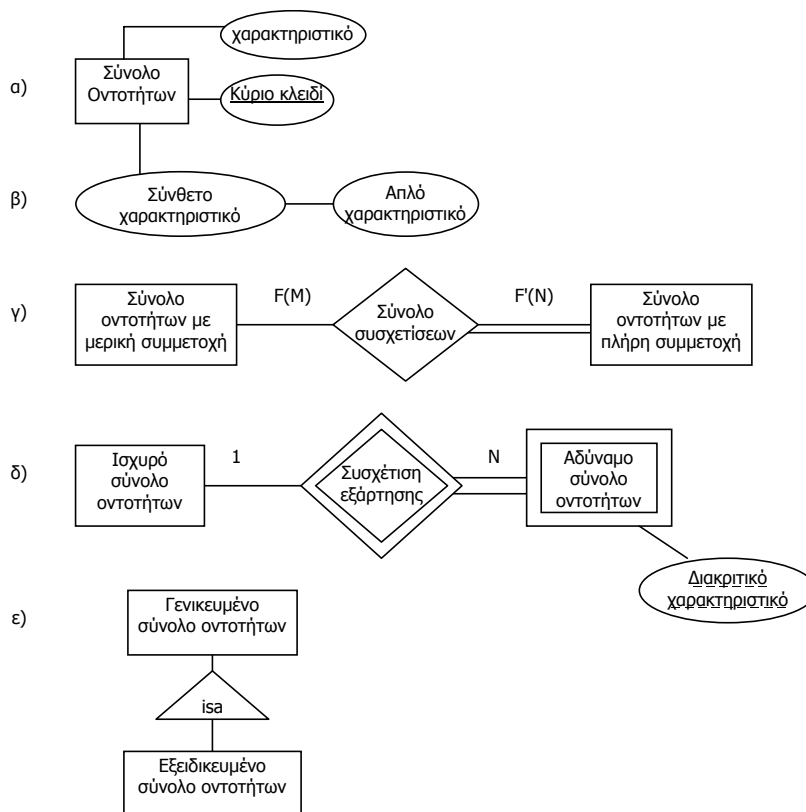
σχεσιακό μοντέλο, όπου ο κωδικός της οντότητας στο 1 γίνεται χαρακτηριστικό των οντοτήτων στο N. Αν θέλουμε να υλοποιήσουμε τη σχέση αντίστροφα ή και από τις δύο κατευθύνσεις το οποίο αντικειμενοστραφές μοντέλο και υποστηρίζεται από την ODL με τη σύνταξη "inverse" ([33]).

Η πληρότητα εξειδίκευσης, δηλαδή ότι οι γενικευμένες οντότητες πρέπει να συμμετέχουν και σε κάποιο σύνολο εξειδίκευσης συμβολίζεται με διπλή γραμμή ανάμεσα στο γενικευμένο σύνολο οντοτήτων και τη σχέση "isa" που το συνδέει με τα εξειδικευμένα σύνολα οντοτήτων.

Υποστηρίζονται αφηρημένοι τύποι οντοτήτων. Οι τύποι αυτοί χρησιμεύουν για την ομαδοποίηση εννοιών του μοντέλου, αλλά δεν υλοποιούνται στη βάση δεδομένων. Συνήθως αποτελούν οντότητες υψηλού επιπέδου σε ιεραρχίες γενίκευσης. Συμβολίζονται με διακεκομμένα περιγράμματα γύρω από τα σύνολα οντοτήτων αλλά και γύρω από τις συσχετίσεις τους.

Comment [Γ.Κ.18]: Στη θεωρία συνόλων δεν έχει νόημα να μιλάμε για αντίστροφη σχέση (ή αντιστοίχιση) γιατί $(\alpha, \beta) = (\beta, \alpha)$ εξ' ορισμού. Στην υλοποίηση έχει σημασία. Π.χ. για μια 1:1 συσχέτιση, ποια είναι η κατεύθυνση; στο σχεσιακό έχει νόημα να πούμε ότι το id του ενός θα μπει στο άλλο ή και τα δύο. Για μια 1:N, στο σχεσιακό υλοποιείται σε μία κατεύθυνση, δηλαδή το id του 1 μπαίνει στο N. Γίνεται το id του N να μπει στο 1; να με multi-valued attribute ή με ξεχωριστό entity (όπως στη M:N). Η M:N όπως υλοποιείται δεν έχει κατεύθυνση, δηλαδή δημιουργείται νέο entity. Θα μπορούσε να υπάρχει ένα multi-valued attribute σε κάθε σύνολο. Μπορεί να είναι θέμα απόδοσης. Δηλαδή αν η M:N έχει πολλά στοιχεία, να μην έχει κατεύθυνση, αλλιώς να έχει. Το ίδιο ισχύει και στο αντικειμενοστραφές.

Comment [Γ.Κ.19]: Πως συμβολίζονται οι περιορισμοί στην "isa" σχέση, δηλ. η γενίκευση με βάση κάποια συνθήκη, η επικάλυψη των εξειδικευμένων συνόλων;



Σχήμα 5-1. Στοιχεία του διαγράμματος οντοτήτων-συσχετίσεων

α) Σύνολο οντοτήτων με κύριο κλειδί και άλλο ένα χαρακτηριστικό. β) Σύνθετο χαρακτηριστικό που αποτελείται από ένα απλό χαρακτηριστικό. γ) Σύνολο συσχετίσεων

ανάμεσα σε δύο σύνολα οντοτήτων. Το αριστερό σύνολο οντοτήτων έχει μερική συμμετοχή στη συσχέτιση, δηλαδή οι οντότητες του συνόλου είναι να μη συσχετίζονται με οντότητες του άλλου συνόλου. Το δεξί σύνολο οντοτήτων συμμετέχει πλήρως στη συσχέτιση. δ) Συσχέτιση εξάρτησης ανάμεσα σε ένα ισχυρό και ένα αδύναμο σύνολο οντοτήτων. ε) Συσχέτιση "είναι" (isa) ανάμεσα σε ένα γενικευμένο και ένα εξειδικευμένο σύνολο οντοτήτων.

Το τρίγωνο που συμβολίζει τη σχέση "isa" έχει την κορυφή στραμμένη προς το γενικευμένο (υψηλού επιπέδου) σύνολο οντοτήτων, εφόσον σημαίνει ότι μια εξειδικευμένη οντότητα χαμηλού επιπέδου "είναι" γενικευμένη οντότητα υψηλού επιπέδου.

Η προκαθορισμένη περίπτωση όσον αφορά στην επικάλυψη των εξειδικευμένων συνόλων οντοτήτων είναι ότι είναι ξένα μεταξύ τους, δηλαδή μια οντότητα δεν είναι δυνατόν να ανήκει σε περισσότερα από ένα σύνολα χαμηλού επιπέδου στην ιεραρχία εξειδίκευσης.

Στο προηγούμενο σχήμα φαίνονται τα βασικότερα στοιχεία της διαγραμματικής αναπαράστασης του μοντέλου οντοτήτων συσχετίσεων.

Στην ενότητα που ακολουθεί περιγράφεται η αποθήκευση των περιγραφών του περιεχομένου στη βάση δεδομένων.

5.3 Αποθήκευση περιγραφών σε βάση δεδομένων

Η αποθήκευση των περιγραφών στη βάση δεδομένων γίνεται σύμφωνα με το λογικό μοντέλο που καθορίζει το είδος, τη δομή και τις συσχετίσεις των δεδομένων. Για την κατασκευή του μοντέλου οντοτήτων-συσχετίσεων της περιγραφής καθορίζονται καταρχήν κάποιες γενικές και ενιαίες αρχές σχεδιασμού, οι οποίες εφαρμόζονται σε οποιοδήποτε σύνολο περιγραφών και οδηγούν στην ανάπτυξη ενός μοντέλου αντιστοίχισης της περιγραφής με το λογικό μοντέλο της βάσης δεδομένων. Το μοντέλο αντιστοίχισης της περιγραφής με το λογικό μοντέλο δεδομένων εξυπηρετεί στην καταγραφή μεταπληροφορίας για την αποθήκευση των περιγραφών στη βάση. Δεν περιέχει τα αποθηκευμένα στοιχεία της γλώσσας AVDL, αλλά ένα σύνολο οντοτήτων που είναι απαραίτητες για την αποθήκευση μιας συγκεκριμένης περιγραφής στη βάση δεδομένων. Το μοντέλο αυτό εκφράζεται επίσης με το μοντέλο οντοτήτων-συσχετίσεων και αποθηκεύεται επίσης στη βάση δεδομένων, όπως περιγράφεται στην ενότητα που ακολουθεί. Στην ενότητα 5.3.2 περιγράφεται η διαδικασία αποθήκευσης της περιγραφής στη βάση δεδομένων με τη χρήση του μοντέλου αντιστοίχισης, ενώ το πλήρες μοντέλο δεδομένων για τη βάση MPEG-4 παρουσιάζεται στην ενότητα 5.5 .

5.3.1 Αντιστοίχιση της περιγραφής στο μοντέλο οντοτήτων-συσχετίσεων

Στόχος αυτής της ενότητας είναι να καθορίσει ένα σύνολο από γενικές αρχές σχεδιασμού του λογικού μοντέλου δεδομένων που υποστηρίζει την αποθήκευση των περιγραφών σε μια βάση, δηλαδή τους κανόνες αντιστοίχισης της περιγραφής στο μοντέλο οντοτήτων-συσχετίσεων. Οι κανόνες αυτοί επίσης αποθηκεύονται στη βάση δεδομένων, δηλαδή εκφράζονται και οι ίδιοι με το μοντέλο οντοτήτων-συσχετίσεων. Έτσι καταγράφεται η διαδικασία αντιστοίχισης και μπορεί να χρησιμοποιηθεί κατά την αντίστροφη φορά, δηλαδή από τη βάση δεδομένων μπορεί να γίνει ανακατασκευή της περιγραφής.

Τα στοιχεία της περιγραφής αντιστοιχίζονται σε οντότητες της βάσης δεδομένων. Η αντιστοίχιση αυτή προκύπτει από το γεγονός ότι στοιχεία και οντότητες αποτελούν τις βασικές έννοιες σε έγγραφα της γλώσσας AVDL και σε βάσεις δεδομένων αντίστοιχα. Επιπλέον παρουσιάζουν ομοιότητα όσον αφορά στο ότι περιέχουν ένα πεπερασμένο σύνολο από χαρακτηριστικά συγκεκριμένων τύπων. Κατά συνέπεια τα χαρακτηριστικά των στοιχείων αντιστοιχίζονται σε χαρακτηριστικά των οντοτήτων της βάσης. Εφόσον τα στοιχεία περιγραφής που έχουν τα ίδια χαρακτηριστικά και την ίδια δομή ομαδοποιούνται σε τύπους στοιχείων (ή ορίζονται μέσα από τύπους στοιχείων) και οι οντότητες της βάσης δεδομένων ομαδοποιούνται σε σύνολα οντοτήτων, υπάρχει αντιστοίχιση ανάμεσα σε τύπους στοιχείων και σύνολα οντοτήτων.

Η δομή των στοιχείων περιγραφής μπορεί να είναι πιο πολύπλοκη από την επιτρεπτή δομή οντοτήτων που υποστηρίζει το μοντέλο οντοτήτων-συσχετίσεων. Ένα στοιχείο μπορεί να είναι σύνθετο, δηλαδή να περιέχει άλλα στοιχεία, ή να περιέχει επιλογή άλλων στοιχείων. Στο μοντέλο οντοτήτων-συσχετίσεων τέτοιου είδους περιπτώσεις μοντελοποιούνται με συσχετίσεις ανάμεσα στις οντότητες της βάσης. Οι συσχετίσεις αυτές ομαδοποιούνται σε σύνολα συσχετίσεων που συνδέουν σύνολα οντοτήτων. Μια πιο λεπτομερής ανάλυση της περιγραφής μπορεί να οδηγήσει σε πιο σαφή διάκριση της σημασιολογίας στη δομή των στοιχείων. Έτσι διακρίνονται οι εξής περιπτώσεις:

- Ένα στοιχείο της γλώσσας AVDL μπορεί να μην έχει πολύπλοκη δομή, δηλαδή να μην περιέχει άλλα στοιχεία, αλλά να αποτελείται μόνο από χαρακτηριστικά.
- Ένα στοιχείο της γλώσσας μπορεί να αποτελείται από άλλα στοιχεία, τα οποία με τη σειρά τους να έχουν πολύπλοκη ή απλή δομή.
- Τα στοιχεία που περιέχονται μέσα σε ένα στοιχείο είναι δυνατόν ή να χρησιμοποιούνται ως περιεχόμενο και άλλων τύπων στοιχείων, είτε να έχουν νόημα ύπαρξης μόνο μέσα από τον συγκεκριμένο τύπο στοιχείων.

- Τα στοιχεία είναι δυνατόν να έχουν ταυτότητα και να γίνονται αντικείμενα αναφοράς μέσω της ταυτότητάς τους, από άλλα στοιχεία.
- Η δομή στοιχείου όταν εκφράζεται ως επιλογή ανάμεσα σε ένα σύνολο άλλων στοιχείων είναι δυνατόν να αντιπροσωπεύει εξειδίκευση της σημασιολογίας του στοιχείου σε μία ή περισσότερες έννοιες.

Σύμφωνα με τα παραπάνω ένας τύπος στοιχείων που αποτελεί μέρος της δομής κάποιου άλλου τύπου στοιχείων είναι δυνατόν να μετατραπεί σε ένα από τα εξής συστατικά του μοντέλου οντοτήτων-συσχετίσεων:

- Σε ένα χαρακτηριστικό του συνόλου οντοτήτων που αντιστοιχεί στον αρχικό τύπο στοιχείου. Η περίπτωση αυτή είναι κατάλληλη όταν ο τύπος στοιχείων δεν έχει σύνθετη δομή, αλλά αποτελείται μόνο από χαρακτηριστικά ή και από απλά στοιχεία τα οποία επίσης αποτελούνται μόνο από χαρακτηριστικά. Αν το στοιχείο αποτελείται από ένα μόνο χαρακτηριστικό, τότε αντιστοιχίζεται σε απλό χαρακτηριστικό. Διαφορετικά αντιστοιχίζεται σε σύνθετο χαρακτηριστικό.
- Σε ένα σύνολο οντοτήτων και ένα σύνολο συσχετίσεων. Το νέο σύνολο οντοτήτων συσχετίζεται με το σύνολο οντοτήτων που αντιστοιχεί στον αρχικό τύπο στοιχείου, ο οποίος περιέχει τον συγκεκριμένο τύπο. Η κατεύθυνση της συσχέτισης εξαρτάται από τη σημασιολογία των συγκεκριμένων τύπων στοιχείων που συμμετέχουν σε αυτήν, οπότε καθορίζεται με ευέλικτο τρόπο. Ο πληθάρθμος της συσχέτισης από το αρχικό σύνολο οντοτήτων προς το σύνολο με το οποίο συσχετίζεται, όπως προκύπτει από την περιγραφή μπορεί να είναι 1, μεγαλύτερος ή ίσος του 0 ή μεγαλύτερος ή ίσος του 1. Ο πληθάρθμος της συσχέτισης από το περιεχόμενο σύνολο οντοτήτων προς αυτό που το περιέχει εξαρτάται επίσης από τη σημασιολογία των στοιχείων.
- Σε ένα αδύναμο σύνολο οντοτήτων, το οποίο εξαρτάται από το σύνολο οντοτήτων που αντιστοιχεί στον αρχικό τύπο στοιχείων που τον περιέχει.
- Σε εξειδικευμένο σύνολο του αρχικού συνόλου οντοτήτων αν είναι επιλογή ανάμεσα σε δύο ή περισσότερους τύπους στοιχείων και η σημασιολογία του είναι τέτοια ώστε να μπορεί να θεωρηθεί εξειδίκευση. Οι τύποι στοιχείων που δεν έχουν χαρακτηριστικά αλλά έχουν δομή η οποία αποτελεί επιλογή άλλων τύπων στοιχείων, είναι δυνατόν να αντιστοιχιστούν σε αφηρημένα σύνολα οντοτήτων που εξειδικεύονται σε ένα ή περισσότερα σύνολα οντοτήτων.

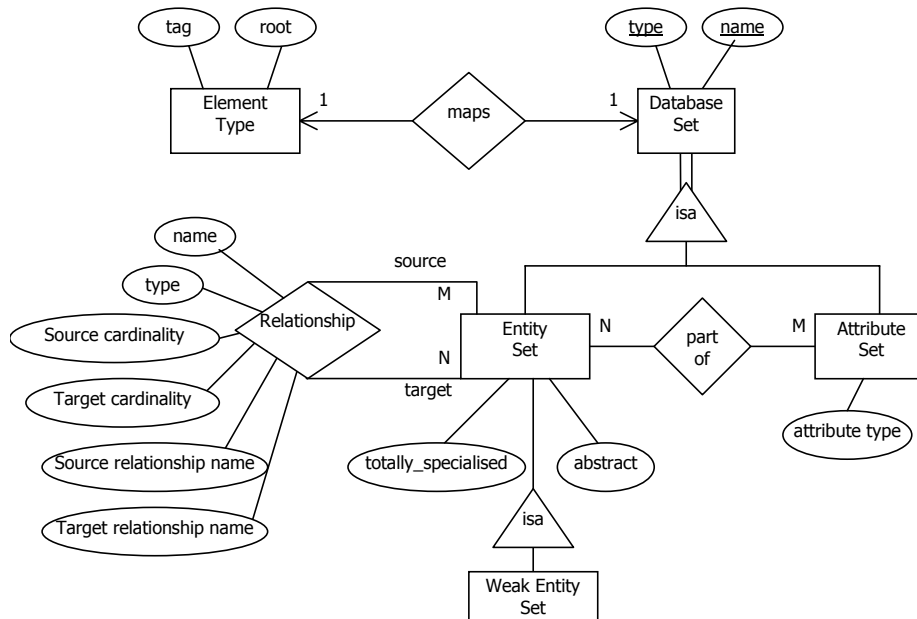
Στην επόμενη ενότητα περιγράφεται με ποιο τρόπο το μοντέλο οντοτήτων-συσχετίσεων της βάσης δεδομένων υποστηρίζει τους παραπάνω κανόνες αντιστοίχισης της περιγραφής σε οντότητες της βάσης δεδομένων.

5.3.2 Μοντέλο αποθήκευσης περιγραφών στη βάση δεδομένων

Όλοι οι τύποι στοιχείων που χρησιμοποιούνται σε μια περιγραφή, δηλαδή σε ένα έγγραφο που περιέχει στιγμιότυπα των τύπων στοιχείων της AVDL, αντιστοιχίζονται σε οντότητες της βάσης δεδομένων. Οι οντότητες αυτές ανήκουν στο σύνολο οντοτήτων `ElementType` (Τύπος Στοιχείου) και περιέχουν τη μεταπληροφορία για τον τύπο στοιχείων. Δηλαδή ένα από τα χαρακτηριστικά των οντοτήτων `ElementType` είναι το όνομα του τύπου στοιχείου της γλώσσας AVDL με τον οποίο αντιστοιχίζονται. Επιπλέον κάθε συγκεκριμένος τύπος στοιχείων της γλώσσας αντιστοιχίζεται και σε ένα ξεχωριστό σύνολο της βάσης δεδομένων το οποίο περιέχει την πληροφορία για τη δομή των στοιχείων του. Το σύνολο αυτό μοντελοποιείται στη βάση δεδομένων από την γενική οντότητα `DatabaseSet` (Σύνολο της βάσης δεδομένων), η οποία εξειδικεύεται σε δύο σύνολα, το `EntitySet` (Σύνολο Οντοτήτων) και το `AttributeSet` (Σύνολο χαρακτηριστικών). Η εξειδίκευση αυτή είναι πλήρης και υποχρεώνει όλες τις οντότητες του συνόλου `DatabaseSet` να ανήκουν και σε ένα από τα δύο εξειδικευμένα σύνολα. Το σύνολο `EntitySet` εξειδικεύεται με τη σειρά του στο σύνολο `WeakEntitySet` (Αδύναμο σύνολο οντοτήτων), ώστε να μοντελοποιηθεί η περίπτωση όπου ένας τύπος στοιχείου αντιστοιχίζεται σε ένα αδύναμο σύνολο οντοτήτων. Τα χαρακτηριστικά των οντοτήτων `DatabaseSet` που κληρονομούνται στα εξειδικευμένα σύνολα είναι ο τύπος και το όνομα. Ο τύπος είναι το χαρακτηριστικό στο οποίο στηρίζεται η εξειδίκευση των οντοτήτων του συνόλου και επομένως δέχεται τις τιμές `entity_set`, `attribute_set` και `weak_entity_set`. Επιπλέον χαρακτηριστικό του συνόλου `EntitySet` αποτελεί το `totally_specialised`, που δηλώνει αν το σύνολο εξειδικεύεται με ολική εξειδίκευση σε άλλα σύνολα. Το χαρακτηριστικό αυτό είναι σημαντικό γιατί επηρεάζει άμεσα την υλοποίηση των οντοτήτων στη βάση δεδομένων.

Το σύνολο `EntitySet` της βάσης δεδομένων συσχετίζεται με τον εαυτό του με την αναδρομική M:N συσχέτιση `Relationship` (Συσχέτιση). Η συσχέτιση αυτή μοντελοποιεί τις σχέσεις ενός τύπου στοιχείου που έχει αντιστοιχιστεί σε ένα σύνολο οντοτήτων της βάσης με άλλους τύπους στοιχείων που έχουν επίσης αποθηκευτεί στη βάση δεδομένων. Στη συσχέτιση διακρίνεται το σύνολο εκκίνησης της σχέσης (`source entity set`) και το σύνολο άφιξης ή στόχος της συσχέτισης (`target entity set`). Ο τύπος της συσχέτισης, ο οποίος μοντελοποιείται από το χαρακτηριστικό `type`, είναι ένας από τους εξής: εξειδίκευση (`specialisation`), γενίκευση (`generalisation`), εξάρτηση (`dependency`), είναι-τμήμα-του (`part-of`) και άλλη συσχέτιση (`other`). Ο πληθώραριθμός είναι M:N διότι μία συγκεκριμένη συσχέτιση μπορεί να εφαρμόζεται ανάμεσα σε διαφορετικά ζεύγη από σύνολα οντοτήτων, δηλαδή να υπάρχει μία συσχέτιση με το ίδιο όνομα και τον ίδιο τύπο ανάμεσα σε δύο ζεύγη διαφορετικών συνόλων. Στα χαρακτηριστικά της συσχέτισης περιλαμβάνονται και τα

ονόματα που χρησιμοποιούνται από την οντότητα πηγής (*source_relationship_name*) και από την οντότητα στόχος (*target_relationship_name*), όταν η συσχέτιση αυτή υλοποιηθεί ανάμεσα στα δύο σύνολα οντοτήτων. Το σύνολο οντοτήτων *AttributeSet* συσχετίζεται με το σύνολο *EntitySet* με τη συσχέτιση "είναι τμήμα του", η οποία μοντελοποιεί την αντίστοιχη συσχέτιση ανάμεσα σε οντότητες και σύνθετα χαρακτηριστικά. Στο παρακάτω σχήμα φαίνεται το μοντέλο αντιστοίχισης των τύπων στοιχείων μιας περιγραφής σε οντότητες της βάσης δεδομένων.



Σχήμα 5-2. Μοντέλο αποθήκευσης τύπων στοιχείων στη βάση δεδομένων

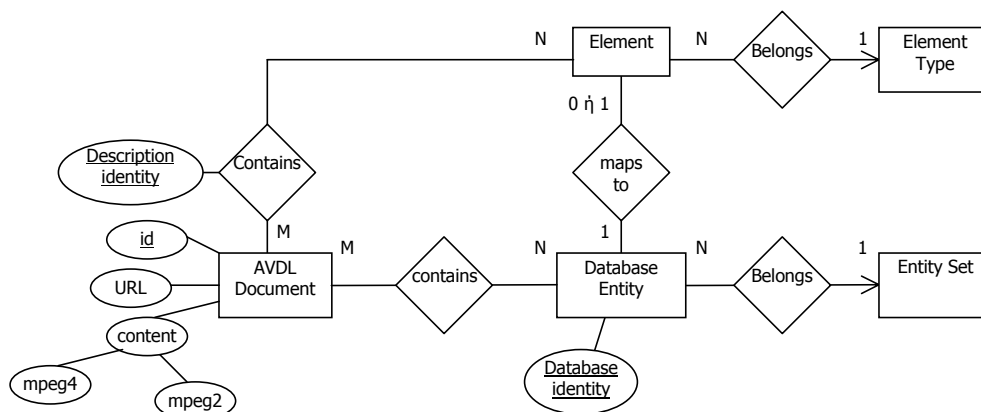
α) Το χαρακτηριστικό **root** του συνόλου **Element** είναι δίτιμο και δηλώνει αν ο τύπος στοιχείου είναι η ρίζα της περιγραφής. β) Το χαρακτηριστικό **type** των συσχετίσεων του συνόλου **Relationship** παίρνει τιμές {*isa*, *partof*, *general*}. "Source relationship name", "Target relationship name" είναι οι ετικέτες της συσχέτισης για κάθε σύνολο που συμμετέχει σε αυτήν. Τα ονόματα αυτά συνδέουν τη συσχέτιση με την υλοποίησή της. Το χαρακτηριστικό **attribute type** του συνόλου **AttributeSet** παίρνει τιμές {*simple*, *composite*}.

Επιπλέον, κάθε σύνολο οντοτήτων που δημιουργείται από την περιγραφή με βάση το παραπάνω μοντέλο, αντιστοιχίζεται και σε ένα σύνολο οντοτήτων που αποτελεί εξειδίκευση του συνόλου **EntitySet**. Το σύνολο αυτό έχει συγκεκριμένα χαρακτηριστικά και συγκεκριμένες συσχετίσεις με άλλα σύνολα, που σχηματίζουν ένα εκτεταμένο μοντέλο στη

βάση δεδομένων. Για την περιγραφή του MPEG-4 το εκτεταμένο μοντέλο οντοτήτων-συσχετίσεων παρουσιάζεται στην ενότητα 5.5 .

Κατά αναλογία με τους τύπους στοιχείων, τα αυτόνομα στοιχεία μιας περιγραφής μοντελοποιούνται ως οντότητες του συνόλου Element της βάσης δεδομένων. Αυτόνομα στοιχεία της περιγραφής θεωρούνται αυτά τα οποία έχουν τη δική τους ταυτότητα περιγραφής. Επομένως ένα από τα χαρακτηριστικά των οντοτήτων του συνόλου Element είναι η ταυτότητα περιγραφής element_id. Κατά κανόνα οι ταυτότητες, δηλαδή τα κλειδιά των αντικειμένων της βάσης δε συμπίπτουν με τις ταυτότητες περιγραφής. Μερικά στοιχεία περιγραφής μπορεί να μην έχουν ταυτότητα περιγραφής. Όταν όμως αποθηκεύονται στη βάση δεδομένων έχουν ταυτότητα-κλειδί. Επίσης ανάμεσα στις οντότητες του συνόλου ElementType και αυτές του συνόλου Element είναι προφανές ότι υπάρχει η συσχέτιση Belongs, δηλαδή ότι ένα στοιχείο ανήκει σε έναν τύπο στοιχείου.

Τα στοιχεία, δηλαδή οι οντότητες του συνόλου Element αντιστοιχίζονται σε αντικείμενα της βάσης δεδομένων που έχουν ταυτότητα στη βάση, όπως φαίνεται στο σχήμα από τη συσχέτιση "maps to" (αντιστοιχίζεται) ανάμεσα στο σύνολο Element και στο σύνολο DatabaseObject.



Σχήμα 5-3. Μοντέλο αποθήκευσης στοιχείων στη βάση δεδομένων

Ένα έγγραφο της γλώσσας AVDL περιέχει στοιχεία περιεχομένου MPEG-4 ή MPEG-2. Καθένα από αυτά τα στοιχεία ανήκει στο σύνολο οντοτήτων Element και αντιστοιχίζεται σε μια οντότητα της βάσης δεδομένων. Τα στοιχεία ανήκουν σε έναν τύπο στοιχείων ενώ οι οντότητες της βάσης ανήκουν σε ένα σύνολο οντοτήτων.

Τα έγγραφα της γλώσσας AVDL είναι αυτά που περιέχουν όλα τα στοιχεία περιγραφών και μοντελοποιούνται στη βάση δεδομένων με το σύνολο οντοτήτων AVDLDocument. Η

πληροφορία που αποθηκεύεται στη βάση δεδομένων για τα έγγραφα είναι το URL τους και το είδος του περιεχομένου τους που μπορεί να είναι συνδυασμός οπτικών και ακουστικών αντικειμένων, σκηνών και περιγραφών αντικειμένων, όπως φαίνεται από το χαρακτηριστικό `content` της οντότητας `AVDLDocument`. Τα έγγραφα συσχετίζονται με (περιέχουν) τα αυτόνομα στοιχεία περιγραφής, αλλά και με τις οντότητες της βάσης που δημιουργούνται κατά την αποθήκευση του εγγράφου. Ενώ μια οντότητα της βάσης μπορεί να μην αντιστοιχίζεται με ένα αυτόνομο στοιχείο περιγραφής, όλες οι οντότητες που δημιουργούνται από ένα έγγραφο συσχετίζονται με αυτό.

Για παράδειγμα, θεωρούμε την περίπτωση της περιγραφής ενός οπτικού αντικειμένου που είναι πρόσωπο. Η περιγραφή του αντικειμένου αποτελείται από ένα στιγμιότυπο του τύπου στοιχείου `VisualObjectSequence` το οποίο σύμφωνα με τη γλώσσα `AVDL` θα περιέχει ένα στιγμιότυπο του τύπου στοιχείου `FaceObject`. Η αντιστοίχιση της περιγραφής του στοιχείου αυτού στο μοντέλο οντοτήτων-συσχετίσεων της βάσης δεδομένων, περιλαμβάνει τη δημιουργία των εξής δύο οντοτήτων που ανήκουν στο σύνολο `EntitySet` για την αντιστοίχιση των τύπων στοιχείων: $\{es1, entity_set, FaceObject\}$ και $\{es2, entity_set, VisualObjectSequence\}$. Επίσης δημιουργείται η παρακάτω συσχέτιση που ανήκει στο σύνολο `Relationship`: $\{ \{es1, entity_set, Face\}, \{es2, entity_set, VisualObjectSequence\}, part-of, M, N, sequence_face, face_is_in_sequence, sequence_has_face \}$. Αυτό σημαίνει ότι θα σχηματιστεί στο μοντέλο μία συσχέτιση με όνομα "sequence_face" ανάμεσα στα σύνολα `Face` και `VisualObjectSequence`. Όταν η συσχέτιση προσπελαίνεται από το σύνολο `Face` ονομάζεται `face_is_in_sequence` ενώ όταν προσπελαίνεται από το σύνολο `VisualObjectSequence` ονομάζεται `sequence_has_face`.

Με βάση τα παραπάνω κατασκευάζουμε έναν αλγόριθμο μετατροπής ενός στοιχείου περιγραφής σε οντότητες της βάσης δεδομένων, ο οποίος αποτελεί μέρος των εργαλείων της υποδομής υποστήριξης οπτικοακουστικών υπηρεσιών. Τα δεδομένα του αλγορίθμου είναι η αναπαράσταση του στοιχείου στη γλώσσα `AVDL` και ο τύπος `t` του στοιχείου. Το αποτέλεσμα του αλγορίθμου είναι ένα σύνολο οντοτήτων της βάσης δεδομένων. Η λειτουργία του αλγορίθμου ακολουθεί την από πάνω προς τα κάτω συντακτική ανάλυση των εγγράφων της γλώσσας `AVDL` και διατηρεί μια προσωρινή στοίβα οντοτήτων, η οποία χρησιμοποιείται για τη μετατροπή της ιεραρχίας των στοιχείων. Για κάθε στοιχείο της γλώσσας, ο αλγόριθμος κρατάει πληροφορία για την οντότητα που αντιστοιχεί στο στοιχείο-πατέρα του. Η πληροφορία αυτή είναι ο τύπος `parent_type` του πατέρα και η οντότητα `P` που αντιστοιχεί στον πατέρα. Στην περίπτωση του στοιχείου ρίζα η τιμή `parent_type` και η οντότητα `P` είναι κενές.

Τα βήματα του αλγορίθμου έχουν ως εξής:

1. Βρες το επόμενο στοιχείο της περιγραφής με τύπο t.
2. Με βάση τον τύπο t του στοιχείου, βρες την οντότητα ET του συνόλου Element Type της βάσης δεδομένων που αντιστοιχεί σε αυτόν τον τύπο στοιχείου.
3. Βρες από τη συσχέτιση "maps to" ανάμεσα στα σύνολα Element Type και Database Set που υπάρχει στη βάση δεδομένων, την οντότητα DS του συνόλου Database Set στην οποία αντιστοιχίζεται η οντότητα ET.
4. Αν ο τύπος (χαρακτηριστικό type) της οντότητας DS που βρέθηκε στο βήμα 2 είναι "entity_set", βρες την οντότητα ES του συνόλου Entity Set στην οποία εξειδικεύεται η οντότητα DS και κάνε τα εξής:
 - 4.1. Αν το χαρακτηριστικό abstract της ES έχει την τιμή TRUE, τότε
 - 4.1.1. δε δημιουργείται καμία οντότητα στη βάση.
 - 4.1.2. Βάλε στη στοίβα με τους πατέρες μια κενή οντότητα.
 - 4.1.3. Πήγαινε στο βήμα 4.5.
 - 4.2. αλλιώς
 - 4.2.1. δημιούργησε μια οντότητα E του συνόλου που έχει όνομα την τιμή του χαρακτηριστικού name της οντότητας ES.
 - 4.2.2. Βρες τα χαρακτηριστικά του στοιχείου και ενημέρωσε τα χαρακτηριστικά της οντότητας E.
- 4.3. Αν η οντότητα P (πατέρας) δεν είναι κενή, δηλαδή δεν πρόκειται για το στοιχείο ρίζα της περιγραφής, κάνε τα εξής βήματα:
 - 4.3.1. Βρες τη συσχέτιση R ανάμεσα στις οντότητες P και DS από το σύνολο συσχετίσεων Relationship.
 - 4.3.2. Με βάση τα χαρακτηριστικά της R, ενημέρωσε τη σχέση ανάμεσα στις οντότητες P και E.
- 4.4. Βάλε την οντότητα E στη στοίβα οντοτήτων με τους πατέρες.
- 4.5. Αν το στοιχείο της περιγραφής, περιέχει και άλλα στοιχεία επανέλαβε το βήμα 1 για καθένα από αυτά,
- 4.6. αλλιώς
 - 4.6.1. βγάλε από τη στοίβα την οντότητα που βρίσκεται στην κορυφή.
 - 4.6.2. Πήγαινε στο βήμα 1.

5. Αν ο τύπος της οντότητας είναι `attribute_set`:

5.1. Βρες τη οντότητα AS του συνόλου Attribute Set η οποία έχει τον τύπο στοιχείου `t`.

5.2. Ενημέρωσε το χαρακτηριστικό με όνομα `AS.name` της οντότητας `P` με τις τιμές των χαρακτηριστικών του στοιχείου.

Ο παραπάνω αλγόριθμος χρησιμοποιεί τη μεταπληροφορία της βάσης για την δημιουργία οντοτήτων της βάσης από τα στοιχεία της περιγραφής και αποτελεί μέρος των εργαλείων της υποδομής υποστήριξης οπτικοακουστικών υπηρεσιών, ώστε η αποθήκευση του περιεχομένου να γίνεται με αυτόματο τρόπο στη βάση δεδομένων. Επίσης ο αλγόριθμος είναι πλήρης, με την έννοια ότι υποστηρίζει την αποθήκευση στοιχείων αλλά και ολόκληρων εγγράφων της γλώσσας AVDL στη βάση δεδομένων.

5.4 Ανάκτηση περιγραφών από τη βάση δεδομένων

Στην προηγούμενη ενότητα παρουσιάστηκε το μοντέλο και ο αλγόριθμος αντιστοίχισης μιας περιγραφής στη γλώσσα AVDL στο μοντέλο οντοτήτων-συσχετίσεων για την αποθήκευσή της σε βάση δεδομένων. Σε πολλές περιπτώσεις υπάρχει ανάγκη να γίνεται αντιστοίχιση ή μετατροπή και κατά την αντίθετη φορά, δηλαδή από τη βάση δεδομένων σε περιγραφή. Η λειτουργία αυτή είναι χρήσιμη όταν γίνεται ανταλλαγή περιεχομένου όπου η μορφή του περιεχομένου είναι η γλώσσα AVDL.

Επομένως ένας από τους στόχους του μοντέλου της βάσης δεδομένων είναι να διευκολύνει αυτήν τη μετατροπή. Η διαδικασία αυτή όπως είναι φυσικό εξαρτάται από το είδος της βάσης δεδομένων στο οποίο έχει γίνει αποθήκευση των περιγραφών, εφόσον αναφέρεται στην επεξεργασία αποθηκευμένων οντοτήτων. Στη συγκεκριμένη ενότητα παρουσιάζεται η κατασκευή περιγραφών ως προς το μοντέλο οντοτήτων-συσχετίσεων.

Η διαδικασία μετατροπής ενός αποθηκευμένου αντικειμένου σε αντικείμενο της γλώσσας AVDL αποτελείται από έναν αλγόριθμο. Τα δεδομένα του αλγορίθμου είναι μια οντότητα `E` της βάσης δεδομένων. Το αποτέλεσμα είναι μια περιγραφή στη γλώσσα AVDL, η οποία περιέχει μια ιεραρχία στοιχείων που ξεκινάει με την συγκεκριμένη οντότητα. Τα βήματα του αλγορίθμου είναι τα εξής:

1. Έστω `s` το σύνολο οντοτήτων στο οποίο ανήκει η αποθηκευμένη οντότητα `E`.
2. Βρες την οντότητα `DS` που περιέχει μεταπληροφορία γι' αυτό το σύνολο.

3. Βρες την οντότητα ET του συνόλου Element Type με την οποία αντιστοιχίζεται η οντότητα DS. Η πληροφορία αυτή υπάρχει στη συσχέτιση maps του συνόλου οντοτήτων Database Set με το σύνολο οντοτήτων Element Type. Κάνε τα εξής:

3.1. Κατασκεύασε ένα στοιχείο περιγραφής με ετικέτα (δηλ. τύπο στοιχείου) την τιμή του χαρακτηριστικού tag της ET.

3.2. Βρες την ταυτότητα περιγραφής του στοιχείου, αν υπάρχει ως εξής:

3.2.1. Βρες την οντότητα L που συνδέεται με την E από το σύνολο οντοτήτων Element.

3.2.2. Πρόσθεσε την ταυτότητα περιγραφής στο στοιχείο που κατασκευάστηκε.

Comment [Γ.Κ.20]: δεν είναι σωστό γιατί η ταυτότητα στο σχήμα ορίζεται μόνο μέσω ενός εγγράφου που περιέχει το στοιχείο.

4. Βρες τα χαρακτηριστικά της οντότητας E που αντιστοιχίζονται σε τύπους στοιχείων. Η αναζήτηση γίνεται μέσω της συσχέτισης ανάμεσα σε EntitySet και AttributeSet και κατόπιν για κάθε χαρακτηριστικό ανακτάται ο τύπος στοιχείου στον οποίο αντιστοιχίζεται. Για καθένα από αυτά επανέλαβε το βήμα 1.

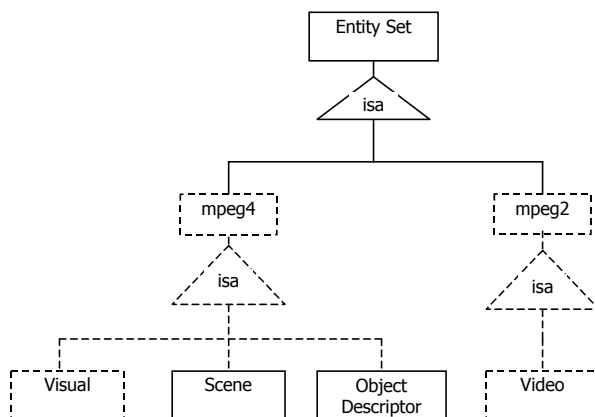
5. Βρες τις συσχετίσεις της E με άλλες οντότητες. Οι συσχετίσεις μιας οντότητας είναι επίσης δυνατόν να αντιπροσωπεύουν στοιχεία που περιέχονται μέσα στο στοιχείο της οντότητας. Όλες οι συσχετίσεις που αφορούν στο σύνολο της οντότητας και έχουν τύπο "part-of" (χαρακτηριστικό Relationship.type), αντιστοιχίζονται στην εισαγωγή ενός επιπλέον στοιχείου μέσα στο αρχικό. Οι συσχετίσεις τύπου "is-a" από την πλευρά τους υπονοούν δομή επιλογής στο περιεχόμενο του στοιχείου.

Η παραπάνω αλγοριθμική διαδικασία επαναλαμβάνεται κάθε φορά που πρέπει να δημιουργηθεί ένα νέο στοιχείο της περιγραφής από τη βάση δεδομένων.

5.5 Μοντέλο οντοτήτων-συσχετίσεων της AVDL

Σε αυτήν την ενότητα παρουσιάζεται το τμήμα του λογικού μοντέλου που αναπαριστά τα στοιχεία της γλώσσας περιγραφής AVDL σε συστήματα διαχείρισης βάσεων δεδομένων. Η μοντελοποίηση έχει αφετηρία το στοιχείο ρίζα της γλώσσας, το οποίο περιέχει ένα από τους 4 βασικούς τύπους οντοτήτων: οπτικά αντικείμενα, ακουστικά αντικείμενα, περιγραφείς αντικειμένων και σκηνές. Το στοιχείο ρίζα μοντελοποιείται με ένα αφηρημένο σύνολο οντοτήτων, γιατί οι οντότητες του συνόλου δεν έχουν χαρακτηριστικά και δεν υπάρχουν στην περιγραφή αυτόνομες χωρίς εξειδικευμένο περιεχόμενο. Το σύνολο οντοτήτων Visual είναι επίσης αφηρημένο γιατί δηλώνει μόνο το διαχωρισμό ανάμεσα στα

είδη οπτικών αντικειμένων, τα οποία εξειδικεύονται περαιτέρω. Η ρίζα της ιεραρχίας του μοντέλου φαίνεται στο παρακάτω σχήμα.



Σχήμα 5-4. Ιεραρχία οπτικοακουστικών αντικειμένων MPEG-4 και MPEG-2 στον μοντέλο βάσης δεδομένων

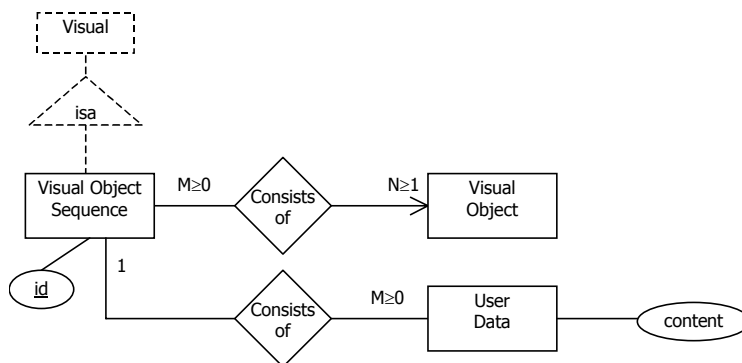
Η ρίζα των στοιχείων της γλώσσας AVDL μοντελοποιείται με τους αφηρημένους τύπους οντοτήτων mpeg4 και mpeg2, οι οποίοι εξειδικεύονται σε 4 σύνολα οντοτήτων. Το σύνολο mpeg4 και mpeg2 αποτελούν εξειδίκευση του συνόλου EntitySet του μοντέλου μεταπληροφορίας για την αντιστοίχιση των περιγραφών στη βάση δεδομένων.

Η επόμενη ενότητα παρουσιάζει το λογικό μοντέλο βάσης δεδομένων για τα οπτικά αντικείμενα του MPEG-4. Ακολουθεί η παρουσίαση του μοντέλου βάσης δεδομένων των οπτικοακουστικών σκηνών και τέλος παρουσιάζεται το μοντέλο οντοτήτων-συσχετίσεων για τις σειρές βίντεο του MPEG-2.

5.5.1 Μοντέλο οπτικών αντικειμένων του MPEG-4

Οι οντότητες οπτικού τύπου περιέχονται σε μία σειρά οπτικών αντικειμένων μαζί με οντότητες του τύπου UserData. Η σειρά οπτικών αντικειμένων έχει χαρακτηριστικά την ταυτότητά της που είναι και το κύριο κλειδί της. Τα οπτικά αντικείμενα μπορεί να υπάρχουν και αυτόνομα στη βάση, επομένως δεν έχουν σχέση εξάρτησης από τη σειρά οπτικών αντικειμένων. Η τελευταία τα ομαδοποιεί για μετάδοση σε μία ροή (που δεν έχει και πολλή σχέση με τη βάση). Εφόσον το σύνολο οντοτήτων Visual είναι αφηρημένο συνδέεται με μια αφηρημένη συσχέτιση "είναι" με το VisualObjectSequence. Η κατεύθυνση της συσχέτισης είναι από τη σειρά προς τα οπτικά αντικείμενα που περιέχει. Δεν είναι συχνή ερώτηση το να γνωρίζουμε την αντίστροφη συσχέτιση, δηλαδή σε ποιες

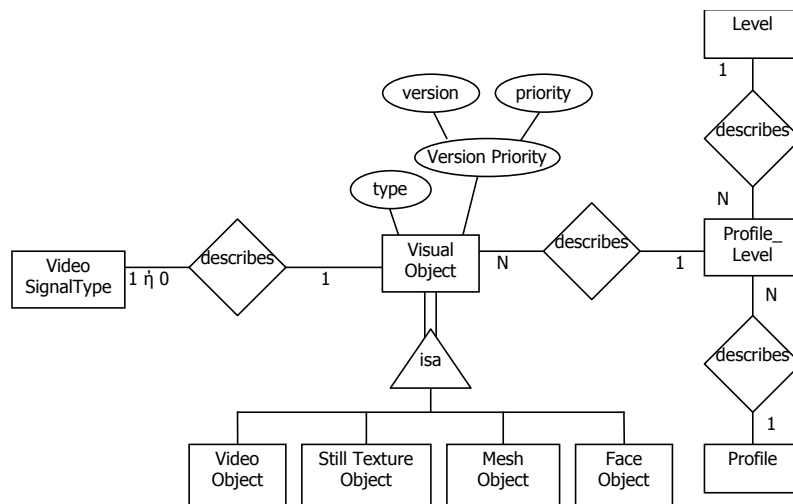
σειρές συμμετέχει ένα οπτικό αντικείμενο. Επίσης από την άποψη της υλοποίησης είναι συμφέρον η πληροφορία να φυλάσσεται μαζί με τη σειρά.



Σχήμα 5-5. Σειρά οπτικών αντικειμένων

Η σειρά οπτικών αντικειμένων του MPEG-4 αποτελείται από ένα ή περισσότερα οπτικά αντικείμενα και ένα σύνολο από οντότητες που περιέχουν δεδομένα του χρήστη τυχαίου περιεχομένου. Η συσχέτιση ανάμεσα στα σύνολα Visual Object Sequence και Visual Object είναι M:N ώστε να επιτρέπει στα οπτικά αντικείμενα να περιέχονται σε περισσότερες από μία σειρές.

Όλα τα οπτικά αντικείμενα έχουν κοινά το προφίλ, την έκδοση και την προτεραιότητά τους, καθώς και δεδομένα του χρήστη (UserData). Το προφίλ και το επίπεδο έχουν N:1 συσχέτιση με το οπτικό αντικείμενο και αποτελούνται από προκαθορισμένες τιμές.



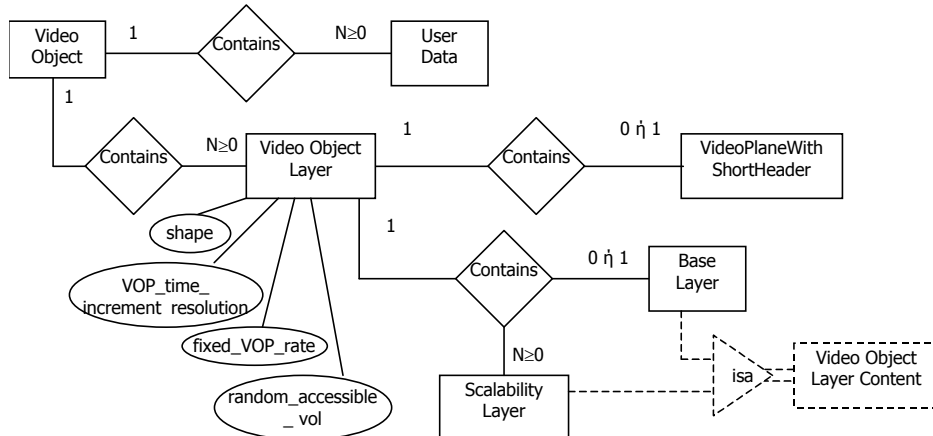
Σχήμα 5-6. Οπτικά αντικείμενα

Τα οπτικά αντικείμενα του MPEG-4 μοντελοποιούνται στη βάση δεδομένων από μια οντότητα που εξειδικεύεται στα τέσσερα είδη οπτικών αντικειμένων με μια συσχέτιση ολικής εξειδίκευσης. Κοινά για όλα τα είδη οπτικών αντικειμένων είναι τα χαρακτηριστικά

έκδοση, προτεραιότητα, προφίλ και επίπεδο που μοντελοποιούνται με σύνθετα χαρακτηριστικά.

Η εξειδίκευση των διαφόρων τύπων οπτικών αντικειμένων γίνεται με βάση το χαρακτηριστικό τύπος και είναι ολική, δηλαδή όλες οι οντότητες του γενικευμένου συνόλου πρέπει να ανήκουν σε κάποιο από τα εξειδικευμένα σύνολα, τα οποία κληρονομούν και όλες τις ιδιότητες. Επίσης τα οπτικά αντικείμενα τύπου βίντεο και ακίνητης συνδέονται οπωσδήποτε με μία οντότητα που μοντελοποιεί τα χαρακτηριστικά του σήματος βίντεο. Επομένως το σύνολο οντοτήτων VisualObject συνδέεται με το σύνολο VideoSignalType.

Η δομή των αντικειμένων βίντεο όπως περιγράφεται στο κεφάλαιο 4 της διατριβής, αποτελείται από δεδομένα του χρήστη (User Data) και βαθμίδες αντικειμένων βίντεο (Video Object Layers). Οι συσχετίσεις αυτές έχουν πληθάρημο 1:N, όπου το N μπορεί να είναι και 0, δηλαδή ένα αντικείμενο βίντεο να μην περιέχει την περιγραφή των βαθμίδων του. Η οντότητα VideoObjectLayer έχει ένα σύνολο από χαρακτηριστικά που προκύπτουν από τα χαρακτηριστικά του αντίστοιχου στοιχείου της γλώσσας AVDL. Ένα από τα χαρακτηριστικά της είναι η ταυτότητα περιγραφής, η οποία καθιστά το στοιχείο αυτόνομο στοιχείο περιγραφής και επομένως οντότητα του συνόλου Element της βάσης. Η οντότητα αυτή επίσης συσχετίζεται με τις οντότητες VideoPlaneWithShortHeader, BaseLayer και ScalabilityLayer. Το περιεχόμενο των βαθμίδων βάσης και κλιμάκωσης έχει κοινά στοιχεία οπότε και στη βάση δεδομένων τα κοινά χαρακτηριστικά γενικεύονται από το αφηρημένο σύνολο οντοτήτων VideoObjectLayerContent.

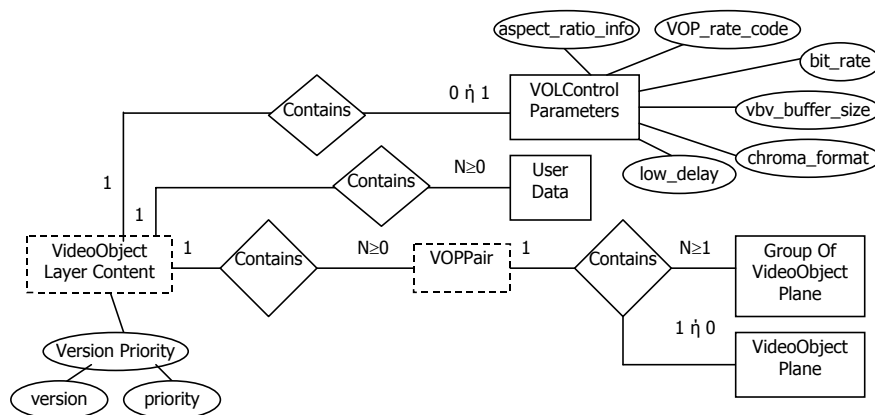


Σχήμα 5-7. Αντικείμενα βίντεο και βαθμίδες βίντεο

Η μοντελοποίηση των αντικειμένων βίντεο στη βάση δεδομένων διατηρεί τη σημασιολογία της διαβάθμισης των δεδομένων σε βαθμίδες κλιμάκωσης. Κάθε βαθμίδα

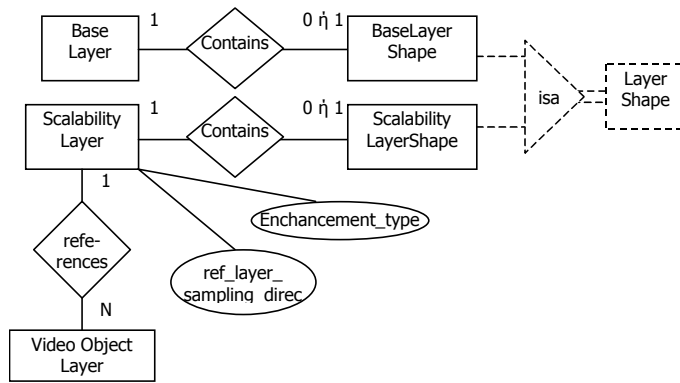
ανήκει στο σύνολο Video Object Layer και περιέχει μία βαθμίδα βάσης και μία ή περισσότερες βαθμίδες κλιμάκωσης.

Το περιεχόμενο των βαθμίδων μοντελοποιείται από το γενικευμένο αφηρημένο σύνολο οντοτήτων Video Object Layer Content όπως φαίνεται στο παραπάνω διάγραμμα οντοτήτων-συσχετίσεων. Στο παρακάτω σχήμα απεικονίζεται αυτό το σύνολο να συσχετίζεται με τα σύνολα VOLControlParameters, UserData και VOPPair και έχει το σύνθετο χαρακτηριστικό έκδοση και προτεραιότητα (Version_Priority). Οι οντότητες VOLControlParameters αποτελούνται από τα αντίστοιχα χαρακτηριστικά με αυτά της γλώσσας περιγραφής. Το αφηρημένο σύνολο VOPPair, το οποίο δεν αποτελεί στοιχείο της γλώσσας περιγραφής, ομαδοποιεί τις οντότητες GroupOfVideoObjectPlane και VideoObjectPlane που είναι δυνατόν να περιέχονται σε μια βαθμίδα βίντεο.



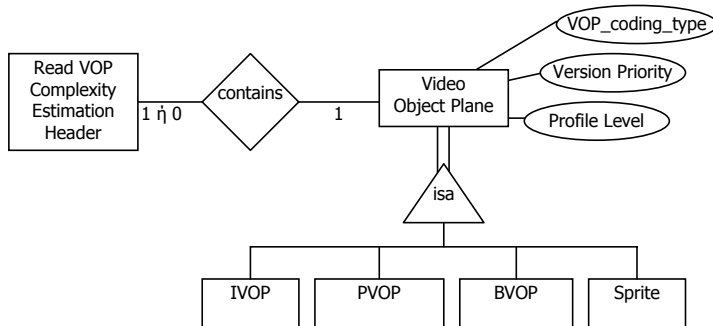
Σχήμα 5-8. Μοντέλο οντοτήτων-συσχετίσεων των βαθμίδων βίντεο

Οι οντότητες των δύο διαφορετικών ειδών βαθμίδων βίντεο συσχετίζονται με οντότητες που μοντελοποιούν το σχήμα και γενικεύονται σε ένα κοινό αφηρημένο σύνολο που ονομάζεται LayerShape, ώστε να κληρονομήσουν από αυτό τα κοινά τους συστατικά και συσχετίσεις. Επιπλέον οι οντότητες ScalabilityLayer συσχετίζονται με μια βαθμίδα αντικειμένου βίντεο από την οποία προκύπτουν με κλιμάκωση.



Σχήμα 5-9. Σχήμα των βαθμίδων βίντεο

Τα επίπεδα αντικειμένου βίντεο όπως φαίνονται στο επόμενο διάγραμμα, μοντελοποιούνται από το σύνολο VideoObjectPlane που εξειδικεύεται με ολική εξειδίκευση στα τέσσερα σύνολα που μοντελοποιούν τα βασικά είδη επιπέδων IVOP, PVOP, BVOP και Sprite.



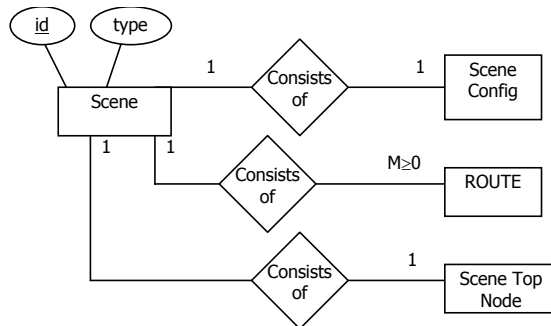
Σχήμα 5-10. Επίπεδα αντικειμένων βίντεο

Εφαρμόζοντας τις γενικές αρχές σχεδιασμού που περιγράφονται στην ενότητα 5.3.1, προκύπτει το πλήρες μοντέλο οντοτήτων-συσχετίσεων για τα υπόλοιπα στοιχεία των οπτικών αντικειμένων αλλά και για τους περιγραφείς αντικειμένων.

5.5.2 Μοντέλο βάσης δεδομένων σκηνών του MPEG-4

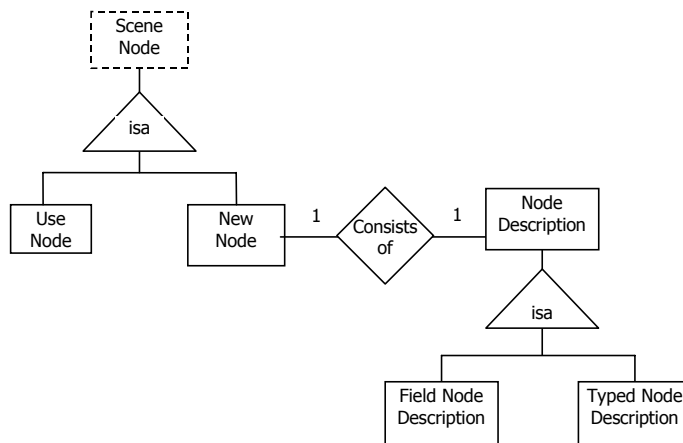
Στην ενότητα αυτή παρουσιάζεται το μοντέλο οντοτήτων-συσχετίσεων για τις σκηνές της γλώσσας AVDL. Οι οντότητες του συνόλου Scene που αποτελούν εξειδίκευση του συνόλου mpeg4 όπως φαίνεται στο Σχήμα 5-4, συσχετίζονται με την οντότητα SceneConfig που

περιέχει χαρακτηριστικά ρύθμισης του αποκωδικοποιητή της, και αποτελείται από μία ή περισσότερες οντότητες του συνόλου ROUTE και μία οντότητα του συνόλου SceneTopNode που μοντελοποιεί τον αρχικό της κόμβο στη βάση δεδομένων.



Σχήμα 5-11. Μοντέλο αναπαράστασης σκηνών

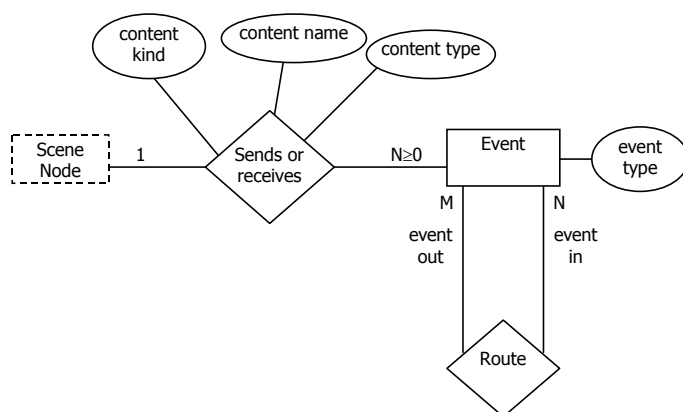
Οι κόμβοι των σκηνών μοντελοποιούνται από την αφηρημένη οντότητα SceneNode, η οποία εξειδικεύεται στα σύνολα UseNode και NewNode. Οι νέοι κόμβοι συσχετίζονται με μια οντότητα που τους περιγράφει, όπως φαίνεται στο παρακάτω διάγραμμα.



Σχήμα 5-12. Κόμβος σκηνής

Τα γεγονότα που δέχονται οι σκηνές μοντελοποιούνται από το σύνολο οντοτήτων Event, το οποίο έχει τύπο, δηλαδή εξερχόμενο ή εισερχόμενο. Οι κόμβοι συσχετίζονται με τα γεγονότα με 1:N συσχέτιση. Τα γεγονότα αφορούν στο περιεχόμενο του κόμβου αποστολέα ή δέκτη, δηλαδή σε ένα από τα χαρακτηριστικά του ή έναν από τους κόμβους

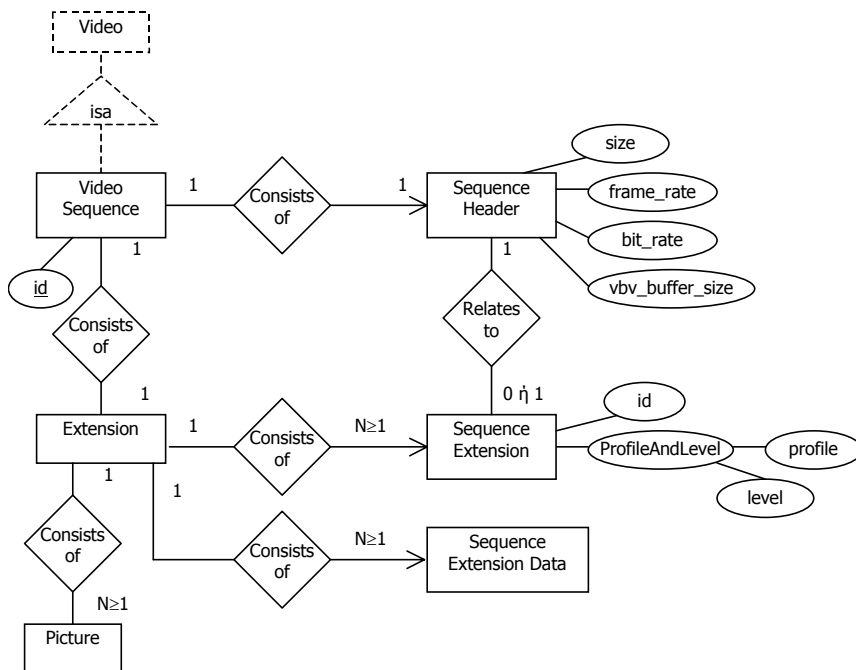
που περιέχει. Επομένως η συσχέτιση ανάμεσα σε κόμβους και γεγονότα χαρακτηρίζεται από το είδος και το όνομα του περιεχομένου (χαρακτηριστικό ή κόμβος παιδί) του κόμβου που στέλνει ή δέχεται το γεγονός καθώς και από τον τύπο αυτού του περιεχομένου. Ο τύπος είναι ένας από τους δυνατούς τύπους χαρακτηριστικών ή κόμβων και είναι χρήσιμος για την αντιστοίχιση εξερχόμενων με εισερχόμενα γεγονότα από έναν κόμβο σε έναν άλλο, εφόσον τα γεγονότα πρέπει να αφορούν οντότητες του ίδιου τύπου.



Σχήμα 5-13. Γεγονότα σκηνών

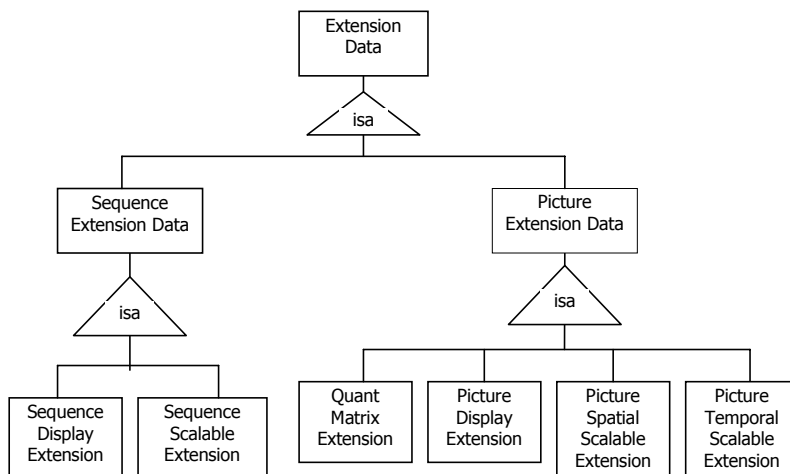
5.5.3 Μοντέλο βάσης δεδομένων MPEG-2 βίντεο

Τα δεδομένα του MPEG-2 βίντεο, όπως φαίνεται στο Σχήμα 5-4. , αποτελούν εξειδίκευση του συνόλου Entity Set του λογικού μοντέλου της βάσης δεδομένων. Το σύνολο αυτό μοντελοποιεί όλες τις οντότητες οπτικοακουστικού περιεχομένου που αποθηκεύονται στη βάση δεδομένων. Οι οντότητες του MPEG-2 εξειδικεύονται σε οντότητες που περιέχουν βίντεο και ανήκουν στο αφηρημένο σύνολο οντοτήτων Video. Αυτές εξειδικεύονται περαιτέρω σε σειρές βίντεο όπως φαίνεται στο ακόλουθο σχήμα. Οι σειρές βίντεο αποτελούνται από μια επικεφαλίδα που μοντελοποιεί δεδομένα για τα χαρακτηριστικά της σειράς όπως μέγεθος, ρυθμό εικόνων και δεδομένων, κλπ. Η σειρά βίντεο επίσης αποτελείται από μια οντότητα του συνόλου Extension, η οποία περιέχει το σύνολο των εικόνων του βίντεο αλλά και επιπλέον πληροφορία που ονομάζονται πληροφορία επέκτασης (extension) και αφορά ολόκληρες σειρές επέκτασης. Οι σειρές επέκτασης (Sequence Extension) αντιπροσωπεύουν τις βαθμίδες βάσης και κλιμάκωσης μιας σειράς βίντεο. Κάθε μία από αυτές τις σειρές έχει τη δική της επικεφαλίδα με τα χαρακτηριστικά της. Επιπλέον οι σειρές επέκτασης περιέχουν πληροφορία για το προφίλ και το επίπεδο της σειράς.



Σχήμα 5-14. Μοντέλο οντοτήτων-συσχετίσεων σειρές MPEG-2 βίντεο

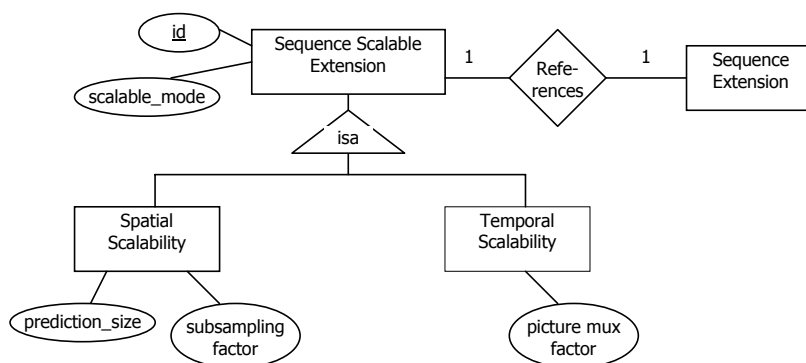
Τα δεδομένα επέκτασης όπως αυτά που αφορούν στις σειρές του προηγούμενου σχήματος, μοντελοποιούνται από μία ιεραρχία γενίκευσης η οποία περιλαμβάνει δεδομένα επέκτασης και για τις εικόνες. Η επέκταση των σειρών αφορά στην παρουσίαση του βίντεο ή στις βαθμίδες κλιμάκωσης της σειράς.



Σχήμα 5-15. Ιεραρχία εξειδίκευσης των δεδομένων της σειράς βίντεο

Η επέκταση των εικόνων περιλαμβάνει τους πίνακες κβαντοποίησης, πληροφορία για την παρουσίαση των εικόνων και πληροφορία για τη χωρική ή χρονική κλιμάκωση μιας εικόνας, όπως απεικονίζεται στο προηγούμενο διάγραμμα οντοτήτων-συσχετίσεων.

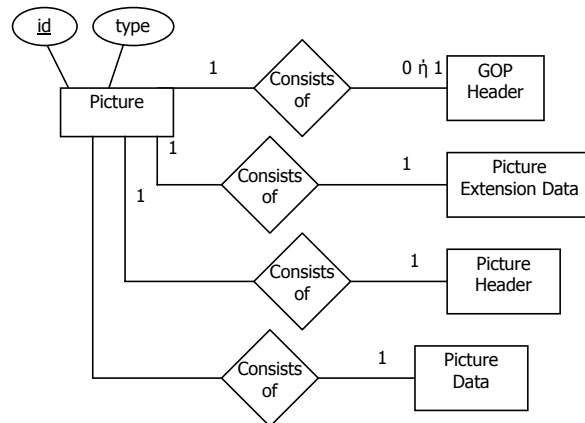
Η επέκταση των βαθμίδων κλιμάκωσης του βίντεο που μοντελοποιούνται από το σύνολο Sequence Scalable Extension του προηγούμενου διαγράμματος, διακρίνονται ανάλογα με το είδος της κλιμάκωσης σε δεδομένα χωρικής και χρονικής κλιμάκωσης. Η κατηγοριοποίηση αυτή εκφράζεται με την ιεραρχία εξειδίκευσης του διαγράμματος που ακολουθεί. Τα χαρακτηριστικά μια βαθμίδας κλιμάκωσης είναι η ταυτότητά της και το είδος της κλιμάκωσης, δηλαδή χωρική, χρονική, τμηματοποίηση δεδομένων και SNR. Κάθε βαθμίδα κλιμάκωσης σχετίζεται με μια βαθμίδα βάσης, της οποίας αποτελεί βελτίωση. Η σχέση αυτή εκφράζεται με τη συσχέτιση References ανάμεσα στα σύνολο οντοτήτων Sequence Scalable Extension και Sequence Extension και η οποία κληρονομείται στις οντότητες των συνόλων Spatial Scalability Temporal και Scalability. Οι βαθμίδες χωρικής κλιμάκωσης περιέχουν πληροφορία για το μέγεθος της βαθμίδας, το οποίο είναι μεγαλύτερο από αυτό της βαθμίδας βάσης και για τον παράγοντα υποδειγματοληψίας της βαθμίδας βάσης. Η χρονική κλιμάκωση περιγράφει τον αριθμό των εικόνων της βαθμίδας κλιμάκωσης, οι οποίες παρεμβάλλονται ανάμεσα στις εικόνες της βαθμίδας βάσης για να αυξήσουν το ρυθμό των εικόνων της βελτιωμένης σειράς. Η πληροφορία αυτή μοντελοποιείται από το χαρακτηριστικό picture_mux_factor του συνόλου Temporal Scalability.



Σχήμα 5-16. Μοντέλο βαθμίδων κλιμάκωσης MPEG-2

Μια εικόνα MPEG-2 βίντεο τελικά χαρακτηρίζεται από την ταυτότητα και τον τύπο της (I,P ή B) και αποτελείται από μια επικεφαλίδα (Picture Header), πληροφορία για την επέκταση της εικόνας και τα ίδια τα κωδικοποιημένα δεδομένα της εικόνας. Επίσης είναι δυνατόν να

συσχετίζεται με μια επικεφαλίδα ομάς εικόνων (GOP Header) στην περίπτωση που η εικόνα είναι η πρώτη μιας ομάδας εικόνων (Group Of Pictures ή GOP) μέσα στη σειρά του βίντεο. Οι ομάδες εικόνων αποτελούν ανεξάρτητες ομάδες προσπέλασης μέσα στη σειρά. Το μοντέλο οντοτήτων-συσχετίσεων για τις εικόνες MPEG-2 βίντεο απεικονίζεται στο Σχήμα 5-17 παρακάτω.



Σχήμα 5-17. Μοντέλο εικόνων MPEG-2 βίντεο

Στην ενότητα αυτή 5.5 παρουσιάστηκε η μοντελοποίηση των οντοτήτων του MPEG-4 και του MPEG-2 σε βάση δεδομένων με χρήση του μοντέλου οντοτήτων-συσχετίσεων. Η μοντελοποίηση αυτή αποτελεί το δεύτερο τμήμα της βάσης δεδομένων που μοντελοποιεί τα σύνολα των οπτικοακουστικών αντικειμένων, ενώ το πρώτο υποστηρίζει την αποθήκευση στη βάση οποιουδήποτε τύπου στοιχείου της γλώσσας περιγραφής, αλλά και αντίστροφα την ανάκτησή του από τη βάση και τη μετατροπή του σε στοιχείο της AVDL. Στην επόμενη ενότητα παρουσιάζεται ένα παράδειγμα περιγραφής αντικειμένου του MPEG-4 στη γλώσσα AVDL και η αποθήκευσή του στη βάση δεδομένων σύμφωνα με τους μηχανισμούς της ενότητας 5.3 και το λογικό μοντέλο δεδομένων της παρούσας ενότητας.

5.6 Σχισιακό μοντέλο δεδομένων

Στην ενότητα αυτή παρουσιάζεται το σχισιακό μοντέλο δεδομένων που προκύπτει από τη μετατροπή του μοντέλου οντοτήτων-συσχετίσεων της βάσης δεδομένων οπτικοακουστικού περιεχομένου. Η κατασκευή του μοντέλου γίνεται σύμφωνα με γνωστούς κανόνες μετατροπής από το μοντέλο οντοτήτων-συσχετίσεων στο σχισιακό

μοντέλο και αποτελείται από πίνακες ή σχέσεις οι οποίοι περιγράφονται με τη σχεσιακή γλώσσα βάσεων δεδομένων SQL. Κάθε σύνολο οντοτήτων μετατρέπεται σε έναν πίνακα του σχεσιακού σχήματος. Οι συσχετίσεις μεταξύ των οντοτήτων μετατρέπονται σε πίνακες, ή χαρακτηριστικά ανάλογα με τον πληθώραριθμο των συσχετίσεων. Η φορά των συσχετίσεων επίσης καθορίζει τον τρόπο μετατροπής τους στο σχεσιακό μοντέλο.

Το πρώτο μέρος του σχεσιακού σχήματος που αναπτύχθηκε αποτελείται από τους πίνακες που μοντελοποιούν τους τύπους στοιχείων της AVDL και των εγγράφων της και φαίνεται στο παρακάτω σχήμα. Ο πίνακας Element Type περιέχει την πληροφορία για τους τύπους στοιχείων της γλώσσας. Τα σύνολα οντοτήτων που δημιουργούνται στη βάση για την αποθήκευση των στοιχείων των εγγράφων μοντελοποιούνται από τον πίνακα Database Set και για τη συσχέτιση ανάμεσα σε τύπους στοιχείων και σύνολα της βάσης δημιουργείται ο πίνακας Element Type References Database Set που αποτελείται από τα κλειδιά των δύο αυτών συνόλων. Για την καλύτερη παρουσίαση των πινάκων αλλά και την ανεξάρτητη υλοποίησή τους σε σχεσιακό σύστημα βάσεων δεδομένων θεωρείται ότι υπάρχει ένα σύνολο τύπων δεδομένων που έχει οριστεί με βάση τους τύπους δεδομένων της SQL και χρησιμοποιείται για τον ορισμό των χαρακτηριστικών των πινάκων. Στο σχήμα αυτό παρουσιάζονται οι πίνακες που μοντελοποιούν τα τρία διαφορετικά είδη συνόλων της βάσης, δηλαδή σύνολα οντοτήτων, σύνολα χαρακτηριστικών και αδύναμα σύνολα οντοτήτων, με τρεις διαφορετικούς πίνακες. Τα σύνολα χαρακτηριστικών σχετίζονται με ένα σύνολο οντοτήτων στο οποίο ανήκουν μέσω του πίνακα Attribute Set Part Of Entity Set, ενώ τα αδύναμα σύνολα οντοτήτων αποτελούν εξειδίκευση των ισχυρών μέσω της ταυτότητάς τους. Η συσχέτιση εξάρτησης δύο συγκεκριμένων ισχυρών και αδύναμων συνόλων υλοποιείται με μια εγγραφή στον πίνακα Entity Set Relationship. Ο πίνακας αυτός γενικά μοντελοποιεί τις συσχετίσεις ανάμεσα σε σύνολα οντοτήτων και χρησιμοποιείται για τη μετατροπή της περιγραφής σε δεδομένα της βάσης.

```
CREATE TABLE ElementType (
    id          t_id PRIMARY KEY,
    tag         t_name,
    root        t_boolean)

CREATE TABLE DatabaseSet (
    id          int primary key,
    name        t_name,
    type        t_type check (type in ("entity_set", "attribute_set",
"weak_entity_set")))

CREATE TABLE ElementTypeReferencesDatabaseSet (
    element_type t_id FOREIGN KEY REFERENCES ElementType(id),
    database_set t_id FOREIGN KEY REFERENCES DatabaseSet(id))

CREATE TABLE EntitySet (
    id          t_id PRIMARY KEY REFERENCES DatabaseSet(id),
    abstract    t_boolean,
    totally_specialised t_boolean)
```

```
CREATE TABLE AttributeSet (
    id          t_id PRIMARY KEY REFERENCES DatabaseSet(id),
    attribute_type t_type)

CREATE TABLE AttributeSetPartOfEntitySet(
    entity_set      t_id FOREIGN KEY REFERENCES EntitySet(id),
    attribute_set    t_id FOREIGN KEY REFERENCES AttributeSet(id),
    PRIMARY KEY (entity_set, attribute_set))

CREATE TABLE WeakEntitySet (
    id          t_id PRIMARY KEY REFERENCES EntitySet(id))

CREATE TABLE EntitySetRelationship(
    source_entity_set      t_id FOREIGN KEY REFERENCES EntitySet(id),
    target_entity_set      t_id FOREIGN KEY REFERENCES EntitySet(id),
    name                   t_name,
    type                   t_type check(type in("isa", "generic",
"dependency")),
    source_cardinality     t_string,
    target_cardinality     t_string,
    source_relationship_name t_name,
    target_relationship_name t_name,
    PRIMARY KEY (source_entity_set, target_entity_set))
```

Σχήμα 5-18. Σχεσιακό μοντέλο των τύπων στοιχείων της AVDL

Τα αυτόνομα στοιχεία της περιγραφής μοντελοποιούνται με τον πίνακα Element και συσχετίζονται με εγγραφές του πίνακα Database Entity όπως φαίνεται στο ακόλουθο σχήμα. Τα έγγραφα αποθηκεύονται στον πίνακα AVDL Document και συσχετίζονται με στοιχεία και οντότητες της βάσης.

```
CREATE TABLE Element (
    id          t_id PRIMARY KEY,
    belongs_to_element_type t_id FOREIGN KEY REFERENCES
    ElementType(id))

CREATE TABLE DatabaseEntity (
    id          t_id PRIMARY KEY,
    belongs_to_entity_set t_id FOREIGN KEY REFERENCES EntitySet(id),
    maps_to_element      t_id FOREIGN KEY REFERENCES Element(id))

CREATE TABLE AVDLDocument (
    id          t_id PRIMARY KEY,
    URL         t_URL,
    content     t_string check (content in ("mpeg4.visual",
"mpeg4.scene", "mpeg4.objectdescriptor", "mpeg2.video" ))

CREATE TABLE AVDLDocumentContainsElement(
    document_id      t_id FOREIGN KEY REFERENCES AVDLDocument(id),
    element_id       t_id FOREIGN KEY REFERENCES Element(id),
    description_identity t_descr_id,
    PRIMARY KEY (document_id, element_id, description_identity) )

CREATE TABLE AVDLDocumentContainsDatabaseEntity (
    document_id      t_id FOREIGN KEY REFERENCES AVDLDocument(id),
    database_entity_id t_id FOREIGN KEY REFERENCES DatabaseEntity(id),
    PRIMARY KEY (document_id, database_entity_id) )
```

Σχήμα 5-19. Σχεσιακό μοντέλο των στοιχείων και των εγγράφων της AVDL

Ο πίνακας Visual Object Sequence που απεικονίζεται στο ακόλουθο σχεσιακό σχήμα, αποτελείται από εγγραφές που έχουν ήδη μοντελοποιηθεί ως οντότητες της βάσης στον πίνακα Database Entity.

```
CREATE TABLE VisualObjectSequence (
    id          t_id PRIMARY KEY REFERENCES DatabaseEntity (id))

CREATE TABLE VisualObject (
    id          t_id PRIMARY KEY REFERENCES DatabaseEntity(id),
    type        t_type check (type in ("video_object",
    "still_texture_object", "mesh_object", "face_object"))
    version     int,
    priority    int,
    profile     t_string,
    level       t_string)

CREATE TABLE VisualObjectSequenceConsistsOfVisualObject (
    visual_object_sequence t_id FOREIGN KEY REFERENCES
    VisualObjectSequence(id),
    visual_object t_id FOREIGN KEY REFERENCES VisualObject(id),
    PRIMARY KEY (visual_object_sequence, visual_object)

CREATE TABLE VideoSignalType (
    id          t_id PRIMARY KEY REFERENCES VisualObject(id),
    describes_visual_object t_id REFERENCES VisualObject(id))

CREATE TABLE VideoObject (
    id          t_id PRIMARY KEY REFERENCES VisualObject(id))

CREATE TABLE UserData (
    id          t_id PRIMARY KEY REFERENCES DatabaseEntity(id),
    contained_in t_id REFERENCES DatabaseEntity(id))
```

Σχήμα 5-20. Σχεσιακό σχήμα οπτικών αντικειμένων του MPEG-4

5.7 Παράδειγμα

Στην ενότητα αυτή παρουσιάζουμε ένα παράδειγμα εγγράφου της γλώσσας AVDL, το οποίο περιέχει την περιγραφή ενός οπτικού αντικειμένου που αντιπροσωπεύει μία μπάλα. Το αντικείμενο αυτό έχει κωδικοποιηθεί ως ένα φυσικό αντικείμενο βίντεο χωρίς κλιμάκωση, δηλαδή με το απλό (Simple) προφίλ και στο επίπεδο 1, δηλαδή σε ρυθμό 64kb/s. Το έγγραφο που περιέχει την περιγραφή του αντικειμένου, βρίσκεται στον Παγκόσμιο Ιστό στη διεύθυνση (URL) <http://www.music.tuc.gr/~georgia/AVDL/ball.AVDL> και το περιεχόμενό του έχει ως εξής:

```
<!DOCTYPE avdl SYSTEM "http://www.music.tuc.gr/~georgia/AVDL/avdl.dtd">
<!ENTITY    % MPEG4.description "INCLUDE">
<!ENTITY    % Visual.description "INCLUDE">
<!ENTITY    % Scene.description "INCLUDE ">
<!ENTITY    % ObjectDescriptor.description " INCLUDE ">

<avdl>
<mpeg4>
```

```

<Visual> <VisualObjectSequence>
<VisualObject type="videoID">
  <Profile_Level profile="simple" level="1"> </>
  <VersionAndPriority version="1" priority="1"></>
  <VideoSignalType>
    <VideoFormat video_format="PAL"> </>
    <VideoRange video_range="Y16-35CbCr-112+122"></>
  </VideoSignalType>
  <VideoObject id="Ball">
    <VideoObjectLayer id="Ball_VOL1" shape="rectangular"
      fixed_VOP_rate=TRUE>
      <BaseLayer>
        <BaseLayerShape>
          <VOLRectangularShape width="150" height="150"></>
        </BaseLayerShape>
        <VideoObjectPlane VOP_coding_type="I" VOP_coded="TRUE">
          <VOPCoded>
            <IVOP> <INotBinaryOnlyShape VOP_quant="10" intra_dc_vlc="1">
              <INoScalability>
                <MotionShapeTexture>
                  ...
                </MotionShapeTexture>
              </INoScalability>
            </INotBinaryOnlyShape></IVOP>
          </VOPCoded>
        </VideoObjectPlane>
      </BaseLayer>
    </VideoObjectLayer></VideoObject>
  </VisualObject>
</VisualObjectSequence></Visual>
</mpeg4>
</avdl>

```

Για την αποθήκευση του αντικειμένου στη βάση γίνεται συντακτική ανάλυση του εγγράφου. Η μετατροπή έχει αφετηρία τα στοιχεία του εγγράφου και πιο συγκεκριμένα το στοιχείο ρίζα. Παρακάτω περιγράφεται η διαδικασία της μετατροπής ως προς το μοντέλο οντοτήτων-συσχετίσεων. Υποθέτουμε ότι υπάρχει συγκεκριμένο μοντέλο υλοποίησης και κάθε βήμα αντιστοιχίζεται με συγκεκριμένο βήμα του μοντέλου υλοποίησης, είτε αυτό είναι σχεσιακό είτε αντικειμενοστραφές. Τα βήματα αποθήκευσης έχουν ως εξής:

- 5) Συντακτική ανάλυση του τύπου και του προφίλ του εγγράφου. Το βήμα αυτό συνεπάγεται δημιουργία και αποθήκευση οντότητας του συνόλου AVDLDocument με τιμή στο χαρακτηριστικό URL τη διεύθυνση του εγγράφου. Στο σύνθετο χαρακτηριστικό content, τοποθετείται η τιμή true μόνο στο visual τμήμα του, εφόσον το προφίλ του εγγράφου είναι μόνο οπτικό.
- 6) Ανάλυση του στοιχείου mpeg4. Γίνεται αναζήτηση του στοιχείου στο σύνολο ElementType της βάσης με βάση το όνομα του τύπου στοιχείου mpeg4. Το αποτέλεσμα της αναζήτησης συνδυάζεται με το σύνολο DatabaseSet μέσω της συσχέτισης "Maps to" για να καθοριστεί με ποιο και τι είδους σύνολο της βάσης έχει αντιστοιχιστεί ο συγκεκριμένος τύπος στοιχείου. Το αποτέλεσμα είναι το σύνολο

οντοτήτων mpeg4, το οποίο είναι αφηρημένο. Για τα αφηρημένα σύνολα οντοτήτων δε δημιουργούνται οντότητες αποθήκευσης στη βάση.

- 7) Η μετατροπή συνεχίζεται με το στοιχείο Visual, το οποίο επίσης αντιστοιχίζεται σε αφηρημένο σύνολο οντοτήτων και δεν αποθηκεύεται στη βάση.
- 8) Στη συνέχεια εμφανίζεται το στοιχείο VisualObjectSequence που δεν αποτελεί αφηρημένο σύνολο οντοτήτων. Το στοιχείο αυτό πρέπει να αποθηκευτεί ως οντότητα της βάσης δεδομένων και μάλιστα του συνόλου VisualObjectSequence. Η οντότητα αυτή δεν έχει άλλα χαρακτηριστικά αλλά συσχετίζεται με την οντότητα AVDLDocument που δημιουργήθηκε στο πρώτο βήμα, εφόσον είναι η πρώτη μη αφηρημένη οντότητα του εγγράφου. Δηλαδή ενημερώνεται η συσχέτιση contains ανάμεσα στο έγγραφο και την οντότητα VisualObjectSequence.
- 9) Το επόμενο στοιχείο είναι του τύπου VisualObject. Βρίσκουμε ότι πρέπει να αντιστοιχιστεί σε οντότητα του συνόλου VisualObject. Από το στοιχείο του συνόλου EntitySet που περιέχει τη μεταπληροφορία για το σύνολο VisualObject προκύπτει ότι το τελευταίο υφίσταται ολική εξειδίκευση και είναι ζήτημα της υλοποίησης του μοντέλου με ποιον τρόπο θα υποστηριχθεί η εξειδίκευση, δηλαδή αν θα δημιουργηθεί μια γενικευμένη οντότητα στη βάση, ή θα δημιουργηθεί ως τμήμα της εξειδικευμένης οντότητας η οποία δεν έχει ακόμα αναλυθεί στο έγγραφο. Η οντότητα που θα δημιουργηθεί (γενικευμένη ή εξειδικευμένη) έχει στο χαρακτηριστικό type την τιμή videoID. Επίσης πρέπει να καταγραφεί η συσχέτισή της με την προηγούμενη οντότητα που την περιέχει, οπότε αναζητούμε τη συσχέτιση ανάμεσα στα σύνολα οντοτήτων VisualObjectSequence και VisualObject. Η συσχέτιση αυτή έχει καταγραφεί ως $\{\{VisualObjectSequence\}, \{VisualObject\}, "Consists\ of", part-of, M, N\}$, όπου M και N είναι οι πληθάρημοι της συσχέτισης.
- 10) Το επόμενο στοιχείο της περιγραφής είναι το Profile_Level που περιέχει το προφίλ και το επίπεδο του αντικειμένου. Ο τύπος στοιχείου Profile_Level αντιστοιχίζεται σε σύνολο σύνθετων χαρακτηριστικών των οντοτήτων VisualObject, όπως προκύπτει από τη μεταπληροφορία για το σύνολο Profile_Level. Από την "πάνω προς τα κάτω" (top-down) συντακτική ανάλυση του εγγράφου προκύπτει ότι η οντότητα που περιέχει αυτό το χαρακτηριστικό είναι αυτή που δημιουργήθηκε στο αμέσως προηγούμενο επίπεδο, δηλαδή η οντότητα VisualObject, οπότε ενημερώνονται τα αντίστοιχα χαρακτηριστικά της.
- 11) Το προηγούμενο βήμα επαναλαμβάνεται και για το στοιχείο VersionAndPriority που είναι επίσης σύνθετο χαρακτηριστικό του συνόλου VisualObject.

- 12) Το στοιχείο `VideoSignalType` αποθηκεύεται σε μια οντότητα του συνόλου `VideoSignalType` της βάσης, με σύνθετα χαρακτηριστικά που προκύπτουν από τα στοιχεία `VideoFormat` και `VideoRange` της δομής του. Επίσης αποθηκεύεται η συσχέτισή του με το αντικείμενο `VisualObject`.
- 13) Κατόπιν αναλύεται το στοιχείο `VideoObject`, το οποίο θα αποθηκευτεί ως οντότητα του συνόλου `VideoObject`. Στο σημείο αυτό και σύμφωνα με την υλοποίηση της βάσης δεδομένων ενημερώνεται και η συσχέτιση εξειδίκευσης της οντότητας αυτής με την οντότητα `VisualObject` του βήματος 5). Ένα επιπλέον σημείο στην αποθήκευση της τρέχουσας οντότητας είναι ότι το στοιχείο της περιγραφής περιέχει και ταυτότητα περιγραφής (`id="Ball"`), δηλαδή αποτελεί αυτόνομο στοιχείο περιγραφής. Οπότε είναι απαραίτητο να δημιουργηθεί μια οντότητα στο σύνολο `Element` με αυτήν την ταυτότητα περιγραφής, καθώς και μια οντότητα στο σύνολο `DatabaseObject` με τύπο (χαρακτηριστικό `type`) `entity`, η οποία αντιστοιχίζεται με την οντότητα `Element`.

Με παρόμοιο τρόπο συμπληρώνονται τα παραπάνω βήματα για τα υπόλοιπα στοιχεία του εγγράφου. Για την αντίστροφη διαδικασία θεωρούμε ότι το ζητούμενο είναι η κατασκευή του εγγράφου <http://www.music.tuc.gr/~georgia/AVDL/ball.AVDL> από τη βάση δεδομένων. Το ζητούμενο θα μπορούσε επίσης να είναι η κατασκευή της περιγραφής του οπτικού αντικειμένου βίντεο "Ball" (η ταυτότητα του αντικειμένου βίντεο `VideoObject` όπως φαίνεται στο παραπάνω σχήμα είναι "Ball"). Οι δύο αυτές περιπτώσεις διαφέρουν στο αν θα δημιουργηθεί ένα σωστό (well-formed) AVDL έγγραφο ή μόνο ένα στοιχείο εγγράφου. Τα βήματα στην πρώτη περίπτωση έχουν ως εξής:

- 1) Ανακτάται η οντότητα `AVDLDocument` που έχει το συγκεκριμένο URL και εξετάζεται το είδος του περιεχομένου της, το οποίο προκύπτει ότι είναι μόνο οπτικό. Για την κατασκευή του εγγράφου μπορούν επομένως να παραχθούν οι δηλώσεις που περιλαμβάνουν τον τύπο και το προφίλ του.
- 2) Κατόπιν ανακτώνται όλες οι οντότητες που περιέχονται στο έγγραφο, μαζί με το σύνολο στο οποίο ανήκουν. Για κάθε μία από αυτές τις οντότητες θα κατασκευαστεί η περιγραφή της σύμφωνα με την ιεραρχία που προκύπτει από τις συσχετίσεις μεταξύ τους και σύμφωνα με την μεταπληροφορία που υπάρχει στη βάση.
- 3) Καταρχήν πρέπει να γίνει η κατασκευή του στοιχείου ρίζα του εγγράφου, το οποίο ανήκει στον τύπο στοιχείου `avdl`, όπως προκύπτει από τις οντότητες του συνόλου `ElementType`. Ο τύπος αυτός αντιστοιχίζεται στο σύνολο της βάσης δεδομένων (`DatabaseSet`) με όνομα `avdl`, το οποίο είναι αφηρημένο σύνολο οντοτήτων και ανάμεσα στις αποθηκευμένες οντότητες του εγγράφου δεν υπάρχει καμία αυτού του συνόλου. Το αποτέλεσμα είναι ότι παράγεται απλώς το tag `<avdl>` στο έγγραφο. Το

σύνολο οντοτήτων `avdl` συσχετίζεται στη βάση δεδομένων με άλλα σύνολα των οποίων τα στοιχεία αποτελούν επίσης περιεχόμενο του εγγράφου. Επειδή όμως το σύνολο `avdl` είναι αφηρημένο, η πληροφορία αυτή δεν υπάρχει σε αποθηκευμένες στοιχεία, αλλά στην μεταπληροφορία της βάσης δεδομένων, οπότε γίνεται αναζήτηση του συνόλου `DatabaseSet` που συνδέονται με το `avdl` με τη συσχέτιση `Relationship`. Η αναζήτηση καταλήγει στα σύνολα `Visual`, `Audio`, `Scene` και `ObjectDescriptor`, αλλά και πάλι διαπιστώνεται δεν υπάρχουν οντότητες αυτών των συνόλων στο έγγραφο. Τα μη αφηρημένα σύνολα από τα παραπάνω επομένως εξαιρούνται από το έγγραφο. Τα αφηρημένα σύνολα πρέπει να εξεταστούν ένα προς ένα μέχρι να βρεθεί το πρώτο μη αφηρημένο σύνολο στην ιεραρχία εξειδίκευσης. Αν υπάρχουν οντότητες αυτού του συνόλου στο έγγραφο, πρέπει να προστεθούν όλα τα `tags` της ιεραρχίας.

- 4) Για το αφηρημένο σύνολο `Visual`, ανακτώνται από συσχέτιση `Entity Set Relationship` τα ονόματα των εξειδικευμένων του συνόλων, δηλαδή το `VisualObjectSequence`. Μέσα στο έγγραφο υπάρχει οντότητα αυτού του συνόλου, οπότε παράγονται τα `tags` `Visual` και `VisualObjectSequence`. Η τελευταία οντότητα δεν περιέχει χαρακτηριστικά για την περιγραφή, αλλά συμμετέχει σε συσχετίσεις με τα σύνολα `VisualObject` και `UserData` όπως φαίνεται στο Σχήμα 5–5. Από αυτά τα δύο σύνολα υπάρχουν μέσα στο έγγραφο μόνο οντότητες του πρώτου.
- 5) Η οντότητα `VisualObject` περιέχει το απλό χαρακτηριστικό `type` το οποίο γίνεται χαρακτηριστικό του στοιχείου περιγραφής. Για τα σύνθετα χαρακτηριστικά `Version_Priority` και `Profile_Level`, δημιουργούνται δύο στοιχεία περιγραφής με τα αντίστοιχα χαρακτηριστικά. Επίσης το σύνολο οντοτήτων `VisualObject` συσχετίζεται με τα σύνολα `VideoSignalType`, `VideoObject`, `StillTextureObject`, `MeshObject` και `FaceObject`, από τα οποία περιέχονται στο έγγραφο μόνο οντότητες των δύο πρώτων. Η οντότητα `VideoSignalType` αποτελείται από δύο σύνθετα χαρακτηριστικά τα οποία μετατρέπονται σε στοιχεία της περιγραφής.

5.8 Ανακεφαλαίωση

Στο κεφάλαιο αυτό επιτεύχθηκε ο στόχος της διατριβής που αφορά στην αποθήκευση οπτικοακουστικού περιεχομένου σε βάσεις δεδομένων. Η επίτευξη αυτού του στόχου αποτελεί μέρος του γενικού στόχου της διατριβής, ο οποίος αφορά στην υποστήριξη μιας γενικής υποδομής για την διαχείριση οπτικοακουστικού περιεχομένου. Η βάση δεδομένων οπτικοακουστικού περιεχομένου αποτελεί το βασικό μέρος της υποδομής που παρέχει αποθήκευση του περιεχομένου και ανάκτησή του βάσει των ιδιοτήτων του για την

υποστήριξη των λειτουργιών ελεγχόμενης παρουσίασης, ανάκτησης βάσει περιεχομένου, διαβαθμισμένης μεταβίβασης και ανταλλαγής εγγράφων.

Η βάση δεδομένων κατασκευάζεται μέσω του λογικού μοντέλου δεδομένων, το οποίο στην παρούσα διατριβή εκφράζεται με διαγράμματα οντοτήτων-συσχετίσεων. Το μοντέλο αυτό καταρχήν επιτρέπει την αποθήκευση των εγγράφων και στοιχείων της γλώσσας AVDL με ομοιόμορφο τρόπο, μέσω ενός συνόλου μεταπληροφορίας που εκφράζει τη μετατροπή τους από στοιχεία της γλώσσας σε οντότητες της βάσης δεδομένων. Εκτός από αυτή τη μεταπληροφορία μοντελοποιούνται στη βάση και οι ίδιες οι οντότητες οπτικοακουστικού περιεχομένου, καθώς και τα χαρακτηριστικά και οι συσχετίσεις μεταξύ τους, οι οποίες μεταξύ άλλων περιλαμβάνουν και ιεραρχίες κληρονομικότητας. Για την πληρέστερη αναπαράσταση της πληροφορίας χρησιμοποιήθηκαν και ορισμένες επεκτάσεις του μοντέλου οντοτήτων συσχετίσεων.

Κεφάλαιο 6

Υποστήριξη εφαρμογών οπτικοακουστικού περιεχομένου

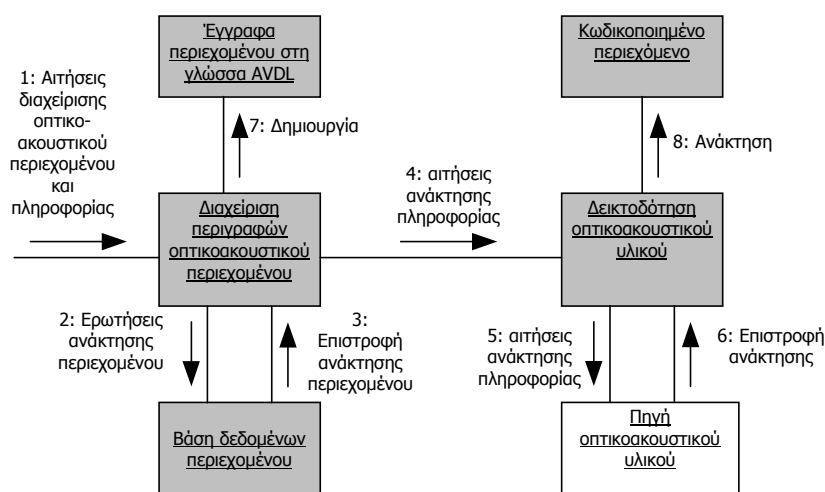
Στο κεφάλαιο αυτό περιγράφονται οι βασικές λειτουργίες της υποδομής υποστήριξης οπτικοακουστικού περιεχομένου σε πληροφοριακά συστήματα. Οι λειτουργίες αυτές περιλαμβάνουν τον έλεγχο της περιήγησης στο οπτικοακουστικό περιεχόμενο, την ανάκτηση βάσει περιεχομένου, την διαβαθμισμένη μεταβίβαση περιεχομένου ανάλογα με το προφίλ του τερματικού και το χρήστη και την ανταλλαγή εγγράφων περιεχομένου στο περιβάλλον του Παγκόσμιου Ιστού. Τέλος παρουσιάζεται ένα παράδειγμα υποστήριξης εφαρμογής εκπαίδευσης από απόσταση με χρήση οπτικοακουστικού περιεχομένου. Για την υποστήριξη της εφαρμογής χρησιμοποιούνται τα βασικά εργαλεία της γενικής υποδομής που αναπτύχθηκε στη διατριβή.

6.1 Εισαγωγή

Στο κεφάλαιο αυτό περιγράφεται η υποστήριξη εφαρμογών που έχουν την ανάγκη λειτουργικότητας διαχείρισης οπτικοακουστικού περιεχομένου. Η λειτουργικότητα των εφαρμογών αυτών υποστηρίζεται από την γενική υποδομή διαχείρισης οπτικοακουστικού περιεχομένου και συγκεκριμένα μέσω των περιγραφών του περιεχομένου, οι οποίες είναι αποθηκευμένες σε βάση δεδομένων. Επομένως το γενικό πλαίσιο ή περιβάλλον λειτουργίας των εφαρμογών αυτών περιλαμβάνει μια βάση δεδομένων περιγραφών, η οποία αποτελεί την πηγή των περιγραφών και ίσως μια πηγή οπτικοακουστικού υλικού. Οι εφαρμογές στέλνουν αιτήσεις στο σύστημα που περιέχει την πηγή των περιγραφών, οι οποίες αφορούν στο περιεχόμενο του οπτικοακουστικού υλικού, δηλαδή στο περιεχόμενο των

περιγραφών. Οι αιτήσεις αυτές μεταφράζονται σε ερωτήσεις για τη βάση δεδομένων και αφού εκτελεστούν παράγουν κάποιο αποτέλεσμα το οποίο μπορεί να αποτελεί ένα υποσύνολο των περιγραφών που μεταβιβάζεται την εφαρμογή, ή μπορεί να επιδρά κατευθείαν στο οπτικοακουστικό υλικό που παρουσιάζεται στο τερματικό.

Το γενικό σχήμα της αρχιτεκτονικής των εφαρμογών αυτών φαίνεται στο παρακάτω σχήμα. Το πληροφοριακό σύστημα αποτελείται από την υποδομή διαχείρισης περιεχομένου και την πηγή οπτικοακουστικού υλικού. Η υποδομή διαχείρισης περιεχομένου περιλαμβάνει τη βάση δεδομένων που χρησιμοποιείται για την αποθήκευση του περιεχομένου, το υποσύστημα των περιγραφών του περιεχομένου και το υποσύστημα δεικτοδότησης του περιεχομένου. Η βάση δεδομένων μαζί με το υποσύστημα διαχείρισης των περιγραφών μπορούν να χρησιμοποιηθούν ανεξάρτητα από την πηγή της οπτικοακουστικής πληροφορίας για ανταλλαγή περιεχομένου στη μορφή που παρέχει η γλώσσα AVDL. Όταν χρησιμοποιούνται μαζί με την πηγή πληροφορίας στηρίζονται στους μηχανισμούς δεικτοδότησης για την ανάκτηση μέρους της ίδιας της πληροφορίας και τη μεταβίβασή της στην εφαρμογή ως αποτέλεσμα αιτήσεων ανάκτησης.



Σχήμα 6-1. Αρχιτεκτονική υποστήριξης εφαρμογών οπτικοακουστικού περιεχομένου

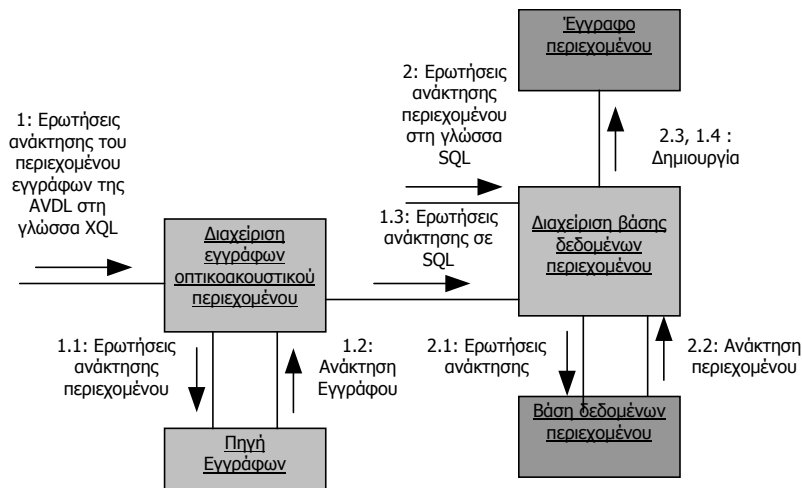
Οι εφαρμογές οπτικοακουστικού περιεχομένου στέλνουν αιτήσεις διαχείρισης οπτικοακουστικού περιεχομένου στο πληροφοριακό σύστημα, το οποίο αποτελείται από την υποδομή διαχείρισης περιεχομένου και πιθανόν την πηγή οπτικοακουστικού υλικού. Η υποδομή διαχείρισης περιεχομένου αποτελείται από τη βάση δεδομένων περιεχομένου που αποθηκεύει το μοντελοποιημένο περιεχόμενο, το υποσύστημα των περιγραφών του περιεχομένου και το υποσύστημα δεικτοδότησης του περιεχομένου.

Οι αιτήσεις ανάκτησης διαχείρισης περιεχομένου που φαίνονται στο σχήμα μπορεί να είναι ερωτήσεις στη γλώσσα XQL που είναι η γλώσσα ερωτήσεων της XML και έχει τη δυνατότητα υποστήριξης ερωτήσεων απευθείας στα έγγραφα της και εξαγωγής περιεχομένου από αυτά. Επίσης είναι δυνατόν οι αιτήσεις αυτές να έχουν τη μορφή SQL ερωτήσεων στο μοντέλο της βάσης δεδομένων, οπότε εκτελούνται από το σύστημα διαχείρισης της βάσης για τη ανάκτηση του περιεχομένου.

Σύμφωνα με τα παραπάνω, οι εφαρμογές διαχείρισης οπτικοακουστικού περιεχομένου μπορούν να διακριθούν σε δύο κατηγορίες, ανάλογα με τις απαιτήσεις διαχείρισης: εφαρμογές διαχείρισης στατικού οπτικοακουστικού περιεχομένου και εφαρμογές διαχείρισης της οπτικοακουστικής πληροφορίας μέσω του περιεχομένου της. Στις πρώτες ανήκουν οι εφαρμογές ανάκτησης, φιλτραρίσματος, ανταλλαγής και σύνθεσης περιεχομένου. Στη δεύτερη κατηγορία εφαρμογών ανήκουν αυτές που επιτρέπουν στο χρήστη αλληλεπίδραση με το ίδιο το οπτικοακουστικό υλικό, όπως έλεγχο της παρουσίας του, διαβαθμισμένη μεταβίβασή του και ανάκτηση βάσει του περιεχομένου του. Στην πρώτη κατηγορία ανήκουν κυρίως οι εφαρμογές βάσεων δεδομένων πολυμέσων, οι ψηφιακές βιβλιοθήκες αρχειακού υλικού και οι εφαρμογές παραγωγής περιεχομένου. Στη δεύτερη κατηγορία ανήκουν οι αλληλεπιδραστικές εφαρμογές και η εκπομπή πολυμέσων. Επομένως για τις δύο αυτές κατηγορίες εφαρμογών, όσον αφορά στην αναπαράσταση του οπτικοακουστικού περιεχομένου με τη γλώσσα AVDL και σε σχεσιακή βάση δεδομένων, μπορούμε να διακρίνουμε δύο τρόπους διαχείρισης του περιεχομένου:

- Διαχείριση στατικού οπτικοακουστικού περιεχομένου.
- Διαχείριση δυναμικού περιεχομένου. Η διαχείριση δυναμικού περιεχομένου περιλαμβάνει τη δυναμική μεταβολή της αναπαράστασης του περιεχομένου ώστε να είναι συγχρονισμένες με τις μεταβολές στην παρουσίαση του υλικού.

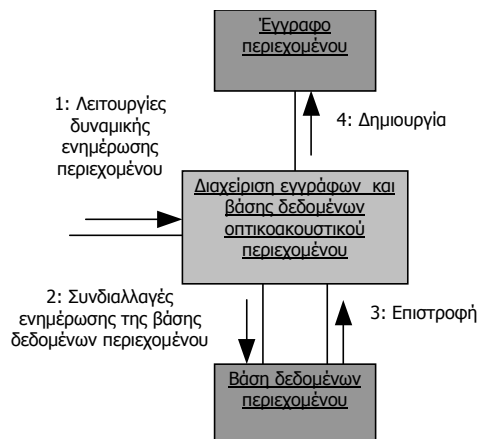
Στο παρακάτω σχήμα φαίνεται η υποστήριξη των εφαρμογών διαχείρισης στατικού περιεχομένου, όπου οι ενέργειες του χρήστη μεταφράζονται σε ερωτήσεις, οι οποίες χρησιμοποιούν τη βάση αποθήκευσης των περιγραφών και επιστρέφουν τα αποτελέσματα των ερωτήσεων στο χρήστη. Αυτό το τμήμα της υποδομής μπορεί να χρησιμοποιηθεί και ανεξάρτητα από την οπτικοακουστική πληροφορία εφόσον δεν υπάρχει ανάγκη συγχρονισμού ανάμεσα σε αυτήν και το περιεχόμενο.



Σχήμα 6-2. Υποστήριξη εφαρμογών διαχείρισης στατικού περιεχομένου

Οι εφαρμογές διαχείρισης στατικού περιεχομένου στέλνουν ερωτήσεις ανάκτησης στην πηγή εγγράφων ή στη βάση δεδομένων του πληροφοριακού συστήματος. Στην πρώτη περίπτωση οι ερωτήσεις είναι στη μορφή της γλώσσας ερωτήσεων XQL η οποία μπορεί να εφαρμοστεί απευθείας πάνω σε έγγραφα της XML. Στη δεύτερη περίπτωση οι ερωτήσεις είναι στη γλώσσα SQL των σχεσιακών βάσεων δεδομένων. Το αποτέλεσμα των ερωτήσεων είναι η δημιουργία ενός εγγράφου της γλώσσας AVDL είτε από την πηγή εγγράφων είτε από τη βάση δεδομένων.

Όταν η παρουσίαση του υλικού, επομένως και η περιγραφή είναι δυναμική, παρακολουθεί τις μεταβολές στο περιεχόμενο, οι οποίες όπως έχει ήδη αναφερθεί προκαλούνται είτε από το χρήστη είτε από δυναμική συμπεριφορά των αντικειμένων ή των σκηνών παρουσίασης. Η αρχιτεκτονική αυτού του είδους εφαρμογών φαίνεται στο παρακάτω σχήμα.



Σχήμα 6-3. Εφαρμογές διαχείρισης δυναμικού περιεχομένου

Οι εφαρμογές διαχείρισης δυναμικού περιεχομένου στηρίζονται στην ενημέρωση του περιεχομένου των σκηνών στη βάση δεδομένων και την πηγή εγγράφων περιεχομένου. Οι αιτήσεις ενημέρωσης που προέρχονται από ενέργειες αλληλεπίδρασης του χρήστη με τη σκηνή, μετατρέπονται σε συνδιαλλαγές ενημέρωσης της βάσης δεδομένων, η οποία σε συνδυασμό με το σύνολο εργαλείων που παρουσιάστηκαν στο προηγούμενο κεφάλαιο, υποστηρίζει την κατασκευή εγγράφων με το νέο περιεχόμενο.

Στις εφαρμογές αυτές, οι ενέργειες του χρήστη επιδρούν στην παρουσίαση των οπτικοακουστικών σκηνών και υπάρχει λογισμικό που αντιλαμβάνεται τις μεταβολές και προκαλεί αντίστοιχες αλλαγές στην περιγραφή της σκηνής.

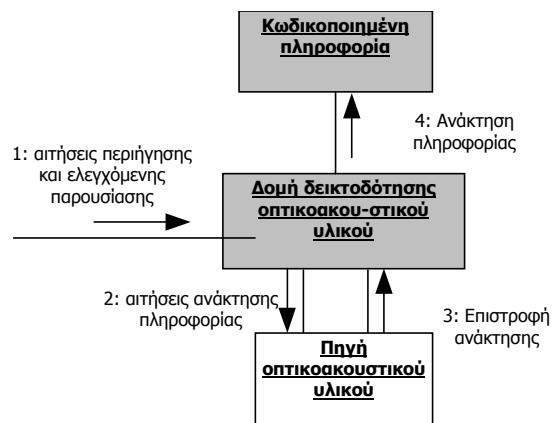
Στις επόμενες ενότητες περιγράφεται η υποστήριξη της λειτουργικότητας συγκεκριμένων εφαρμογών. Στην ενότητα 6.2 παρουσιάζεται η υποστήριξη των υπηρεσιών περιήγησης και διαβαθμισμένης μεταβίβασης της πληροφορίας, ενώ στην ενότητα 6.3 περιγράφεται πως υποστηρίζεται η ανάκτηση περιεχομένου με βάση το προφίλ του τερματικού (υποενότητα 6.3.1) και στην υποενότητα 6.3.2 περιγράφεται η υποστήριξη προφίλ χρηστών. Στις ενότητες 6.4 και 6.5 παρουσιάζονται οι υπηρεσίες ανάκτησης βάσει περιεχομένου και ανταλλαγής του σε ετερογενή περιβάλλοντα. Τέλος στην ενότητα 6.6 δίνεται ένα παράδειγμα εφαρμογής εκπαίδευσης από απόσταση με χρήση οπτικοακουστικών δεδομένων και ο τρόπος που υποστηρίζεται η λειτουργικότητά της από την υποδομή διαχείρισης που αναπτύχθηκε στη διατριβή.

6.2 Υπηρεσίες περιήγησης και ελεγχόμενης παρουσίασης

Οι υπηρεσίες περιήγησης και ελεγχόμενης παρουσίασης του υλικού υποστηρίζονται από τις δομές δεικτοδότησης της κωδικοποιημένης πληροφορίας. Οι δομές αυτές παρέχουν τη δυνατότητα τυχαίας προσπέλασης σε τμήματα του υλικού μέσα στα μεμονωμένα αντικείμενα μέσων. Επομένως οι δομές δεικτοδότησης υποστηρίζουν τις εξής λειτουργίες:

- Παρουσίαση τυχαίων σημείων μέσα στην πληροφορία.
- Δημιουργία ροών γρήγορης επισκόπησης με εξαγωγή κωδικοποιημένου περιεχομένου από την αρχική ροή.
- Έλεγχο της παρουσίασης με λειτουργίες σταμάτημα, παύση, συνέχιση και εναλλαγή γρήγορης επισκόπησης και κανονικής παρουσίασης του περιεχομένου.

Στο παρακάτω σχήμα απεικονίζεται η αρχιτεκτονική διαχείρισης των αιτήσεων αυτών από το πληροφοριακό σύστημα.



Σχήμα 6-4. Διάγραμμα υποστήριξης υπηρεσιών διαβαθμισμένης μεταβίβασης περιεχομένου

Η περιήγηση και ελεγχόμενη παρουσίαση της πληροφορίας στηρίζονται στην επεξεργασία των δομών δεικτοδότησης του κωδικοποιημένου οπτικοακουστικού υλικού ή και των ρών γρήγορης επισκόπησης που είναι πιθανόν να έχουν κατασκευαστεί από το υλικό.

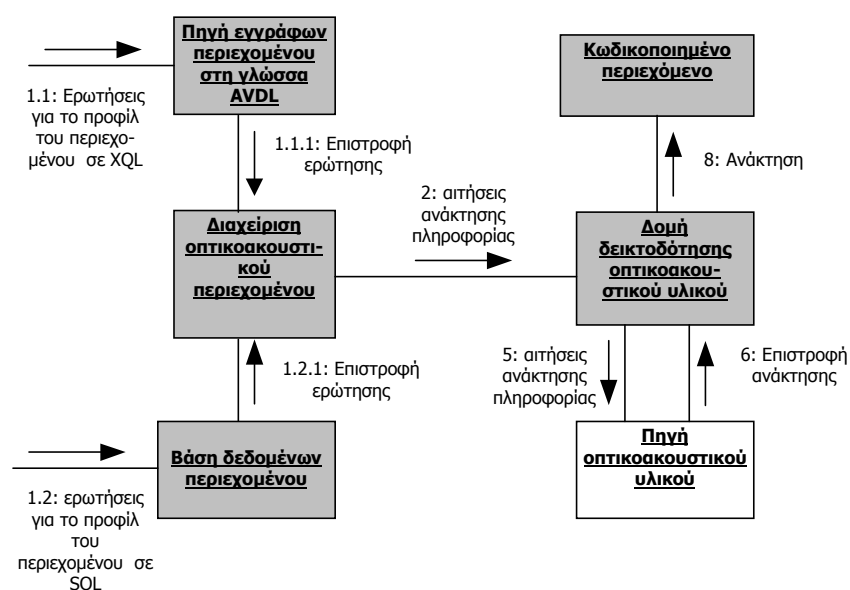
Οι υπηρεσίες διαβαθμισμένης μεταβίβασης είναι επεκτάσιμες διότι μπορεί να υποστηριχθεί η αναπαράσταση και αποθήκευση επιπλέον πληροφορίας για το περιεχόμενο, η οποία να επιτρέπει την περιήγηση βάσει περιεχομένου σε συνδυασμό με τις υπάρχουσες δομές δεικτοδότησης.

6.3 Υπηρεσίες διαβαθμισμένης μεταβίβασης περιεχομένου

Οι υπηρεσίες διαβαθμισμένης μεταβίβασης του περιεχομένου στηρίζονται σε αιτήσεις ανάκτησης πληροφορίας για τις απαιτήσεις μετάδοσης και αποκωδικοποίησης των σκηνών ή αντικειμένων παρουσίασης που πρόκειται να μεταδοθούν στο τερματικό, αλλά και για το είδος και τα χαρακτηριστικά της πληροφορίας που περιέχεται σε μια παρουσίαση. Οι απαιτήσεις μετάδοσης και αποκωδικοποίησης αφορούν στις δυνατότητες του δικτύου και του τερματικού και είναι απαραίτητο να γίνει διαπραγμάτευσή τους με το πληροφοριακό σύστημα πριν αρχίσει η μεταβίβαση του υλικού. Το είδος και τα χαρακτηριστικά της πληροφορίας ελέγχονται ως προς το προφίλ των χρητών, δηλαδή τις ιδιαίτερες απαιτήσεις ή προτιμήσεις τους.

Στο ακόλουθο σχήμα φαίνεται η αρχιτεκτονική υποστήριξης των υπηρεσιών διαβαθμισμένης μεταβίβασης, από το σύνολο της υποδομής διαχείρισης MPEG περιεχομένου. Το πληροφοριακό σύστημα δέχεται αιτήσεις ή ερωτήσεις που αφορούν στο προφίλ του περιεχομένου. Οι ερωτήσεις αυτές είναι δυνατόν να εκφράζονται είτε στη

γλώσσα SQL της σχεσιακής βάσης δεδομένων, είτε στη γλώσσα XQL, που αποτελεί τη γλώσσα ερωτήσεων της XML και υποστηρίζει απευθείας ανάκτηση εγγράφων με ερωτήσεις. Οι ερωτήσεις αυτές επιστρέφουν έγγραφα που περιέχουν την επιθυμητή πληροφορία για το προφίλ του περιεχομένου. Κατόπιν είναι δυνατό να συνδυαστούν με τις δομές δεικτοδότησης για την ανάκτηση πληροφορίας. Η λειτουργία αυτή βρίσκεται εφαρμογή κυρίως στην περίπτωση όπου αυτή περιέχει κωδικοποιημένες βαθμίδες κλιμάκωσης, οι οποίες δεν είναι επιθυμητό να μεταδοθούν στο τερματικό είτε γιατί είναι παραπάνω από τις δυνατότητές του, είτε γιατί απαιτούν μεγαλύτερους ρυθμούς μετάδοσης από ότι μπορεί να υποστηρίξει το δίκτυο.



Σχήμα 6-5. Διάγραμμα υποστήριξης υπηρεσιών διαβαθμισμένης μεταβίβασης περιεχομένου

Η διαβαθμισμένη μεταβίβαση περιεχομένου στηρίζεται στην εκτέλεση ερωτήσεων για το προφίλ του περιεχομένου σε σχέση με το προφίλ του τερματικού ή με το προφίλ των χρηστών. Οι ερωτήσεις αυτές εκφράζονται είτε στη γλώσσα ερωτήσεων της βάσης δεδομένων, είτε απευθείας στα έγγραφα περιεχομένου με τη γλώσσα ερωτήσεων της XML. Η ανάκτηση της ίδιας της πληροφορίας μετά την επιστροφή των αποτελεσμάτων των ερωτήσεων μπορεί να γίνει με χρήση των δομών δεικτοδότησης που παρέχουν προσπέλαση σε τμήματα του περιεχομένου, όπως για παράδειγμα στην κωδικοποιημένη βαθμίδα βάσης ενός βίντεο με βαθμίδες κλιμάκωσης.

Στις επόμενες υποενότητες παρουσιάζεται η υποστήριξη των υπηρεσιών διαβαθμισμένης μεταβίβασης με βάση το προφίλ τερματικών και χρηστών, όπως παρέχεται από την υποδομή που παρουσιάστηκε στη διατριβή. Οι υποστήριξη των υπηρεσιών αυτών

περιγράφεται με συνδιαλλαγές ανάκτησης περιεχομένου από τη βάση δεδομένων του πληροφοριακού συστήματος.

6.3.1 Προφίλ τερματικού

Το προφίλ του τερματικού αποτελείται από παραμέτρους ή χαρακτηριστικά που καθορίζουν τις δυνατότητες του τερματικού όσον αφορά στην αποκωδικοποίηση οπτικοακουστικού υλικού. Το προφίλ του τερματικού μπορεί να ελεγχθεί με τα χαρακτηριστικά του υλικού για να καθοριστεί αν το τερματικό έχει τη δυνατότητα να το αποκωδικοποιήσει. Η περιγραφή του οπτικοακουστικού περιεχομένου υποστηρίζει χαρακτηριστικά που μπορούν να συγκριθούν με το προφίλ του τερματικού ως προς τα τέσσερα βασικά είδη οπτικοακουστικών αντικειμένων, δηλαδή οπτικά, ακουστικά αντικείμενα, σκηνές και περιγραφείς αντικειμένων. Τα στοιχεία που υποστηρίζουν το προφίλ του τερματικού ως προς τα οπτικά αντικείμενα είναι:

- Το είδος του οπτικού αντικειμένου, δηλαδή βίντεο, ακίνητη υφή, πλέγμα ή πρόσωπο, το οποίο περιγράφεται από το χαρακτηριστικό type του στοιχείου VisualObject.
- Το προφίλ και το επίπεδο οπτικών αντικειμένων. Το προφίλ και το επίπεδο καθορίζουν το είδος των οπτικών αντικειμένων και την πολυπλοκότητά τους. Από το προφίλ καθορίζεται αν το περιεχόμενο αποτελείται I, P, B επίπεδα, δυαδικό σχήμα, κλιμάκωση, πεπλεγμένη κωδικοποίηση κλπ. Ο συνδυασμός προφίλ σε επίπεδο που περιγράφεται από το στοιχείο ProfileAtLevel του στοιχείου VisualObject περιγράφει το μέγιστο ρυθμό μετάδοσης του αντικειμένου.
- Το είδος του σήματος βίντεο αντικειμένων βίντεο ή ακίνητης υφής, δηλαδή το στοιχείο VideoSignalType που περιέχεται μέσα στο στοιχείο VisualObject και πιο συγκεκριμένα η μορφή του βίντεο (VideoFormat) και τα χαρακτηριστικά του χρώματος (ColorPrimaries).
- Σταθερός ή μεταβλητός ρυθμός εικόνων ανά επίπεδο αντικειμένου βίντεο που περιγράφεται από το χαρακτηριστικό fixed_VOP_rate του στοιχείου VideoObjectLayer.
- Λόγος των διαστάσεων μιας βαθμίδας (aspect_ratio_info), ρυθμός δεδομένων (bit_rate), μέγεθος μνήμης αποκωδικοποιητή (vbn_buffer_size), μορφή χρώματος (chroma_format) και καθυστέρηση (low_delay) που περιέχονται στο στοιχείο VOLControlParameters των BaseLayer και ScalabilityLayer.
- Διαστάσεις και σχήμα των οπτικών αντικειμένων (BaseLayerShape, ScalabilityLayerShape, Sprite).

Στο ακόλουθο σχήμα παρουσιάζονται οι συνδιαλλαγές της σχεσιακής βάσης δεδομένων, οι οποίες υποστηρίζουν την υπηρεσία διαβαθμισμένης μεταβίβασης οπτικών αντικειμένων με βάση το προφίλ του τερματικού.

```

<!-- Συνδιαλλαγή ανάκτησης οντοτήτων περιεχομένου βίντεο -->
BEGIN TRANSACTION
SELECT      *
FROM        DatabaseEntity
WHERE       type like "%Video%" or type like "%video%"
COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης των αντικειμένων βίντεο του MPEG-4 -->
BEGIN TRANSACTION
SELECT      *
FROM        VisualObject
WHERE       type="video"
COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης των αντικειμένων βίντεο του MPEG-4 που
περιέχουν επίπεδα τύπου B (μέγιστη συμπίεση και πολυπλοκότητα
αποκωδικοποίησης) και μέγιστους ρυθμούς μικρότερους ή ίσους από 64
kbits/sec -->

BEGIN TRANSACTION
SELECT      *
FROM        VisualObject, Profile_Level, Profile, Level
WHERE       VisualObject.type="video" and
            VisualObject.profile_level = ProfileAtLevel.id and
            Profile_Level.profile = Profile.id and
            Profile.BVOP=true and
            Level.id = ProfileAtLevel.level and
            Level.maxbitrate < 64
COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης των αντικειμένων βίντεο του MPEG-4 που έχουν
κωδικοποιηθεί από βίντεο τύπου PAL -->
BEGIN TRANSACTION
SELECT      *
FROM        VisualObject, VideoObjectLayer
WHERE       VideoObject.id=VideoObjectLayer.video_object and
            VideoObjectLayer.fixed_VOP_rate= true
COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης των αντικειμένων βίντεο του MPEG-4 που έχουν
σταθερό ρυθμό εικόνων στις βαθμίδες βίντεο -->
BEGIN TRANSACTION
SELECT      *
FROM        VideoObject, VideoObjectLayer
WHERE       VideoSignalType.visual_object = VisualObject.id and
            VideoSignalType.video_format="PAL"
COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης των αντικειμένων βίντεο του MPEG-4 που έχουν
βασική βαθμίδα με ρυθμό δεδομένων μικρότερο των 2 Mbits/sec και μέγεθος
απομονωτή 500 Kbytes -->
BEGIN TRANSACTION
SELECT      *
FROM        VideoObject, BaseLayer, VOLControlParameters
WHERE       VideoObject.id = BaseLayer.video_object and
            BaseLayer.id = VOLControlParameters.bit_rate*400/1000 < 2000
            And VOLControlParameters.vbv_buffer_size < 500

```

```

COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης των αντικειμένων βίντεο του MPEG-4 που έχουν
τυχαίο (δυναμικό) σχήμα και μία ή περισσότερες βαθμίδες κλιμάκωσης -->
BEGIN TRANSACTION
SELECT      *
FROM        VideoObject, VideoObjectLayer VOL
WHERE       VideoObject.id = VideoObjectLayer.video_object and
            VideoObjectLayer.shape = "binary" and
            EXISTS (SELECT      *
                     FROM        ScalabilityLayer
                     WHERE       ScalabilityLayer.vol = VOL.id)
COMMIT TRANSACTION

```

Η περιγραφή των σκηνών μπορεί να συγκριθεί με το προφίλ κάποιου τερματικού με βάση τα εξής συστατικά της στοιχεία:

- Μέγεθος (πλάτος και ύψος) παρουσίασης της σκηνής στην οθόνη, που αποτελεί χαρακτηριστικό του στοιχείου SceneSize.
- Είδος των περιεχομένων της σκηνής που καθορίζεται από την ύπαρξη μέσα σε αυτήν κόμβων μέσων (MediaNode), οπότε μπορεί σε πρωταρχικό στάδιο να αποφασιστεί το είδος της σκηνής χωρίς να ελεγχθούν τα χαρακτηριστικά των ίδιων των αντικειμένων.
- Είδος της γεωμετρίας της σκηνής που καθορίζεται από την ύπαρξη δυοδιάστατων και τρισδιάστατων κόμβων στην περιγραφή της.

Οι συνδιαλλαγές υποστήριξης της ανάκτησης σκηνών με βάση τα παραπάνω φαίνονται στο ακόλουθο σχήμα.

```

<!-- Συνδιαλλαγή ανάκτησης σκηνών με διαστάσεις 200 x 300 -->
BEGIN TRANSACTION
SELECT      *
FROM        Scene
WHERE       pixel_width = 200 and
            pixel_height = 300
COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης σκηνών που περιέχουν συνθετικά αντικείμενα
προσώπων και μέχρι ένα φυσικά αντικείμενα βίντεο-->
BEGIN TRANSACTION
SELECT      *
FROM        Scene S
WHERE       EXISTS (SELECT      *
                     FROM        Face, ObjectDescriptor, FaceObject
                     WHERE       Face.scene = S.id and
                                Face.object_descriptor=ObjectDescriptor.id
                                And ObjectDescriptor.media = FaceObject.id)

            And COUNT(SELECT      *
                        FROM        VideoObject2D
                        WHERE       VideoObject2D.scene = S.id) <= 1
COMMIT TRANSACTION

```

6.3.2 Προφίλ χρηστών

Το προφίλ των χρηστών κάποιας εφαρμογής, εκτός από το προφίλ του τερματικού που διαθέτουν για την κατανάλωση οπτικοακουστικού υλικού είναι δυνατόν να περιλαμβάνει επιπλέον πληροφορία για το περιεχόμενο που τους ενδιαφέρει ή για την πρόσβαση σε αυτό. Η γλώσσα AVDL περιέχει στοιχεία που μπορούν να χρησιμοποιηθούν σε εφαρμογές για την υποστήριξη εξατομικευμένου (personalized) περιεχομένου.

Καταρχήν η περιγραφή περιέχει πληροφορία για το είδος του περιεχομένου. Η πληροφορία αυτή μπορεί να προέρχεται από τη σκηνή, τα ίδια τα αντικείμενα μέσω των ή τους περιγραφείς αντικειμένων. Η περιγραφή της σκηνής περιέχει κόμβους που αναπαριστούν τα αντικείμενα μέσω των περιγράφων το είδος του καθενός. Έτσι μπορεί να υποστηριχθεί το προφίλ κάποιου χρήστη που ενδιαφέρεται για ακουστικές σκηνές που περιέχουν:

- πολυκαναλικό ήχο, όπως περιγράφεται από το στοιχείο AudioSource της περιγραφής της σκηνής και από το στοιχείο Sound,
- βίντεο ή εικόνες όπως περιγράφεται από τα στοιχεία VideoObject2D, Image2D, ή
- αλληλεπιδραστικές σκηνές των οποίων οι περιγραφές περιέχει στοιχεία των τύπων SensorNode.

Στο παρακάτω σχήμα παρουσιάζονται συνδιαλλαγές υποστήριξης του προφίλ των χρηστών ως προς το είδος και τα χαρακτηριστικά του περιεχομένου και τη δυνατότητα αλληλεπίδρασης του χρήστη με αντικείμενα της σκηνής:

```
<!-- Συνδιαλλαγή ανάκτησης σκηνών που περιέχουν πολυκαναλικό ήχο -->
BEGIN TRANSACTION
SELECT      Scene.*
FROM        Scene, AudioSource
WHERE       AudioSource.scene = Scene.id and
            AudioSource.num_channels > 1
COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης σκηνών που περιέχουν αλληλεπίδραση με ένα
γραφικό αντικείμενο-κύλινδρο της σκηνής -->

BEGIN TRANSACTION
SELECT      Scene.*
FROM        Scene, CylinderSensor
WHERE       CylinderNode.scene = Scene.id
COMMIT TRANSACTION
```

Τα ίδια τα αντικείμενα μέσω των επίσης περιγράφουν παρόμοια χαρακτηριστικά της παρουσίας τους που είναι πιθανόν να ενδιαφέρουν από ποιοτική άποψη τους χρήστες των εφαρμογών. Οι περιγραφείς αντικειμένων συμπληρώνουν το προφίλ των χρηστών με στοιχεία που αφορούν στα εξής:

- Είδος περιεχομένου όπως περιγράφεται από τον περιγραφέα IP_IdentificationDataSet, ο οποίος περιέχει ένδειξη της κατηγορίας του περιεχομένου (οπτικοακουστικό, βιβλίο, σειρά, κείμενο, μουσικό κομμάτι, εικόνα, κλπ).
- Γλώσσα και τίτλο του περιεχομένου όπως περιγράφονται από το στοιχείο SupplContentIdentifier.
- Ποιότητα υπηρεσιών που περιγράφεται από το στοιχείο QoSDescriptor και περιέχει ένα σύνολο από παραμέτρους αυτού του είδους.

Οι συνδιαλλαγές υποστήριξης του προφίλ των χρηστών με βάση τα παραπάνω φαίνονται στο σχήμα:

```

<!-- Συνδιαλλαγή ανάκτησης σκηνών που περιέχουν ένα αντικείμενο που έχει
ταξινομηθεί διεθνώς ως μουσικό κομμάτι -->
BEGIN TRANSACTION
SELECT      Scene.*
FROM        Scene, AudioSource
WHERE       AudioSource.scene=Scene.id and
           AudioSource.object_descriptor IN
           (SELECT      ES_Descriptor.id
            FROM        ES_Descriptor, IP_IdentificationDataSet,
                        ContentIdentifier
            WHERE       ES_Descriptor.id =
                        IP_IdentificationDataSet.es_descriptor and
                        IP_IdentificationDataSet.id =
                        ContentIdentifier.ip_id_dataset and
                        ContentIdentifier.content_type="musical_work" )

COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης σκηνών που περιέχει βίντεο στην ελληνική
γλώσσα -->
BEGIN TRANSACTION
SELECT      Scene.*
FROM        Scene, VideoObject2D
WHERE       VideoObject2D.scene = Scene.id and
           VideoObject2D.object_descriptor IN
           (SELECT      ES_Descriptor.id
            FROM        ES_Descriptor, IP_IdentificationDataSet,
                        SupplContentIdentifier
            WHERE       ES_Descriptor.id =
                        IP_IdentificationDataSet.es_descriptor and
                        IP_IdentificationDataSet.id =
                        SupplContentIdentifier.ip_id_dataset and
                        SupplContentIdentifier.language="Greek" )

COMMIT TRANSACTION

<!-- Συνδιαλλαγή ανάκτησης σκηνών που περιέχουν βίντεο με μέγιστη
καθυστερήση για να αρχίσει η αναπαραγωγή του, λιγότερο από 5 δευτερόλεπτα
-->
BEGIN TRANSACTION
SELECT      Scene.*
FROM        Scene, VideoObject2D
WHERE       VideoObject2D.scene = Scene.id and
           VideoObject2D.object_descriptor IN
           (SELECT      ES_Descriptor.id
            FROM        ES_Descriptor, QoSDescriptor
            WHERE       ES_Descriptor.id = QoSDescriptor.es_descriptor
                        And QoSDescriptor.metric_type=MAX_DELAY and

```

```
And QosDescriptor.value*1000 < 5000 )
COMMIT TRANSACTION
```

Οι υπηρεσίες διαβαθμισμένης μεταβίβασης περιεχομένου είναι εύκολα επεκτάσιμες ως προς το είδος της πληροφορίας που περιέχεται στο σύστημα για την υποστήριξή τους. Αυτό οφείλεται στην ευελιξία και επεκτασιμότητα της γλώσσας αναπαράστασης περιεχομένου και της βάσης δεδομένων για την αποθήκευσή του.

6.4 Υπηρεσίες ανάκτησης βάσει περιεχομένου

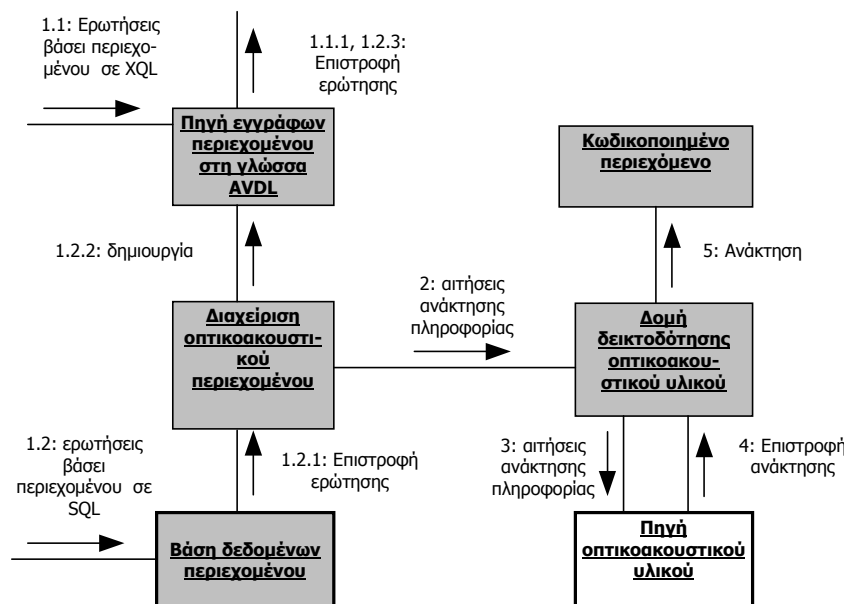
Οι υπηρεσίες ανάκτησης βάσει περιεχομένου υποστηρίζονται από το σύνολο των υποδομών που παρουσιάστηκαν στα προηγούμενα κεφάλαια της διατριβής με στόχο την υποστήριξη εφαρμογών που έχουν ανάγκη από αυξημένες ανάγκες ανάκτησης βάσει περιεχομένου. Παραδείγματα τέτοιων εφαρμογών είναι οι ψηφιακές βιβλιοθήκες, οι βάσεις αρχειακού υλικού, τα εκπαιδευτικά πολυμέσα, κλπ.

Στο σχήμα που ακολουθεί απεικονίζεται η αρχιτεκτονική υποστήριξης των υπηρεσιών ανάκτησης βάσει περιεχομένου στον πυρήνα ενός πληροφοριακού συστήματος που αποτελείται από τη βασική υποδομή υποστήριξης εφαρμογών, δηλαδή από μια βάση δεδομένων οπτικοακουστικού περιεχομένου, μια πηγή εγγράφων περιεχομένου στη γλώσσα AVDL, μια πηγή οπτικοακουστικής πληροφορίας και τα κατάλληλα τμήματα διαχείρισης της πληροφορίας και του περιεχομένου.

Η ανάκτηση του περιεχομένου, όπως φαίνεται στο σχήμα, μπορεί να γίνει από το πληροφοριακό σύστημα με δύο τρόπους: α) με αιτήσεις ανάκτησης στη γλώσσα ερωτήσεων SQL της βάσης δεδομένων ή και β) με αιτήσεις ανάκτησης απευθείας στα έγγραφα περιεχομένου με χρήση της γλώσσας XQL που είναι η γλώσσα ερωτήσεων της XML. Η γλώσσα XQL υποστηρίζει την ανάκτηση και σύνθεση εγγράφων που είναι γραμμένα στη γλώσσα XML ή κάποια εφαρμογή της. Το αποτέλεσμα των δύο διαφορετικών ειδών ερωτήσεων είναι έγγραφα περιεχομένου στη γλώσσα AVDL τα οποία ανακτώνται από την εφαρμογή. Η δημιουργία εγγράφων της γλώσσας εξυπηρετεί την άμεση επεξεργασία τους από την εφαρμογή ή και από τον ίδιο το χρήστη τους, εφόσον τα έγγραφα αυτά είναι σε απλή και κατανοητή μορφή. Επιπλέον, η ανάκτηση βάσει περιεχομένου μπορεί να υποστηριχθεί στον Παγκόσμιο Ιστό, γεγονός που διευκολύνεται από τη γλώσσα AVDL.

Εκτός από την ανάκτηση περιεχομένου, με βάση τον τρόπο που αυτό έχει μοντελοποιηθεί και αποθηκευτεί στο πληροφοριακό σύστημα, είναι δυνατή και η ανάκτηση πληροφορίας

από την πηγή οπτικοακουστικού υλικού. Η ανάκτηση αυτή υποστηρίζεται από την υποδομή δεικτοδότησης της πληροφορίας, η οποία παρέχει τυχαία προσπέλαση σε οποιοδήποτε σημασιολογικό τμήμα της, δηλαδή σε αντικείμενα και τμήματα αυτών.



Σχήμα 6-6. Διάγραμμα υποστήριξης υπηρεσιών ανάκτησης βάσει περιεχομένου

Η υποστήριξη υπηρεσιών ανάκτησης βάσει περιεχομένου, οι αιτήσεις που απευθύνονται στο πληροφοριακό σύστημα μπορεί να είναι ερωτήσεις στη γλώσσα SQL ή στη γλώσσα XQL που είναι η γλώσσα ερωτήσεων της XML. Οι ερωτήσεις αυτές απαντώνται είτε με ανάκτηση εγγράφων, με ανάκτηση περιεχομένου από τη βάση και δημιουργία εγγράφων ή και με επεξεργασία της πληροφορίας με βάση τις δομές δεικτοδότησης.

Οι υπηρεσίες ανάκτησης βάσει περιεχομένου είναι ευέλικτες και επεκτάσιμες τόσο ως προς την πληροφορία που μπορεί να αποθηκευτεί στο σύστημα και να χρησιμοποιηθεί για την απάντηση σε ερωτήσεις ανάκτησης, όσο και ως προς τις τεχνικές δεικτοδότησης και ανάκτησης της πληροφορίας.

6.5 Υπηρεσίες ανταλλαγής περιεχομένου

Οι υπηρεσίες ανταλλαγής περιεχομένου στηρίζονται στη γλώσσα AVDL, η οποία υποστηρίζει την ανταλλαγή οπτικοακουστικού περιεχομένου σε ετερογενή περιβάλλοντα και ιδιαίτερα στον Παγκόσμιο Ιστό. Η ανταλλαγή περιεχομένου στηρίζεται είτε στην ανάκτησή του μέσω ερωτήσεων με τον τρόπο που περιγράφηκε στην προηγούμενη ενότητα, είτε με την ανταλλαγή εγγράφων περιεχομένου από την πηγή τους που είναι

δημοσιευμένη στον Παγκόσμιο Ιστό. Η ανταλλαγή των εγγράφων μπορεί να υποστηριχθεί με τη δημιουργία τους από το αποθηκευμένο περιεχόμενο στη βάση δεδομένων του πληροφοριακού συστήματος.

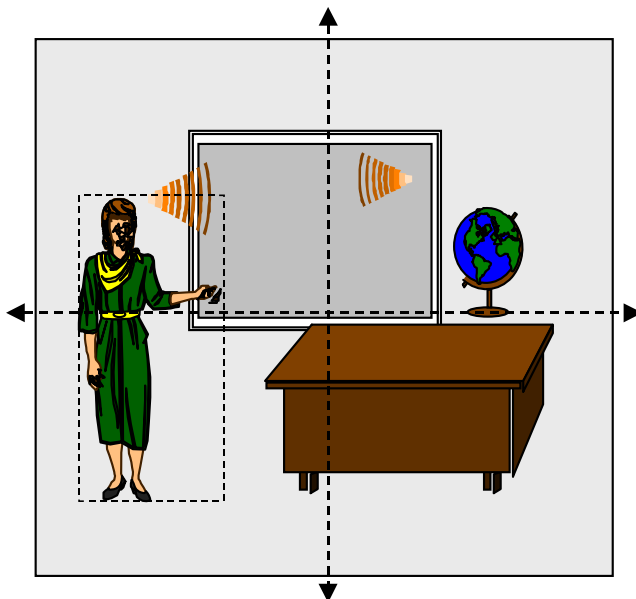
6.6 Παράδειγμα εφαρμογής

Στην ενότητα αυτή παρουσιάζονται ένα παράδειγμα εφαρμογής που υποστηρίζεται με την υποδομή διαχείρισης περιεχομένου που παρουσιάστηκε στην διατριβή. Η εφαρμογή αυτή είναι εκπαίδευση από απόσταση και περιγράφεται ως προς το είδος του περιεχομένου και τη λειτουργικότητα που παρέχεται στον τελικό χρήστη. Η υποστήριξη της εφαρμογής με την υποδομή διαχείρισης περιεχομένου περιλαμβάνει την αναπαράσταση του περιεχομένου με τη γλώσσα AVDL, την ενημέρωση της βάσης δεδομένων περιεχομένου και την κατασκευή των δομών δεικτοδότησης. Η επιθυμητή λειτουργικότητα υποστηρίζεται από συγκεκριμένες ενέργειες, αιτήσεις ανάκτησης, ερωτήσεις και συνδιαλλαγές που αφορούν στο περιεχόμενο.

6.6.1 Εκπαίδευση από απόσταση

Στην εφαρμογή εκπαίδευσης από απόσταση, το σενάριο είναι ότι οι μαθητές ή φοιτητές βρίσκονται στο σπίτι τους ή σε κάποια αίθουσα με υπολογιστές στο πανεπιστήμιο και παρακολουθούν τα μαθήματα μέσω υπολογιστή από απόσταση. Στο σπίτι του η υπόθεση είναι ότι ένας φοιτητής συνδέεται μέσω Ίντερνετ για να παρακολουθήσει με ρυθμούς μετάδοσης βίντεο μέχρι 64 kbps. Στο πανεπιστήμιο συνδέεται με δίκτυο υψηλών ταχυτήτων με μεγαλύτερες δυνατότητες ρυθμών μετάδοσης ως 4 Mbps ή και περισσότερο.

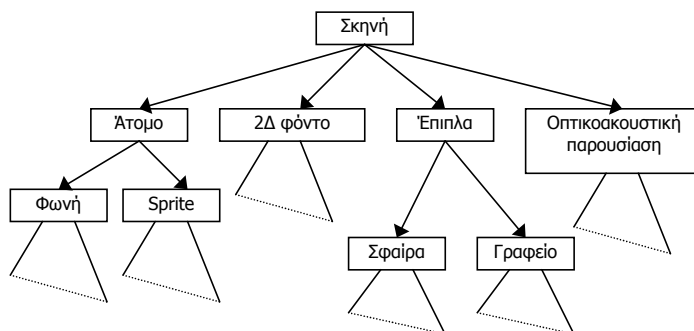
Το μάθημα είναι μαγνητοσκοπημένο και αποτελείται από τη σκηνή μιας καθηγήτριας που δίνει μια διάλεξη μέσα στην τάξη. Η σκηνή αποτελείται από την άποψη του σημείου της αίθουσας στο οποίο βρίσκεται η καθηγήτρια και περιλαμβάνει μεταξύ άλλων κάποια έπιπλα και συγκεκριμένα το γραφείο και ένα αντικείμενο πάνω σε αυτό (υδρόγειος σφαίρα), ένα δυσδιάστατο φόντο που δε μεταβάλλεται κατά τη διάρκεια της παρουσίασης και την εικόνα της καθηγήτριας μαζί με τη φωνή της. Επιπλέον η σκηνή αποτελείται από μια περιοχή όπου προβάλλεται ένα εκπαιδευτικό βίντεο μέσα στην αίθουσα. Η σκηνή αυτή απεικονίζεται στο παρακάτω σχήμα.



Σχήμα 6-7. Σκηνή εφαρμογής εκπαίδευσης από απόσταση

Η σκηνή της εφαρμογής εκπαίδευσης από απόσταση αποτελείται από την εικόνα και τη φωνή της καθηγήτριας που δίνει τη διάλεξη, την εικόνα της αίθουσας που αποτελείται από ένα γραφείο, ένα αντικείμενο (υδρόγειο σφαίρα) και το φόντο, καθώς και από μια περιοχή της σκηνής στην οποία προβάλλεται μια παρουσίαση βίντεο.

Η κωδικοποίηση της πληροφορίας της σκηνής και των αντικειμένων της γίνεται με το πρότυπο MPEG-4. Η σκηνή αποτελεί μια ιεραρχία που στο δεύτερο επίπεδο περιέχει τέσσερις κόμβους-παιδιά. Οι κόμβοι αυτοί παριστάνουν την καθηγήτρια, το δυσδιάστατο φόντο, τα έπιπλα και την οπτικοακουστική παρουσίαση. Η καθηγήτρια παριστάνεται με έναν κόμβο που ομαδοποιεί τη φωνή και την εικόνα της. Η φωνή είναι ένα αντικείμενο ήχου, ενώ η εικόνα είναι ένα στατικό αντικείμενο βίντεο τύπου sprite. Το δυσδιάστατο φόντο αποτελεί έναν κόμβο γραφικών. Η υδρόγειος σφαίρα και το γραφείο είναι αντικείμενα ακίνητης υφής. Η οπτικοακουστική παρουσίαση του εκπαιδευτικού υλικού αποτελείται από ένα αντικείμενο βίντεο και ένα αντικείμενο ήχου. Η λογική δομή της σκηνής φαίνεται στο ακόλουθο σχήμα.



Σχήμα 6-8. Λογική δομή της οπτικοακουστικής σκηνής της εφαρμογής εκπαίδευσης από απόσταση

Η οπτικοακουστική σκηνή του σχήματος αποτελείται από 4 σύνθετα αντικείμενα που φαίνονται στο δεύτερο επίπεδο της ιεραρχίας κάτω από τη ρίζα: την καθηγήτρια, το δυσδιάστατο φόντο της σκηνής, τα έπιπλα και μια οπτικοακουστική παρουσίαση. Η καθηγήτρια παριστάνεται από τη φωνή του και ένα στατικό οπτικό αντικείμενο.

Επίσης το οπτικοακουστικό υλικό του μαθήματος περιέχει κωδικοποιημένα επίπεδα κλιμάκωσης που παρέχουν τους διαφορετικούς ρυθμούς μετάδοσης όταν η εκπαίδευση γίνεται από το σπίτι, δηλαδή μέσω Ίντερνεντ ή μέσω τοπικού δικτύου στο Πανεπιστήμιο. Τα επίπεδα κλιμάκωσης αφορούν στην αναπαράσταση της καθηγήτριας όπου το στατικό αντικείμενο αποτελείται από ένα βασικό επίπεδο και ένα επίπεδο χρονικής κλιμάκωσης για τη λεπτομερέστερη αναπαράσταση της κίνησης του σώματός της. Επίσης το βίντεο της οπτικοακουστικής παρουσίασης είναι δυνατόν να περιέχει ένα ή περισσότερα επίπεδα χρονικής ή και χωρικής κλιμάκωσης.

Η λειτουργικότητα που παρέχεται από την εφαρμογή στο χρήστη δηλαδή στον φοιτητή αποτελείται από τα εξής:

- Διαβαθμισμένη μεταβίβαση του περιεχομένου ανάλογα με το προφίλ του δικτύου (Ίντερνεντ ή τοπικό δίκτυο υψηλών ταχυτήτων)
- Περιήγηση και τυχαία προσπέλαση στο βίντεο της καθηγήτριας
- Περιήγηση και τυχαία προσπέλαση στο εκπαιδευτικό βίντεο
- Ανάκτηση βάσει περιεχομένου στο εκπαιδευτικό βίντεο.

Η προσέγγιση που ακολουθείται στη διατριβή για την υποστήριξη της λειτουργικότητας της εφαρμογής μέσω της υποδομής διαχείρισης περιεχομένου περιλαμβάνει τα εξής στάδια:

- Αναπαράσταση του περιεχομένου της σκηνής με τη γλώσσα AVDL για την κατασκευή εγγράφων που αναπαριστούν το περιεχόμενο.

- Αποθήκευση των εγγράφων οπτικοακουστικού περιεχομένου στη βάση δεδομένων με χρήση των εργαλείων μετατροπής.
- Δεικτοδότηση του υλικού για την κατασκευή των κατάλληλων δομών.
- Χρήση των παραπάνω για την υποστήριξη της λειτουργικότητας της εφαρμογής.

Στο πρώτο στάδιο παράγεται η περιγραφή των οπτικών αντικειμένων στη γλώσσα AVDL. Η περιγραφή αυτή φαίνεται στο παρακάτω σχήμα, όπου περιγράφονται χωριστά τα αντικείμενα καθηγήτρια, γραφείο, υδρόγειος σφαίρα και βίντεο παρουσίασης. Το προφίλ του αντικειμένου βίντεο της καθηγήτριας είναι Κύριο (Main) διότι είναι το μόνο που επιτρέπει τα στατικά αντικείμενα sprite. Το σχήμα της καθηγήτριας είναι τυχαίο και το είδος της κλιμάκωσης που επιτρέπεται είναι χρονική κλιμάκωση. Η βαθμίδα βάσης έχει ρυθμό δεδομένων 64 kbps που εκφράζεται σε μονάδες των 400 bits/sec και έχει την τιμή 160. Η βαθμίδα κλιμάκωσης έχει επίσης ρυθμό 64 kbps.

```
<!-- Περιγραφή οπτικού αντικειμένου για την καθηγήτρια -->

<!-- Τύπος εγγράφου -->
<!DOCTYPE avdl SYSTEM "http://www.music.tuc.gr/~georgia/AVDL/avdl.dtd">

<!-- Προφίλ εγγράφου -->
<!ENTITY % MPEG4Visual.Description "INCLUDE">

<avdl>
<Visual>
  <VisualObjectSequence>
    <VisualObject type="video" id="video1">

      <Profile><MPEG4Visual.Main></></Profile>
      <ProfileAtLevel> <Core_at_L3> </> </ProfileAtLevel>

      <VideoSignalType>
        <VideoFormat video_format="PAL"></>
      </VideoSignalType>

      <VideoObject>

        <VideoObjectLayer id="Teacher" shape="binary">

          <BaseLayer>
            <VOLControlParameters aspect_ratio_info="1:2" bit_rate="160" >
            </VOLControlParameters>
            <BaseLayerShape>
              <SpriteEnable > </SpriteEnable>
            </BaseLayerShape>
            <VideoObjectPlane>
              <Sprite> </Sprite>
            </VideoObjectPlane>
          </BaseLayer>

          <ScalabilityLayer ref_layer_id="Teacher"
            ref_layer_sampling_direct=HIGHER >
            <VOLControlParameters bit_rate="160"> </VOLControlParameters>
          </ScalabilityLayer>
        </VideoObjectLayer>
      </VisualObject>
    </VisualObjectSequence>
  </Visual>
</avdl>
```

Comment [Γ.Κ.21]:

```

    </VideoObjectLayer>
  </VideoObject>
</VisualObject>
</VisualObjectSequence>
</Visual>
</avdl>

```

Σχήμα 6-9. Περιγραφή του οπτικού αντικειμένου για την καθηγήτρια

Η οπτικοακουστική παρουσίαση που προβάλλεται μέσα στην αίθουσα αποτελείται από ένα αντικείμενο βίντεο και ένα αντικείμενο ήχου. Το αντικείμενο βίντεο αποτελείται από μια βαθμίδα βάσης και μια βαθμίδα κλιμάκωσης. Η περιγραφή του περιεχομένου του με τη γλώσσα AVDL φαίνεται στο σχήμα που ακολουθεί. Ο ρυθμός του βίντεο είναι 1 Mbps που μεταφράζεται σε 2500 μονάδες των 400 bits/sec. Η κλιμάκωση είναι χωρική και επιτρέπει την παρουσίαση σε μέγεθος 3/2 του αρχικού βίντεο. Η βαθμίδα κλιμάκωσης έχει ρυθμό δεδομένων 1200 μονάδες των 400 bits/sec.

```

<!-- Περιγραφή οπτικού αντικειμένου για την οπτικοακουστική παρουσίαση -->
<!-- Τύπος εγγράφου -->
<!DOCTYPE avdl SYSTEM "http://www.music.tuc.gr/~georgia/AVDL/avdl.dtd">

<!-- Προφίλ εγγράφου -->
<ENTITY % MPEG4Visual.Description "INCLUDE">

<Visual>
  <VisualObjectSequence>
    <VisualObject type="video" id="Video2">

      <!-- Προφίλ αντικειμένου βίντεο -->
      <Profile><MPEG4Visual.Main></></Profile>
      <ProfileAtLevel> <Core_at_L3> </> </ProfileAtLevel>

      <!-- Χαρακτηριστικά του σήματος βίντεο -->
      <VideoSignalType>
        <VideoFormat video_format="PAL"></>
      </VideoSignalType>

      <VideoObject>
        <VideoObjectLayer id="Presentation" shape="rect">

          <!-- Βαθμίδα βάσης -->
          <BaseLayer>
            <VOLControlParameters aspect_ratio_info="3:2" bit_rate="2500" >
          </VOLControlParameters>
            <BaseLayerShape>
            </BaseLayerShape>
            <VideoObjectPlane>
            </VideoObjectPlane>
          </BaseLayer>

          <!-- Βαθμίδα χωρικής κλιμάκωσης -->
          <ScalabilityLayer>
            <VOLControlParameters aspect_ratio_info="3:2" bit_rate="1200" >
          </VOLControlParameters>
            <ScalabilityLayerShape hor_sampling_factor_numerator="3"
                                hor_sampling_factor_denominator="2"
                                vert_sampling_factor_numerator="3"

```

```

        hor_sampling_factor_denominator="2">
    </ScalabilityLayerShape>
</ScalabilityLayer>

    </VideoObjectLayer>
</VideoObject>
</VisualObject>
</VisualObjectSequence>
</Visual>

```

Σχήμα 6–10. Περιγραφή του αντικειμένου της οπτικοακουστικής παρουσίασης

Οι περιγραφείς των αντικειμένων συνδέουν τα παραπάνω αντικείμενα μέσω με τη σκηνή και περιέχουν επιπλέον πληροφορία για τα αντικείμενα, όπως φαίνεται στο παρακάτω σχήμα με τους περιγραφείς του βίντεο της καθηγήτριας και της οπτικοακουστικής παρουσίασης. Ο περιγραφέας του αντικειμένου βίντεο της καθηγήτριας περιέχει και ένα στοιχείο του τύπου `ExtensionDescriptor`, το οποίο περιγράφει το περιεχόμενο του αντικειμένου με μια φράση. Και οι δύο περιγραφείς

```

<!-- Περιγραφείς αντικειμένων που θα χρησιμοποιηθούν για την αναφορά σε
ροές δεδομένων -->

<!-- Τύπος εγγράφου -->
<!DOCTYPE avdl SYSTEM "http://www.music.tuc.gr/~georgia/AVDL/avdl.dtd">

<!-- Προφίλ εγγράφου -->
<!ENTITY % MPEG4ObjectDescriptors.Description "INCLUDE">

<!-- Εξωτερικές ταυτότητες αντικειμένων στα οποία αναφέρονται οι
περιγραφείς -->

<!-- Εγγραφο που περιέχει το αντικείμενο βίντεο της καθηγήτριας -->
<!ENTITY TeacherVideoDoc PUBLIC
http://www.music.tuc.gr/~georgia/AVDL/teachervideo.avdl >

<!-- Εγγραφο που περιέχει το αντικείμενο βίντεο της παρουσίασης -->
<!ENTITY PresentationVideoDoc PUBLIC
http://www.music.tuc.gr/~georgia/AVDL/presentationvideo.avdl >

<!-- Ταυτότητα του αντικειμένου βίντεο της καθηγήτριας -->
<ExternalId id="Video1">
  <DocRef doc_ref=TeacherVideoDoc> </DocRef>
</ExternalId>

<!-- Ταυτότητα του αντικειμένου βίντεο της παρουσίασης -->
<ExternalId id="Video2">
  <DocRef doc_ref=PresentationVideoDoc> </DocRef>
</ExternalId>

<avdl> <mpeg4>

<!-- Περιγραφέας για το βίντεο της καθηγήτριας -->

<ObjectDescriptor object_descriptor_id="OID1">
  <ES_Descriptor ES_id="Video1">
    <DecoderConfigDescriptor stream_type="visual" max_bitrate="120 kbps"
avg_bitrate="64 kbps">

```

```

    <ObjectProfileIndication profile_indication="main_visual">
    </ObjectProfileIndication>
  </DecoderConfigDescriptor>
  <ExtensionDescriptor description_data="The video of the teacher in the
distant learning classroom"> </ExtensionDescriptor>
  <LanguageDescriptor language="Greek"> </LanguageDescriptor>
</ES_Descriptor>
</ObjectDescriptor>

<!-- Περιγραφέας για το βίντεο της οπτικοακουστικής παρουσίασης -->

<ObjectDescriptor object_descriptor_id="OID2">
  <ES_Descriptor ES_id="Video2">
    <DecoderConfigDescriptor stream_type="visual" max_bitrate="2 Mbps"
avg_bitrate="1 Mbps">
      <ObjectProfileIndication profile_indication="main_visual">
      </ObjectProfileIndication>
    </DecoderConfigDescriptor>
    <ExtensionDescriptor description_data="The video of the audiovisual
presentation in the distant learning classroom"> </ExtensionDescriptor>
    <LanguageDescriptor language="Greek"> </LanguageDescriptor>
  </ES_Descriptor>
</ObjectDescriptor>

</mpeg4> </avdl>

```

Η σκηνή αποτελεί τη σύνθεση όλων των αντικειμένων σε μία παρουσίαση και η περιγραφή της με τη γλώσσα AVDL αποτελείται από την ιεραρχία των στοιχείων που περιγράφουν τους κόμβους της. Οι κόμβοι που αντιστοιχούν σε αντικείμενα μέσω, όπως το βίντεο της καθηγήτριας και της οπτικοακουστικής παρουσίασης, περιέχουν αναφορά στον αντίστοιχο περιγραφέα αντικειμένου που βρίσκεται σε ξεχωριστό έγγραφο. Οι δηλώσεις του εξωτερικού εγγράφου με τους περιγραφείς περιέχονται στην αρχή της περιγραφής της σκηνής.

```

<!-- Περιγραφή της σκηνής -->

<!-- Τύπος εγγράφου -->
<!DOCTYPE avdl SYSTEM "http://www.music.tuc.gr/~georgia/AVDL/avdl.dtd">

<!-- Προφίλ εγγράφου -->
<ENTITY % MPEG4Scene.Description "INCLUDE">

<!-- Έγγραφο που περιέχει τους περιγραφείς στους οποίους αναφέρονται οι
κόμβοι της σκηνής -->
<ENTITY DescriptorsDoc PUBLIC
http://www.music.tuc.gr/~georgia/AVDL/descr1.avdl >

<!-- Εξωτερικές ταυτότητες περιγραφών που χρησιμοποιούνται στη σκηνή -->

<!-- Ταυτότητα περιγραφέα για το φόντο -->
<ExternalId id="Background">
  <DocRef doc_ref=DescriptorsDoc> </DocRef>
</ExternalId>

<!-- Ταυτότητα περιγραφέα για τη φωνή της καθηγήτριας -->
<ExternalId id="Teacher voice">
  <DocRef doc_ref=DescriptorsDoc> </DocRef>

```

```

</ExternalId>

<!-- Περιγραφή των κόμβων της σκηνής -->

<avdl> <mpeg4>

<CommandFrame>
<BIFSCfg pixel_metric="TRUE" has_size="TRUE">
    <Size pixel_width="600" pixel_height="600"></>
</BIFSCfg>

<Command>
<SceneReplacement>
<Scene id="DLScene" >
    <SceneConfig>
        <IsCommandScene pixel_metrics="TRUE"> </>
        <SceneSize pixel_width="600" pixel_height="600"> </>
    </SceneConfig>
</TopNode>

    <!-- Αντικείμενα για την καθηγήτρια -->

    <Group2D DEF="Teacher">
        <Group2D.Children>
            <BBoxCenter> <Vector2f f1="120" f2="320"></> </BBoxCenter>
            <BBoxSize> <Vector2f f1="160" f2="320"> </BBoxSize>

            <!-- Φωνή της καθηγήτριας -->
            <Sound2D>
                <AudioSource startTime="0" stopTime="299">
                    <ObjectDescriptorRef od_ref="Teacher voice">
                        </ObjectDescriptorRef>
                </AudioSource>
            </Sound2D>

            <!-- Βίντεο της καθηγήτριας -->
            <VideoObject2D>
                <ObjectDescriptorRef od_ref="OID1">
                    </ObjectDescriptorRef>
            </VideoObject2D>

        </Group2D.Children>
    </Group2D>

    <!-- Δυσδιάστατο φόντο της σκηνής -->
    <Background2D>
        <ObjectDescriptorRef od_ref="Background" >
            </ObjectDescriptorRef>
    </Background2D>

    <!-- Επιπλα -->
    <Layer3D>
        <Layer3D.Children>

        </Layer3D.Children>
    </Layer3D>

    <!-- Οπτικοακουστική παρουσίαση -->
    <Group2D DEF="AudioVisual Presentation">
        <Group2D.Children>
            <BBoxCenter> <Vector2f f1="280" f2="200"></> </BBoxCenter>
            <BBoxSize> <Vector2f f1="300" f2="200"> </BBoxSize>

            <!-- Ήχος οπτικοακουστικής παρουσίασης -->

```

Comment [Γ.Κ.22]: Εδώ μπαίνουν γραφείο και υδρόγειος σφαίρα

```

        <Sound2D>
        <AudioSource startTime="300" stopTime="480">
        <ObjectDescriptorRef od_ref=" ">
        </ObjectDescriptorRef>
        </AudioSource>
        </Sound2D>
    </Group2D.Children>
</Group2D>

</TopNode>

</Scene>
</SceneReplacement>
</Command>
</CommandFrame>

</mpeg4> </avdl>

```

Σχήμα 6-11. Περιγραφή σκηνής της εφαρμογής εκπαίδευσης από απόσταση

Το περιεχόμενο του δυσδιάστατου φόντου περιγράφεται από τον κόμβο Background2D που φαίνεται στο σχήμα. Ο τύπος του κόμβου καθορίζεται στο Κεφάλαιο 3, ενότητα 3.9.3.3 και αποτελείται από ένα στοιχείο που περιέχει την ταυτότητα του περιγραφέα αντικειμένου που δείχνει στη ροή των δεδομένων της εικόνας του φόντου.

Όσον αφορά στο ακουστικό αντικείμενο που περιέχει τη φωνή της καθηγήτριας, όπως φαίνεται από τα χαρακτηριστικά startTime και stopTime του στοιχείο AudioSource, αρχίζει από την αρχή παρουσίασης της σκηνής και έχει διάρκεια 5 λεπτά (300 δευτερόλεπτα). Ο ήχος της οπτικοακουστικής παρουσίασης αρχίζει μετά την ομιλία της καθηγήτριας και διαρκεί 3 λεπτά (180 δευτερόλεπτα).

Η βάση δεδομένων του περιεχομένου περιέχει τις οντότητες που προκύπτουν από τη μετατροπή των στοιχείων της παραπάνω περιγραφής με βάση τα εργαλεία που παρουσιάστηκαν στο κεφάλαιο 5. Για καθένα από τα αντικείμενα μέσω των δημιουργείται μια οντότητα στη βάση, η οποία σχετίζεται με ένα σύνολο άλλων οντοτήτων οι οποίες προκύπτουν από τη μετατροπή των στοιχείων που περιέχουν τα αντικείμενα. Η σκηνή αποθηκεύεται ως ένα σύνολο οντοτήτων κόμβων που συνδέονται με την οντότητα της σκηνής αλλά και μεταξύ τους με ιεραρχική οργάνωση.

Οι δομές δεικτοδότησης κατασκευάζονται για να παρέχουν τυχαία προσπέλαση στην κωδικοποιημένη πληροφορία των αντικειμένων και αποτελούνται από:

- Μία δομή δεικτοδότησης για το αντικείμενο της καθηγήτριας, η οποία παρέχει τυχαία προσπέλαση στα Επιπέδα Αντικειμένου του βίντεο.
- Μία δομή δεικτοδότησης για το βίντεο της οπτικοακουστικής παρουσίασης που παρέχει τυχαία προσπέλαση στα Επίπεδα και τις βαθμίδες της παρουσίασης.

- Μία έκδοση γρήγορης επισκόπησης που αποτελείται από ένα υποσύνολο των Επιπέδων της παρουσίασης μαζί με το δείκτη της και τη δομή αντιστοίχισής της με την κανονική ροή.

Στις επόμενες υποενότητες περιγράφεται η υποστήριξη της λειτουργικότητας της εφαρμογής με βάση το περιεχόμενό της, την αναπαράστασή του, τη βάση δεδομένων και τις δομές δεικτοδότησης.

6.6.1.1 Διαβαθμισμένη μεταβίβαση

Η διαβαθμισμένη μεταβίβαση περιεχομένου αφορά στην εκτέλεση ερωτήσεων για το περιεχόμενο και το προφίλ των αντικειμένων. Οι ερωτήσεις αυτές αφορούν στην πολυπλοκότητα του αποκωδικοποιητή, σε παραμέτρους που θα συγκριθούν με τις δυνατότητες του τερματικού και στο προφίλ του χρήστη. Μπορούν να συνοψιστούν στα εξής:

- Ερώτηση για τον αριθμό των αντικειμένων βίντεο που περιέχονται στη σκηνή.
- Ερώτηση για το ρυθμό κάθε αντικειμένου βίντεο.
- Ερώτηση για το μέγιστο ρυθμό των ροών που παρουσιάζονται ταυτόχρονα στη σκηνή και για το μέγεθος της μνήμης του αποκωδικοποιητή.
- Ερώτηση για τη γλώσσα ομιλίας της καθηγήτριας.

Οι ερωτήσεις στη βάση δεδομένων που μπορούν να απαντήσουν στα παραπάνω, φαίνονται στο παρακάτω σχήμα μέσα σε μία συνδιαλλαγή:

```
<!-- Συνδιαλλαγή ανάκτησης πληροφορίας για το προφίλ της σκηνής και των
αντικειμένων της -->
BEGIN TRANSACTION

SELECT      COUNT(*)
FROM        VideoObject2D
WHERE       where VideoObject2D.scene="DLScene"

SELECT      DecoderConfigDescriptor.max_bitrate,
DecoderConfigDescriptor.avg_bitrate
FROM        VideoObject2D, ObjectDescriptor, ES_Descriptor
WHERE       VideoObject2D.scene="DLScene" and
            VideoObject2D.od_ref=ObjectDescriptor.id and
            ES_Descriptor.od_ref=ObjectDescriptor.id

COMMIT TRANSACTION
```

6.6.1.2 Περιήγηση και ελεγχόμενη παρουσίαση

Η περιήγηση και ελεγχόμενη παρουσίαση της σκηνής υποστηρίζεται από την κατασκευή δομών δεικτοδότησης των αντικειμένων και της σκηνής. Η κατασκευή δομών

δεικτοδότησης για το βίντεο της καθηγήτριας και της οπτικοακουστικής παρουσίασης επιτρέπει την προσπέλαση σε τυχαία χρονικά σημεία του περιεχομένου. Οι δομές δεικτοδότησης του βίντεο μπορούν να συνδυαστούν με κατάλληλες δομές για τον ήχο, ώστε να υπάρχει αντιστοίχιση των λειτουργιών τυχαίας προσπέλασης και στις δύο ροές.

Η δομή δεικτοδότησης της σκηνής μπορεί να χρησιμοποιηθεί για τυχαί προσπέλαση σε ολόκληρη τη σκηνή, όπου ο δείκτης χρησιμοποιείται για την ανάκτηση του τμήματος της ροής που περιέχει την πληροφορία για την τυχαία χρονική στιγμή.

6.7 Ανακεφαλαίωση

Στο κεφάλαιο αυτό παρουσιάστηκε η υποστήριξη των εφαρμογών οπτικοακουστικού περιεχομένου με χρήση της γενικής υποδομής διαχείρισής του σε οποιοδήποτε πληροφοριακό σύστημα. Η υποστήριξη των εφαρμογών επιτυγχάνεται μέσα από τις λειτουργίες που παρέχουν διαχείριση και έλεγχο της παρουσίασης του περιεχομένου, ανάκτηση βάσει περιεχομένου, διαβαθμισμένη μεταβίβασή του και ανταλλαγή περιεχομένου στον Παγκόσμιο Ιστό. Οι υπηρεσίες αυτές υποστηρίζονται με ενιαίο τρόπο από το πληροφοριακό σύστημα και στηρίζονται στην περιγραφή του περιεχομένου με τη γλώσσα AVDL, στη βάση δεδομένων, στις δομές δεικτοδότησης της πληροφορίας και στα εργαλεία που έχουν περιγραφεί για τη διαχείρισή τους. Επιπλέον η λειτουργικότητα που υποστηρίζεται περιγράφεται με συνδιαλλαγές της σχεσιακής βάσης δεδομένων. Οι υπηρεσίες υποστήριξης των εφαρμογών, όπως και τα βασικά τμήματα της υποδομής είναι επεκτάσιμα, ώστε να μπορούν να εμπλουτιστούν με επιπλέον πληροφορία αλλά και μεθόδους διαχείρισής της.

Επίσης στο κεφάλαιο αυτό παρουσιάστηκε ένα παράδειγμα εφαρμογής εκπαίδευσης από απόσταση, όπου έγινε πλήρης περιγραφή του περιεχομένου της και των υπηρεσιών παρουσίασης, μεταβίβασης και ανάκτησης του περιεχομένου.

Κεφάλαιο 7

Ανακεφαλαίωση

Έχοντας ως στόχο την υποστήριξη υπηρεσιών διαχείρισης οπτικοακουστικής περιεχομένου στον πυρήνα ενός πληροφοριακού συστήματος με ενιαίο και ευρέως αποδεκτό τρόπο κατάλληλο για μεγάλο εύρος εφαρμογών, η παρούσα διατριβή αποτελείται από την ανάπτυξη μηχανισμών αναπαράστασης, δεικτοδότησης, αποθήκευσης και διαχείρισης περιεχομένου. Οι μηχανισμοί αυτοί υποστηρίζουν έναν πυρήνα λειτουργικότητας στις εφαρμογές οπτικοακουστικής πληροφορίας για τον έλεγχο της παρουσίασης του περιεχομένου και την περιήγηση σε αυτό, για την ανάκτηση βάσει περιεχομένου, τη διαβαθμισμένη μετάδοσή του ανάλογα με το προφίλ του χρήστη, του δικτύου και της εφαρμογής και την ανταλλαγή του σε ετερογενή περιβάλλοντα όπως αυτό του Παγκόσμιου Ιστού.

Το είδος του οπτικοακουστικού περιεχομένου στο οποίο εστιάζει η διατριβή αποτελείται από κωδικοποιημένα δεδομένα των προτύπων MPEG-4 και MPEG-2. Με αυτόν τον τρόπο εξασφαλίζεται η δυνατότητα υποστήριξης εφαρμογών με μεγάλο εύρος απαιτήσεων ως προς τη δίκτυο μετάδοσης, το τερματικό και τη λειτουργικότητα της εφαρμογής.

Η αναπαράσταση του οπτικοακουστικού περιεχομένου που είναι κωδικοποιημένο με τα πρότυπα MPEG-4 ή MPEG-2 αποτελεί βασικό τμήμα της γενικής υποδομής υποστήριξης

υπηρεσιών διαχείρισης οπτικοακουστικού περιεχομένου και επιτυγχάνεται με τη Γλώσσα Περιγραφής Οπτικοακουστικού Περιεχομένου ή AVDL (Audio Visual Description Language). Η γλώσσα AVDL περιγράφει τη δομή και τη σημασιολογία του περιεχομένου με απλό και ευέλικτο τρόπο, ο οποίος στηρίζεται στη μεταγλώσσα αναπαράστασης εγγράφων του Παγκόσμιου Ιστού XML (eXtensible Markup Language). Η χρήση της XML εξασφαλίζει την αναπαράσταση του περιεχομένου με ευρέως αποδεκτό τρόπο, ο οποίος επιτρέπει την ανταλλαγή του σε οποιοδήποτε περιβάλλον συμπεριλαμβανομένου και του Παγκόσμιου Ιστού.

Η αναπαράσταση ή περιγραφή του περιεχομένου των MPEG-4 και MPEG-2 με την γλώσσα AVDL, χρησιμοποιείται στον πυρήνα του πληροφοριακού συστήματος εξυπηρέτησης σε συνδυασμό με την κατασκευή βάσεων δεδομένων για την αποθήκευση οπτικοακουστικού περιεχομένου. Η βάση δεδομένων οπτικοακουστικού περιεχομένου αποτελεί το βασικό μέρος της υποδομής που παρέχει αποθήκευση του περιεχομένου και ανάκτησή του βάσει των ιδιοτήτων του για την υποστήριξη των λειτουργιών ελεγχόμενης παρουσίασης, ανάκτησης βάσει περιεχομένου και διαβαθμισμένης μεταβίβασης. Το μοντέλο δεδομένων της βάσης επιτρέπει την αποθήκευση των εγγράφων και στοιχείων της γλώσσας AVDL με ομοιόμορφο τρόπο και υποστηρίζει τη μετατροπή της πληροφορίας από τη γλώσσα σε οντότητες της βάσης δεδομένων. Το μοντέλο της βάσης δεδομένων υποστηρίζει επίσης τους μηχανισμούς μετατροπής του περιεχομένου από την αποθηκευμένη μορφή του στη γλώσσα AVDL, εξασφαλίζοντας έτσι την ανταλλαγή του ανεξάρτητα από το πληροφοριακό σύστημα διαχείρισης.

Η διαχείριση του MPEG-2 περιεχομένου περιλαμβάνει και την περίπτωση του εξυπηρετητή πολυμέσων ΚΥΔΩΝΙΑ-D. Η διαχείριση αυτή αποτελείται από τη μοντελοποίηση του περιεχομένου στην αντικειμενοστρεφή βάση δεδομένων του συστήματος, τη δεικτοδότηση του περιεχομένου και την υποστήριξη της Υπηρεσίας Περιεχομένου του DAVIC. Η αποθήκευση του περιεχομένου στη βάση δεδομένων του συστήματος το αντιμετωπίζει με ενιαίο τρόπο με όλα τα αντικείμενα της βάσης και γίνεται μέσα από συνδιαλλαγές που εξασφαλίζουν τη συνέπεια των δεδομένων. Η δεικτοδότηση του περιεχομένου εξασφαλίζει την αποδοτική υποστήριξη των υπηρεσιών περιήγησης στο περιεχόμενο του MPEG-2. Η Υπηρεσία Περιεχομένου του DAVIC παρέχει στους ιδιοκτήτες του περιεχομένου έναν τρόπο εγκατάστασής του στον απομακρυσμένο εξυπηρετητή για να χρησιμοποιηθεί από τις εφαρμογές, αλλά και διαχείρισής του.

Τέλος παρουσιάστηκε η υποστήριξη των εφαρμογών οπτικοακουστικού περιεχομένου με χρήση της γενικής υποδομής διαχείρισής του σε οποιοδήποτε πληροφοριακό σύστημα. Η υποστήριξη των εφαρμογών επιτυγχάνεται μέσα από τις λειτουργίες που παρέχουν διαχείριση και έλεγχο της παρουσίασης του περιεχομένου, ανάκτηση βάσει περιεχομένου, διαβαθμισμένη μεταβίβασή του και ανταλλαγή περιεχομένου στον Παγκόσμιο Ιστό.

Βιβλιογραφία

- [1] S. Christodoulakis, N. Pappas, G. Kyriakaki, Y. Maragoudakis, Y. Mavarganis, C. Tsinaraki, *The KYDONIA multimedia information server*, Multimedia Application Services and Techniques - ECMAS'T'97, Proc. of 2nd European Conference, Milan 1997.
- [2] C. Brandt, G. Kyriakaki, W. Lamotte, R. Roling, et. Al., *The SICMA Multimedia Server and Virtual Museum Application*, Multimedia Application Services and Techniques - ECMAS'T'98, Proc. of 3rd European Conference, 1998.
- [3] ISO/IEC JTC1/SC29/WG11 N1909, *MPEG-4 Overview*, October 1997, Fribourg.
- [4] ISO/IEC 14496-2 Committee Draft, *Information Technology – Coding of Audio-Visual Objects: Visual*, March 1998, Tokyo.
- [5] ISO/IEC JTC1/SC29/WG11 N2201, *MPEG-4 Systems*, Text for ISO/IEC FCD 14496-1 Systems, May 1998.
- [6] O.Avaro, C.Herpel, J.Signes, *Tutorial on MPEG-4 Systems*, May 1998, <http://garuda.imag.fr/MPEG4/syssite/syspub/docs/tutorial>
- [7] ISO/IEC JTC1/SC29/WG11 N2323, *Overview of the MPEG-4 Standard*, July 1998, Dublin.
- [8] ISO/IEC JTC1/SC29/WG11 N2459, *Overview of the MPEG-4 Standard*, October 1998, Atlantic City.
- [9] ISO/IEC JTC1/SC29/WG11 N2527, *MPEG Systems (1-2-4-7) FAQ, Version 7.0a*, October 1998.
- [10] ISO/IEC JTC1/SC29/WG11 MPEG98/N2457, *MPEG-4 Applications*, October 1998, Atlantic City.
- [11] *MPEG-4 Standard Promoted to Final Committee Draft*, February 1998.
- [12] Rob Koenen, "MPEG-4 Multimedia for our time", IEEE Spectrum, vol. 36 No 2, 1999.
- [13] O.Avaro, P.A.Chou, A.Eleftheriadis, C.Herpel, C.Reader, J.Signes, *The MPEG-4 Systems and Description Languages: A Way Ahead in Audio Visual Information Representation*, 1996.
- [14] R.Koenen, *MPEG-4: The Representation Framework for a New Content Paradigm*, AccenTS Art.1, Vol.2, February 1997.
- [15] Thomas Sikora, *MPEG-4 Video Standard for Multimedia Applications*, 1997.
- [16] *Overview of Flavor*, <http://www.ee.columbia.edu/flavor> .
- [17] J.Signes, Y.Fisher, A.Eleftheriadis, *MPEG-4 Scene Representation and Interactivity*, <http://flavor.ee.columbia.edu/docs/BIFS.pdf> .
- [18] *MPEG Systems (1-2-4) FAQ Version 4.0*, October 1997.
- [19] ISO/IEC 14772-1:1998, *The Virtual Reality Modeling Language*.
- [20] J.Hartman, J.Wernecke, *The VRML 2.0 Handbook*, Addison Wesley Dev. Press, 1996, ISBN 0-201-47944-3.
- [21] F.Pereira, *MPEG-7: A Standard for Content-Based Audiovisual Description*, .
- [22] ISO/IEC JTC1/SC29/WG11 N2207, *MPEG-7 Context and Objectives*, March 1998.
- [23] ISO/IEC JTC1/SC29/WG11 N2209, *MPEG-7 Applications*, March 1998.
- [24] ISO/IEC JTC1/SC29/WG11 N1686, *Report on MPEG-7 Seminar in Bristol*, April 1997.

- [25] ISO/IEC JTC1/SC29/WG11 N2326, *MPEG-7 Context and Objectives*, July 1998. ISO/IEC JTC1/SC29/WG11 N2327, *MPEG-7 Requirements Document V.6*, July 1998.
- [26] ISO/IEC JTC1/SC29/WG11 N2228, *MPEG-7 Applications Document v.6*, July 1998.
- [27] ISO/IEC JTC1/SC29/WG11 N2461, *MPEG-7 Requirements Document v.7*, October 1998
- [28] *Extensible Markup Language (XML) 1.0*, <http://www.w3.org/TR-REC-xml-19980210>
- [29] *Vector Markup Language (VML)*, <http://www.w3.org/TR/NOTE-VML-19980513>.
- [30] OMF Developer's Desk, *Open Media Framework™ Interchange Specification*, Version 2.0, Avid Technology Inc, November 1995.
- [31] J.Harris, I.Ruben, *Bento Specification*, Revision 1.0d5, Apple Computer Inc, July 1993.
- [32] A.Silnerschatz, H.F.Korth, Sudarshan, *Database System Concepts*, 3rd edition, McGraw Hill, ISBN 0-07-114810-8, 1997.
- [33] *The Object Database Standard: ODMG-93*, edited by R. Catell, Release 1.1, Morgan Kaufman, 1994, ISBN 1-55860-302-6.