Visual Importance-biased Image Synthesis Animation

Ross Brown
Faculty of Information Technology
QUT
r.brown@qut.edu.au

Binh Pham
Faculty of Information Technology
QUT
b.pham@qut.edu.au

Anthony Maeder
Faculty of Built Environment and Engineering, QUT
a.maeder@qut.edu.au

Abstract

Present ray tracing algorithms are computationally intensive, requiring hours of computing time for complex scenes. Our previous work has dealt with the development of an overall approach to the application of visual attention to progressive and adaptive ray-tracing techniques. The approach facilitates large computational savings by modulating the supersampling rates in an image by the visual importance of the region being rendered.

This paper extends the approach by incorporating temporal changes into the models and techniques developed, as it is expected that further efficiency savings can be reaped for animated scenes. Applications for this approach include entertainment, visualisation and simulation.

CR Categories: I.3.7 (Three-Dimensional Graphics and Realism): Animation, Raytracing; I.3.3 (Picture/Image Generation): Antialiasing, Display algorithms.

Keywords: image synthesis, animation techniques, motion importance.

1. Introduction

Research indicates that motion is a strong attractor of visual attention [Senders 1976; Stelmach, Tam et al. 1991; Niebur and Koch 1995; Wolfe 1996; Osberger and Rohaly 2001; Yee, Pattanaik et al. 2001]. There is also physiological evidence for the pre-eminence of motion in the hierarchy of visual features, due to the presence of receptors sensitive to moving contours [Bruce and Green 1990]. These results are consistent with psychophysical evidence showing that motion strongly attracts visual attention [Wolfe 1996].

Experiments have shown that in non-attentive modes the sudden onset of stimuli within the periphery brings about attentional capture [Yantis and Jonides 1990]. Research has also indicates the high correlation of points of regard between viewers when observing images containing movement [Stelmach, Tam et al. 1991].

It can be concluded that much rendering effort can be saved by further exploiting visual attention concepts for the motion of objects, where rendering resources are concentrated on the most visually important regions.

Given the attention capturing ability of moving objects [Hillstrom and Yantis 1994], it is expected that the best results will be gained from adding motion to a previously developed visual attention model [Brown, Pham et al. 1999; Brown, Pham et al. 2000; Brown, Pham et al. 2001]. This motion importance model also extends previous motion models by incorporating relative rather than absolute effects.

Furthermore, new animation techniques are developed to exploit the visual importance evaluation offered by the temporal change model, including more effective motion estimation techniques. The paper details the theoretical basis and the design of the major components of the approach. Implementation issues are also discussed at the end of this paper.

Structurally the rest of the paper is organised as follows. An analysis of present research into modelling visual attention is presented in Section 2. Section 3 details the development of extensions to the visual attention model. Section 4 then details the incorporation of the new temporal change model into an adaptive rendering system. Finally, the paper concludes with a discussion of achievements in the design of the model in Section 5.

2. Previous Work

Previously, the majority of the application research work has been carried out into detecting changes in an image for compression purposes, in order to reduce the amount of data needing to be sent for low bandwidth video applications [LeGall 1991]. Recently, in addition to this raw detection and compensation for change in an image, there has been the application of the previous psychophysical experimental results to the determination of the importance and visibility of changes occurring within a video stream [De Vleeschouwer, Marichal et al. 1997; Osberger and Rohaly 2001]. Models have been developed to simulate the visual importance of motion within the application areas of video processing and image synthesis, in order to further reap efficiency gains not possible through raw change detection.

A multiresolution motion model has been developed by Yee [Yee 2000; Yee, Pattanaik et al. 2001], as an addition to the visual saliency model of Koch and Ullman. [Koch and Ullman 1985] and Itti and Koch [Itti and Koch 2001]. The model uses a magnitude value to ascertain the visual importance of pixels in the region, using an object ID-based method of pixel displacement calculation developed by Agrawala [Agrawala, Beers et al. 1995].

A simple fuzzy logic motion importance model has also been developed by Marichal et al. [Marichal, Delmot et al. 1996] and De Vleeschouwer et al. [De Vleeschouwer, Delmot et al. 1997; De Vleeschouwer, Marichal et al. 1997] for applications to low-bandwidth video. The motion estimation is based upon absolute magnitudes of motion, which do not account for global quantities of motion. In addition, the model does not account for the direction of motion.

Region-based approaches have also been used to model motion importance within a series of images. Osberger [Osberger and Rohaly 2001] has devised an effective threshold model to incorporate motion into a video processing system. While the motion model reports good results for determining motion importance, the parameters are not referenced to any psychophysical data. Motion importance in this approach is calculated by the absolute magnitude of the motion. Even though this approach allows for global distribution of motion, it does not allow for local difference effects due to directional or magnitude values.

These region-based models only treat motion as a magnitude, and do not include its vector component in their calculations. The models also do not differentiate between motion and abrupt onset in any fashion. Therefore, these approaches can be improved by incorporating a region-based measure of local differences, which accounts for global suppression and enhancement effects.

3. A Visual Attention Model Incorporating Temporal Changes

Temporal changes involve two major categories: actual motion of objects in a scene, and sudden changes occurring due to luminance changes unrelated to object motion. Both of these have been characterised within this temporal change model, to accommodate most effects occurring within an animated scene. In particular, the new model improves on others by incorporating the following factors:

- relative forms of motion are used to ascertain the importance of the region, as apposed to absolute measures used in present models;
- magnitude and directional factors are treated as separate factors contributing to the final motion-based visual importance of a region;
- global effects are also incorporated, to model the enhancing and suppressing influences of surrounding motion in a scene;
- parameters for the model are gained from psychophysical research, instead of using arbitrary values;
- differentiation of onset effects from those caused by the motion of objects in the scene–for example, lighting changes.

Figure 1 illustrates how both the magnitude and direction of the differences in motion can contribute to a region standing out.

We have chosen to use fuzzy logic in our modelling of such perceptual phenomena dues to its efficacy in processing imprecise data and its ability to model human reasoning [Berkan and Trubatch 1997]. It also enables the fine-tuning of parameters in an intuitive fashion, due to the ease of modelling human expert rules. Fuzzy systems implement truth as a continuum between true and false. Therefore, a statement can be partially true or false, so its *fulfilment* value may be quantified to be anywhere between 0.0 and 1.0 inclusive. The fulfilment value is modelled using fuzzy membership functions, which match the human estimation of truth to the fuzzified term.

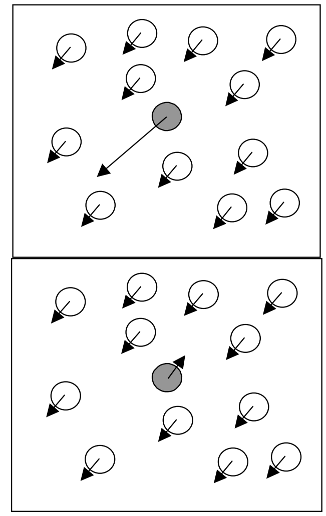

Figure 1 Illustration of the concepts of motion magnitude importance and motion direction importance. Both images show regions with vectors attached, indicating their direction and magnitude of motion. The left diagram shows a grey region standing out due to a difference in velocity magnitude. The right diagram shows a grey region standing out because of its relative difference in direction.

In our system the motion difference is fuzzified to three terms *Low*, *Medium* and *High*. The membership functions indicate the truth of these terms, with respect to the quantified variable being fuzzified (refer to Figure 2). In the case of motion magnitude, the values are the degrees subtended in the visual field per second (deg/sec)–this assumes a fixed viewing distance. The direction difference is derived from the difference in degrees of the direction vectors of the regions being considered. The abrupt onset function is derived from the ratio of the magnitude of the frame to frame luminance change.

In addition, results from experiments by Nothdurft also indicate a sigmoidal pop-out effect from local motion differences, with a saturation effect past a high level of local motion difference [Nothdurft 1993; Nothdurft 1993]. As with luminance and colour, the effect is suppressed by surrounding motion differences [Nothdurft 1993]. These facts have inspired the shape and functionality of the membership functions used. In a similar fashion to the membership functions in our previous spatial visual attention system, the motion membership functions are adaptive to the magnitude and direction of motion detected within the whole visual field. This is modelled by an adaptive threshold that draws its value from the average value of motion direction differences and magnitudes in the whole image–th_{dir} or th_{hmag} and th_{lmag}. The latter two thresholds represent the lower and higher points at which the magnitude of motion is tracked by the HVS. Both of these factors contribute to the motion importance of regions.

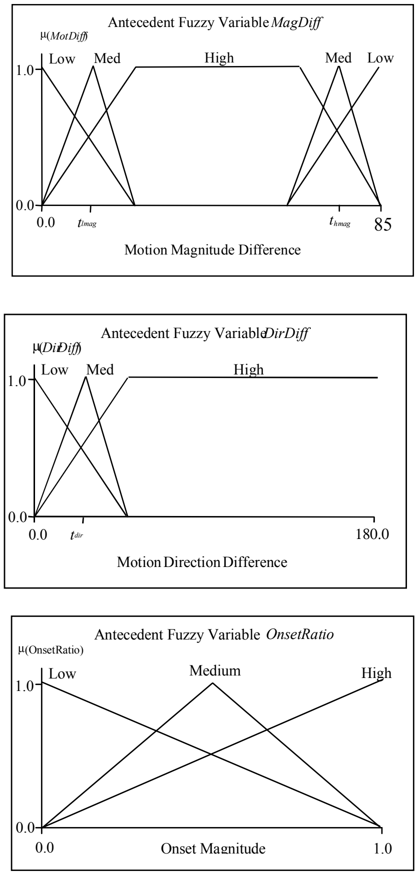

Figure 2 Diagram of the motion evaluation membership functions for the magnitude of the motion (top left) and the direction of the motion (top right) and object Onset (bottom).

The object tracking ability of the HVS is known as smooth pursuit, and has been analysed by a number of researchers [Westheimer 1954; Robinson 1965; Girod 1988; Daly 1998; Verstraten, Cavanagh et al. 2000]. These experimental results become important when evaluating the importance of a region within the visual field of a viewer. The motion perceived must be able to be tracked, in order to attract attention. Daly [Daly 1998] reports a value of 80 deg/sec as an upper threshold of the smooth pursuit capabilities of a viewer.

Therefore, the absolute motion value of the region being examined is thresholded to 80 deg/sec before being processed for relative motion analysis, to prevent objects above the tracking capabilities of the HVS influencing the visual importance by relative differences. In a similar manner, only moving regions are considered for the motion importance calculations. This removes the case of stationary objects having large relative motion differences causing inappropriate pop-out.

These membership functions are used in the following rule base within the motion importance module:

IF *MagDiff*	IS High	THEN *FinImp* IS High
IF *MagDiff*	IS Med	THEN *FinImp* IS Med
IF *MagDiff*	IS Low	THEN *FinImp* IS Low
IF *DirDiff*	IS High	THEN *FinImp* IS High
IF *DirDiff*	IS Med	THEN *FinImp* IS Med
IF *DirDiff*	IS Low	THEN *FinImp* IS Low
IF *OnsRatio*	IS High	THEN *FinImp* IS High
IF *OnsRatio*	IS Med	THEN *FinImp* IS Med
IF *OnsRatio*	IS Low	THEN *FinImp* IS Low

The membership functions are used in the same way as the other spatial importance rules. The other rules in the system model spatial importance based upon the differences luminance, colour, size, and position of segmented regions in the image. The membership functions are similar to the motion importance functions and use the same implication process to establish an estimate of the stationary visual importance of the region (refer to Figure 3).

Both the modules within the visual importance system aggregate the fulfilment values in a multiple-additive manner [Berkan and Trubatch 1997], with a weighting of 0.6 for the temporal rules and 0.4 for the spatial rules [Niebur and Koch 1995; Osberger and Rohaly 2001]. These are then defuzzified to form a visual importance estimate for the region being examined.

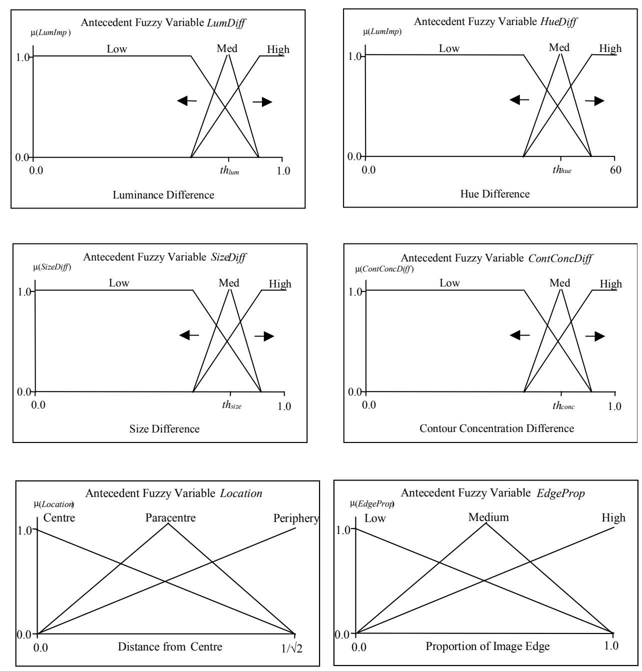

Figure 3 Illustration of the spatial feature rules for the visual importance module.

We show in Figure 4 a processing example for the rule *IF DirDiff is High THEN FinImp is High.* This example is for the direction difference variable, which evaluates the visual importance caused by a difference in direction of a region with regards to surrounding objects. The first diagram in the figure shows the direction difference value *D* being *fuzzified* into fulfilment level $\mu(D)$. The second diagram shows the implication process forming a fuzzy set for the consequent term *FinImp*. This process is repeated for all the rules in the fuzzy rule set contributing to Low, Medium or High final importance values for the region. The last diagram shows the *aggregated* fuzzy sets for the *FinImp* membership functions Low and High (Medium omitted for clarity) after all the rules have been evaluated. From this aggregated set, the final *defuzzified* value *I* is calculated using the *Weighted Fuzzy Mean* method [Leekwijck and Kerre 1999]. We now describe how this value *I* is used in a newly developed animation approach.

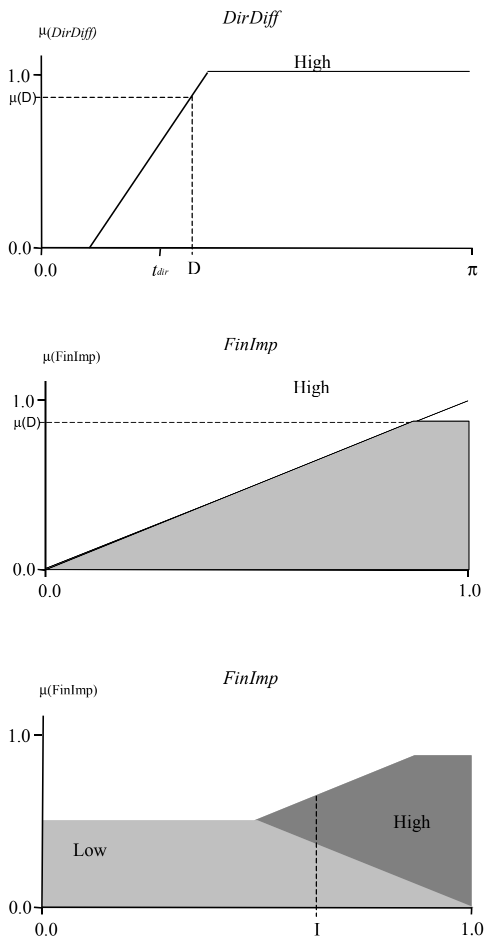

Figure 4 Illustration of the implication process used in motion importance rules IF DirDiff IS HighTHEN FinImp IS High.

4. A Motion-based Adaptive Rendering Approach

In order to incorporate the above model into an adaptive rendering approach a number of stages must take place. The system must make a segmentation of the scene based upon motion information, using previous frames and region importance maps. The approach must also compensate for ego motion caused by camera movement. Next, the temporal model is applied to the motion vector estimates from the segmented regions to produce a relative visual importance value for the moving regions. Finally, the importance value is used to control the adaptive rendering system.

The major components of this approach are depicted in Figure 5.

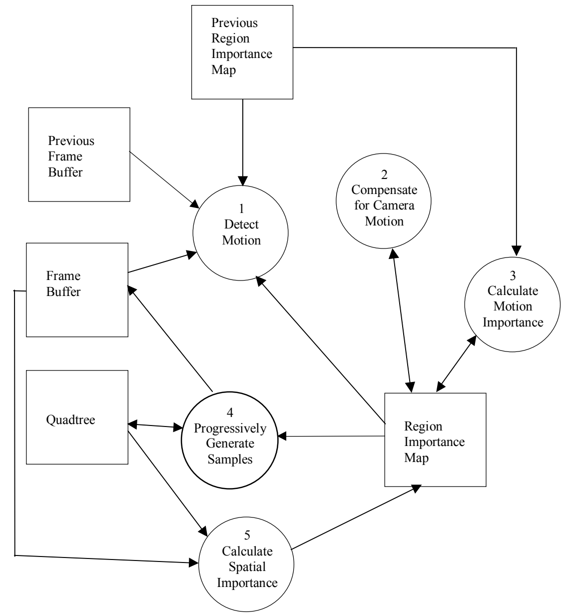

Figure 5 Flow diagram of the major stages in the temporal change approach.

In the newly developed approach, the calculations are performed from frame to frame, this requires the storage of the *Previous Frame Buffer* to facilitate change analysis. While the image is progressively sampled to the level of one sample per pixel, the present and previous frame buffers are analysed for motion. A previous region segmentation is stored in order to facilitate the motion importance calculations. Once region motion vectors have been derived for the regions in the scene, then these vectors are processed to remove any camera motion. The resultant vectors and other spatial information is then fed to the motion and spatial membership functions that are detailed in Section 3. The resultant importance values are stored in the *Region Importance Map*. The region importance value is used to modulate the supersampling performed within each pixel.

We now detail the motion estimation techniques used within this approach.

4.1. Motion Estimation Technique

There are two major methods for motion estimation within the area of image synthesis. The motion estimation can be performed in an image-based manner, similar to video systems [LeGall 1991], or by using object-based techniques. Object-based methods exploit the object-space geometry and associated transformation matrices to estimate where a geometry segment will be translated to on the screen [Agrawala, Beers et al. 1995] [Guenter and Tumblin 1996] [Wallach, Kunapalli et al. 1994] [Yun, Guenter et al. 1997]. Object-based approaches perform better than image-based methods in 3D animation applications, due to the unambiguous nature of the object ID information. A number of motion estimation techniques have been developed for image synthesis to facilitate compression of synthetic movies. A region-based motion detection approach is developed which has the following features:

- the ability to account for gross and local motion effects of regions explicitly–eg. both translation and internal rotation of objects;

- the ability to account for non-affine transformations of the regions being analysed–ie. non-linear region deformations;
- the ability to remove camera motion effects from the derived region motion vectors.

A motion estimation process used by Guenter et al. [Guenter, Yun et al. 1993; Yun, Guenter et al. 1997] is pixel-based, utilising pixel RGB colour, object ID and depth buffer values. The estimation technique uses these parameters as an aid in determining, in the forward direction from frame n to frame $n + 1$, the position of a four pixel (2×2) square. While Wallach et al. [Wallach, Kunapalli et al. 1994] use hardware Gouraud shading and texturing techniques to garner information about optical flow within the image being generated. The mode vector of the 16×16 pixel optical flow block is used as the centre of a brute force search. Agrawala et al. use back projection to ascertain the location of a pixel in the previous frame to the one being examined [Agrawala, Beers et al. 1995]. The transformation and projection matrices are used to obtain the position of a pixel in object-space in the previous frame. The difference between the two gives an object-space accurate optical flow motion vector for the pixel.

As the temporal change approach developed here continues the region-based paradigm, it is appropriate that the motion estimation technique will be developed from a region-based perspective. Furthermore, it can be argued that the perceived motion in a scene is region-based in nature, due to a person focusing on regions in an image, and not pixels or blocks [Marr 1982; Wolfe 1996; Wolfe 2000].

The motion estimation scheme detailed here obtains regions by segmenting the scene using the object ID as a basis for the comparison operations in the segment merging stage. Once the image is segmented by object ID, then the regions are further segmented by luminance and hue features. This provides a two level hierarchy, facilitating the detection of gross region motion effects within the top-level, while internal motion effects are detected within the second segmentation level.

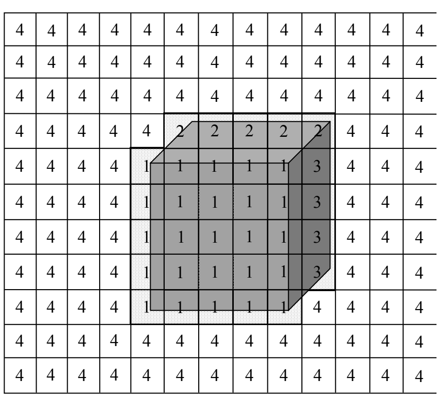

Figure 6 An example of the hierarchy of segmentation used in the motion estimation system. The colour of the segment represents object ID segmentations. The dotted area represents one object ID, while the white background represents another object ID. The numbers represent the segmented regions.

As well as aiding the correct segmentation of regions for more accurate motion calculation, the use of object IDs speeds up the merge segmentation algorithm. The segmentation algorithm is now divided into two main processes. The first is the merging of segments that have the same object IDs. This can be performed in a serial fashion by simply scanning the segments from top to bottom, left to right, placing them in segment lists identified by the mode of the object ID samples within the segments.

The lists of segments are then processed using a simple region merge algorithm. The segments are divided up into regions based upon hue and luminance differences (refer to Figure 6). This segmentation approach is incremental in nature. If a segment changes object ID, luminance or hue through further sampling, then the segment is reallocated to another region, triggering new importance calculations to update the importance map.

Due to the inherent correlation of the object ID with the motion in a scene, the segmentation based on object IDs provides effective search windows for further internal motion estimation. These windows are more accurate than arbitrary sized square regions, which do not necessarily contain the blocks causing the perceived motion. This brings about better matches when performing motion prediction within an object region.

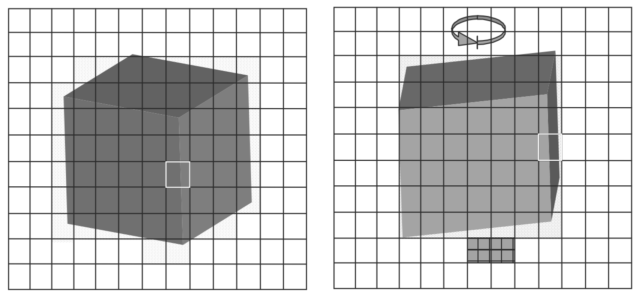

Figure 7 Illustration of the internal motion search method over two frames (frame *n* on the left, *n + 1* on the right), within the regions segmented at the level of object IDs– dotted regions surrounding cube. A segment which changes from frame to frame is highlighted in white. Example segment s which change across two object IDs are highlighted towards the bottom of the diagram by a cross hatch pattern in the second frame.

For video compression systems, motion vectors must be collated across the whole scene for every pixel and block [LeGall 1991]. Transmitting the change vector, instead of the actual image data, reaps efficiency savings. As this application is the computation of region-based visual attention using motion differences, there is not such a need to search for block motion outside of the area segmented by an object ID. This approach only requires estimates of the motion of the segmented region across the image, and the internal motion of a region. Therefore, the algorithm works within the object ID region to search for internal motion. This internal motion difference can be processed in the same manner as the gross region motion, to gain a measure of internal motion importance for each region.

Segment object ID changes may occur across different region segmentations, as shown in the lower portion of Figure 7 by the hatched segments. They can be classified as the appearance of a new object within the segmentation, and therefore are treated as an abrupt onset change, as per the model developed in Section 2.

The possible object motions can be divided into two categories: rigid and non-rigid. A combination of these two may occur for any object in the scene. The method developed here will process both forms by performing a hierarchy of motion calculations. The first level is the gross motion of the object ID segmented region, while the second is the internal motion of the hue and luminance segmented regions.

The motion to be calculated for both the region levels is translational in nature. This will still effectively model the internal motion of the regions–for example, the spinning of a cube (refer to Figure 7). The object ID level of segmentation will compute importance for the translation of the cube through the scene. The second level, which represents the segmentation of the gross region into similar hue and luminance regions, will give an estimate of the internal motion importance of the object.

To identify the motion of the gross object ID regions is straightforward. The method searches for a corresponding region object ID from frame $n + 1$ in the previous region segmentation for frame n. The centroids of the regions in the two frames are then subtracted to form a motion difference vector M_r. This vector is then used to compute the motion importance for the gross region motion. If the object ID cannot be found in the previous region importance map, then the motion vector M_r is set to zero. An object entering a scene will, on the first frame, be treated as a sudden onset region. Frames occurring afterwards containing this object will then process its movement in a normal fashion.

The internal motion of the regions uses a region matching technique, similar in nature to the block matching techniques used in video compression [LeGall 1991]. In this case the method is modified to match regions, not blocks, for efficiency purposes. This motion and onset information is then passed to the camera compensation module to remove any camera motion from the vector.

4.2. Camera Compensation Technique

Before a motion importance value can be computed, the component of motion due to the camera must be removed from the computed region motion values. Motion in an image may be caused by movements within the scene, and the spatial transformation of the camera viewing the scene. Camera motion forms a background motion noise, which needs to be suppressed in order to ascertain correctly the true changes in the scene for motion importance purposes [Osberger and Rohaly 2001].

The advantage with image synthesis camera compensation is the availability of the world to camera space transformation T_{wc}, and the projection matrix P. These two transformation matrices enable a complete model of the contribution of the camera to region motion estimates. Before the region being examined is searched for in the previous frame region list, its centroid is transformed by the opposite of the difference in world to camera transformation matrices, thus removing any image-plane motion produced by the view camera.

The difference between the two camera transformations ΔT_{wc} is formed by analysing the world to camera space transform matrix T_{wc}. The following equation provides the camera transformation matrix between two frames n and $n + 1$:

$$\Delta T_{wc} = T_{wc,\, n+1} \cdot T_{wc,\, n}^{-1} \qquad (1)$$

where: ΔT_{wc} is the camera motion transformation from frame n to frame $n + 1$; $T_{wc,\, n+1}$ is the world to camera space transformation for the frame $n + 1$ in the animation; $T_{wc,\, n}^{-1}$ is the inverse of the world to camera space transformation for the frame n in the animation.

The inverse of ΔT_{wc} matrix, $\Delta T_{wc\text{-}1}$, is then applied to the centroid $C_{r,\, n+1}$ of the region being camera compensated, to remove the motion caused by the camera. This means that the final 2D motion vector $M_{r,\, n+1}$ for a region r, for frame $n + 1$ is:

$$C'_{r,\, n+1} = C_{r,\, n+1} \cdot P^{-1} \cdot \Delta T^{-1}_{wc,\, n+1} \cdot P \qquad (2)$$

$$M_{r,\, n+1} = C'_{r,\, n+1} - C'_{r,\, n} \qquad (3)$$

where: $C'_{r,\, n+1}$ is the camera compensated centroid for the region being examined in frame $n + 1$; $C'_{r,\, n}$ is the camera compensated centroid of the region in the previous frame n; P, P^{-1} are the view projection matrix and its inverse; $M_{r,\, n+1}$ is the final 2D motion vector (Δx, Δy, z ignored) computed for region r in frame $n + 1$; $\Delta T^{-1}_{wc,\, n+1}$ is the inverse world to camera space transformation for frame $n + 1$ in an animation.

The final calculation is the addition of the internal region motion vector with the object ID region vector to produce the overall motion of the internal region. This is accomplished by the following equation:

$$MFin_{r,\, n+1} = MObjectID_{r,\, n+1} + MInternal_{r,\, n+1} \qquad (4)$$

where: $MFin_{r,\, n+1}$ is the final vector combining the gross and internal region motion; $MObjectID_{r,\, n+1}$ is the camera corrected gross region motion vector for frame $n + 1$; $MInternal_{r,\, n+1}$ is the camera corrected internal region motion vector for frame $n + 1$.

The final 2D motion vector $MFin_{r,\, n+1}$ is passed to the motion evaluation component of the temporal change model to derive a motion-based visual importance value.

4.3. Adaptive Image Synthesis Animation

In order to exploit these temporal importance values drawn from the region importance map, a new framework for animation rendering must be fabricated.

In the past, adaptive rendering for motion has been handled in a number of ways. Distributed rendering is a technique used to simulate motion blur caused by shutter speed effects in cameras [Glassner 1986]. Other temporal motion detection models have been used to perform motion compensation for video compression of image synthesis animations [Guenter, Yun et al. 1993; Wallach, Kunapalli et al. 1994; Agrawala, Beers et al. 1995; Yun, Guenter et al. 1997]. They typically detect changes in pixels and create pixel flow vectors by identifying which object has been intercepted at the pixel level and then tracking the transformation of the pixel with the object ID to the next frame, using the 3D transformation matrices contained within the animation script. This method while accurate, is restrictive, as it requires the transformation of every pixel within the image to ascertain pixel flow vectors for the next scene. In addition, the methods may require hardware support in order to be efficient, due to the overhead of performing the calculations for every pixel [Wallach, Kunapalli et al. 1994]. The region-based techniques developed in this chapter are much more efficient due to the eschewing of pixel-based motion estimation, in favour of region-based motion estimation.

A multiresolution model of temporal importance has been implemented [Yee 2000; Yee, Pattanaik et al. 2001] in order to control sampling rates in a ray tracing system. The approach uses a pixel-based absolute value motion model, which lacks the ability to deal with the relative motion of regions and requires a hardware prerendering of the scene to provide information to the motion model used. From these observations it seems that no work has been developed for the region-based processing of motion for image synthesis efficiency purposes. In addition, the region-based method developed here is more in line with present psychophysical thinking on object-based visual attention, and should be more efficient being region-based rather than pixel-based in its calculations. Furthermore, in the same manner as the spatial model, the motion importance animation technique is truly progressive. The approach uses the early samples of the scene to make estimates of region importance, and does not require a hardware prerendering to ascertain motion importance values.

The adaptive and progressive methods used in this approach will modify the supersampling rate of a region according to its visual importance. The supersampling techniques to be implemented include both constant and perceptual methods of pixel subdivision control, modulated by the importance of the region containing the pixel. Furthermore, the region importance algorithm within the progressive rendering approach needs to be modified in order to obtain frame-to-frame changes in luminance and region motion.

Even though there are costs involved with maintaining the region motion information, the algorithm still scales well, being linear in nature in both time and space complexity with regard to regions and segments, for both the region segmentation and the region importance calculations. This is due in the major part to the algorithm being reliant on image-space information, in which the number of segments and regions varies linearly with the size of the image.

5. Conclusions and Future Work

This paper has detailed the development of a novel region-based temporal importance model. The major achievements are:

- The development of a region-based temporal change model that uses region motion differences, not just absolute motion values. This more closely follows psychophysical models of visual pop-out.
- The development of a motion model that more fully characterises region motion as being a combination of the gross regional motion and the internal regional motion. This allows the model to produce an accurate estimation of region motion for both translational motion and the internal effects from rotation as well as non-rigid deformation of the object in the scene.
- The development of an improved region segmentation algorithm, which utilises the object ID information returned from the rendering system. This removes ambiguity problems caused by the coarse segmentation of the scene, as the object ID has an unambiguous relationship to the image-plane region segmentation. Furthermore, this facilitates more accurate and efficient calculation of region motion.
- The modification of supersampling techniques to accommodate region-based motion importance values.
- The development of a novel image synthesis-based camera compensation model for motion estimation. Camera compensation has been used in video motion importance calculations, but this method is novel due to the use of object-space transformation information to remove camera motion from the derived motion vectors.

At this stage, the model and the associated techniques have been fully designed and analysed. The next process is their implementation and incorporation into the present visual importance rendering system [Brown, Pham et al. 2001]. This task is relatively straightforward due to a number of factors. Firstly, the frame buffer and region importance map data structures already exist and have been implemented. Finally, the framework for still image supersampling modulation has already been implemented. Therefore, the process of performing frame-to-frame modification of supersampling rates is, again, an incremental implementation process.

Furthermore, due to the good results from work performed with still images in previous work [Brown, Pham et al. 2001], it is expected that the application of motion to both the visual importance model and the rendering techniques should give the same, if not better, results.

References

AGRAWALA, M., BEERS, A. AND CHADDHA, N. 1995. Model-Based Motion Estimation for Synthetic Animations. In *Third International Conference on Multimedia*, 477-488.

BERKAN, R. AND TRUBATCH, S. 1997. *Fuzzy Systems Design Principles, Building Fuzzy If-Then Rule Bases*. New York, U.S.A., IEEE Press.

BROWN, R., PHAM, B., AIDMAN, E. AND MAEDER, A. 2000. Efficient Image Rendering Using a Fuzzy Logic Model of Visual Attention. In *Advances in Intelligent Systems: Theory and Applications (AISTA)*, IOS Press, Amsterdam, Holland, 314-319.

BROWN, R., PHAM, B. AND MAEDER, A. 1999. A Fuzzy Model for Scene Decomposition Based on Preattentive Visual Features. In *Proceedings of SPIE, San Jose, USA*, 3644, 461-472.

BROWN, R., PHAM, B. AND MAEDER, A. 2001. A Fuzzy Logic Model of Visual Importance for Efficient Image Synthesis. In *Fuzz-IEEE 2001*, 1400-1403.

BRUCE, V. AND GREEN, P. 1990. *Visual Perception: Physiology, Psychology and Ecology*. Hove, UK, Lawrence Erlbaum.

DALY, S. 1998. Engineering Observations from Spatiovelocity and Spatiotemporal Visual Models. In *Proceedings of Human Vision and Electronic Imaging III*, 3299, SPIE - The International Society for Optical Engineering, Bellingham, USA, 179-206.

DE VLEESCHOUWER, C., DELMOT, T., MARICHAL, X. AND MACQ, B. 1997. A Fuzzy Logic System for Content-Based Bit-Rate Allocation. *Signal Processing. Image Communication* 10, 1-3, 115-141.

DE VLEESCHOUWER, C., MARICHAL, X., DELMOT, T. AND MACQ, B. 1997. A Fuzzy Logic System Able to Detect Interesting Areas of a Video Sequence. In *Human Vision and Electronic Imaging II*, 3016, The International Society for Optical Engineering, Bellingham, USA, 234-245.

GIROD, B. 1988. Eye Movements and Coding of Video Sequences. In *Visual Communications and Image Processing*, 1001, 398-405.

GLASSNER, A. 1986. Adaptive Precision in Texture Mapping. In *Computer Graphics (Proceedings of ACM SIGGRAPH 86)*, 20, ACM, 297-306.

GUENTER, B. AND TUMBLIN, J. 1996. Quadrature Prefiltering for High Quality Antialiasing. *ACM Transactions on Graphics* 15, 4, 332- 353.

GUENTER, B., YUN, H. AND MERSEREAU, R. 1993. Motion Compensated Compression of Computer Animation Frames. In *ACM SIGGRAPH 1993*, Proceedings of ACM Press / ACM SIGGRAPH, New York, 279-288.

HILLSTROM, A. AND YANTIS, S. 1994. Visual Motion and Attentional Capture. *Perception and Psychophysics* 55, 4, 399-411.

ITTI, L. AND KOCH, C. 2001. Computational Modelling of Visual Attention. *Nature Reviews: Neuroscience* 2, 3, 194-203.

KOCH, C. AND ULLMAN, S. 1985. Shifts in Selective Visual Attention: Towards the Underlying Circuitry. *Human Neurobiology* 4, 219-227.

LEEKWIJCK, W. AND KERRE, E. 1999. Defuzzification: Criteria and Classification. *Fuzzy Sets and Systems* 108, 2, 159-178.

LEGALL, D. 1991. Mpeg: A Video Compression Standard for Multimedia Applications. *Communications of the ACM* 34, 4, 30-44.

MARICHAL, X., DELMOT, T., DE VLEESCHOUWER, C., WARSCOTTE, V. AND MACQ, B. 1996. Automatic Detection of Interest Areas of an Image or of a Sequence of Images. In *IEEE International Conference on Image Processing*, 371-374.

MARR, D. 1982. *Vision*. San Francisco, USA, W. H. Freeman.

NIEBUR, E. AND KOCH, C. 1995. Control of Selective Visual Attention: Modelling the "Where" Pathway. In *Advances in Neural Information Processing Systems*, 8, The MIT Press, Cambridge, USA, 802-808.

NOTHDURFT, H. 1993. The Conspicuousness of Orientation and Motion Contrast. *Spatial Vision* 7, 341-363.

NOTHDURFT, H. 1993. The Role of Features in Preattentive Vision: Comparison of Orientation, Motion and Color Cues. *Vision Research* 33, 14, 1937-1958.

OSBERGER, W. AND ROHALY, A. 2001. Automatic Detection of Regions of Interest in Complex Video Sequences. In *Human Vision and Electronic Imaging VI*, 4299, SPIE - The International Society for Optical Engineering, Bellingham, USA, 361-372.

ROBINSON, D. 1965. The Mechanics of Human Smooth Pursuit Eye Movements. *Journal of Physiology* 180, 569-591.

SENDERS, J. 1976. Speculations and Notions. *Eye Movements and Psychological Processes*. M. R. AND S. J. Hillsdale, USA, Lawrence Erlbaum.

STELMACH, L., TAM, W. AND HEARTY, P. 1991. Static and Dynamic Spatial Resolution in Image Coding: An Investigation of Eye Movements. In *Human Vision, Visual Processing and Digital Display II*, 1453, 147-152.

VERSTRATEN, F., CAVANAGH, P. AND LABIANCA, A. 2000. Limits of Attentive Tracking Reveal Temporal Properties Fo Attention. *Vision Research* 40, 3651-3664.

WALLACH, D., KUNAPALLI, S. AND COHEN, M. 1994. Accelerated Mpeg Compression of Dynamic Polygonal Scenes. In *Proceedings of ACM SIGGRAPH 1994*, ACM Press / ACM SIGRAPH, New York, 193-196.

WESTHEIMER, G. 1954. Eye Movement Responses to Horizontally Moving Visual Stimuli. *Arch. Opthamology* 52, 932-941.

WOLFE, J. 1996. Visual Search: A Review. *Attention*. H. PASHLER. London, UK, University College London Press, 13-73.

WOLFE, J. 2000. The Deployment of Visual Attention: Two Surprises. *Search and Target Acquisition*. NATO-RTO. Utrecht, Netherlands, NATO-RTO. 45, 20.21-20.11.

YANTIS, S. AND JONIDES, J. 1990. Abrupt Visual Onsets and Selective Attention: Voluntary Versus Automatic Allocation. *Journal of Experimental Psychology: Human Perception and Performance* 16, 121-134.

YEE, H., PATTANAIK, S. AND GREENBERG, D. 2001. Spatiotemporal Sensitivity and Visual Attention for Efficient Rendering of Dynamic Environments. *ACM Transactions on Graphics* 20, 1, 47-54.

YEE, Y. 2000. *Spatiotemporal Sensitivity and Visual Attention for Efficient Rendering of Dynamic Scenes*. Masters Thesis. Program of Computer Graphics, Ithaca, USA, Cornell.

YUN, H., GUENTER, B. AND MERSEREAU, R. 1997. Lossless Compression of Computer Generated Animation Frames. *ACM Transactions on Graphics* 16, 4, 359-396.