



Where people look when watching movies: Do all viewers look at the same place?

Robert B. Goldstein*, Russell L. Woods, Eli Peli

The Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA

Received 29 July 2005; received in revised form 11 August 2006; accepted 14 August 2006

Abstract

Magnification around the most important point of a movie scene (center of interest—COI) might aid people with visual impairments that cause resolution loss. This will be effective only if most people look at the same place when watching a movie. We recorded the eye movements of 20 normally sighted subjects as each watched six movie clips, totaling 37.5 min. More than half of the time the distribution of subject gaze points fell within an area statistic that was less than 12% of the movie scene. Male and older subjects were more likely to look in the same place than female and younger subjects, respectively. We conclude that the between-subject agreement is sufficient to make the approach practical. © 2006 Elsevier Ltd. All rights reserved.

Keywords: Eye movements; Magnification; Scanpath; Video; Visual aids; Low vision; Visual impairment; Television

1. Introduction

People who, due to eye diseases, suffer loss of visual resolution (low vision) could benefit from modified information displays. The most commonly used modification is magnification. While magnification is effective in improving resolution it inherently restricts the field of view. Thus, magnification may impede the acquisition of information attained in normal vision by the use of scanning eye movements. This limitation may be addressed by dynamic control of the displayed information. Dynamic control of large text presentation is helpful for people with low vision [1–4]. We propose a similar approach to improve access to movies and television programs.

Magnifying moving images using electronic zoom [5] would enable users to select and vary the desired level of magnification from time to time. However, only part of the magnified scene can be presented on the screen. Consequently, large parts of the scene become invisible. To be useful the magnified image should be centered on the screen around the center of interest (COI)—an area in the scene that is most relevant for the

viewer. Manual zoom-and-roam devices are available in commercial television systems (e.g. DVD players). However, the rapid changing of scenes in most movies may not allow for effective manual control of the position of the magnified section of the image or even of the desired magnification. We proposed pre-selecting the COI in the scene and providing that position with each frame for automatic centering [6,7]. This selection should maintain the most relevant details in view at any magnification level set manually by the user, or alternately at a level deemed appropriate by some other automated method.

Together with DigiVision (San Diego, CA), we have developed a computer controlled zoom-and-roam device for playback of movies on a television monitor. The computer plays a DVD and simultaneously reads the COI. These coordinates are sent to the zoom-and-roam device so that the magnified image is centered on the COI coordinates. We proposed using eye movement recordings from normally sighted observers watching the movie to determine the desired COI. Although other methods of determining the COI can be envisioned, eye movement recording is automatic and objective.

Choosing the COI using eye movements is akin to finding the scanpath for a movie sequence. Much work has been done regarding the scanpath of still images [8,9] but little is known about viewing moving images. With the exception of a few studies [10–12] most developments that depend on knowing

* Corresponding author. Tel.: +1 617 912 2512; fax: +1 617 912 0112.
E-mail address: bob.goldstein@schepens.harvard.edu (R.B. Goldstein).

where the gaze is directed (e.g. compression schemes [13] and transmission of images for limited screen space [14]) assume that most people look at the same place all the time while watching movies. To our knowledge this assumption has not been verified experimentally. Here we quantify the proportion of the time a group of people look at the same place while watching a movie, and begin to examine the effects of age and gender on this behavior.

Film editors have used assumed knowledge of viewer's eye movements, and even blinks, to assemble movies [15]. Stelmach et al. [10] recorded 24 subjects viewing 15 forty-five second clips to determine if viewing behavior can be incorporated into video coding schemes. They found that there was substantial agreement among subjects in terms of where they looked. In a follow-up experiment related to gaze-contingent processing techniques [11], recorded eye movements of eight subjects were used to create a "predicted gaze position". Tosi et al. [12] recorded the eye movements of 10 subjects watching a variety of clips totaling about 1 h and reported that, qualitatively, individual differences in scanpaths were relatively small. Theoretical saliency models [16,17] predict where people will look based on spatial and temporal properties of the scene and thus, make no assumptions regarding individual differences in predictions of regions of interest. Top-down models [18–20] that predict eye movements based on scene context, seem to provide a more obvious source for individual differences. We hope to compare our data with predictions from both these types of models, in collaboration with colleagues developing such models.

Here we address three specific questions relevant to our proposed low-vision aid for viewing television. (1) To what extent do people look at the same place when watching a movie? (2) Does that vary with age and gender? (3) Does the position of the COI differ from the center of the screen? Answers to these questions will guide us in the development of dynamic controlled magnification.

2. Methods

Six movie clips were selected to span a broad range of scene activity, from stationary newscasters to athletes in motion, and to appeal to both younger and older audiences (Table 1). The video clips from DVDs were presented in a 16×9 movie format (the "movie scene area") on a 27 inch diagonal viewable

Table 1
The lengths of the six movie clips and the categories they were selected to represent

Title	Time (min:s)	Category
Big (1988) (Big)	6:29	Comedy
Any Given Sunday (1999) (Sunday)	4:12	Sports
Network (1976) (Network)	4:02	News
Blue Planet (2001) (Planet)	8:14	Documentary
Shakespeare in Love (1998) (Shakes)	7:06	Drama
Quiz Show (1994) (Quiz)	6:40	Game Show
Total per subject	37:29	

area NTSC (4×3) monitor as interlaced video at 30 frames per second (60 fields per second).

Twenty-six normally sighted subjects were seated 46 inches from the movie scene which spanned a $26.3^\circ \times 14.8^\circ$ visual angle. Subjects viewed movie clips binocularly while eye movements were recorded with an ISCAN model RK726PCI eye tracking system. The ISCAN had a nominal accuracy of 0.3° over a $\pm 20^\circ$ range and a sampling rate of 60 Hz. Thus, we could acquire two eye samples per video frame. The ISCAN compensated for modest head movements, permitting gaze monitoring without head restraint, thus allowing a comfortable and natural viewing situation. Before each subject viewed the videos, the ISCAN was calibrated to an area larger than the viewable area of the monitor (calibration area was $30.1^\circ \times 22.7^\circ$), using its five-point calibration scheme. This was done to minimize loss of tracking and to ensure that we obtained valid, calibrated data near the edge of the screen. To further control the quality of the eye tracking [21,22], we performed and analyzed additional five-point, pre-clip calibrations (independent of the ISCAN) before each movie clip was viewed. If we could not obtain satisfactory data yield for the pre-clip calibration we repeated the ISCAN calibration until satisfactory pre-clip calibration was achieved. A post-clip calibration was also recorded and the final analysis program averaged the pre- and post-clip calibration data in implementing the calibration parameters.

The ISCAN reports zero data during blinks and during loss of tracking. During the recording phase, immediate feedback was available regarding the amount of non-zero data acquired from each clip. If fewer than 80% were valid for any clip, the subject's data were excluded, and recording stopped. We did not repeat movie clips, as we preferred to record the subject's eye movements during their first viewing of a clip. People may view a movie clip differently on subsequent occasions, as occurs in scanpath studies when instructions are altered for viewing static images [8]. Inadequate eye sample yield happened with only one subject. Of the remaining 25 subjects, the five subjects with the highest eye sample yields in each of the four gender and age groups were selected. The 20 subjects were grouped as: younger female 18–29 y, younger male 16–36 y; and older female 51–62 y, older male 42–66 y. Eye sample data yields from these groups are shown in Table 2.

Table 2
Percent yield of acceptable eye sample data (number of non-zero out of total recorded) obtained from each movie clip for the four groups (OM—older males, YM—younger males, OF—older females, YF—younger females)

	OM	YM	OF	YF
Big	97	95	95	93
Sunday	97	95	93	93
Network	96	96	93	93
Planet	98	94	95	93
Shakes	96	95	94	92
Quiz	96	94	94	91

The yield did not vary significantly between clips. Yield was greater for male subjects ($F_{1,16} = 7.8$, $p = 0.01$) and slightly greater for older subjects ($F_{1,16} = 3.2$, $p = 0.09$).

3. Analysis

3.1. Preprocessing of individual records

The individual subjects' eye movement recordings were processed to apply the calibration to the raw eye sample data and remove recording artifacts caused by blinks and other failures. Recording could fail if the head moved too quickly to be tracked by the system or when specular reflections, such as from the tear film menisci, were erroneously detected by the ISCAN as the corneal reflection. Blinks and loss of tracking were filtered from the file by removal of records containing zero value data or frames where the pupil diameter was out of a pre-determined range.

Video presentation was controlled with a Visual Basic program that used the Microsoft DirectX 8.1 DVD interface, and eye position data were collected through RS-232 interface at 19,600 baud rate. The DVD interface only interrupted the processor every 0.4–1.0 s with a timing (frame number) information request. Frame numbers between such interrupts were calculated from the elapsed time, assuming a 30 frame per second rate. This procedure resulted, on occasion, in non-monotonic or duplicate frame numbers. Non-monotonic frames were discarded and duplicates were re-assigned in the following manner. Because ISCAN data were recorded asynchronously with the video, each assigned frame could be associated with one, two or three eye samples. We designated these multiple eye samples per frame as “subframes” (note—these subframes are not video fields). When there were more than two subframes, we re-assigned eye samples to adjacent video frames that had less than two subframes associated with them. Subframes that could not be re-assigned were discarded. For the 20 subjects watching 37 min 27 s of video, recording eye data at 60 Hz, there were potentially 2.7×10^6 eye samples. Blinks and loss of tracking during the recording phase reduced this to 2.35×10^6 (94.5%). Of these, 3.6×10^4 eye samples (1.5%) could not be re-assigned and were discarded.

3.2. Merging of eye recordings of multiple subjects' records to find extent of overlap

The 120 eye movement recording files (20 subjects \times 6 clips) were processed to count how many of those subjects had valid eye sample data for each subframe (Fig. 1). For 1.25×10^5 of the potential 1.35×10^5 subframes (93%), at least 15 subjects had valid eye samples.

For each subframe, the calibrated (x, y) coordinates of subjects' eye samples, the gaze points, were distributed over the calibration area. Various methods have been applied by other investigators to compute the coincidence between the gaze points of multiple subjects [10,23–27]. To quantify the spatial coincidence of the gaze points of all the subjects with valid eye samples, we calculated the area of the best-fit bivariate contour ellipse (BVCEA) [28,29], BVCEA has been used in the past to quantify fixation eye movement stability [2,29].

$$BVCEA = 2k\pi\sigma_H\sigma_V(1 - \rho^2), \quad (1)$$

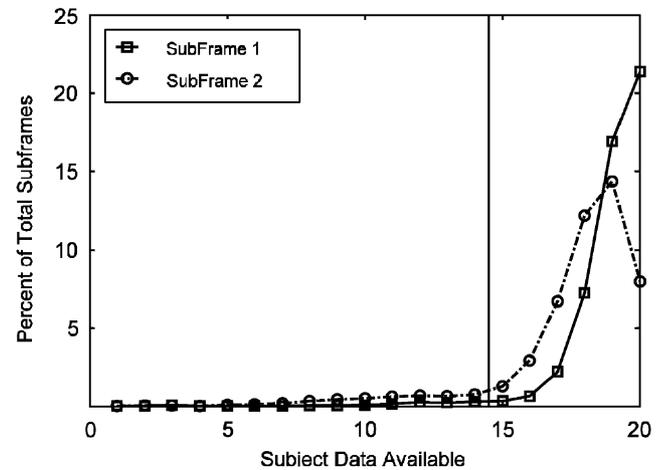


Fig. 1. The percentage of the total number of subframes for which the number of subjects had valid eye samples. For 1.25×10^5 of the potential 1.35×10^5 subframes (93%), 15 or more subjects contributed valid data (to the right of the vertical line).



Fig. 2. The gaze points (calibrated eye movement data) of the 19 (out of 20) subjects with valid eye samples for a representative subframe (number 1700) of the movie clip “Network” are shown as black and white circles. The positions were superimposed on a monochrome rendering of this video frame that was modified to improve visibility of features of the scene. Removal of the outlier (circle with the gray \times) reduced the BVCEA (measure of the spread of the gaze points) from 31.1 deg^2 (degrees square) to 19.5 deg^2 (37%). The data are plotted on the $26.3^\circ \times 14.8^\circ$ movie scene. The gray and white crosses mark the prior and subsequent COIs (mean of the gaze points), which moved by only 1.0° on removal of the outlier. Note that this COI would be outside the screen for most magnification levels if magnification were to be centered at the center of the movie scene.

where σ_H is the standard deviation of the point locations over the horizontal meridian, σ_V is the standard deviation of the point locations over the vertical meridian, and ρ is the correlation coefficient of the horizontal and vertical values. The k parameter of the BVCEA determines the enclosure of the ellipse. We set $k = 1$, for which 63% of the points would have been enclosed by the ellipse. The calculation of the BVCEA does not require that an ellipse be fitted to the data (gaze points).

Before calculation of the BVCEA, gaze-point outliers were removed from each subframe's data. To determine if a particular gaze point was an outlier, the x and y distributions of the gaze points in a subframe were considered separately. If a gaze point had an x or a y value that fell outside the 99.5% probability range (t distribution), the point was defined as an outlier. As an example of outlier removal, Fig. 2 shows outlier removal in a

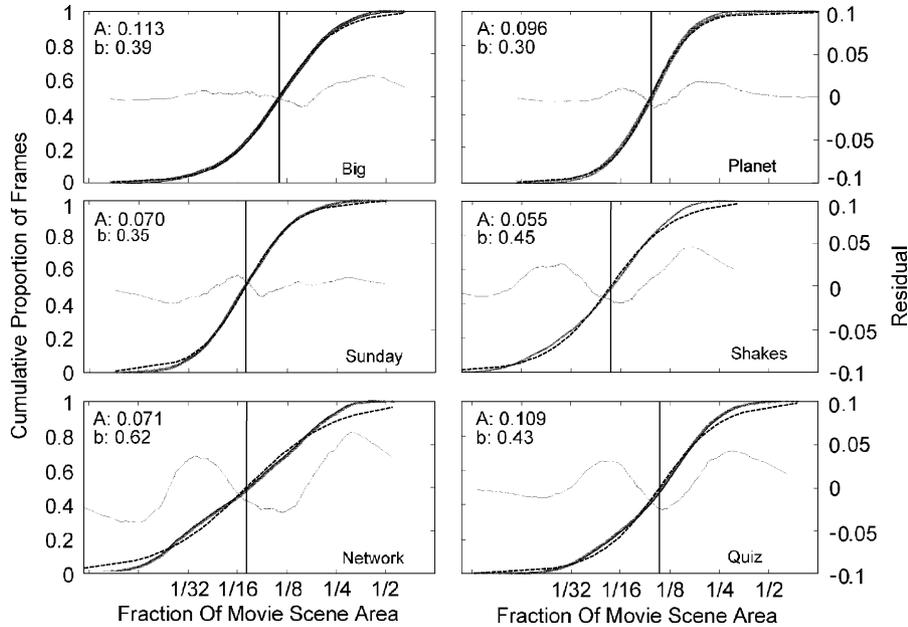


Fig. 3. For each movie clip subframe, after removing outliers (see Fig. 2), the spread of the gaze points was estimated using the BVCEA, reported here as a proportion of the movie scene area. Only those subframes where 15 or more subjects had valid eye samples were used (see Fig. 1). The cumulative curves (heavy lines) show the actual proportion of the total subframes for which the BVCEA was less than a given fraction of movie screen area. Log_e transforms of the distributions were fitted to a logistic function (dashed lines) that were then used to calculate the screen fraction for which half of the samples had a smaller BVCEA (A : vertical line and the value indicated by the inset). The residuals of the fits are shown (thin lines—see right y -axis for scale).

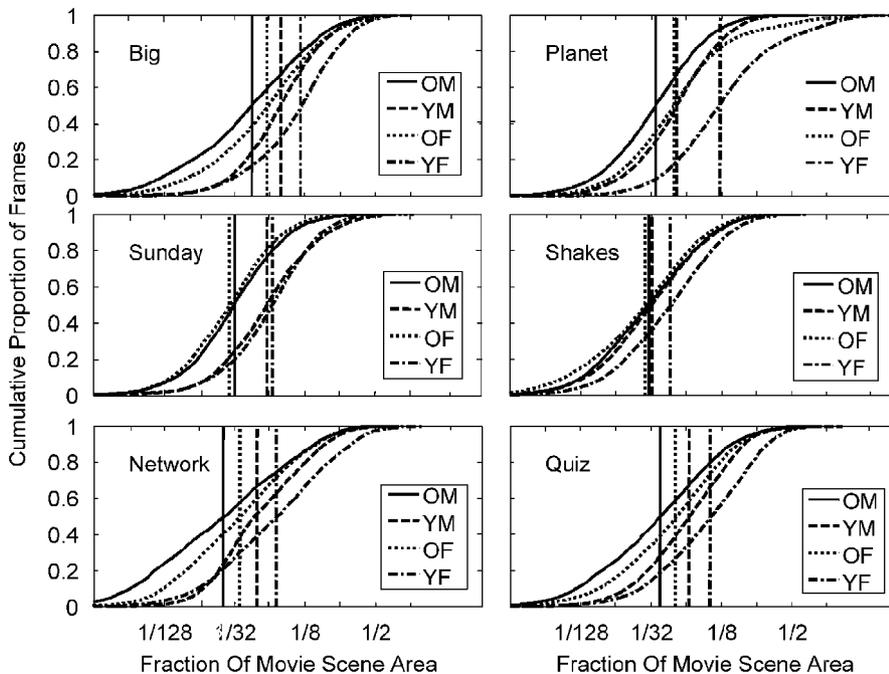


Fig. 4. Similar to Fig.3, the cumulative distributions of the BVCEA and positions of A are shown for each movie clip and for each age–gender group (for subframes where at least four out of five subjects had useable data). Older and male groups had smaller spreads of their gaze points (smaller BVCEA) than younger and female groups, respectively. The fits used to derive the positions of the vertical lines are not shown.

single subframe of data. Based on these two 99.5% confidence intervals, outliers were removed from 22.5% of subframes, and for those subframes the BVCEA was reduced by a median of 48% (inter-quartile range 40–60%) and the location of the COI

(the mean x and y coordinates of the gaze points of the group) changed by a median of 2.4° (inter-quartile range 1.2 – 4.1°). As we used a statistical definition of outliers (two independently applied 99.5% confidence intervals), if all eye samples were

drawn only from a normal distribution we expected to exclude about 1% of the eye samples. The number of outliers removed, 1.3%, was higher than that prediction ($z = 40$, $p < 0.0001$), suggesting that some of the eye samples that were removed were true outliers, like that illustrated in Fig. 2.

For subframes containing at least 15 eye samples, the cumulative distributions of the BVCEA found for each movie clip (Figs. 3 and 4) were fit with a logistic function,

$$y = 1/(1 + e^{-(x-a)/b}), \quad (2)$$

where x is $\log_e(\text{BVCEA}/\text{movie scene area})$, a is the mid-point of the function and b is the slope parameter. To quantify effects of age and gender, the BVCEA was calculated for every sub-frame for which there were gaze points for at least four subjects from each group of five subjects and fitted with a logistic function. In subsequent data analyses we used these fit values of a and b . Data were evaluated using analysis of variance, with movie clip treated as a within-subject factor. Later, for ease of interpretation, instead of reporting values of a , we report values of A , defined as e^a . A is a proportion of the movie scene area.

4. Results

People do tend to look in the same place (here represented by a small BVCEA). As shown in Fig. 3, for all six movie clips, when there were 15 or more gaze points available, more than half of the time the distribution of subject gaze points fell within an area statistic (BVCEA, $k = 1$) that was less than 12% of the movie scene. This represents an area equivalent to a circle with a diameter of about 8° . To examine the effects of age and gender on whether people look in the same place we performed analyses of variance on a and b of the fits to the BVCEA data shown in Fig. 4. As shown in Fig. 5, male and older subjects were more likely to look in the same direction (smaller a) than female ($F_{1,20} = 6.1$, $p = 0.02$) and younger ($F_{1,20} = 21$, $p < 0.001$) subjects, respectively. Older subjects were slightly more variable (slower rise—larger b) than younger subjects ($F_{1,20} = 3.9$, $p = 0.06$). Between the movie clips, there were significant differences in b ($F_{5,18} = 5.4$, $p = 0.003$) but not of a , indicating that the area within which most subjects looked was more variable for some movies. The area was more variable when watching *Network* than *Planet* ($p = 0.01$), *Big* ($p = 0.016$) and *Sunday* ($p = 0.002$). It appears that, as might be expected, movies with high levels of motion more tightly controlled the subject gaze locations than movies with relatively static scenes.

For those subframes where there were gaze-point data for 15 or more of the 20 subjects, the position of the COI was calculated. The distributions of those COIs (0.94° horizontal by 0.71° vertical bins over the movie scene area) are shown in Fig. 6 for each of the six movie clips. In general, the peaks of the COI distributions were approximately in the center of the movie scene, though they varied by as much as $\frac{1}{4}$ of the width or height from center.

The COI distributions were tested on a pair-wise basis using the one-factor independence test for similarities of distributions [30]. The distributions shown in Fig. 6 (clip by clip comparisons), as well as distributions from the different gender-age

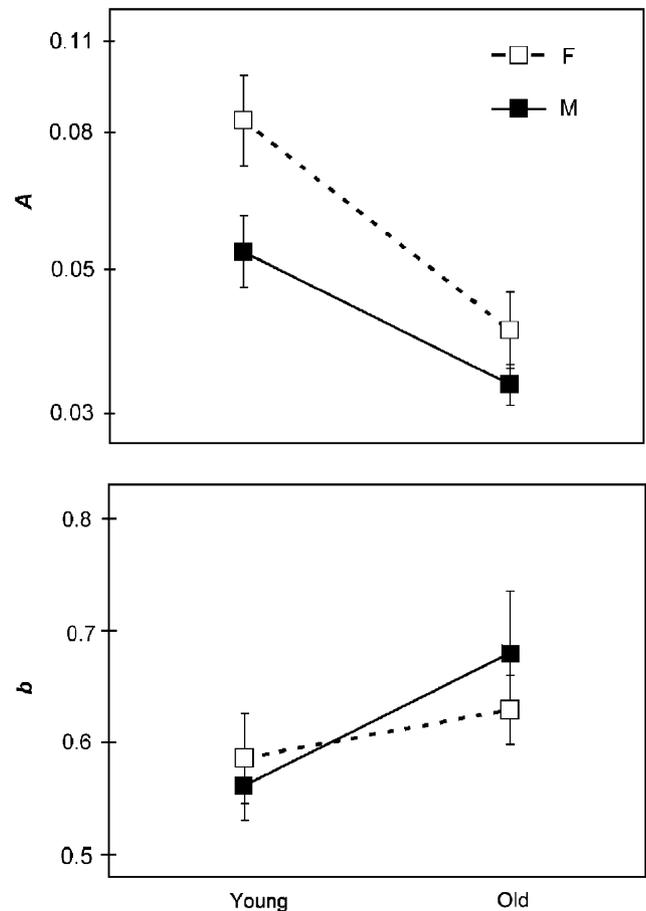


Fig. 5. The effects of age and gender on the parameters A and b of the fits of the logistic function to the cumulative distributions. There were statistically significant effects of gender and age on the likelihood that the subjects in a group looked in the same direction (smaller A for older and male subjects); and there was a small, non-significant effect of age on the variability of direction of gaze (smaller b for younger subjects). Note that the ordinate axis for the upper panel (A) is logarithmic, since the fit and subsequent statistics were done to a , ($\log_e A$). Error bars indicate SEM.

groups were tested. All distributions were statistically different from each other ($\chi^2 \geq 1271$, $p < 0.0001$). The practical significance of these statistically significant differences is not clear, as our sample sizes were very large.

5. Discussion

Measuring and providing the coordinates of the COI along with each frame may allow better use of electronic magnification as a low-vision aid for watching movies (and other television programs). The eye movement method presented here is a natural and efficient way of determining these COIs. We envisage that, just as video programs are now being provided with “closed captions for the hearing impaired” and audio description formats [31], movies can be provided with these COIs encoded.

We have demonstrated that it is possible to determine the COI in a movie scene by recording the eye movements of normally sighted subjects while they watched a movie. Over half

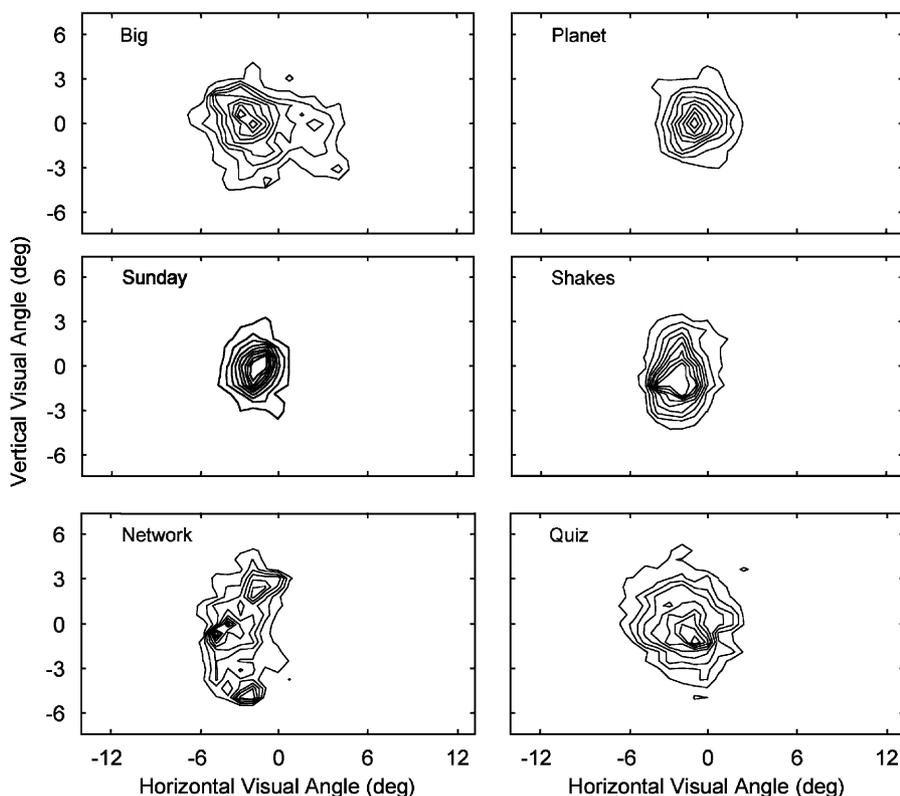


Fig. 6. COI distributions computed for all movie clips for all subjects (for subframes with eye samples available for at least 15 subjects) are shown as a topographic map (derived from 28×21 bins). The borders represent the movie scene area. Data were normalized so that the maximum was 1.0 and levels were drawn at 0.1 intervals. Although generally, the distributions peaked near the center of the screen, the spread indicates that a large proportion of the time, subjects did not look at the center of the movie scene.

of the time, the distribution of gaze points of most subjects fell within a BVCEA ($k = 1$) that was less than about 12% of the movie scene (Fig. 3) or for the smaller age–gender sub-groups to about 5% of the movie scene (Fig. 4). This is crucial for our application. We rarely expect to need to magnify the image by greater than a factor of 4 (showing $\frac{1}{16}$ th (6.25%) of the movie scene). Higher magnification might cause too much loss of context and may appear blurred. The distribution of COIs (Fig. 6) illustrates that magnification centered on the COI would provide more information than magnification simply centered on the center of the movie scene. If magnification were arbitrarily applied around the center of the movie scene, important information may be out of view in certain scenes in which the COI was not near the center of the screen. For example, if a COI was at the middle of a face and that COI fell close to the edge of the screen once magnified, half of the face would not be visible. To assess the impact of any particular level of magnification (e.g. $4\times$), one approach is to consider how often the COI would fall outside of the screen for magnification that is slightly higher (e.g. $5\times$) than the magnification of interest. Thus, for a magnification of interest of $4\times$, we found the number of the COIs that were outside the screen for a magnification of $5\times$ (for which only $\frac{1}{25}$ th, or 4% of the screen area would be visible). Seventy-three per cent of COIs were outside the central 4% of the movie scene area ($5\times$ magnification area). Similarly, for a magnification of interest of $3\times$, 50% were outside the central 6.25% of

the movie scene area ($4\times$ magnification area). Thus magnifications of $3\times$ and $4\times$ around the center would be unsatisfactory and would severely limit the utility of that simple approach. In our experience, $4\times$ is the maximum practical magnification.

Also, we found that there are some significant differences in the observation behaviors between gender and age groups. The current analysis only found that the older and male subjects' COIs were more tightly grouped than the younger and female subjects, respectively (Figs. 4 and 5). Most people with low vision are over 65 years of age (and many are over 75), so our “older” subjects may not be representative of most (older) people with low vision who might use magnification when watching TV. We still need to determine if the COI locations varied with gender and age. The conditions or scenes that resulted in a large BVCEA (spread of gaze points) might be just as interesting as the condition of a small BVCEA. The BVCEA could be used to determine an advised magnification. For example, a frame with a small BVCEA can be effectively magnified (without losing important information) more than a frame with a large BVCEA.

We need to emphasize that we are not considering using the raw data (analyzed here) directly to control the video presentation. We are planning to heavily process it and adjust for various effects. For example, due to latencies in the saccadic system, after a scene cut, the eye will remain fixated in the pre-cut COI for 200–500 ms until a saccade is made to the new

COI. If these locations are different, this delay would introduce a brief pulse of irrelevant material on the screen. In pilot trials of controlling the zoom-and-roam device using the raw data, we observed extremely jittery presentations. Most of the apparent jitter was probably due to small frame-to-frame variations in the computed position of the COI. We were able to eliminate most of that jitter by keeping the center of magnification at the forward and backward average of the COI. The COI shifted only when this average changed by $\frac{1}{8}$ th the screen width. This heuristic algorithm resulted in a much smoother video, but other processing algorithms will be investigated in the future.

The area within which most subjects looked was more variable for some movies. This is most likely due to the content of the watched movie (fast motion, multiple relevant objects, etc.). Excerpts from clips illustrating different levels of movement and the varied responses can be found on our website at <http://www.eri.harvard.edu/faculty/peli/lab/videos/videos.htm>.

In addition to our interest in the application of this technique to our movie (or television) magnification device, we see this work as a beginning of an interesting examination of the nature and characteristics of the motion scanpath of dynamic environments—the movie environment being one that is simpler to study—perhaps followed by the dynamic real world of a mobile observer.

6. Summary

Magnification around the most important point of the scene on a display (center of interest—COI) might be an effective aid for people with visual impairments that cause resolution loss (low vision). This requires that a COI exist for most video frames. Operationally, we defined the COI by recording the eye movements of normally sighted subjects as they watched movies. We address three specific questions relevant to our proposed low-vision aid for viewing television. (1) To what extent do people look at the same place when watching a movie? (2) Does that vary with age and gender? (3) Does the position of the COI differ from the center of the screen?

We analyzed the eye movements (at 60 Hz) of 20 normally sighted subjects watching 37.5 min of six movie clips of various types (sports, comedy, documentary, news, game show and drama). The subjects were divided into four age–gender groups (younger and older; males and females). The 120 subject data files were pre-processed to apply calibrations and to remove recording artifacts. The coincidence between the gaze points of multiple subjects was determined by calculating the area of the best-fit bivariate contour ellipse expressed as percentage of the movie screen area. The position of the COI was determined as the mean x and y coordinates of each group.

The 20 subjects did tend to look in the same direction. Over half of the time, the distribution of gaze points of most subjects was contained within an area that was less than about 12% of the movie scene, or for smaller age–gender sub-groups, of about 5% of the movie scene area. This is crucial for our application since we rarely expect to need to magnify the image by greater than a factor of 4 (showing about 6% of the movie scene). Male and older subjects were more likely to look in the same

direction than female and younger subjects, respectively. This inter-group variability in gaze points also varied between the movie clips. The peaks of the COI spatial distributions were approximately in the center of the movie scene, though they varied by as much as $\frac{1}{4}$ the width or height from the center, and the distributions varied between movie clips.

Providing the coordinates of the COI along with each frame may allow magnification to be used more successfully as a low-vision aid for watching movies (and other television programs). The eye movement method presented here is a natural and efficient way of determining these COIs. We envisage that, just as programs are now being provided in “closed captions” and described video formats, movies can be provided with these COIs encoded.

Acknowledgment

This work was supported in part by NIH Grants EY05957 and EY12890. Assistance was provided by Miguel A. García-Pérez, Shabtai Lerner and Gang Luo.

References

- [1] G. Legge, G. Rubin, D. Pelli, M. Schleske, Psychophysics of reading II, Low vision, *Vision Res.* 25 (1985) 253–266.
- [2] G. Rubin, K. Turano, Reading without saccadic eye movements, *Vision Res.* 32 (1992) 895–902.
- [3] E. Fine, E. Peli, Computer display of dynamic text, in: *Vision 96: International Conference on Low Vision, Madrid, Spain*, vol. Book 1, ONCE, Madrid, Spain, 1996, pp. 259–267.
- [4] E. Fine, E. Peli, Benefits of rapid serial visual presentation (RSVP) over scrolled text vary with letter size, *Optom. Vision Sci.* 75 (1998) 191–196.
- [5] S. Kazuo, J. Shimizu, Image scaling at rational ratios for high-resolution LCD monitors, in: J. Morreale (Ed.), *Society for Information Display Digest of Technical Papers*, vol. 31, Society for Information Display, San Jose, CA, 2000, pp. 50–53.
- [6] E. Peli, Vision multiplexing—an engineering approach to vision rehabilitation device development, *Optom. Vision Sci.* 78 (2001) 304–315.
- [7] R.B. Goldstein, H. Apfelbaum, G. Luo, E. Peli, Dynamic magnification of video for people with visual impairment, in: *2003 SID International Symposium, Digest of Technical Papers*, vol. 34, Society for Information Display, Baltimore, MD, 2003, pp. 1152–1155.
- [8] A.L. Yarbus, *Eye Movements and Vision*, Plenum, New York, 1967.
- [9] D. Noton, L. Stark, Scanpaths in eye movements during pattern perception, *Science* 171 (1971) 308–311.
- [10] L. Stelmach, W.J. Tam, P. Hearty, Static and dynamic spatial resolution in image coding: an investigation of eye movements, in: *Human Vision, Visual Processing and Digital Display II*, vol. 1453, SPIE, Bellingham, WA, 1991, pp. 147–152.
- [11] L. Stelmach, W.J. Tam, Processing image sequences based on eye movements, in: B.E. Rogowitz, J.P. Allebach, (Eds.), *Human Vision, Visual Processing and Digital Display V*, vol. 2179, Proceedings of SPIE, SPIE, Bellingham, WA, 1994, pp. 90–98.
- [12] V. Tosi, L. Mecacci, E. Pasquali, Scanning eye movements made when viewing film: preliminary observations, *Int. J. Neurosci.* 92 (1997) 47–52.
- [13] W.S. Geisler, H.L. Webb, A foveated imaging system to reduce transmission bandwidth of video images from remote camera systems, 19990025482, AD-A358811, AFRL-SR-BL-TR-98-0858, NASA, Austin, TX, 1998.
- [14] U. Rauschenbach, H. Schumann, Demand-driven image transmission with levels of detail and regions of interest, *Comput. Graphics* 23 (1999) 857–866.

- [15] E. Dmytryk, *On Film Editing*, Focal Press, Boston, 1984.
- [16] L. Itti, C. Koch, Computational modelling of visual attention, *Nat. Rev. Neurosci.* 2 (2001) 194–203.
- [17] E.M. Reingold, L.C. Loschky, Saliency of peripheral targets in gaze-contingent multiresolutional displays, *Behav. Res. Methods Instrum. Comput.* 34 (2002) 491–499.
- [18] E. Moores, L. Laiti, L. Chelazzi, Associative knowledge controls deployment of visual selective attention, *Nat. Neurosci.* 6 (2003) 182–189.
- [19] A. Oliva, A. Torralba, M.S. Castelhana, J.M. Henderson, Top down control of visual attention in object detection, in: *IEEE Proceedings of the International Conference on Image Processing*, vol. I, 2003, pp. 253–256.
- [20] M.B. Neider, G.J. Zelinsky, Scene context guides eye movements during visual search, *Vision Res.* 46 (2006) 614–621.
- [21] A.J. Hornof, T. Halverson, Cleaning up systematic error in eye-tracking data by using required fixation locations, *Behav. Res. Methods Instrum. Comput.* 34 (2002) 592–604.
- [22] D.M. Stampe, Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems, *Behav. Res. Methods Instrum. Comput.* 25 (1993) 137–142.
- [23] J. Goldberg, J. Sclavy, Eye-gaze-contingent control of the computer interface: methodology and example for zoom detection, *Behav. Res. Methods Instrum. Comput.* 27 (1995) 338–350.
- [24] L. Stark, Y. Choi, Experimental metaphysics: the scanpath as an epistemological mechanism, in: W.H. Zangemeister, H.S. Stiehl, C. Freska (Eds.), *Visual Attention and Cognition*, Elsevier, Amsterdam, 1996, pp. 3–69.
- [25] C.M. Privitera, L.W. Stark, Algorithms for defining visual regions-of-interest: comparison with eye fixations, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (2000) 970–982.
- [26] D. Salvucci, J. Goldberg, Identifying fixations and saccades in eye-tracking protocols, in: *Eye Tracking Research & Applications Symposium*, ACM Press, New York, 2000, pp. 71–78.
- [27] W. Osberger, A.M. Rohaly, Automatic detection of regions of interest in complex video sequences, in: *Proceedings of SPIE, Human Vision and Electronic Imaging VI*, vol. 4299, 2001, pp. 361–372.
- [28] M.D. Crossland, G.S. Rubin, The use of an infrared eyetracker to measure fixation stability, *Optom. Vision Sci.* 79 (2002) 735–739.
- [29] G.T. Timberlake, M.K. Sharma, S.A. Grose, D.V. Gobert, J.M. Gauch, J.H. Maino, Retinal location of the preferred retinal locus (PRL) relative to the fovea in scanning laser ophthalmoscope (SLO) images, *Optom. Vision Sci.* 82 (2005) 177–185.
- [30] T.D. Wickens, *Multivariate Contingency Tables Analysis for the Social Sciences*, LEA, Hillsdale, 1989.
- [31] E. Peli, E.M. Fine, A.T. Labianca, Evaluating visual information provided by audio description, *J. Visual Impairment Blindness* 90 (1996) 378–385.

Robert B. Goldstein obtained his BS from Brooklyn College of the City University of New York in 1969 and his Ph.D. in Nuclear Physics from MIT in 1975. Dr. Goldstein is currently employed as a Senior Computer Analyst at Schepens Eye Research Institute and is an Associate in Ophthalmology at Harvard Medical School. His major interest is the application of Computer Science and technology to scientific research. Dr. Goldstein is most interested in areas of vision research, bioengineering, software development, image processing, clinical trials, software development and database systems.

Russell Woods completed optometry training in 1981 at University of New South Wales, Sydney, and obtained a Ph.D. in vision science at The City University, London in 1992. Currently, Dr. Woods is an Investigator at the Schepens Eye Research Institute and an Instructor in Ophthalmology at Harvard Medical School. His interests include the impact of, and rehabilitation of visual impairment, human spatial vision, visual optics and ocular aberrations.

Eli Peli received BSEE and MSEE in 1976 and 1978, respectively, from Technion-Israel Institute of Technology, Israel, a Doctor of Optometry Degree in 1983 from the New England College of Optometry, Boston, and an honorary DSc from SUNY in 2006. Dr. Peli is the Moakley Scholar in Aging Eye Research at The Schepens Eye Research Institute, and Professor of Ophthalmology at Harvard Medical School. Since 1983 he has been caring for visually impaired patients as the director of the Vision Rehabilitation Service at the New England Medical Center Hospitals in Boston. Dr. Peli's principal research interests are image processing in relation to visual function and clinical psychophysics in low-vision rehabilitation, image understanding and evaluation of display–vision interaction. He also maintains an interest in oculomotor control and binocular vision. He is a Fellow of the American Academy of Optometry, the optical Society of America, and the Society for Information Display.