

Eye movements of large populations: II. Deriving regions of interest, coverage, and similarity using fixation maps

DAVID S. WOODING
University of Derby, Derby, England

The analysis of eye movement traces (i.e., the patterns of fixations in a search) is more complex than that of such parameters as mean fixation duration, and as a result, previous attempts have focused on a qualitative appraisal of the form of an eye movement trace. In this paper, the concept of the *fixation map* is introduced. Its application to the quantification of similarity of traces and the degree of coverage by fixations of a visual stimulus is discussed. The ability of fixation maps to aid in the understanding and communication of large numbers of eye movement traces is examined.

Over the winter of 2000/2001, the Applied Vision Research Unit at the University of Derby conducted the world's largest eye movement experiment. An automated eye tracker was left running in a room of the National Gallery, London, as part of the millennium exhibition "Telling Time." During the 3 months of the exhibition, over 5,000 participants had their eye movements successfully recorded while they viewed digitized images of paintings from the National Gallery collection. The quantity of data (well in excess of 1 million fixations) was unprecedented and posed many challenges in terms of both academic analysis and communication of the results to the public. Aside from the traditional measures, such as mean fixation duration, number of fixations, and so on, it became clear that novel methods of manipulation, analysis, and representation of the large amount of data would be required.

Before free-viewing each of three images, participants had been presented with a series of questions that put them into different categories on the basis of, for example, their sex, age, and experience of art. *Where* individuals and groups of participants had looked in an image would clearly be an important part of the analysis. It was felt that the concept of *similarity* between the eye movement patterns of different individuals and groups should be ex-

plored, along with the degree to which their fixations had *covered* the image.

Lastly, it was clear that a means of visualization and communication of these unwieldy data sets was necessary, to satisfy both public and academic interest in the work.

These issues are, to a large extent, addressed by the use of *fixation map* analysis, which represents a versatile analytical tool that retains objectivity and minimizes the number of assumptions made in the process. The methodology of the technique and its applications are described in the following sections.

FIXATION MAP ANALYSIS

At its simplest, the fixation map is a two-dimensional (2-D) record of the locations of all the fixations being analyzed, whether in an individual eye movement trace or in a selection of traces. However, the maps developed for the analysis in this paper are three dimensional (3-D), with the third dimension being the quantity of property d obtained from that fixation. The precise definition of d is deliberately left vague in this paper. The definition is not essential to the concept of the fixation map; as a parameter, d can be defined to suit the analysis required. It can be thought of as a low-level measure of the amount or quality of information obtained from a single fixation or, at a higher level, the degree of discrimination, detection, or perception achieved.

Since the property d falls off with distance from the center of the fixation, the (3-D) form of an individual fixation is approximated to a (3-D) Gaussian or, less elegantly, a *lump* (Figure 1A). In this way, it is an extension of the concept of the soft-shelled *visual lobe* (Overington, 1976), *functional visual field* (Nelson & Loftus, 1980), *useful field of view* (Sekuler & Ball, 1986), or *perceptual span* (Rayner, 1998).

This Gaussian representation should not be taken to mean that little or no information is available to the visual system in more peripheral regions. The periphery contributes

The National Gallery "Telling Time" exhibit was conceived and constructed by David Wooding, Mark Muggleston, Kevin Purdy, and Alastair Gale of the Applied Vision Research Unit, University of Derby. I am indebted to the members of the team for their helpful advice and encouragement in the preparation of this manuscript. The exhibit was conducted in cooperation with the National Gallery, London (exhibition curator, Alexander Sturgis) and with the support of Derby University Enterprises Limited (DUEL). The analysis of data was funded in part by the British Academy (Grant 31757). The author acknowledges with gratitude the efforts and insightful comments of the reviewers of this journal. Correspondence concerning this article should be addressed to D. S. Wooding, Applied Vision Research Unit, Kingsway House (West), University of Derby, Derby DE22 3HL, England (e-mail: d.wooding@derby.ac.uk).

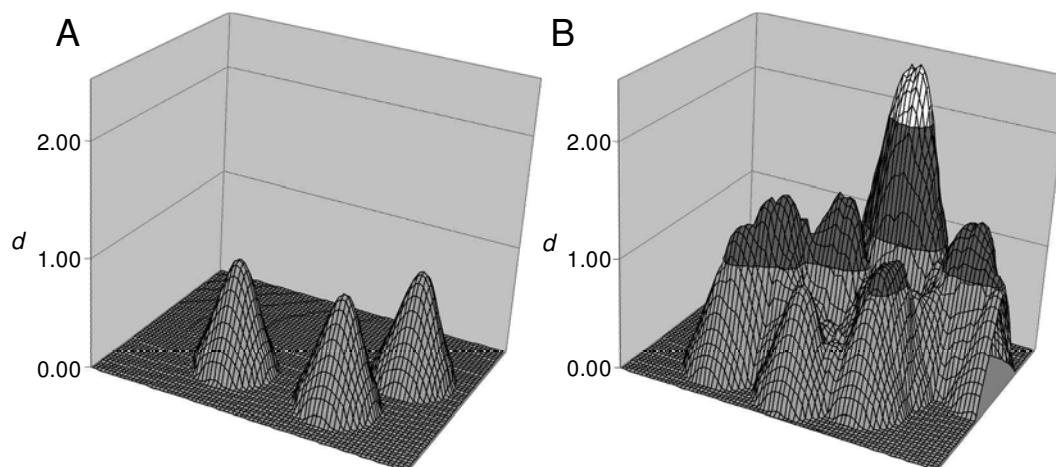


Figure 1. Building a fixation map: (A) three individual fixations on a fixation map and (B) the appearance of the fixation map after 17 fixations, some having overlapped.

essential information, such as that utilized in the planning of subsequent fixations. A typical Gaussian width might be the size of the fovea projected onto the stimulus. The *true* width depends on the area over which a fixation can be said to exist, and this is determined by the properties of the stimulus that the visual system utilizes in its examination. As with the visual lobe or useful field of view, the width of the Gaussian is therefore stimulus and task dependent. A simple detection task may well be modeled best by a large width, whereas a task involving discrimination of small detail would require a smaller width. Less well defined tasks, such as the free examination of artistic works, will require the calculation of an appropriate lobe size from analysis of the data, and this will be explored in future publications.

Creation

Building a fixation map. The creation of a fixation map begins with a blank map of the stimulus presented. For digital images, it is reasonable to use a blank map of the same dimensions in pixels as the original stimulus, with one value of d per pixel. This arrangement will be assumed in the following description, with the individual locations in the fixation map being termed *map pixels*. The scale of the map is, though, essentially arbitrary and under the control of the experimenter.

For each fixation location, an identical 3-D Gaussian of unit height is dropped onto the map (Figure 1A). If this overlaps with an existing fixation, the height at any map pixel is added to the existing height at that point. With increasing numbers of fixations, a landscape is built up (Figure 1B), indicating the variation in d across the image.

With large numbers of fixations, the fixation may be approximated to a cylinder, rather than to the 3-D Gaussian. As more cylinders overlap and superimpose, the fixation map effectively “softens.”

The fixation map itself is a 2-D array of the property d . The depictions of the fixation map used in Figure 1B

might be more descriptively termed a *landscape* or *terrain*, since the value at any point on the surface shown indicates the *height* or *amount* of property d at that point. Since the terrain is the most easily accessible visualization of the fixation map, it will be described as a fixation map for the purposes of this paper.

It should be noted that in analyzing fixation data in this way, the only assumptions that have been made have been (1) to represent the property d associated with the fixation as a 3-D Gaussian and (2) to determine a width for the Gaussian (i.e., over what area it is said to act). These assumptions are well defined and repeatable and can be quoted when any results are reported.

The choice of Gaussian width will, to a degree, be reflected in the final map: A small width will result in a more “spikey” surface, whereas a larger width will produce a more gently undulating one. The final appearance of the map is not in itself an indicator of the “correctness” of the Gaussian width, since discrimination of a very coarse pattern might be possible over a large area, resulting in an appropriately large width. Having said that, it may be possible to obtain a value of the Gaussian width from the mean separation of fixations or by maximizing the similarity in fixation maps between two halves of a large population, and this will be the subject of a future paper.

Since the author’s first publication on the subject (Wooding, 2002), it has been noted that Pomplun, Ritter, and Velichkovsky (1996) adopted an initially similar treatment of fixation data in order to illustrate and investigate conflicting views of ambiguous figures. In their methodology, they constructed what they termed *attentional landscapes* along similar lines to those described here. They used 3-D Gaussians of widths approximating to the projected width of the fovea (a standard deviation of 2° of visual angle from Hood & Finkelstein, 1986). In an attempt to represent the attention deployed by their experimental participants, they adjusted the height of their 3-D Gaussian in proportion to the duration of the fixation (although they noted

that “not every fixation is ‘filled with attention’”). Their landscape therefore became a type of “dwell map,” representing not only the areas fixated, but also the amount of time spent there.

Although the attentional landscape makes for an interesting analysis, the present paper is not concerned with fixation duration, concentrating instead on whether regions of the stimulus received a fixation, without making any assumptions about the nature of that fixation. This assumes that all fixations are “equal”—that is, that a constant amount of information is acquired in each fixation over an identical area. In reality, each fixation will be different—for example, (1) each fixation may fall on a different part of the image, resulting in the processing of different visual information, (2) fixations of different durations *may* indicate the acquisition of differing information, (3) the processing required of or undertaken by the visual system may well vary during examination of the stimulus, (4) later fixations may have a different character owing to information already obtained about that region of the stimulus or the stimulus as a whole, or (5) some fixations, such as those preceding corrective saccades, may contribute less in terms of information about the stimulus.

It should be noted that adaptation of the Gaussian to take into account any or all of the issues above will necessitate making certain assumptions. It might be possible to include in the analysis only fixations with properties within an “acceptable” range (e.g., excluding short duration fixations, such as those in Point 5 above). Since the processing that takes place within a fixation is only poorly understood, even equating fixation duration with the amount of information obtained may be an inappropriate step. Needless to say, constantly to vary the properties of the Gaussian would also result in a considerably more complex analysis. In this paper, to simplify the analysis and to limit the number of assumptions, each fixation makes the same contribution to the analysis and is, therefore, represented by a Gaussian of unit height.

All the fixation maps in this paper have been prepared on the basis of real data from the National Gallery exhibit, and a Gaussian with a standard deviation of 2° (i.e., approximating the size of the fovea) has been used throughout. Details of the images are given below in the Appendix. The generation of the maps, together with all visualizations and calculations, was undertaken with software written by the author, using Microsoft Visual Basic v6.0.

Normalization. The map can represent the fixations of one trace by one individual or, just as easily, those of many traces by one or many individuals. As has been described, the map is a representation of the cumulative fixations used in its construction.

If the map is to be compared with another map to determine differences in location of fixations, it is desirable to normalize the map so that the maximum value of d (the highest peak) is given a value of 1. If the comparison is made in order to determine whether one map contains more clusters of fixations and another a more even distribution of fixations, the absolute totals of d may well be

important, and it would be desirable to leave the map unnormalized.

Applications

Graphical descriptions. The principal goal in undertaking fixation map analysis might be to answer the question, Where in the image did people tend to look? This is particularly the case when results are communicated to a nonacademic audience. Fixation maps are particularly well suited to the communication of eye movement data from a large number of participants, where the alternative (i.e., the representation of large numbers of individual eye movement traces) would soon lead to a complex and confusing result. Fixation maps afford a consistent representation regardless of the number of observers.

The simplest and most direct approach to a graphical description is to change the image in some way so as to reflect the variation in height of the fixation map. The chosen method was to change the luminance of each pixel in the original image to reflect the value of d from the fixation map at the point. In other words, those parts of the image that received more fixations appeared brighter than those areas that received fewer fixations. Coincidentally, this treatment of images was also that preferred by Pomplun et al. (1996), who used the eye movements of experimental participants to alter the appearance of the presented stimuli before presenting these new versions in a new experiment. The disadvantage of this treatment of the image is that the final luminance of a pixel in the modified image depends both on the value of d and on the initial luminance at that point. There is clearly a tradeoff between the faithful representation of data values and a clear representation of the original image. It was felt that this method provided the most accessible integration of fixation map and original image.

Various other methods were considered. The degree of blurring of parts of the image can be altered depending on the value of d at that point, with those areas receiving the highest density of fixations appearing more clear. A similar method was used by Shiori and Ikeda (1989), although this was not derived from eye movement data. In the context of the National Gallery experiment, it was felt that this method was open to misinterpretation, particularly with fixation maps created from multiple traces, since the graphical representation might be construed as a representation of visual function. A false-color spectrum could also be used to represent the range of values of d , although the representation of the original image would suffer as a result, making interpretation difficult.

Results of the luminance method are shown for two images: Figure 2 for Image N0931 and Figure 3 for Image N1313. In each block of four images, the top left panel is the original image, the top right is the fixation map for around 130 individual traces, the bottom left panel is the contour plot corresponding to the fixation map, and the bottom right panel is the original image modified so that areas receiving larger numbers of fixations appear brighter. Figure 2D illustrates how the distribution of fixations with Image N0931 mainly occurred in a very small informative

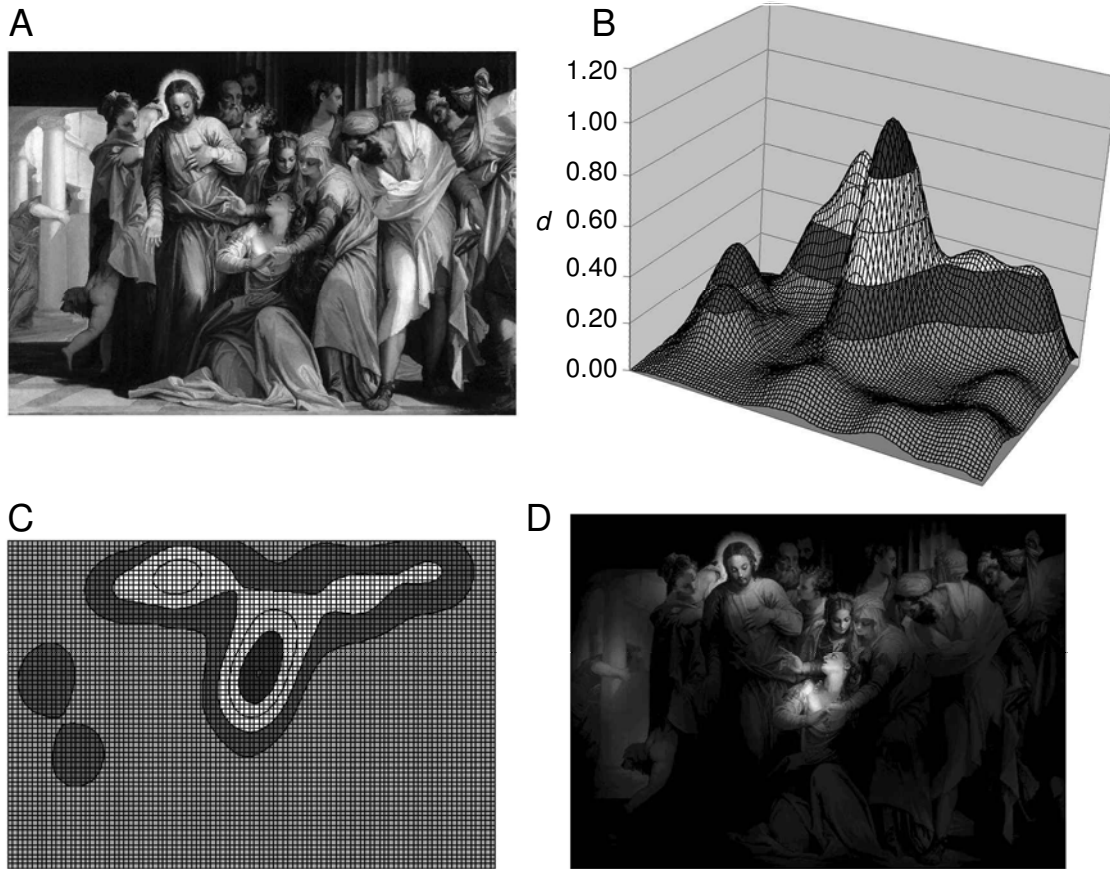


Figure 2. (A) The original image N0931 (copyright the National Gallery Co. Ltd; see the Appendix for details). (B) Map of fixations of 131 traces. (C) The contour plot corresponding to the map in panel B. (D) The image redrawn with areas receiving higher numbers of fixations appearing brighter.

part of the image, leaving the remainder of the image in darkness. Figure 3D illustrates that Image N1313 provided a greater number of regions of interest to the observer and that these were well distributed around the image.

In part, this paper contributes to the growing body of work on information visualization. Thomas et al. (2001) described the advantages of such descriptions of data: Researchers use their own visual systems to detect emerging patterns and trends in graphical representations of data sets, determining where best to employ further statistical analysis. Unlike the mapping of complex text data by Thomas et al., the spatial relationship between different fixations is not a construction of the analysis used; rather, it is their real spatial separation relative to the stimulus used. This makes the analysis of fixation data by fixation maps both powerful and more straightforward. Visualization is only one *application* of fixation map analysis and is the least quantitative of the applications described here. It does, however, provide a useful and intuitive means of describing and communicating the underlying patterns of otherwise overwhelming data sets.

Regions of interest. The graphical representations of the data described above are an immediate and powerful

demonstration of a key feature of eye movement data: that the fixations are not spread evenly or randomly over the stimulus but cluster into regions of interest according to the features of the stimulus.

Fixation map analysis affords an opportunity to define objectively the principal regions of interest of observers when they view an image. This can be achieved on the basis of all the areas in which the value of d is greater than a critical value, d_{crit} . To visualize this, if we return to the *landscape* metaphor, the landscape is flooded to the level of d_{crit} , leaving only the highest peaks (the regions of interest) as islands. Alternatively, the “top five” areas (for example) can be determined by gradually increasing d_{crit} from zero until the required number of areas remains on the map. Figure 4 illustrates the initial map (top panel) and the same map “flooded” with two values of d_{crit} . Figure 4D is a 2-D contour map of the same data as those for the terrain in Figure 4C, to illustrate the regions of interest.

The regions of interest in an image are usually defined a priori, with the experimenter using his/her own understanding of the stimulus or of the regions that are of interest to the investigation at hand. It is then determined whether these areas received fixation. Avoiding the subjective bias

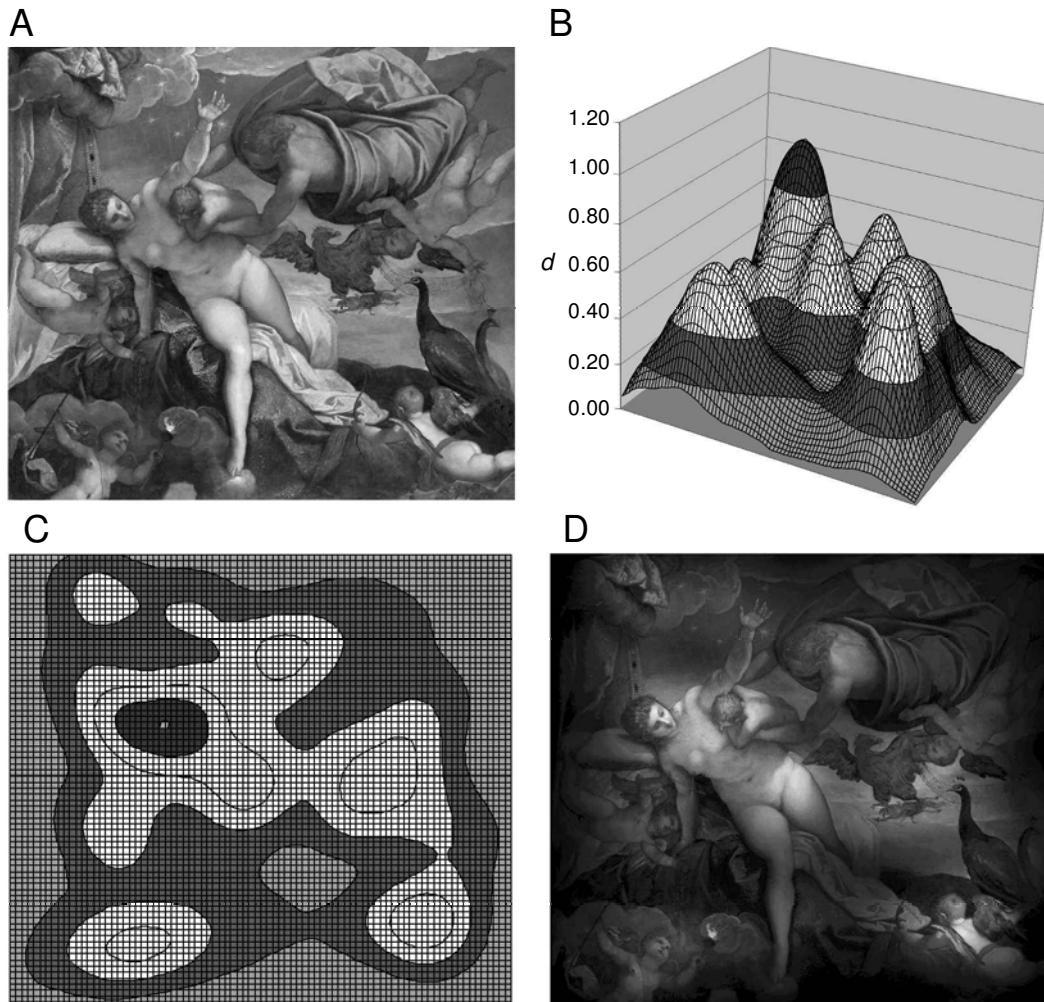


Figure 3. (A) The original image N1313 (copyright the National Gallery Co. Ltd; see the Appendix). (B) Map of fixations of 130 traces. (C) Contour plot corresponding to the map in panel B. (D) The image redrawn with areas receiving higher numbers of fixations appearing brighter.

of this approach, others (e.g., Privitera & Stark, 1998) have developed algorithms with which to process the stimuli and suggest regions according to the presence of symmetry, edges, and so on. Another approach is to use the eye movement fixations themselves as the locations of regions of interest (e.g., Privitera & Stark, 2000). Although this may be useful in the examination of an individual trace, the analysis becomes extremely complex if multiple traces are being examined with a view to establishing the shared regions of interest of a population.

Coverage. A measure of the amount of the original stimulus covered by the fixations included in the analysis is easily obtained from the fixation map. The analysis requires that a critical threshold value of d , d_{crit} , is defined, as in the previous section. Clearly, the value of d_{crit} is in some way implied when setting a width for the 3-D Gaussian that represents a fixation. As d falls off from the center of the fixation, it will reach a critical level at which it is not

high enough for the task being undertaken (such as detection).

To determine coverage, for each map pixel, if d is less than d_{crit} , d is set to zero, and if it is greater than d_{crit} , d is set to 1. The sum of values of d across the map, divided by the area of the map, therefore gives the proportion of the map effectively covered by fixations. This is equivalent to the fractional area occupied by the “islands” in Figure 4D (a coverage of 17%).

Figure 5 offers some examples of the calculation of coverage at different values of d_{crit} . The value of coverage will of course depend on the value of d_{crit} at which the calculation is made, and so it may be best used as a comparative measure between two fixation maps at a set value of d_{crit} . Both Figures 5B and 5D illustrate that at a d_{crit} of 50%, only a small fraction of the image is covered (6.4% and 2.8%, respectively). The area in Figure 5D is half that of Figure 5B, illustrating how the participants have con-

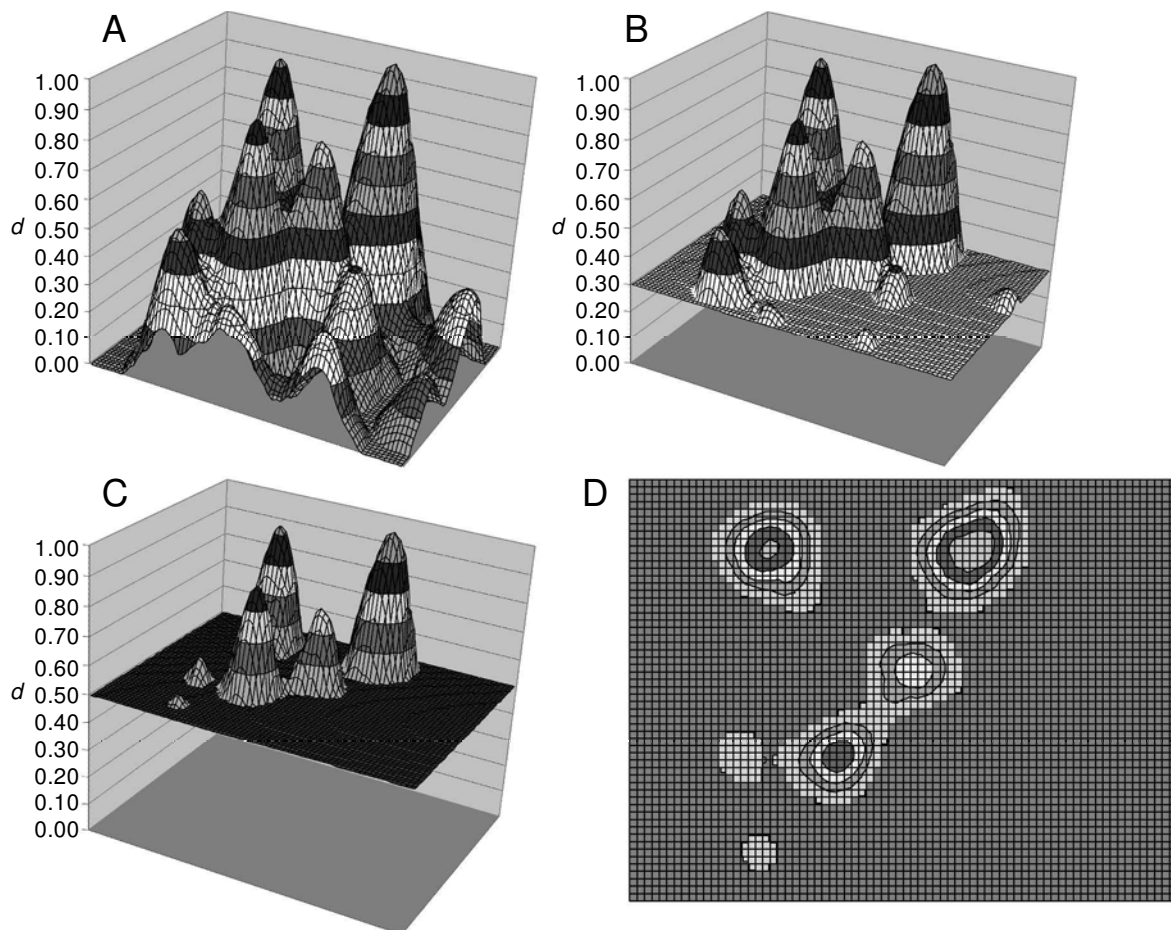


Figure 4. The original fixation map (A) is “flooded” to different values of d : 0.3 (B) and 0.5 (C). The contour map (D) corresponding to panel C is also shown.

centrated their fixations around the smaller main region of interest in that image (i.e., the train in N0538). As d_{crit} is lowered to 10%, the value of coverage gets greater at 44.9% but does not approach the value of 88.1% for N4240, where the uniform distribution of detail and objects in the image have led to a more evenly distributed pattern of coverage.

Similarity. Another application of fixation map analysis is the quantification of the similarity between eye movement traces. Nearly all previous studies in which similarity in eye movement traces has been examined have relied on a subjective decision by the authors as to the degree of similarity or, more usually, the general family of similar traces to which a particular pattern belongs.

The quantification of the similarity between one eye movement trace and another initially appears to be a simple task. It is easy, however, to create a number of artificial traces that do not *appear* similar, although one trace comprises fixations with identical *locations* to those in the other trace. This raises two important questions: (1) How important is the path taken (i.e., the temporal and spatial arrangement of fixations) in assessing similarity of traces? (2) What importance should be given to differences in the

path? For example, if one trace involves two saccades in travel between two locations, how similar is this to a trace in which the journey is made with only one saccade?

The answers to these questions will, of course, depend on the analysis being undertaken. However, for many purposes, it is the *location* of fixations, not the *order*, that is important. This is fortunate, since the addition of order would markedly increase the complexity of the analysis, assuming it is possible to come by a solution that is sufficiently objective. A subjective interpretation, unavoidably taking into account the path and drawing on an unknown internalized set of weightings in the mind of the individual undertaking the analysis, will lead to a misleading and probably unrepeatable answer.

Attempts at *quantifying* similarity are rare, presumably because the complexity of the undertaking means that solutions are not easily yielded and the compromises, which are made in deriving a workable method, result in coarse methods or methods limited in applicability to other settings.

A popular approach (Brandt & Stark, 1997; Eigen, Winklerowatitsch, & Dress, 1988; Hacisalihzade, Stark, & Allen, 1992) has been to impose a grid of cells on the

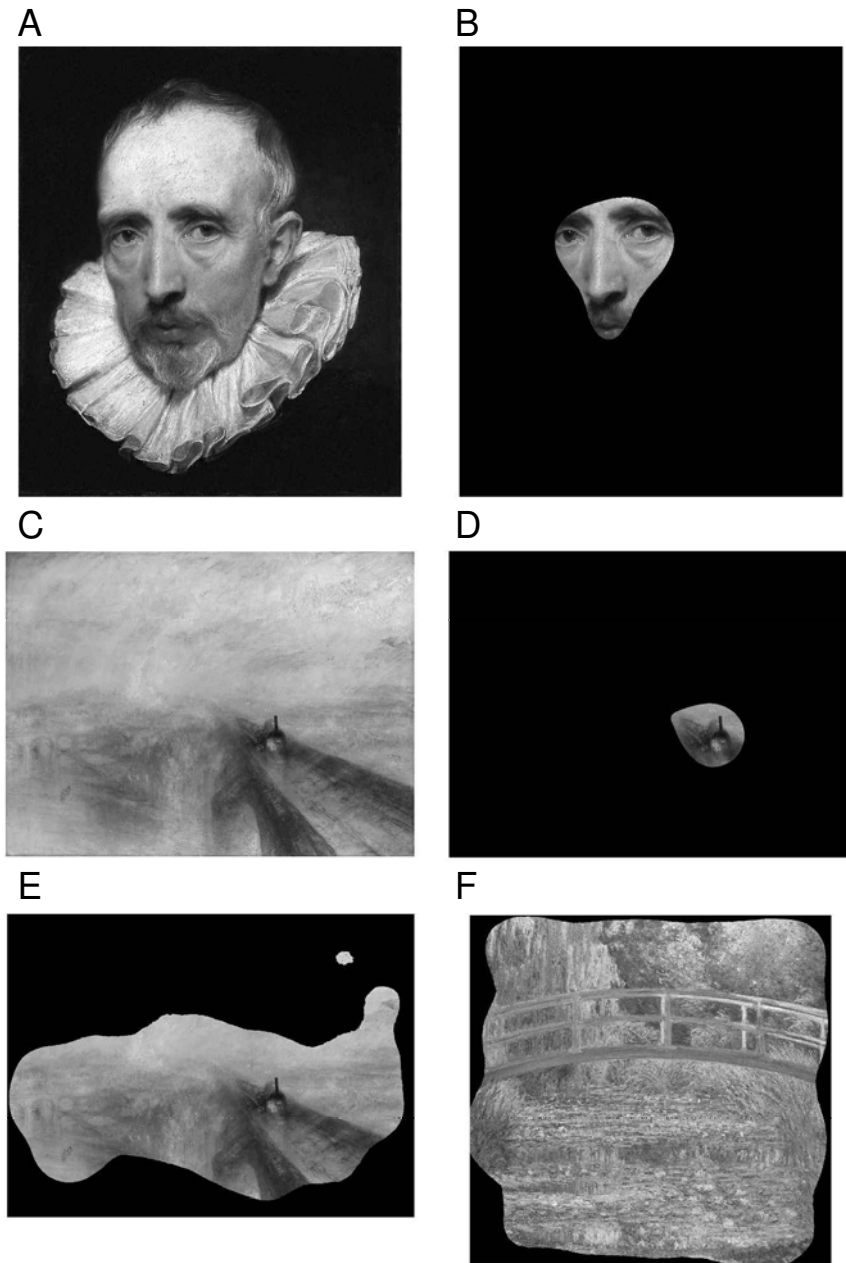


Figure 5. Calculation of coverage: (A) original N0052, (B) $d_{crit} = 50\%$, coverage = 6.4%; (C) original N0538, (D) $d_{crit} = 50\%$, coverage = 2.8%, (E) $d_{crit} = 10\%$, coverage = 44.9%; (F) N4240, $d_{crit} = 10\%$, coverage = 88.1%. (Original images are copyright the National Gallery Co. Ltd; see Appendix for image details.)

spatial distribution of fixations, with each cell assigned a unique letter. An eye movement trace can then be transformed into a letter string according to the sequence of cells in which the fixations landed. Two letter strings can then be compared. For example, a string-editing algorithm can be used to calculate the minimum editing cost of transforming one letter string into another. The cost can then be normalized by dividing by the length of the analyzed strings to give the relative distance and, therefore, a

measure of similarity between the eye movement patterns. This method provides a measure of the similarity in sequence of the locations. Stark and Ellis (1981) derived Markov matrices from such letter strings, demonstrating the existence of a few structured processes, although the size of the higher order Markov matrices was a limiting factor in this analysis.

Mannan, Ruddock, and Wooding (1995) concentrated on fixation location without regard to the order in which

the fixations were made. They used a least-squares index of similarity, I_s , to calculate, for each fixation, the distance between the center of a fixation in one trace and its nearest neighbor in the other trace. Their index compared a single trace with another. In order to give similarity measures for populations of traces (e.g., in the calculation of intra- or intergroup similarity), it was necessary to average the similarities of separate trace pairs. In later work, Mannan, Ruddock, and Wooding (1997) adapted this method to examine sequences of fixations by the similarity of fragments of traces. Privitera and Stark (2000) also examined similarity in location of fixations but simplified the analysis by setting a threshold separation below which points were said to be coincident and above which they were not.

Pomplun et al. (1996) imposed a grid of dimensions n_x and n_y on the fixation pattern, calculating the sum s_n of the durations of fixations falling on each square. By this method, each fixation set generated a distribution vector of the values of s_n that consisted of $n_x \times n_y$ dimensions. They then calculated the angle between the resultant distribution vectors from each set. This measure of similarity was averaged over a range of grid size determined experimentally in order to overcome stimulus-dependent effects.

Fixation map analysis promises to be a powerful tool in this area, since it allows the comparison of fixation maps of individuals and of groups. Tompa (2001) used a map comparison method to compare the locations of image features with those of fixations. The degree of overlap between fixations and image features (both represented as “hard-shelled” objects, rather than the “soft-shelled” fixations in the fixation maps described in this paper) were used as a measure of the correspondence between the two.

Although a reliable method of calculating the similarity of two fixation maps is still being refined, at the simplest level two normalized maps can be subtracted from each other. A third, *dissimilarity* map is thereby created, in which the d value for each map pixel is the magnitude of the *difference* of the d values for the corresponding map pixels in the two original maps. The value of d in the dissimilarity map will range from zero (for identical) to 1 (for completely dissimilar).

In the example given in Figure 6, the 130 participants who had viewed the same image were randomly assigned to two groups of 65. Figure 6A is the fixation map for the individual traces of the participants in the first group, with Figure 6B being the fixation map for the participants in the sec-

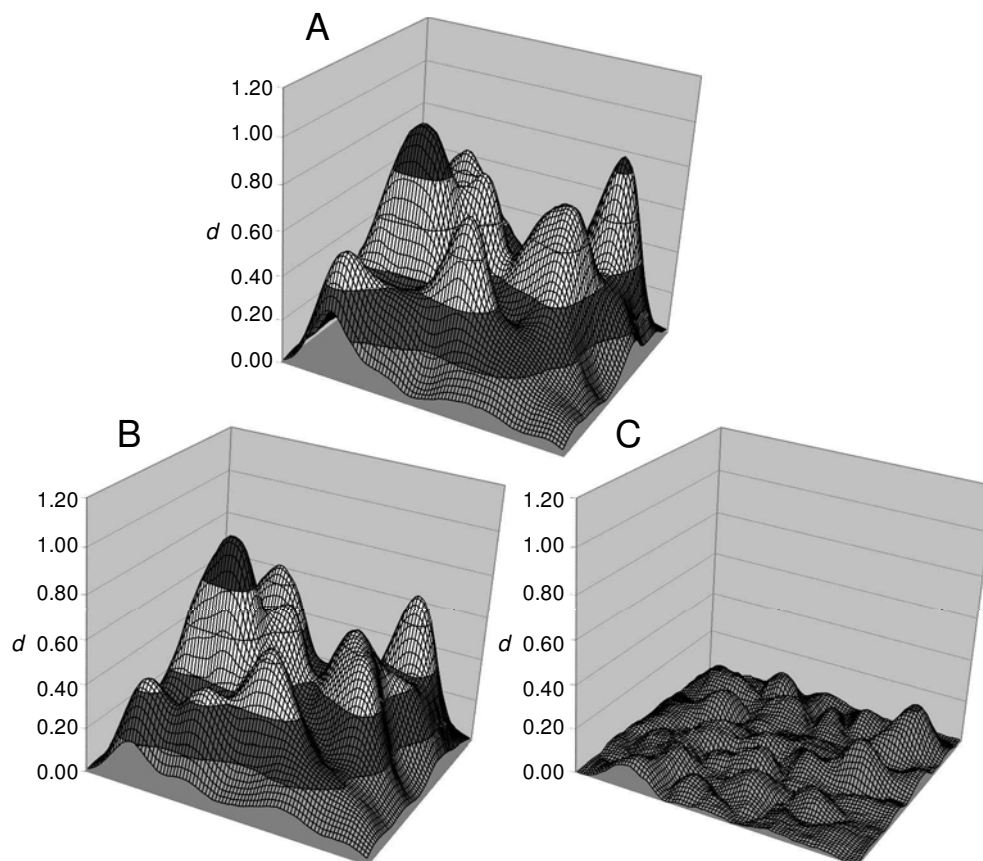


Figure 6. Fixation maps (A and B) for two different groups of 65 participants viewing the same image and (C) the similarity map obtained by subtracting one map from the other.

ond. Figure 6C (the magnitude of the difference between the two maps in Figures 6A and 6B) shows that there is little difference in regions fixated by the two groups and that the mean value of d per map pixel is appropriately low at 0.04.

The normalization of the two initial maps ensures that the analysis is between the two distributions of fixations. This is an analysis made on the basis of eye movement *patterns* and is deliberately insensitive to the relative numbers of fixations (which can be compared by using other methods). Unlike the vector method of Pomplun et al.

(1996) and the *letter string* approach, the *grid size* of the fixation map is not critical, since each fixation contributes to the final value of many map pixels, and indeed it is desirable that the fineness of the grid size should be maximized. The smooth decrease in the value of d from the fixation's center also ensures that the measure is insensitive to the precise location of fixations.

Map sequences. Fixation maps can be produced to represent a particular time slice in the presentation, and multiple slices can then be combined to produce a se-

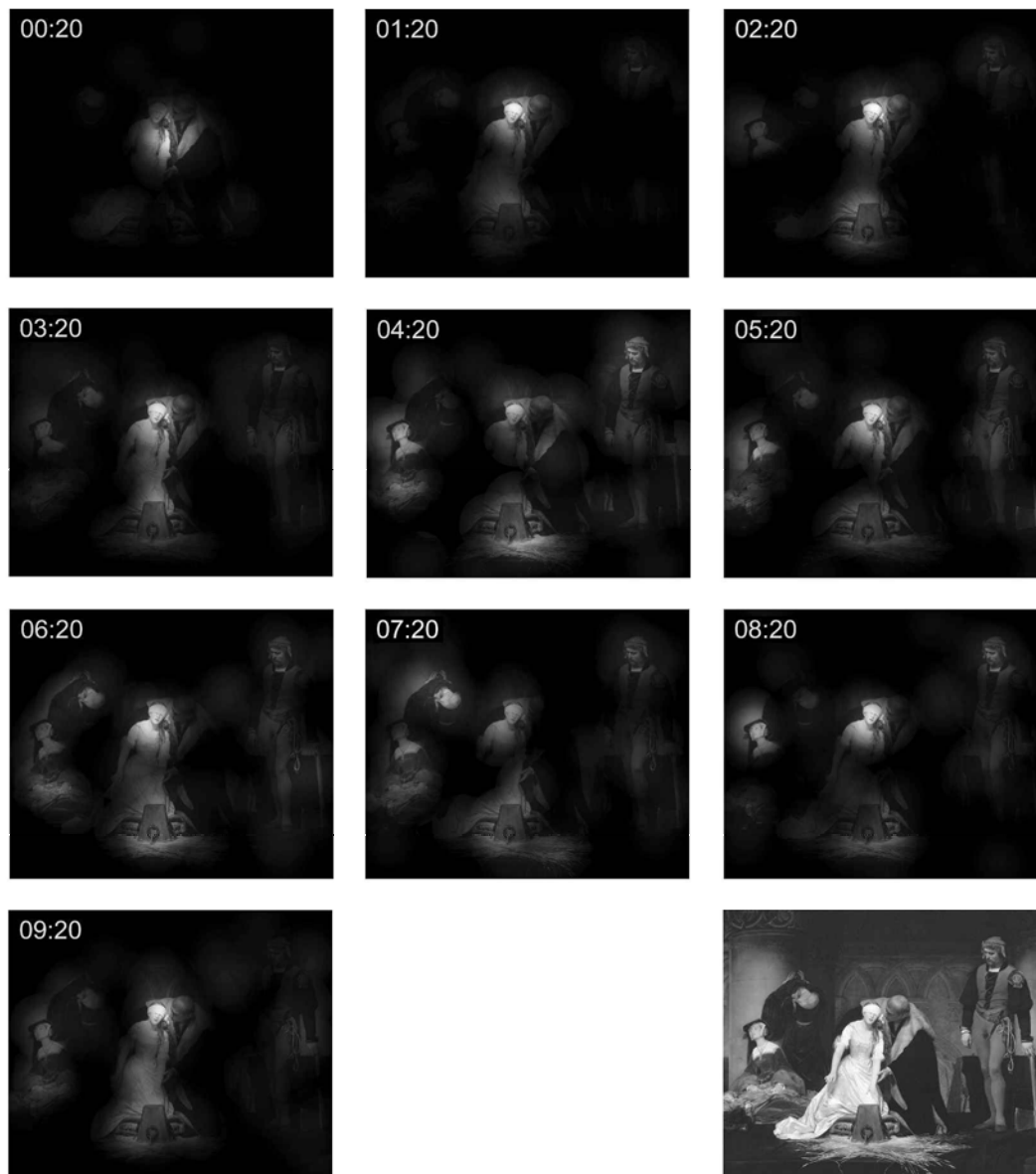


Figure 7. Principal locations of fixations for 139 participants viewing Image N1909 (copyright the National Gallery Co. Ltd; see the Appendix for details) in time slices of 200 msec. Each panel depicts the principal locations of all fixations occurring within the first 200 msec of each second. The digits in the top left-hand corner of each panel indicate the total presentation time that has elapsed (in seconds and hundredths of a second). For example, the second frame represents the fixations in the interval of 1.00–1.20 sec. The presented stimulus is given in the bottom right for reference.

quence of fixation locations for all observers as the stimulus presentation progressed. Figure 7 shows fixation data for 139 participants for the first 10 sec of the presentation of Image N1909, in time slices of 200 msec. For example, the second panel is constructed from the fixations that were occurring in the time interval of 1.00–1.20 sec. The panel is therefore a depiction of where the participants were looking during that part of the presentation.

This method can be used to examine the way in which regions of interest changed over the duration of the presentation. In Figure 7, the initial regions of interest are concentrated around the central figure of Lady Jane Grey, with some interest being paid to the block at her feet. It is only after a few seconds that the main regions of interest become more distributed around the stimulus, and there is then little change in distribution for the remainder of the presentation.

Although this is only one example, it is clear that the technique could be used to examine the prioritization of regions of interest in a search and the existence of different phases of a search and may be useful in determining the importance of temporal order to fixations. One of the unknown factors in the analysis of patterns of fixations is the relevance of temporal order, and although there is evidence that it is not important (e.g., Mannan et al., 1997), this technique might offer useful new insights.

CONCLUSIONS AND FURTHER WORK

Fixation map analysis represents an objective method of quantifying aspects of eye movement traces. As has been described above, it can be used to define various parameters of the eye movement trace, including the degree of coverage and regions of interest, and can provide a method of comparing the similarity of traces or of populations of traces. It has also proved itself to be a useful technique in the visualization and communication of large eye movement data sets.

Fixation maps can be utilized in a predictive way, identifying not only the regions of interest, but also their relative attractiveness to the eye movement system. It should be possible to use the fixation map as a probability map to anticipate the likely location of fixations of a viewer. This is well illustrated in Figure 6, in which two populations of participants have very similar fixation maps. The degree to which an individual conforms to this group pattern will be the subject of future research.

Clearly, the concept of a fixation map can be developed. Further work will identify appropriate widths for the Gaussians and will investigate how these vary with stimulus and task. This will, in turn, increase our understanding of the perceptual processes that occur during such tasks. Through selection of particular parameter values (such as Gaussian width), it may be possible to distinguish between subgroups of observers or to test a particular hypothesis.

Finally, in addition to their usefulness in applied situations, fixation maps have the potential to illuminate some

of the more basic questions of research into visual search and eye movements. Figure 7 shows a sequence of fixation maps representing the distribution of fixations at various times in the stimulus presentation. This sequence can be used as a probe to examine differences in eye movement behavior throughout a search, monitoring the phases of different activity and the development of different strategies.

Once we are able to say *where* people look in an image, *why* they looked there becomes an easier question to answer.

REFERENCES

- Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, *9*, 27-38.
- Eigen, M., Winkler, R., & Dress, A. (1988). Statistical geometry in sequence space: A method of quantitative comparative sequence-analysis. *Proceedings of the National Academy of Sciences*, *85*, 5913-5917.
- Hacisalihsade, S. S., Stark, L. W., & Allen, J. S. (1992). Visual-perception and sequences of eye-movement fixations: A stochastic modeling approach. *IEEE Transactions on Systems, Man & Cybernetics*, *22*, 474-481.
- Hood, D. C., & Finkelschtein, M. A. (1986). Sensitivity to light. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception & human performance: Vol. 1. Sensory processes and perception* (pp. 5.1-5.66). New York: Wiley.
- Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision*, *9*, 363-386.
- Mannan, S., Ruddock, K. H., & Wooding, D. S. (1997). Fixation sequences made during visual examination of briefly presented 2-D images. *Spatial Vision*, *11*, 157-178.
- Nelson, W. W., & Loftus, G. R. (1980). The functional visual field during picture viewing. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 391-399.
- Overington, I. (1976). *Vision and acquisition*. London: Pentech Press.
- Pomplun, M., Ritter, H., & Velichkovsky, B. (1996). Disambiguating complex visual information: Towards communication of personal views of a scene. *Perception*, *25*, 931-948.
- Privitera, C. M., & Stark, L. W. (1998). Evaluating image processing algorithms that predict regions of interest. *Pattern Recognition Letters*, *19*, 1037-1043.
- Privitera, C. M., & Stark, L. W. (2000). Algorithms for defining visual regions-of-interest: Comparison with eye fixations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *22*, 970-982.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422.
- Sekuler, R., & Ball, K. (1986). Visual localization: Age and practice. *Journal of the Optical Society of America*, *3*, 864-868.
- Shiori, S., & Ikeda, M. (1989). Useful resolution for picture perception as a function of eccentricity. *Perception*, *18*, 347-361.
- Stark, L. W., & Ellis, S. R. (1981). Scanpaths revisited: Cognitive models direct active looking. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 193-226). Hillsdale, NJ: Erlbaum.
- Thomas, J., Cowley, P., Kuchar, O., Nowell, L., Thomson, J., & Wong, P. C. (2001, July). *Discovering knowledge through visual analysis*. Paper presented at the International Conference on Knowledge Management, Graz.
- Tompa, D. (2001). *Perceptual importance maps for visual attention in images*. Unpublished master's thesis, University of Waterloo, Ontario.
- Wooding, D. S. (2002). Fixation maps: Quantifying eye-movement traces. In A. T. Duchowski, R. Vertegaal, & J. W. Senders (Eds.), *Eye tracking research and applications symposium proceedings* (pp. 31-36). New York: ACM Press.

APPENDIX
Images Used in This Paper

The images of paintings used in this study all appear with the permission of the National Gallery, and remain copyrighted by the National Gallery, London, with annotations copyrighted by IBS, University of Derby. All images were digitized and then scaled to a maximum resolution of $1,024 \times 768$ pixels and 16-bit color and were presented at a maximum size of $22.5^\circ \times 16.2^\circ$ for 20 sec.

N0052: *Portrait of Cornelis van der Geest*. Anthony van Dyck (about 1620). Scaled to 653×768 pixels. 137 participants.

N0538: *Rain, Steam and Speed – The Great Western Railway*. Joseph Mallord William Turner (before 1844). Scaled to $1,024 \times 764$ pixels. 179 participants.

N0931: *Christ Addressing a Kneeling Woman*. Paolo Veronese (about 1546). Scaled to $1,024 \times 734$ pixels. 131 participants.

N1313: *The Origin of the Milky Way*. Jacopo Tintoretto (probably 1575–1580). Scaled to 858×768 pixels. 130 participants.

N1909: *The Execution of Lady Jane Grey*. Paul Delaroche (1833). Scaled to 926×768 pixels. 139 participants.

N4240: *The Water-Lily Pond*. Claude-Oscar Monet (1899). Scaled to 788×768 pixels. 192 participants.

(Manuscript received March 20, 2002;
revision accepted for publication October 15, 2002.)