

# A Schema-based Selective Rendering Framework

Alexandros Zotos

Technical University of Crete, Greece

[azotos@ced.tuc.gr](mailto:azotos@ced.tuc.gr)

Katerina Mania

Technical University of Crete, Greece

[k.mania@ced.tuc.gr](mailto:k.mania@ced.tuc.gr)

Nicholaos Mourkoussis

University of Sussex, UK

[n.mourkoussis@gmail.com](mailto:n.mourkoussis@gmail.com)

## Abstract

Perception principles have been incorporated into rendering algorithms in order to optimize rendering computation and produce photorealistic images from a human rather than a machine point of view. In order to economize on rendering computation, selective rendering guides high level of detail to specific regions of a synthetic scene and lower quality to the remaining scene, without compromising the level of information transmitted. Scene regions that have been rendered in low and high quality can be combined to form one complete scene. Such decisions are guided by predictive attention modeling, gaze or task-based information. We propose a novel selective rendering approach which is task and gaze-independent, simulating cognitive creation of spatial hypotheses. Scene objects are rendered in varying polygon quality according to how they are associated with the context (schema) of the scene. Experimental studies in synthetic scenes have revealed that consistent objects which are expected to be found in a scene can be rendered in lower quality without affecting information uptake. Exploiting such expectations, inconsistent items which are salient require a high level of rendering detail in order for them to be perceptually acknowledged. The contribution of this paper is an innovative x3D-based selective rendering framework based on memory schemas and implemented through metadata enrichment.

## 1. Background

Computer graphics algorithms have for long dealt with simulation of physics: simulation of the geometry of a real-world space, simulation of the light propagation in a real environment and simulation of motor actions with appropriate tracking [Brooks 1999]. Perception principles have subsequently been incorporated into rendering algorithms [McNamara 2001], in order to save rendering computation, mainly following the generic idea of ‘do not render what we can not see’ [Marmitt and Duchowski 2002; Itti et al. 1998; Loftus and Mackworth 1978]. However, with Virtual Environment (VE) simulator technologies aiming at simulating real-world task situations, the research community is challenged to produce a much more complex system. We do not necessarily require accurate simulation of physics to induce reality. Much less detail is often adequate [Watson et al. 1997; Ferwerda 2001; Mania and Chalmers 2001; Mania et al. 2004; Mania and Robinson 2004].

In order to economize on rendering computation, previous research dealing with interactive synthetic scenes has been focused on either rendering in high quality the 2-3 degrees foveal region of vision and with less detail the periphery of vision based on gaze information [McConcie and Loschky 1997], or rendering in high quality the foveal area based on a-priory knowledge of the viewer’s task focus [Cater et al. 2003; Sundstedt et al. 2008]. Gaze-dependant rendering encounters difficulties of maintaining display updates free of visual artefacts after a fast (~4ms) eye saccade. Such processes are quite computationally demanding, however, if the speed gaze-to-rendering issue is resolved, task

performance results are indistinguishable to a fully fledged, high resolution real-time environment. It has also been proposed to assign selective high quality rendering in the visual angle of the fovea ( $2^\circ$ ) centred on the users’ task focus [Cater et al. 2003; Sundstedt et al. 2008]. This approach, however, cannot be applied when there is no overt task to be conducted. Moreover, there is no acceptable model of comparing or predicting task-relevant saccades. Following a different approach, Haber et al. 2001 suggested rendering the informative areas of a scene in varying quality based on saliency models. Such models aim to predict the visual features that involuntarily attract visual attention such as object edges, sudden colour changes or movements. It was proposed that the most noticeable areas as derived from saliency modelling should be rendered in higher quality. Bottom-up visual attention models are not shown to predict attention regions successfully [Marmitt and Duchowski 2002]. Correlation between actual human and computationally-derived scan-paths was found to be much lower than predicted when carrying out a real-world task such as making a cup of tea [Land 1999].

A comprehensive approach should be task and gaze-independent, simulating cognitive processes rather than predicting attention employing bottom-up processes such as saliency models. In this paper, we will present a novel selective rendering approach which exploits existing research on memory ‘schemata’ which could ultimately guide selective rendering based on spatial cognition processes. Schemata are knowledge structures based on the notion that an individual’s prior experience will influence how he or she perceives, comprehends and remembers new information. When participants are exposed to a large amount of information in a scene, cognitive psychologists have suggested that schemata are used to guide the search for information in memory [Brewer & Treyens, 1981]. A general premise derived from this research is that information which is not related to the schema being used in retrieval will be harder to recall than information which is schema related. In terms of real world scenes, schemata represent the general meaning of a scene such as ‘office’, ‘theatre’ etc. Schemata influence memory of the objects in a given context according to their association with the schema in place. When being exposed to a synthetic environment, similar information should be transmitted between the simulated scene and the real-world scene, both depicting a specific schema. This would, in due course, indicate which objects or areas in a synthetic scene could be rendered in lower quality without affecting information uptake but at the same time reducing computational complexity. Such a paradigm will be presented in this paper and an implementation framework will be proposed.

## 2. THE ROLE OF SCHEMATA IN MEMORY

We present a novel selective rendering framework which is exploiting classic findings from memory research in which schemata influence memory retrieval. Schemata are knowledge structures of cognitive frameworks drawn from experience that

facilitate the interpretation of events. Memory schemata were introduced to psychology by the work of Barlett 1932. A scene is thought to be comprised by consistent objects, e.g. objects expected to be found in a specific context and inconsistent objects, e.g. objects that are salient or not normally existent in a given scene.

Schema theory proposes that memory for objects in scenes is influenced by the degree to which the objects appear to ‘belong’ to a particular scene (consistent objects – i.e., a computer in an office) or not (inconsistent objects – i.e., a skull in an office). However, the degree to which an object will be consistent with a scene is likely to depend on whether the perception of the scene is itself consistent with our previous experience. Thus, for example, if a familiar scene is presented in an unfamiliar manner, the effect of object consistency on memory is likely to change. Our goal is to investigate whether this kind of effect can be used to investigate the degree to which a virtual environment (VE) is perceived as “normal”, i.e. consistent with our usual experience of similar environments, or “abnormal”, i.e. markedly different from our previous experience.

Whether memory performance is better for present consistent objects, or present inconsistent objects, is where experimental studies tend to disagree. In a comprehensive meta-analysis, Rojahn & Pettigrew 1992 conducted a review of 60 independent studies and concluded that there is a small difference between the number of studies supporting the consistency effect which advocates better memory performance for schema-consistent objects and the inconsistency effect which reflects better memory performance for schema-inconsistent objects. Additionally, they reported that the contradictory results could be explained by: a) the different memory measures used; and b) the existence of several moderating variables in the design of the experiments such as the difficulty of the task and the number of the inconsistent objects present.

Flannery and Walles 2003 investigated how schema theories apply to real versus virtual memories. Participants were instructed to explore either a virtual or a similar real environment for 20 seconds, without prior knowledge that their memory of the space would be subsequently assessed. Participants then completed a recognition task. Recognition scores revealed that participants had better recognition for consistent objects, but were more confident for the recognition of the inconsistent objects.

Previous work [Mania and Robinson 2003; Mania et al. 2005] included a preliminary investigation of the effect of object type (consistent vs. inconsistent) and shadows (flat-shaded scene vs. radiosity scene) on object memory recognition in a VE. The computer graphics simulation was displayed on a Head Mounted Display (HMD) utilizing stereo imagery and head tracking. Thirty-six participants across three conditions of varied rendering quality of the same space were exposed to the computer graphics environment and completed a memory recognition task. The high-quality and mid-quality conditions included a pre-computed radiosity simulation of an academic’s office (with 80% and 40%

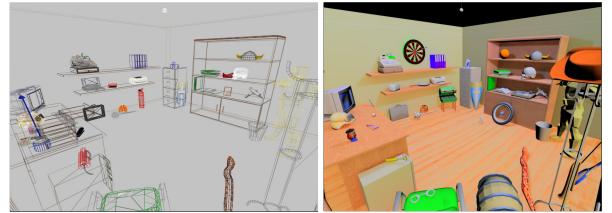


Figure 1. Schema memory experimental studies [Troscianko et al. 2007, Mourkoussis et al. 2009].

radiosity iterations computed respectively). The low-quality condition consisted of a flat-shaded version of the same office. Results revealed that schema consistent elements of the scene were more likely to be recognized than inconsistent information. Overall, higher confidence ratings were assigned to consistent rather than inconsistent items. Total object recognition was better for the scene including rough shadows (mid-quality condition) compared to the flat-shaded scene. The presence of accurate shadow information, though, did not affect recognition of consistent or inconsistent objects, therefore lower quality of rendering was adequate for better memory recognition of consistent objects. This study was limited to the investigation of subtle shadow variations. Another experimental study employed a more extreme set of rendering types: wireframe with added color, and full radiosity [Troscianko et al. 2007, Mourkoussis et al. 2009] (Figure 1). The proportion of inconsistent/consistent objects was varied, and object recognition tests ensured that all objects were easily recognized in all conditions. The results showed a significant interaction between rendering type, object type, and consistency ratio. This suggests that inconsistent objects are only preferentially remembered if the scene looks “normal” or if there are many such objects in an “abnormal” scene such as in the wireframe condition. It was also shown that memory performance is better for the inconsistent objects in the radiosity rendering condition compared to the wireframe condition. We conclude that memory for objects can be used to assess the degree to which the context of a VE appears close to expectations.

Despite contradictory results in literature as detailed above, it seems that perceptual information can be complemented by involuntary knowledge based on past experience. Experimental studies in synthetic scenes have revealed that consistent objects which are expected to be found in a scene can be rendered in lower quality without affecting information uptake taking advantage of such expectations, whereas inconsistent items which are salient would require a high level of rendering detail in order for them to be perceptually acknowledged [Mania et al. 2005]. Therefore, by exploiting schema theory, it is possible to reduce computational complexity, producing scenes from a cognitive point of view without affecting information uptake and resulting in an entirely novel and interdisciplinary approach which is gaze, task and saliency-model independent.

### 3. THE SELECTIVE RENDERER FRAMEWORK

We present a novel selective rendering system that will exploit schema theory by identifying the perceptual importance of scene regions. Objects that have been rendered in low and high quality are incorporated in a scene based on schema expectations. The

rendered quality of these objects will change in real time, dependent on user navigation and interaction.

The **Selective Renderer Software Engineering Framework** has one major objective, to economize on rendering computation. This will be accomplished by initially producing objects of varying polygon count based on whether an object is consistent with the content of the scene. We assign high quality of polygon count to inconsistent objects in the scene, e.g. objects which are salient in a given scene and lower quality to consistent objects, e.g. objects which are expected to be found in a given context.

The **Selective Renderer Framework** (Figure 2) is responsible of modifying and enriching an X3D document with appropriate metadata information. Such a document describes the geometry of a given scene in relation to whether an object is consistent or inconsistent. It exploits this type of information in real-time in order to apply differing polygon quality to specific scene's regions. The architecture of the selective renderer module comprises of two main components: the **Selective Renderer Preprocess** component and the **Selective Renderer Real Time** component. Input to the system is an X3D document that describes the geometry of the scene. Validation and modification of the X3D document is achieved via the Scene Access Interface (SAI) which is a JAVA API that provides all the appropriate functions for the processing of the document.

The **Selective Renderer Preprocess** component is responsible for the pre-processing and enrichment of the X3D document with congruous metadata information available during the real-time process. During the pre-process phase, varied polygon count versions of the scene's objects are created. Consistent objects which are expected to be found in a scene are rendered in lower quality to the point of being recognizable whereas inconsistent items which are salient would require a high level of rendering. Every X3D node that describes the geometry of an object for which different versions have been produced, is replaced by a switch node that points to varied polygon quality objects through inline nodes. The *WhichChoice* field of every switch node is able to show which version of the object should be rendered. Thus, by altering the value of this field, the appropriate polygon count version of each object model can be loaded in real-time, as the user navigates any 3D scene.

The **Metadata Document Enrichment** stage exploits the X3D's capability of defining metadata information for every node of an X3D document. Metadata information (MetadataSet, MetadataString) describe the level of consistency that every object entails in relation to the specific context of a scene's region. Metadata information is defined for every Proximity node. This information describes which schema(s) is activated when the avatar enters a specific region (*enterTimeSchemaActivation* meta-field) as well as when it exits that region (*exitTimeSchemaActivation* meta-field). Similarly, metadata information is defined for every Switch node. This information describes the level of consistency of each version of the object in relation to the context of the scene (*degreeOfConsistency* meta-field, 3 levels of consistency: low, med, high), the area description that corresponds to the positioning of the objects (*schemaBelonginess* meta-field) and whether or not this object is a

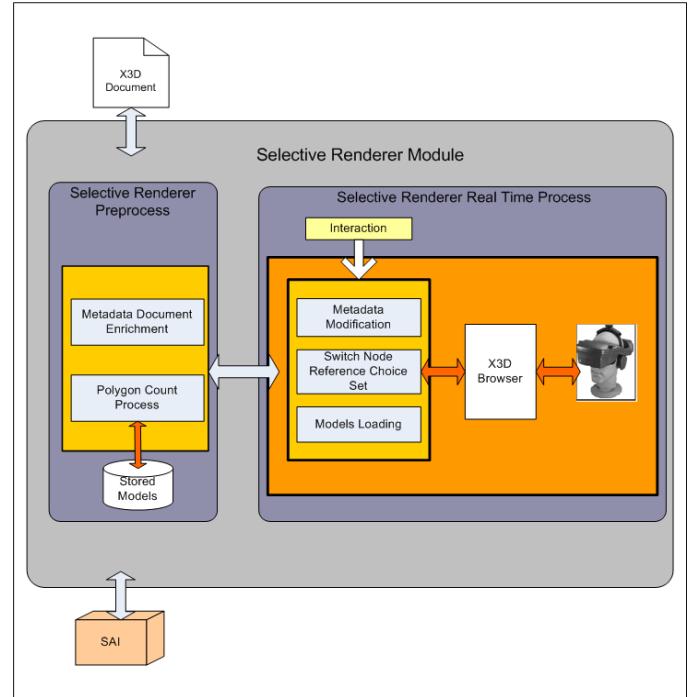


Figure 2. Selective Renderer Architecture

frame object, e.g. wall, ceiling, lights fixture, door, etc (*typeOfConsistency* meta-field).

We assume that a scene (i.e. a house) comprises of areas of varied context such as a kitchen, an office, and other sub-areas. Similarly, these sub-areas can also be associated with sub-schemata: an academic office area could contain a desk sub-area and a bookcase subarea. It is crucial to devise a method that can incorporate information related to the level of consistency of objects in each sub-area, in the X3D input document. At the time that a sub-area is 'activated' as the user navigates through it, information concerning the level of consistency or inconsistency of objects will define the polygon quality version of the objects in that area which is loaded and visualized. A Proximity Node is created for every area and sub-area in order to track user interaction in these regions. Metadata information is attached to every proximity node which describes whether objects are consistent or inconsistent in relation to the context of each sub-area. When the navigation avatar approaches a scene's region, this action is acknowledged and by exploiting relevant metadata information it is feasible to know which knowledge schema should be activated. 'Activation' will occur when the user navigates closely to each area.

The next step of the Metadata Document Enrichment stage is to devise a method that can relate the activated 'memory schema' with the appropriate polygon quality version of each object. The X3D's capability of defining metadata information makes it possible to implement such a strategy. Switch nodes indicate which version of the object should be rendered depending on user navigation. For example, let's assume an area *areaA* that activates the memory schema *schemaA*. This area contains sub-area

*subAreaB* that activates the memory schema *subSchemaB*. An object of the scene *Obj* is consistent in relation to the *area*, however, it could be inconsistent in relation to the *subAreaB* (Figure 3).

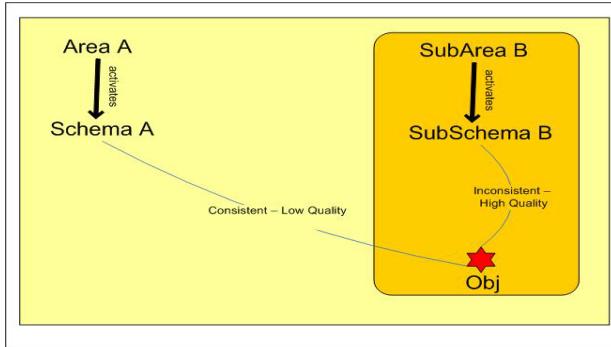


Figure 3: Example of an object that belongs to different schemas with different levels of consistency

The Metadata Document Enrichment component will define the metadata information attached to the proximity nodes that represent *areaA* and *subAreaB* by associating these areas with the memory schema to be activated. So, *areaA* should activate *schemaA* and *subAreaB* should activate *subSchemaB* (Figure 4).

```
<Transform DEF="ProxArea" translation="-0.1614 0.0 -2.053">
  <ProximitySensor DEF="RoomProximity" size="4.613 2.768 4.843">
    <MetadataSet DEF="MetadataSet_Example" containerField="metadata">
      <MetadataString name="enterTimeSchema" value="SubSchemaB"/>
      <MetadataString name="exitTimeSchema" value="SchemaA"/>
    </MetadataSet>
  </ProximitySensor>
</Transform>
```

Figure 4: Metadata example code of a proximity sensor. Schema A is activated when the avatar exits the proximity sensor. SubSchema B is activated at the entrance.

```
<Switch DEF="Obj" whichChoice="0">
  <MetadataSet DEF="MetadataSet_Obj" containerField="metadata">
    <MetadataSet containerField="value" >
      <MetadataString name="consistencyType" value="consistent"/>
      <MetadataString name="degreeOfConsistency" value="low"/>
      <MetadataString name="schemaBelonginess" value="SubSchemaB"/>
    </MetadataSet>
    <MetadataSet containerField="value" >
      <MetadataString name="consistencyType" value="consistent"/>
      <MetadataString name="degreeOfConsistency" value="high"/>
      <MetadataString name="schemaBelonginess" value="SchemaA"/>
    </MetadataSet>
  </MetadataSet>
</Switch>
```

Figure 5: Metadata example code of a switch node. Object *Obj* belongs to *schemaA* with high degree of consistency and to *SubSchemaB* with low degree of consistency

Metadata information is also included in the switch node of the object *Obj* (Figure 5). This information should associate each

version of the object with the appropriate level of consistency in relation to each context. High *degreeOfConsistency* could be related to *schemaA* or low *degreeOfConsistency* could be activated by *subSchemaB*. If an object is consistent, a low quality polygon version of this object will be loaded. If the same object is inconsistent in relation to a different sub-area, then a high quality version of the same object will be loaded. Eventually, by tracking which ‘schema’ or context is activated by user interaction, the appropriate polygon quality version of the object *Obj* is loaded and visualized.

The **Selective Renderer Real-Time Process** component is responsible for the 3D rendering of the synthetic scene as well as the user navigation and the real-time processing of varied polygon quality objects. Rendering and navigation are achieved with the help of the Xj3D browser while users are immersed in the synthetic scene through a stereo capable, head-tracked Head Mounted Display (HMD). At this stage, the previously modified and metadata enriched document of the pre-process stage is exploited. Several methods have been implemented that can track user navigation when a proximity node is activated (*enterTime*, *exitTime*) and update the corresponding metadata information attached to each node. In this manner, decisions are taken concerning the memory schema (context) activated and therefore, which polygon count version of the object should be loaded. Consequently, switch nodes that have been enriched with metadata information relevant to the activated schema are detected. By identifying the degree of consistency of every object of the scene when a sub-area is entered, it is possible to change its level of detail in real-time.

## 4. THE MAIN EXPERIMENT

In order to assess the functional fidelity of the resulting scene in relation to a fully-fledged rendering solution, a scene was designed including several areas and sub-areas (Figure 6) and a memory recognition task was completed. The scene included a kitchen, a lounge and an office area and comprised of consistent and inconsistent objects relevant to each context. The office area included two sub-regions, the desk area and the bookcase area. Similarly, the lounge area included a sofa area. The scene was designed so that it contains objects with differing levels of scene consistency based on the area and sub-area they were placed. The polygon quality of such objects could, therefore, change dynamically. For example, according to pilot studies, a fire extinguisher is consistent in relation to the office schema but it is inconsistent in relation to the desk schema. Thus, this object will be consistent and rendered with low detail when the office area is navigated, however, it will be inconsistent and rendered with high detail when the desk area is approached.

### 4.1 Method

#### 4.1.1 Design

The experiment had two independent variables. The first between-subjects variable was ‘polygon count’ comprised of three levels:

- ‘high poly’ - all objects of the scene were rendered in high quality
- ‘high-low poly’ - consistent objects were rendered in low quality, inconsistent objects were rendered in high quality and objects of varying level of consistency based on sub-area activated were rendered in high quality

- ‘selective poly’ - consistent objects were rendered in low quality, inconsistent objects were rendered in high quality and objects of varying level of consistency based on sub-area activated were rendered either in low or high quality based on their degree of consistency in relation to the sub-area entered.

The second within-subjects independent variable was ‘object type’ comprised of two levels: ‘consistent’ and ‘inconsistent’. The dependant variables were ‘hit rate’, i.e. rate of successful recognition of present objects and ‘false alarm rate’, i.e. rate of successful recognition of absent objects.

According to the group they were assigned to, participants were exposed to an interactive pre-computed radiosity simulation of a synthetic scene in one of the ‘polygon count’ conditions. The scene viewed contained:

- A kitchen area including 23 objects from which 17 were consistent and 6 inconsistent
- A lounge area comprising of a sofa sub-area including 21 objects from which 12 were consistent, 4 inconsistent and 5 of a different level of consistency in relation to the sub-area entered
- An academic office area comprising of a desk sub-area and a book case sub-area including 26 objects from which 9 were consistent, 1 inconsistent and 16 of a different level of consistency in relation to the sub-area entered.



Figure 6: Experimental rooms

#### 4.1.2 Participants

54 participants were recruited belonging to the research population of the Technical University of Crete. They were given course credits for their participation. Age ranged mainly from 18–25. The 54 participants formed 3 balanced groups of 18, corresponding to the three viewing conditions.

#### 4.1.3 Procedure

Participants completed a memory recognition questionnaire after 135 seconds (45 seconds in each area) of exposure to either the ‘high’ or ‘high-low’ or ‘selective poly’ viewing condition. The time of exposure was defined after detailed pilot studies.

The memory recognition questionnaire consisted of 6 sections, e.g. 3 sections containing objects in relation to the kitchen, lounge and academic office areas and 3 sections in relation to the sofa, desk and book case sub-areas. Each section contained of a list of objects which fell under 4 distinct categories:

1. Schema-consistent objects that were present in the scene (e.g. saucepan, toaster).
2. Schema-inconsistent objects that were present in the scene (e.g. shovel).
3. Schema-consistent objects that were absent (e.g. forks).
4. Schema-inconsistent objects that were absent (e.g. kite).

The list of objects was assembled based on an initial pilot study which explored which objects were expected to be found in each area and which were not. 25 participants subsequently ranked the objects on the list. The consistency of each item was rated on a scale from 1 to 6 according to whether each object was expected to be found in each area or not, with 6 being the most expected, and 1 being the least. Based on these ratings, consistent objects were selected from the high end of the scale, and the inconsistent ones from the low end.

#### 4.2 Results

The hit and false alarm rates obtained were used to calculate the signal detection constituents  $d'$  (sensitivity) and  $\beta$  (bias), as outlined in by [T. D. Wickens, 2001], with  $d'$  being the distance between signal and noise distributions in participant responses, and  $\beta$  being the bias towards replying ‘yes’ or ‘no’. The  $d'$  is calculated by transforming hit and false alarm rates into z-scores, which give the position of a score in relation to the number of standard deviations it is above or below the mean. The  $d'$  score equals the hit ratio z-score minus the false alarm ratio z-score:  $d' = z(HIT) - z(FA)$ . The  $\beta$  value is the ratio of the height of the Hit and False Alarm rate distribution curves:  $\beta = h(HIT)/h(FA)$ . These values were calculated for each condition in the experiment, and scores were compared using Analysis of Variance (ANOVA), using  $\log \beta$ , as the  $\beta$  distribution is not statistically normal.

When participants were asked to recognize which items were present in the scene, they were also asked to rate how confident they were of their responses on a scale of 1 to 5. These confidence ratings were subsequently used to correct the recognition responses for possible guessing. This provided three sets of  $d'$  and  $\beta$  results for each area/sub area: one of uncorrected confidence; one of moderate confidence (i.e. corrected so that only recognition scores with corresponding confidence ratings of 3, 4 or 5 were used); and one set of high confidence (i.e. only confidence scores of 5 were used).

The descriptive statistics for the  $d'$  and  $\beta$  scores were calculated. Six tables (one for each area) contain three sets of  $d'$  and  $\beta$ , e.g., uncorrected, moderate and high confidence. Note that the greater the  $d'$  score, the easier it was for participants to distinguish between signal and noise (i.e. recognize correctly), whilst the closer a  $\beta$  score is to 0, the less bias was found in participant responses, with positive scores indicating a bias towards responding ‘no’, and negative scores a bias towards responding ‘yes’. In some areas when objects’ quality changes in real-time such as for the office and lounge area, it is not possible to categorize objects as firmly consistent or inconsistent. The bookcase contained mainly inconsistent objects apart from books.

Table 3. Desk sub-area confidence results

Table 1. Kitchen area confidence results

Kitchen Area					
Consistent Objects					
N	d'un	b'un	d'mod	b'mod	d'high
54	(O=12)	(O=12)	(O=12)	(O=12)	(O=12)
H	0.97(.52)	-0.27(.58)	0.90(.42)	0.11(.54)	0.88(.50)
HL	1.33(.62)	0.10(.59)	1.03(.44)	0.23(.62)	0.95(.47)
S	1.13(.44)	0.12(.63)	0.97(.34)	0.26(.55)	0.82(.47)
T	1.14(.54)	-0.01(.62)	0.97(.40)	0.20(.56)	0.89(.48)
					0.40(.43)
Inconsistent Objects					
N	d'un	b'un	d'mod	b'mod	d'high
54	(O=6)	(O=6)	(O=6)	(O=6)	(O=6)
H	0.96(.64)	0.47(.36)	0.96(.63)	0.52(.37)	0.97(.70)
HL	1.09(.50)	0.59(.34)	1.01(.51)	0.67(.27)	0.96(.53)
S	1.11(.56)	0.45(.50)	1.06(.63)	0.45(.46)	1.03(.65)
T	1.05(.56)	0.50(.40)	1.01(.58)	0.54(.38)	0.99(.62)
					0.64(.28)

$d'$  and  $\beta$  scores and standard deviations (in brackets) as a functional of viewing condition (N= Total Number of Participants, O=total number of objects, H=High Poly, HL=High Low Poly, S>Selective Poly, T=Total, un=uncorrected mod=moderate)

Table 2. Office area confidence results

Office Area					
N	d'un	b'un	d'mod	b'mod	d'high
54	(O=20)	(O=20)	(O=20)	(O=20)	(O=20)
H	0.86(.50)	0.20(.34)	0.91(.49)	0.35(.31)	0.97(.64)
HL	0.82(.33)	0.10(.27)	0.83(.28)	0.24(.22)	0.91(.37)
S	1.08(.46)	0.28(.57)	1.07(.50)	0.43(.51)	1.02(.54)
T	0.92(.44)	0.19(.41)	0.94(.44)	0.34(.37)	0.97(.52)
					0.70(.52)

$d'$  and  $\beta$  scores and standard deviations (in brackets) as a functional of viewing condition (N= Total Number of Participants, O=total number of objects, H=High Poly, HL=High Low Poly, S>Selective Poly, T=Total)

Desk Sub Area						
Consistent Objects						
N	d'un	b'un	d'mod	b'mod	d'high	b'high
54	(O=5)	(O=5)	(O=5)	(O=5)	(O=5)	(O=5)
H	0.70(.70)	0.08(.33)	0.61(.62)	0.20(.39)	0.48(.56)	0.27(.52)
HL	0.55(.75)	-0.06(.34)	0.55(.50)	0.10(.29)	0.63(.52)	0.36(.39)
S	0.65(.76)	0.13(.43)	0.77(.56)	0.29(.43)	0.82(.52)	0.42(.39)
T	0.63(.72)	0.02(.41)	0.69(.56)	0.20(.38)	0.64(.54)	0.34(.43)

Inconsistent Objects						
N	d'un	b'un	d'mod	b'mod	d'high	b'high
54	(O=4)	(O=4)	(O=4)	(O=4)	(O=4)	(O=4)
H	0.65(.69)	0.32(.30)	0.52(.69)	0.27(.30)	0.46(.69)	0.26(.36)
HL	0.43(.56)	0.26(.34)	0.34(.54)	0.20(.34)	0.29(.46)	0.20(.31)
S	0.65(.65)	0.28(.27)	0.59(.63)	0.31(.27)	0.36(.52)	0.21(.26)
T	0.58(.63)	0.30(.30)	0.48(.62)	0.26(.30)	0.37(.56)	0.23(.31)

$d'$  and  $\beta$  scores and standard deviations (in brackets) as a functional of viewing condition (N= Total Number of Participants, O=total number of objects, H=High Poly, HL=High Low Poly, S>Selective Poly, T=Total)

Table 4. Bookcase area confidence results

Bookcase Area						
N	d'un	b'un	d'mod	b'mod	d'high	b'high
54	(O=7)	(O=7)	(O=7)	(O=7)	(O=7)	(O=7)
H	0.27(.72)	0.10(.41)	0.39(.70)	0.17(.49)	0.33(.65)	0.15(.44)
HL	0.61(.61)	0.23(.27)	0.59(.56)	0.34(.34)	0.52(.53)	0.36(.33)
S	0.50(.54)	0.17(.45)	0.65(.58)	0.28(.52)	0.58(.63)	0.26(.49)
T	0.46(.63)	0.16(.38)	0.54(.62)	0.26(.45)	0.48(.60)	0.26(.43)

$d'$  and  $\beta$  scores and standard deviations (in brackets) as a functional of viewing condition (N= Total Number of Participants, O=total number of objects, H=High Poly, HL=High Low Poly, S>Selective Poly, T=Total)

Table 5. Lounge area confidence results

Lounge Area						
N	d'un	b'un	d'mod	b'mod	d'high	b'high
54	(O=17)	(O=17)	(O=17)	(O=17)	(O=17)	(O=17)
H	0.57(.62)	0.18(.32)	0.58(.69)	0.30(.54)	0.67(.58)	0.49(.47)
HL	0.40(.43)	0.03(.21)	0.46(.54)	0.12(.32)	0.65(.46)	0.51(.48)
S	0.75(.73)	0.26(.49)	0.70(.72)	0.31(.58)	0.82(.75)	0.62(.59)
T	0.57(.57)	0.15(.37)	0.58(.65)	0.25(.49)	0.71(.60)	0.54(.51)

$d'$  and  $\beta$  scores and standard deviations (in brackets) as a functional of viewing condition (N= Total Number of Participants, O=total number of objects, H=High Poly, HL=High Low Poly, S>Selective Poly, T=Total)

Table 6.Sofa sub-area confidence results

Sofa Sub Area						
	Consistent Objects					
N	d' un	b' un	d' mod	b' mod	d' high	b' high
54	(O=4)	(O=4)	(O=4)	(O=4)	(O=4)	(O=4)
H	0.84(.95)	0.01(.35)	0.72(.93)	0.06(.32)	0.69(.86)	0.08(.19)
HL	1.24(.83)	-.01(.31)	0.98(.79)	0.14(.28)	0.69(.62)	0.26(.29)
S	1.30(.82)	0.13(.36)	1.07(.82)	0.15(.35)	1.07(.76)	0.25(.32)
T	1.13(.88)	0.04(.34)	0.93(.85)	0.12(.31)	0.81(.76)	0.20(.28)
Inconsistent Objects						
N	d' un	b' un	d' mod	b' mod	d' high	b' high
54	(O=4)	(O=4)	(O=4)	(O=4)	(O=4)	(O=4)
H	0.37(.82)	0.07(.31)	0.44(.75)	0.10(.31)	0.25(.40)	0.12(.25)
HL	0.43(.36)	0.07(.39)	0.42(.31)	0.12(.32)	0.39(.36)	0.03(.28)
S	0.73(.56)	0.16(.35)	0.67(.58)	0.22(.20)	0.37(.49)	0.20(.25)
T	0.51(.62)	0.10(.35)	0.51(.58)	0.15(.28)	0.33(.42)	0.12(.27)

$d'$  and  $\beta$  scores and standard deviations (in brackets) as a functional of viewing condition (N= Total Number of Participants, O=total number of objects, H=High Poly, HL=High Low Poly, S=Selective Poly, T=Total)

ANOVA analysis of the  $d'$  and  $\beta$  results revealed no main effect of viewing condition in all cases (e.g. office  $d'$  score moderate confidence analysis:  $F(2,51)=1.370$ ,  $p=0.263>0.05$ ). This indicates that memory performance was similar across conditions.

Significant main effects of object type (consistent/ inconsistent) were revealed. In relation to  $\beta$  scores, the following significant main effects of object type appeared:

- Kitchen area (Table 1), uncorrected confidence,  $F(1,51)=28.743$ ,  $p<0.05$
- Kitchen area (Table 1), moderate confidence,  $F(1,51)=14.158$ ,  $p<0.05$
- Kitchen area (Table 1), high confidence,  $F(1,51)=12.278$ ,  $p<0.05$
- Desk area (Table 3), moderate confidence,  $F(1,51)=16.812$ ,  $p<0.05$

These results revealed a significant main effect of object type ( $\beta$  inconsistent  $>$   $\beta$  consistent) but no main effect of viewing condition, nor a significant interaction. This indicates that participants, regardless of viewing condition were significantly more biased to responding that inconsistent objects were not present in the scene compared to recognition of consistent objects for which there was close to no bias at all. This is consistent with memory psychology literature [Rojahn & Pettigrew 1992] and indicates that such effects could be simulated when immersed in synthetic scenes.

Concerning  $d'$  scores, the following significant main effects of object type were revealed:

- Sofa area (Table 6), uncorrected confidence,  $F(1,51)=16.225$ ,  $p<0.05$
- Sofa area (Table 6), moderate confidence,  $F(1,51)=7.526$ ,  $p<0.05$

- Sofa area (Table 6), high confidence,  $F(1,51)=15.481$ ,  $p<0.05$
- Desk sub-area (Table 3), high confidence,  $F(1,51)=7.893$ ,  $p<0.05$

These results revealed a significant main effect of object type ( $d'$  consistent  $>$   $d'$  inconsistent) but no main effect of viewing condition, nor a significant interaction. The consistent objects regardless of viewing condition (high poly, high – low poly, selective poly), have a significantly greater distance between signal and noise distribution (thus better recognition) than the inconsistent objects.

### 4.3 Discussion

Memory research has established that schema-based information is drawn in the process of retrieving information from memory. The scope of this work has been to identify regions of a scene that could be rendered in lower quality without affecting spatial information uptake. Recognition of objects of varying quality could be affected according to whether they ‘fitted’ the context of the scene. Contrary to the previously analyzed approaches based on task or gaze, consistent objects that have been rendered in low quality and inconsistent objects that have been rendered in high quality can be combined in the same scene. A selective rendering framework based on metadata enrichment handles dynamic polygon quality variations based on ‘activated’ areas which are triggered by user proximity to invisible space boundaries.

Participants’ memory performance was assessed after exposure to scenes of varying rendering quality through recognition of present objects in a scene as well as inferred objects. It has been shown that a scene which incorporates objects of high and low quality provokes similar memory recognition performance and associated confidence levels when compared to a fully-fledged rendering solution. Thus, in relation to the recognition task utilized, the selective rendering system presented may be of similar functional fidelity but of significant less computational complexity than a scene including objects of a high polygon count. Future work should focus on whether the proposed metric is sensitive to object quality assignment variations which are not schema-related as opposed to quality variations based on the proposed schema-based rationale. How cognitive factors can be investigated empirically in relation to selective rendering algorithms is a challenging research proposition and one that future work will explore.

### References

- BARTLETT, F.C. 1932. *Remembering*. Cambridge University Press, Cambridge, England, p-213.
- BREWER, W.F. & TREYENS, J.C. 1981. Role of Schemata in Memory for Places. *Cognitive Psychology*, 13, 207-230.
- BROOKS JR, F.P. 1999. What’s Real About Virtual Reality. *IEEE Computer Graphics and Applications*, 16-27.
- CATER, K., CHALMERS, A., WARD G. 2003. Detail to Attention: Exploiting Visual Tasks for Selective Rendering. In *Proceedings of the Eurographics Workshop on Rendering*.

- FERWERDA, J. 2001. Hi-Fi rendering. *ACM Siggraph Eurographics campfire on perceptually adaptive graphics*. <http://isg.cs.tcd.ie/campfire/jimferwerda2.html>.
- FLANNERY, K.A. WALLES, R. 2003. How does schema theory apply to real versus virtual memories? *Cyberpsychology and Behavior*, 6, 2, 151-159.
- HABER, J., MYSZKOWSKI, K., YAMAUCHI, H., SEIDEL, H.P. 2001. Perceptually guided corrective splatting, *Computer Graphics Forum*, 20, 3, 142-152.
- ITTI, L., KOCH, C., NEIBUR, E. 1998. A Model of Saliency-based Visual Attention for Rapid Scene Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254-1259.
- LAND, M. F. 1999. Motion and vision: why animals move their eyes. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioural Physiology*, 185, 4, 341-352, 1999.
- LOFTUS, G. R., MACKWORTH, N. H. 1978. Cognitive determinants of fixation location during picture viewing. *Educational Psychology: Human Perception and Performance*, 4, 4, 565-572, 1978.
- MACONCIE, G.W., LOSCHKY, L.C. 1997. Human performance with a gaze linked multi-resolutional display. In *Proceedings of the Advanced Displays and Interactive Displays First Annual Symposium 1997*, 25-34. ARL Federated laboratory advanced displays and interactive displays consortium.
- MANIA, K. & CHALMERS, A. 2001. The Effects of Levels of Immersion on Presence and Memory in Virtual Environments: A Reality Centred Approach. *Cyberpsychology & Behavior*, 4(2), 247-264.
- MANIA, K., ADELSTEIN, B., ELLIS, S.R., HILL, M. 2004. Perceptual Sensitivity to Head Tracking Latency in Virtual Environments with Varying Degrees of Scene Complexity. In *Proceedings of the Symposium on Applied Perception in Graphics and Visualization 2004*, ACM, 39-47.
- MANIA, K., ROBINSON, A. 2003. Simulating Spatial Assumptions. Technical Sketch. *ACM SIGGRAPH 2003*, San Diego, USA.
- MANIA, K., ROBINSON, A. 2004. The Effect of Quality of Rendering on User Lighting Impressions and Presence in Virtual Environments. In *Proceedings of the ACM Siggraph International conference on Virtual Reality Continuum and its Applications in Industry 2004*, Singapore. ACM, 200-205.
- MANIA, K., TROSCIANKO, T., HAWKES, R., CHALMERS, A. 2003. Fidelity Metrics for Virtual Environment Simulations based on Human Judgments of Spatial Memory Awareness States. *Presence, Teleoperators and Virtual Environments*, 12(3), 296-310, MIT Press.
- MANIA, K., ROBINSON, A., BRANDT, K. 2005. The effect of memory schemata on object recognition in virtual environments, *Presence, Teleoperators and Virtual Environments*, 14, 5, 606-615.
- MARMITT, G., DUCHOWSKI, A.T. 2002. Modeling visual attention in VR: measuring the accuracy of predicted scanpaths. In *Proceedings of Eurographics 2002*, Short Presentations, 217-226.
- MCNAMARA, A. 2001. Visual Perception in Realistic Image Synthesis', *Computer Graphics Forum*, Volume 20, #4.
- MOURKOSSIS, N., RIVERA, F., TROSCIANKO, T., DIXON, T., HAWKES, R., MANIA, K. Accepted for publication 2009. Quantifying Fidelity for Virtual Environment simulations employing memory schema assumptions. *ACM Transactions on Applied Perception*.
- ROJAN, K., PETTIGREW, T. F. 1992. Memory for schema-relevant information: a meta-analytic resolution. *British Journal of Social Psychology*, 31, 2, 81-109.
- RUSHMEIER, H., WARD, G., PIATCO, C., SANDERS, P., RUST, B. 1995 Comparing Real and Synthetic Images: Some ideas about metrics. In *Proceedings of EGRW 1995*.
- SUNDSTEDT, V., STAVRAKIS, E., WIMMER, M., REINHARD, E. 2008. A Psychophysical Study of Fixation Behavior in a Computer Game. In *Proceedings of APGV 2008 - Symposium on Applied Perception in Graphics and Visualization 2008*, 43-50.
- TROSCIANKO, T., MOURKOSSIS, N., RIVERA, F., MANIA, K., DIXON, T., HAWKES, R. 2007. Memory for objects in virtual environments, *Journal of Vision*, Volume 7, Number 9, 763a.
- WATSON, B.A., WALKER, N., HODGES, L.F., REDDY, M. 1997. An evaluation of level of detail degradation in head-mounted display peripheries. *Presence: Teleoperators and Virtual Environments*, 6, 6, 630-637, 1997.
- T. D. WICKENS. 2001. *Elementary Signal Detection Theory*. New York: Oxford University Press.