Gaze-Contingent Video Resolution Degradation*

Andrew T. Duchowski† and Bruce H. McCormick
{*andrewd* | *mccormick*}*@cs.tamu.edu*
Department of Computer Science
Texas A&M University
College Station, TX, 77843-3112

ABSTRACT

Experimental results are presented evaluating two wavelet-based gaze-contingent video resolution degradation methods under three foveal Region Of Interest (ROI) placement strategies. ROI placement is described by the introduction of a novel visualization of viewers' scanpaths, termed *Volumes Of Interest (VOIs)*. VOIs represent foveal loci of gaze in 3D space-time. Three ROI placement strategies, *ideal*, *preattentive*, and *aggregate*, are used to determine the location of an unprocessed, dynamic spatial resolution region corresponding to the projected dimension of the fovea. Results indicate imperceptible degradation effects of both *ideal* and *preattentive* strategies under a visual tracking paradigm.

Keywords: Eye Movements, Volumes Of Interest, Gaze-Contingent Video, Spatial Resolution, Wavelets.

1 INTRODUCTION

The goal of the experiment presented in this paper is to test the perceptibility of peripheral spatial image degradation of video sequence frames. Specifically, two variable resolution mappings, generated by wavelet-based gaze-contingent video frame reconstruction, are tested against an unprocessed sequence for perceptible peripheral degradation effects. The resolution mappings are denoted by LIN (for linear), HVS (for human visual system), and ORG (for original, or unprocessed).[3] HVS degradation follows a resolution function derived from empirical quantification of human visual acuity. LIN is a lower bound function (in terms of relative resolution) generating more rapid degradation of the periphery. Details of the wavelet-based multiple-ROI degradation method are described elsewhere.[2]

2 VIDEO SEQUENCES

Three 8-second video sequences are used as stimulus. Each sequence is degraded by the LIN, HVS, and ORG mappings contingent upon expected locations of intra-frame ROIs. Three different intra-frame ROI localization strategies are used in each of four experimental sessions, namely the *ideal*, *preattentive*, and *aggregate*, described

*This research was supported in part by the National Science Foundation, under Infrastructure Grant CDA-9115123 and CISE Research Instrumentation Grant CDA-9422123, and by the Texas Advanced Technology Program under Grant 999903-124.

† As of 01/98 the author can be reached at *andrewd@cs.clemson.edu*, Department of Computer Science, 451 Edwards Hall, Clemson University, Clemson, SC 29634-1906.

below. The *flight* sequence was chosen for its suitability for visual tracking. In the sequence, a Navy fighter chase jet, in the lower right portion of the image, follows a smaller target plane which starts at the top right portion of the first frame. The target plane performs a half-roll evasion maneuver arcing from the top-right to the top-left screen regions through the bottom of the screen. Sequence *flight* is used in two sessions, subject to a different ROI placement strategy in each. The *brain2* sequence simulates a virtual environment, *Exploring the Brain Forest*, being developed at the Scientific Visualization Laboratory, Department of Computer Science, Texas A&M University.[4] *Exploring the Brain Forest* presents hierarchical views of the brain at several levels of scale from a global overview to immersion within its forest of neurons and glial cells. The *brain2* video sequence simulates a "fly-through" of such a neural forest. To provide a visual tracking cue, a crosshair is generated in the center of the video frame. The *cnn* sequence is chosen for its content of a human face. The sequence contains relatively little motion, but is representative of potential imagery found in telephony. The sequence is also well suited for visual attention and eye movement studies since it contains several distinguishable features which may serve as attentional attractors.

3 VOLUMES OF INTEREST

In general, the VOI model of eye movements amalgamates distinct (i.e., multiple viewers') scanpaths into a consolidated spatio-temporal description. Figure 1 shows the abstract conceptualization of the aggregate VOI model. Vertical lines represent time slices, e.g., uniformly sampled video frames, as a temporal reference. Figure 1

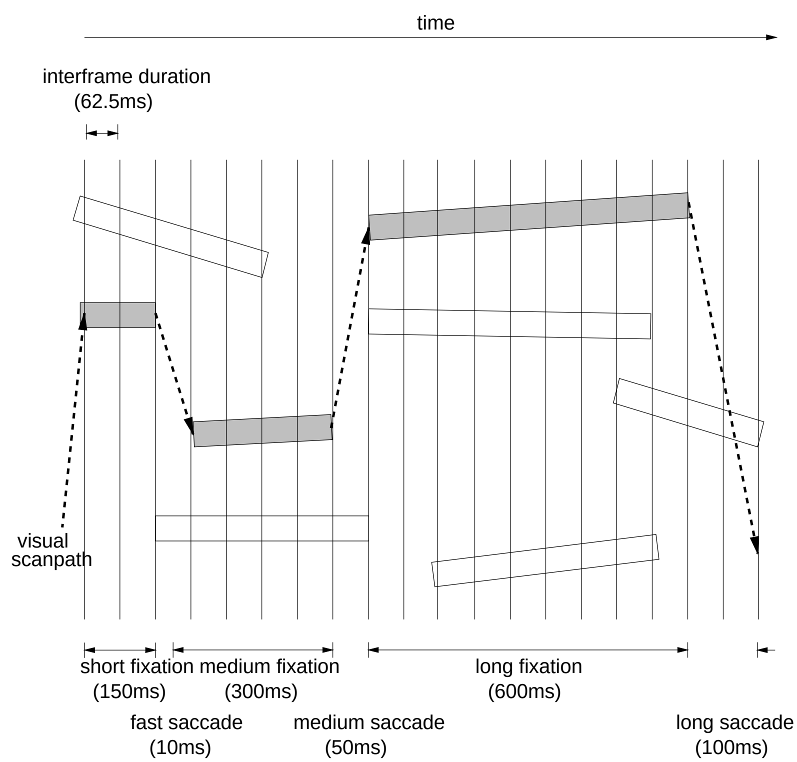

Figure 1: Graphical Volume Of Interest model.

highlights a hypothetical individual scanpath included in the VOI collection. This scanpath may correspond to a past observation of actual view patterns, or may present a candidate future scanpath. That is, observed eye movement patterns, represented by Volumes Of Interest, are concatenated to form a history of multiple viewers' scan patterns over a particular sequence. The three-dimensional eye movement visualization technique relies on the identification of dynamic fixations which, represented by discrete pixel regions in a video sequence, match foveal intra-frame loci of gaze, referred to as Regions Of Interest (ROIs). Inter-frame ROIs are merged to visualize foveal *Volumes Of Interest* (VOIs), providing a depiction of dynamic foveal vision.

3.1 Aggregate Volumes Of Interest

In their early work on eye movements, Noton and Stark investigated viewing patterns of multiple subjects.[5,6] Although recorded scanpaths exhibited significant variability in how different individuals view a scene (inter-subject variability), or for that matter how an individual's scanpaths differ from session to session (intra-subject variability), the authors identified "informative details" as common fixation points. To show this graphically, several images were used to illustrate scanpath variability and the common location of interesting features. Figure 2 shows scanpaths recorded from 7 subjects over the latter 4 seconds of the *cnn* sequence. The Volume Of Interest

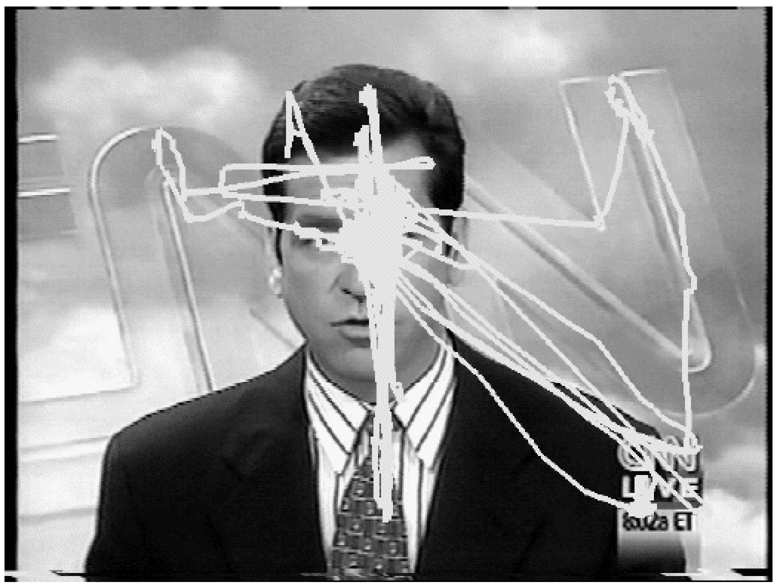

Figure 2: Aggregate scanpaths.

visualization extends the representation of individual scanpaths in space-time to the display of aggregate eye movement trajectories. Figure 3(a) depicts the same scanpaths shown in Figure 2 in three dimensions. The interactive nature of the visualization permits closer inspection of VOI-frame intersections yielding two-dimensional Regions Of Interest at particular points in time, as shown in Figure 3(b). The aggregate representation shows the convergence of multiple scanpaths over the anchorman's face through most of the sequence. Peripheral regions such as the timebox are fixated sporadically propounding the region as one of diminished relevance.

Volumes Of Interest representing spatio-temporal segments of the stimulus constitute potential attractors of visual attention. After its consolidation, the individual scanpath loses its significance in the aggregate model. At any point in time the intersections of multiple VOIs and a video frame constitute potential spatio-temporal candidates for attentive inspection as evidenced by their historical selection by previous viewers. In this sense, aggregate VOIs are similar to Noton and Stark's informative details identified over still images except VOIs depict these details in space-time.

(a) Aggregate VOIs. (b) Aggregate VOI-frame intersections.

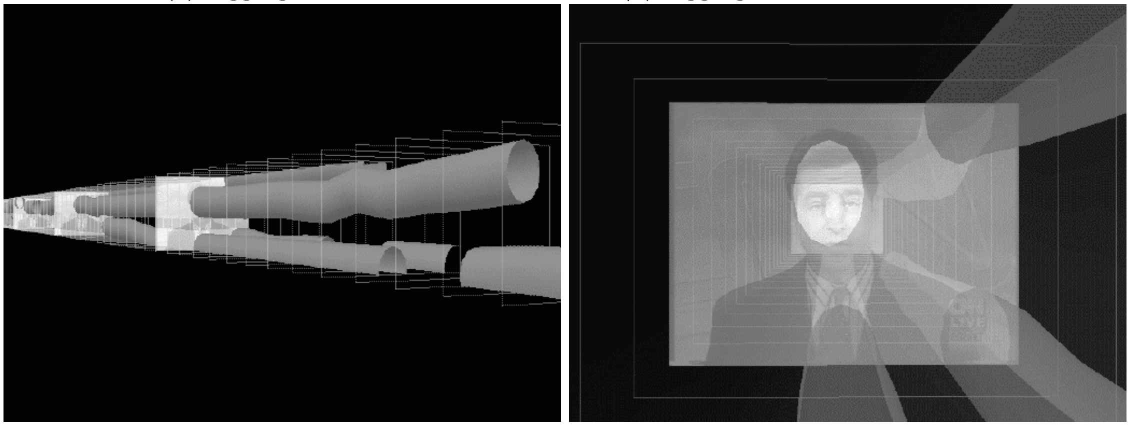

Figure 3: Volumes Of Interest.

4 ROI PLACEMENT STRATEGIES

Three strategies are used to vary the ROI localization within video sequences: (1) The *ideal* strategy utilizes a VOI based on an "ideal observer". The location of the visual target (ROI) is manually measured in each video frame of the sequence and the resulting VOI is assembled from this ROI list. (2) The *preat* (for preattentive) strategy utilizes a VOI obtained from one subject's observed eye movement patterns. The chosen subject provided eye movement patterns considered slightly better than average, in terms of gaze position error relative to the expected stimulus target. When asked whether the subject had any previous visual tracking training, the subject answered in the negative but noted hunting as an enjoyable hobby. The subject is referred to as "hunter" and provides model eye movement patterns for preattentive video sequence processing. This ROI placement strategy is described in further detail below. (3) The *agg* (for aggregate) strategy utilizes an aggregate VOI constructed from eye movement patterns of several viewers. The rationale for this strategy is the evaluation of the aggregate VOI construction as a gaze prediction method. The expectation here is that multiple viewers will identify all the potential spatio-temporal segments of the video stream that need to be presented at full resolution for complete perception of the sequence. Following this argument, it is hypothesized that perception will not be impeded if and only if these segments are displayed at high resolution. Unlike *ideal* and *preat*, however, the *agg* strategy does not explicitly rely on a common visual scanpath. Note that both the *ideal* and *preat* strategies are limited to the choice of viewing paradigm where it can be reasonably expected that viewers will follow a similar viewing pattern. That is, the viewing paradigm must be based on a suitable viewing task, e.g., visual tracking.

4.1 Preattentive ROI Placement

The preattentive ROI placement strategy anticipates the viewer's future point of regard. To clarify, a human subject's eye movement patterns typically include "breaks" or "holes" in the resulting VOI due to saccades, or blinks. In contrast, the VOI of an "ideal observer" is a continuous dynamic fixation. Intra-frame ROIs are defined as intersections of VOIs and video frames. In the case of discontinuous VOIs derived from a human subject, certain frames occur between continuous VOI segments resulting in a null VOI-frame intersection. The lack of an intra-frame ROI results in overall degradation of the frame; inclusion of such a frame in the video sequence results in a sudden, brief loss of resolution. Since human vision is practically blind during saccades and blinks, this resolution loss theoretically does not pose a problem. The prediction of a saccade or blink, however, is not currently possible. For this reason, since it is believed that a sudden loss of resolution would impede perception, VOI "holes" need to be filled in by extending past or future VOI segments in the temporal VOI stream.

Extension of future VOI segments to the past results in an anticipatory video stream, in terms of foveal ROI placement. Consider frame f which happens to occur within a VOI "hole", as depicted in Figure 4. Assume also

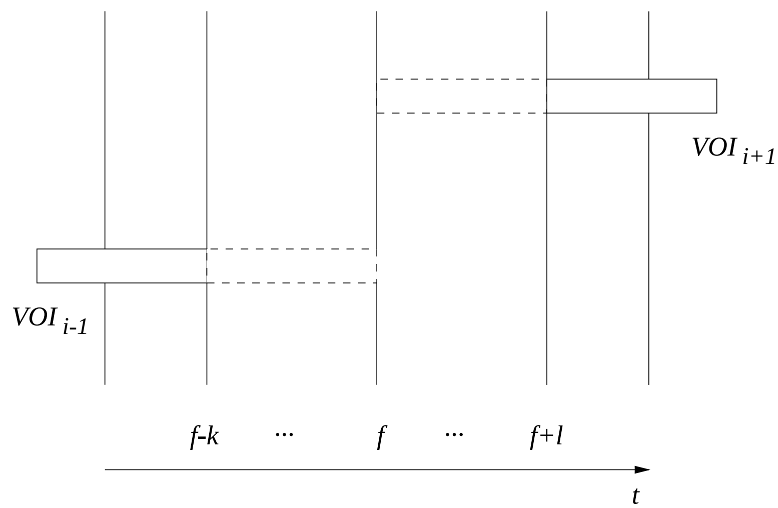

Figure 4: Schematic VOI extension.

that the last VOI (VOI_{i-1}) terminates at frame $f-k$, and that the next VOI (VOI_{i+1}) commences at frame $f+l$. Since the frames $f \in [f-k, f+l]$ constitute a discontinuity (e.g., due to a saccade), the spatial locations of VOI_{i-1} and VOI_{i+1} need not correspond. To ensure that perception is not disrupted, an artificial Region Of Interest must be inserted into frame f. Determination of the location of the ROI at frame f is the source of the current problem. A pragmatic solution to the VOI "hole" problem is the extension of either VOI_{i-1} or VOI_{i+1} over frame f. Extension of VOI_{i-1} into the future is a reactive solution since it assumes that the fixation will persist until frame f. Extension of VOI_{i+1} into the past is an anticipatory strategy since a region of high resolution is introduced into the video stream prior to the expected arrival of a fixation. In essence, this strategy is based on the assumption of attention preceding a fixation change and is adopted as the so-called *preat* VOI strategy.

5 EXPERIMENTAL STIMULUS

Four experimental sessions were conducted based on the following combinations of video sequence and ROI placement strategy: (1) *flight* (ideal), (2) *flight* (preat), (3) *brain2* (ideal), and (4) *cnn* (agg). Ideal observer VOIs were created for both the *flight* and *brain2* sequences. For the *flight* sequence, the VOI follows the target plane, as illustrated in Figure 5. The *flight* sequence is also used in combination with the preattentive VOI strategy. Figure 6 shows the VOI of subject "hunter" over the *flight* sequence. Missing VOI sections are "holes" corresponding to discontinuities identified in the subject's eye movement patterns, presumably corresponding to saccades or eye blinks. In the *brain2* sequence, the ideal observer's VOI follows an artificially overlayed crosshair in the center of the video frame, as shown in Figure 7. The crosshair serves as a visual cue easily located by subjects. Aggregate VOIs (shown in Figure 3) used to generate the *cnn* sequence were composed using eye movement data collected from 7 subjects selected for relatively small average calibration error ($\sim$1°-2° visual angle). The aggregate model does not extend VOIs as in the *preat* model, hence high resolution intra-frame regions are "turned off" within VOI "holes".

(a) Left perspective. (b) Right perspective with zoom.

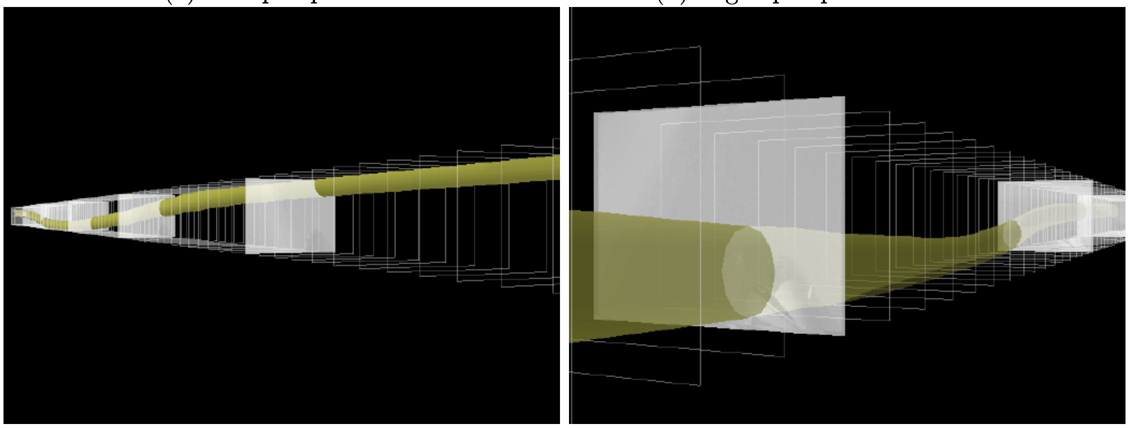

Figure 5: Expected VOIs: *flight* (ideal observer).

(a) Left perspective. (b) Right perspective with zoom.

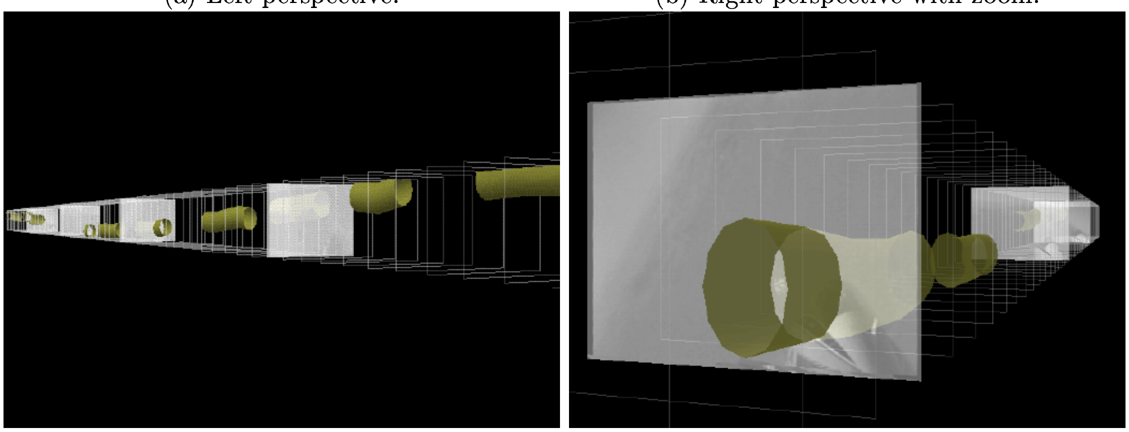

Figure 6: Expected VOIs: *flight* (preattentive).

(a) Left perspective. (b) Left perspective with zoom.

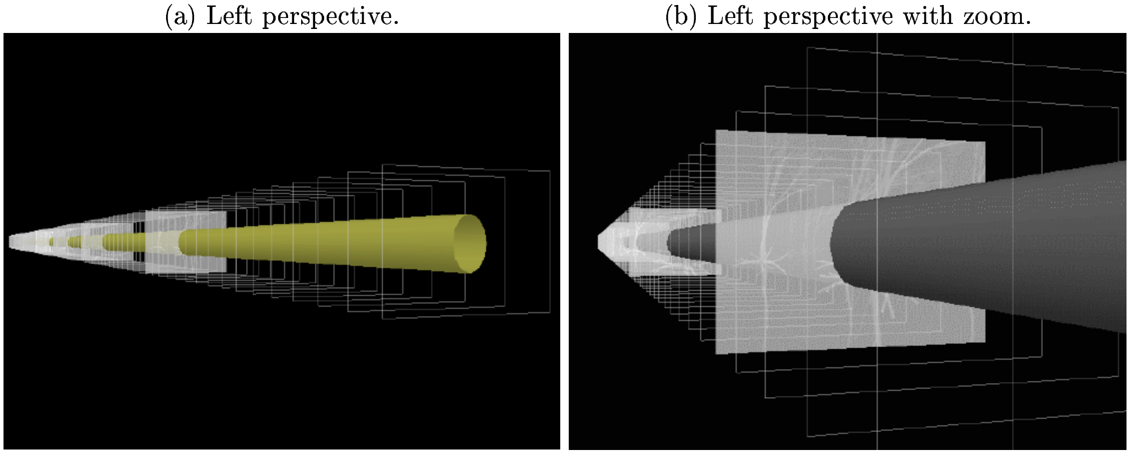

Figure 7: Expected VOIs: *brain2* (ideal observer).

6 EXPERIMENTAL APPARATUS

The gaze-contingent video experiment was conducted via a *gaze-contingent video* (gcv) system developed in the Virtual Environments Laboratory, Department of Computer Science, Texas A&M University. The gcv system is composed of a real-time (60Hz) eye tracker, an SGI 2-processor Onyx® RealityEngine2™ equipped with a Sirius Video™ broadcast-quality video board, and a 21″ standard NTSC television.[1] The front end of the eye tracker is shown in Figure 8. In the present configuration, the eye tracker, treated as a black box, delivers fixation (x, y)

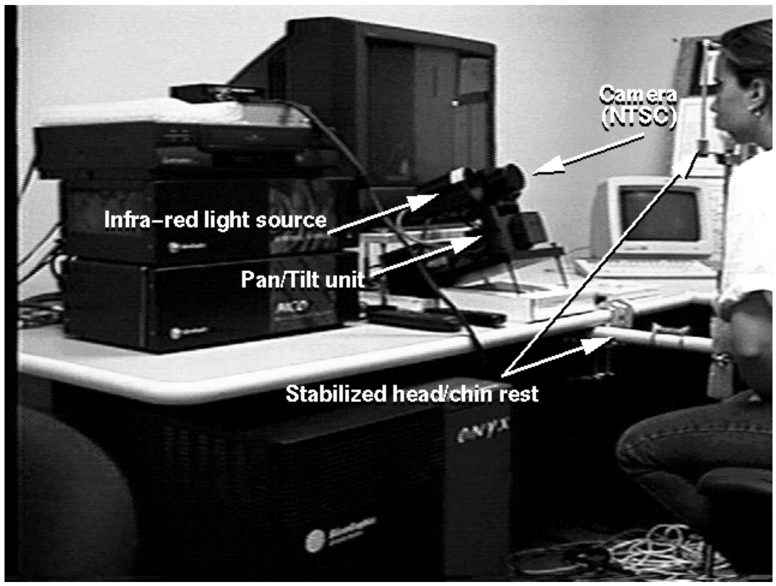

Figure 8: Virtual Environments Laboratory: eye-tracking apparatus.

coordinates over a 19.2 Kbaud RS-232 serial line, while video frames are simultaneously transferred through the video board. The apparatus hardware and software are described in detail elsewhere.[1] System limitations dictate an empirically optimum video sequence duration and display rate for experimentation. Video sequences composed of 128 NTSC (640 ×480) frames are shown at 16fps providing 8 seconds of stimulus duration. Since the primary goal of the experiments was the testing of the gaze-contingent variable resolution display strategy, monochrome (greyscale) video was used.

7 EXPERIMENTAL DESIGN

Processed video sequences (LIN, HVS, and ORG) were presented in random order. The stimulus video was shown twice to each subject after viewing a training video. This constitutes a modification of the CCIR Recommendation 500 double stimulus impairment method (sequence ABAB where A is the unprocessed sequence). The resulting sequence order used is ABB.

7.1 Subjects

A total of 16 subjects (7 female, 9 male) participated in the experiment. The age distribution was mean 19.4, minimum 18, and maximum 22. The subjects were recruited from the Department of Psychology undergraduate

[1]Silicon Graphics, Onyx, RealityEngine2, are registered trademarks of Silicon Graphics, Inc. Sirius Video is a trademark of Silicon Graphics, Inc. As of this writing, see URL: `http://www.sgi.com/Misc/external.list.html` for a complete list of trademarks.

pool. All subjects had good vision with 3 subjects wearing glasses, 2 wearing contact lenses. Subjects were randomly divided into equal groups of 4 for each of 4 experimental sessions. Each session pertained to a particular video sequence (*flight* (ideal), *flight* (preat), *brain2* (ideal), *cnn* (agg)) following a block-randomized experimental design.

7.2 Experimental Trials

Each experimental session consists of multiple subjects tested individually in experimental trials. Each trial consists of three presentations of one video sequence processed in the LIN, HVS, and ORG manner. Trials are limited to three video presentations since loading of the video into memory requires 7 minutes and the experiment is designed to process each individual in roughly 30 minutes. Each trial consists of the following steps: (1) Brief introduction. The equipment is described with emphasis on safety. Subjects are assured of the unobtrusive nature of the experiment. (2) Training. Prior to the display of LIN, HVS, and ORG sequences, the ORG sequence is shown three times from video tape. (3) Stimulus presentation. External calibration of the eye tracker is performed before stimulus presentation. The video sequence is presented twice in succession; eye movements are recorded each time. Internal software calibration is performed before and after stimulus presentation to record the initial and final accuracy of the eye tracker. (4) Stimulus judgement. The subject is asked to mark the perceived level of impairment on a 5-point impairment scale (derived from the CCIR Recommendation 500 rating scale,[7] where values of 1 and 5 are assigned to IMPERCEPTIBLE and VERY ANNOYING judgments, respectively), relative to the ORG sequence.

8 RESULTS

The objective of the experiment is to evaluate whether there is any perceptible difference in the different resolution mappings under visual tracking and free viewing conditions. A specific hypothesis being tested is whether the spatial resolution degradation of the HVS mapping is imperceptible under either or both conditions. The resolution mapping hypothesis is examined through the ANalysis Of VAriance (ANOVA) of subjective quality ratings. If there is no statistically significant difference between mean ratings of the HVS and ORG sequence, then it is reasoned that there is no perceptible effect of the HVS mapping.

Usefulness of subjective ratings of the video sequences are contingent on three factors of the eye tracking experiment: (1) accuracy of the eye tracker, (2) accuracy of the eye tracker during the viewing task, and (3) accuracy of gaze position over the intended Regions Of Interest. Eye tracker accuracy was measured by internal calibration procedures. Measurements were taken before and after stimulus viewing trials. These measurements provide the basis for two statistical measures: (1) the overall accuracy of the eye tracker, and (2) the amount of instrument slippage during stimulus viewing. The latter measurement gives an indication of the instrument accuracy during the viewing task, i.e., by recording loss of accuracy between the before- and after-viewing calibration procedures.

8.1 Verification of Eye Tracker Accuracy

Eye tracker readings were obtained over 30 internal calibration points. Each calibration point measurement consists of eye tracker samples about the calibration point over an 800ms period (approximately 44 individual data points). Raw sample points falling in the exterior 10-pixel wide borders are ignored. This is due to the eye tracker's property of generating (0,0) values during blinks (confirmed by the vendor). An average of valid data points (centroid) is obtained and the error between the centroid and calibration point is calculated. Each two-dimensional euclidian distance measurement is converted to the the full visual angle dependent on the viewing distance and the calculated resolution of the television screen. Thus each calibration run contains 30 average deviations at each calibration point measured in terms of visual angle. A graphical example of this measurement is shown in Figure 9. The internal calibration locations are represented by +, sample measurements are

(a) Calibration before stimulus viewing (avg. error: 0.86°). (b) Calibration after stimulus viewing (avg. error: 1.46°).

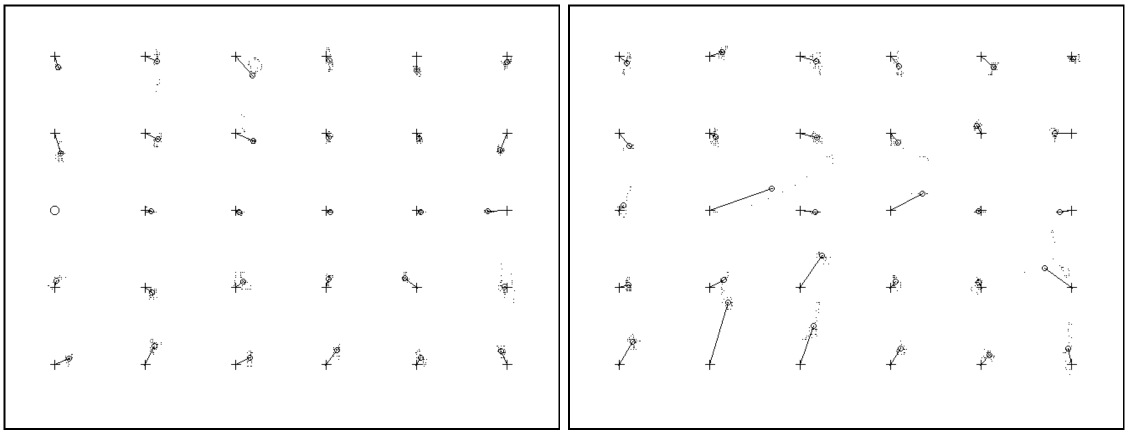

Figure 9: Typical per-trial calibration data (subject # 11).

represented by individual pixel dots, and centroid gaze positions are represented by small circles, joined with the corresponding calibration point by a line. The length of the line is the average error deviation in pixels. A large circle indicates loss of data at the corresponding calibration point.

To quantify the overall eye tracker accuracy succinctly, the average calibration error is obtained from each set of calibration points in order to calculate an overall average statistic of the eye tracker. The resulting instrument error measure is an average statistic over all calibration runs performed in the experiment. The analysis of average errors shows a skewed distribution, suggesting the use of the median value as a more robust measure. Its value is 1.42°. Using similar reasoning for reporting a dispersion statistic, the interquartile range (iqr) is utilized instead of the standard deviation, and its value is 1.10°. These findings indicate an overall acceptable performance, not far off from the vendor's claimed accuracy (roughly 1° visual angle).

8.2 Verification of Eye Tracker Slippage

Quantification of the before- and after-viewing eye tracker error provides a measure of instrument accuracy during the viewing task. Notice that measured eye positions in relation to calibration points coincide well overall. To quantify this correspondence, a one-way ANOVA was performed on the means of the before- and after-viewing average error measures. On average, no significant slippage is detected by this statistic.

To examine eye tracker slippage on a per-trial basis, differences of average errors were calculated between the before- and after-viewing calibration runs on a per-run basis. Difference measurements fit a skewed distribution, suggesting the use of statistical measures robust to outliers. The median error is -0.18°, and interquartile range is 0.54°. Over some calibration points, accuracy improves while over others it degrades. Since the peak frequency is close to 0, generally the accuracy before and after viewing the stimulus remains stable overall.

8.3 Verification of Gaze Position

To verify subjects' gaze position with respect to expected intra-frame Regions Of Interest (ROIs), eye gaze position error was calculated on a per-frame basis over all participating subjects' data. Estimated Volumes Of Interest (VOIs) from the raw Point Of Regard (POR) data identified subjects' fixation locations. These locations were then compared to expected (ideal) VOI locations. Reference (ideal) VOIs were compared to each subject's observed VOIs by dismantling the VOIs into VOI-frame intersections (ROIs). The error measure was calculated as the distance between ideal and observed ROIs for each frame and converted to degrees visual angle. In the case of the aggregate VOI stimulus, subjects' intra-frame fixation locations are compared to the closest expected intra-frame ROI (based on a Euclidian measure). A median error value from each subject's VOI data is used as a representative measure of the gaze error over the entire sequence. Outlier measured error values corresponding to suspected blinks were dropped from the median calculation.

Overall gaze position error was obtained over each sequence. Not surprisingly, the *brain2* (ideal) sequence incurred the least amount of gaze positional error. This was expected since subjects were instructed to maintain steady gaze position at a central location of the display. Two-way ANOVA of the medians indicates no significant difference in mean values (means of medians, $p = 0.7178$). However, the difference between experimental VOI paradigms is significant ($p < 0.0000$), indicating that the VOI presentation strategy (ideal, preat, agg) is a factor in observed gaze position error. To investigate further, one-way analysis of variance was performed between the mean of median gaze errors within each experimental VOI paradigm.

Of all viewing conditions, the aggregate VOI strategy, utilizing multiple intra-frame ROIs, appears to present the most difficult visual tracking task. On the one hand, qualitative observations of this task suggest that multiple regions, when easy to differentiate from background imagery, tend to disrupt normal viewing patterns. Subjects exhibited some confusion as to which ROI to fixate, and in general, found these sequences annoying. Subjective quality ratings support this observation and are presented in the next section. On the other hand, quantitative gaze error measurements suggests that, on average, viewers were capable of fixating a particular intra-frame ROI with surprising accuracy (overall median gaze error $2.98°$ visual angle). Interestingly, this median of medians error measurement is not significantly different from the ideal VOI condition, and is only somewhat worse (in term of statistical significance) from the preattentive condition.

Overall, the analysis of gaze error suggests two important points: (1) viewers have little difficulty in matching gaze to expected locations within fixation (*brain2* (ideal) sequence) and tracking (both *flight* sequences) paradigms, and (2) the degree of error indicates close proximity to the intended region of interest. On average, recorded gaze positions did not vary more than roughly the foveal dimension as projected on the stimulus display. Interestingly, performance is somewhat better over the preattentive sequence than over the ideal observer sequence. This is surprising since viewers were expected to foveally match the ideal target. Qualitatively, gaze of the subject ("hunter"), whose VOI data was used to generate the (preat) sequence, tends to slightly fall behind the moving target. It is interesting to note that in this experiment, "hunter"'s viewing patterns provide a better prediction of scan patterns than the ideal observer.

8.4 Perception Impairment Analysis Over All Conditions

Mean results for the 16 subjects were analyzed by two-way ANOVA. The analysis finds no evidence of variability across resolution mappings (LIN, HVS, ORG), but there is strong indication ($p < 0.0000$) of variability across different viewing conditions (ideal, preat, agg). Furthermore, there appears to be evidence of interaction between resolution mapping and viewing condition factors ($p < 0.03$). At first glance these findings suggest no perceived difference in mappings. However, due to the strong evidence of variability across viewing conditions, interactions between each condition must be further analyzed to gain insight into the results. The goal of the analysis is to ascertain whether there is any perceived difference between the LIN, HVS, and ORG mappings. Overall analysis in this case is meaningful only in the sense that it indicates a statistical difference between sequence viewing conditions.

8.5 Perception Impairment Analysis Between Conditions

The overall analysis was broken down to identify statistically significant interactions between viewing conditions. Two-way ANOVA statistics were examined to test for significant differences between viewing conditions. The analysis suggests no significant difference between the preattentive and ideal viewing conditions, and no significant difference between sequence types over the ideal (and hence preattentive) viewing condition. By deductive reasoning, the condition that appears to stand out is the aggregate condition. Analysis within conditions is carried out below in an attempt to narrow down the effect that appears to set the aggregate experiment apart from the other conditions. Analysis within conditions is also performed to draw conclusions regarding resolution mappings.

8.6 Perception Impairment Analysis Within Conditions

The primary interest in performing the within-condition analysis is to test for significance between impairment ratings of the three resolution mapping factors (LIN, HVS, ORG). One-way ANOVA was performed over results from each viewing condition. In the ideal and preattentive cases, no significant difference between subjective quality ratings was found. In the aggregate case, a siginifcant difference ($p < 0.01$) between subjective ratings emerged. The analysis suggests that the LIN mapping is significantly different from the HVS and ORG mappings. A comparison between each pair of resolution mapping effects supports this finding.

9 DISCUSSION

Interpretation of the statistical analysis suggests the HVS resolution mapping produced imperceptible degradation of the video stream under ideal and preattentive visual tracking conditions. This result was expected since the degradation strategy was designed to match human visual acuity. Surprisingly, the linear mapping (LIN) also generated imperceptible degradation results (under ideal and preattentive conditions). Results between the ideal and preattentive conditions suggest that VOI-based prediction of eye movements is effective for visual tracking tasks. Before this contention was tested, an aggregate VOI was compiled from several subjects' eye movement patterns during a visual tracking task. Since the aggregate VOI qualitatively resembled the ideal observer's, the goal of this experiment was to test the minimal condition of using only one subject's patterns as a predictor for resolution degradation in the video stream (*flight*). Results suggest that one viewer's eye movement pattern is sufficient for this purpose.

VOI-based prediction of eye movements may not be effective for natural (free) viewing tasks. This hypothesis was tested through the degradation of the *cnn* (agg) sequence. Results show that under natural viewing conditions (no imposed viewing task), aggregate VOIs impair perception (or perhaps more fairly natural eye movements). It is suspected that the problem resides in the discontinuous nature of the VOIs. In this experiment, VOIs were not extended as in the preattentive case. The resulting video sequence possessed a bubbling quality of high resolution regions. That is, in a low resolution peripheral region, a high resolution inset suddenly appears.

10 CONCLUSION

Evidence suggests that under certain conditions, a significant amount of information may be withheld in gaze-contingent visual displays with little perceptible effect. Surprisingly, experimental results suggest that under the visual tracking paradigm, effects of both linear and acuity-matching resolution mapping are imperceptible. Dyadic linear mapping results in roughly 50% resolution degradation over 97% of the image (50% resolution beyond the 105×105 foveal ROI over a 640×480 image frame). This information reduction has significant implications for gaze-contingent image and video compression. Further research is required to test HVS-matching multidimensional digital compression of imagery, e.g., degradation of color, contrast, and motion.

The visual tracking paradigm utilized in human subject experiments limits the generalizability of results. Visual tracking assumes *a priori* positional gaze information. In free-viewing, the effects of the linear mapping degradation are clearly visible. In fact, using this degradation, the sudden onset of displayed high resolution areas tends to distract natural viewing patterns, as demonstrated by the aggregate VOI condition. Further research is needed to determine whether (a) a milder form of degradation can offer imperceptible results with significant savings in terms of compression, and/or (b) linear (or stronger) degradation effects can be used to induce eye movement patterns affecting perception without impeding performance. Current results provide positive support for both avenues of research.

11 REFERENCES

[1] Andrew T. Duchowski. Incorporating the Viewer's Point-Of-Regard (POR) in Gaze-Contingent Virtual Environments. In *The Engineering Reality of Virtual Reality'98*, Bellingham, WA, January 1998. SPIE.

[2] Andrew T. Duchowski. Representing Multiple Regions Of Interest with Wavelets. In *Visual Communications and Image Processing'98 (VCIP)*, Bellingham, WA, January 1998. SPIE.

[3] Andrew T. Duchowski and Bruce H. McCormick. Simple Multiresolution Approach for Representing Multiple Regions of Interest (ROIs). In *Visual Communications and Image Processing'95 (VCIP SPIE vol. 2501)*, pages 175–186, Bellingham, WA, May 23-26 1995. SPIE.

[4] Bruce H. McCormick, David A. Batte, and Andrew T. Duchowski. A Virtual Environment: Exploring the Brain Forest. Presentation at the CENAC Conference, October 1996, Mexico City, Mexico (preprint available from principal author, Department of Computer Science, Texas A&M University, College Station, TX).

[5] David Noton and Lawrence Stark. Eye Movements and Visual Perception. *Scientific American*, 224:34–43, 1971.

[6] David Noton and Lawrence Stark. Scanpaths in Saccadic Eye Movements While Viewing and Recognizing Patterns. *Vision Research*, 11:929–942, 1971.

[7] Lew B. Stelmach and Wa James Tam. Processing Image Sequences Based on Eye Movements. In *Conference on Human Vision, Visual Processing, and Digital Display V*, pages 90–98, San Jose, CA, February 8-10 1994. SPIE.