

Analysis of Accretion and Deletion at Boundaries in Dynamic Scenes

KATHLEEN M. MUTCH, MEMBER, IEEE, AND WILLIAM B. THOMPSON, MEMBER, IEEE

Abstract—In dynamic scenes, the presence of object boundaries is often signaled by the appearance or disappearance of occluded surfaces over time. Such regions of surface accretion or deletion can be found using matching techniques similar to those used to determine optical flow in an image sequence. Regions in one frame that are not adequately matched by any region in previous frames correspond to accretion. Regions that have no matches in subsequent frames correspond to deletion. In either case, an occlusion boundary is present. Furthermore, by associating accretion or deletion regions with a surface on one side of a boundary, it is possible to determine which side of the boundary is being occluded. This association can be based purely on visual motion—the accretion or deletion region moves with the same image velocity as the remaining visible surface to which it is attached.

Index Terms—Dynamic scene analysis, edge detection, occlusion, optical flow, segmentation.

I. INTRODUCTION

LOCATING object boundaries in images is an important but difficult problem. Intensity-based edge detection provides ambiguous or misleading boundary information in many situations, such as textured regions. Motion-based techniques can provide more reliable results in these cases. At object boundaries where occlusion occurs, surface regions will typically appear or disappear over time when motion is present. These regions of changing visibility may be used to indicate both object boundaries and the side of the boundary corresponding to the occluded surface.

At a typical object boundary, one surface will be blocking the view of another more distant surface. In the presence of motion, regions of the more distant surface will often either appear or disappear from view over time. Such regions are called areas of *accretion* or *deletion*, respectively. A similar situation arises in stereo vision, where a region of the more distant surface near an occlusion edge will be visible in one image of the pair but invisible in the other image. Thus, recognition of accretion/deletion regions is a means of locating object boundaries in image sequences. In addition, accretion and deletion regions will belong to the occluded surface, providing sufficient information to determine which of the two surfaces

at a boundary is being occluded. To recover the information available from such regions, it is necessary to determine both how regions of accretion and deletion in the imagery may be identified, and what characteristics of such regions permit identification of the occluded surface.

This paper describes a scheme to locate regions of accretion and deletion, and to identify occluding surfaces at a boundary using these regions. A technique which matches image features in two frames is used to determine feature displacement on the image plane. Areas in the image with a high percentage of features which are unmatchable in a previous or subsequent image are identified as accretion or deletion regions, respectively. These regions indicate the presence of an occlusion boundary. Since the accretion/deletion region belongs to the occluded surface, it will be displaced on the image plane in the same fashion as nearby areas of that surface. The occluded surface is then identified by determining which of the two surfaces adjacent to the accretion/deletion region displays a similar displacement on the image plane. This identification combines information about accretion and deletion with optical flow to produce a description of the occlusion boundary more complete than any existing technique based purely on flow alone.

II. PREVIOUS WORK

Several research efforts in computational vision have utilized motion information to recover object boundaries. The basic idea behind most motion-based approaches is that image plane motion, or optical flow, across the object surface will be constant or slowly varying, and discontinuities in flow will occur only at object edges. Previous approaches either search for discontinuities in the optical flow, or group together regions of similar flow. Nakayama and Loomis [1] propose a local, center-surround operator for detecting object boundaries in flow fields. Clocksin [2] shows that zero-crossings will occur at edge locations in the Laplacian of the magnitude of the optical flow field when an observer translates through an otherwise static environment. Thompson *et al.* [3], [4] demonstrate that the Laplacian is useful as an edge detector in the more general case of unconstrained motion. After obtaining point velocities by template matching, Potter [5] groups all points with the same velocity into single object regions. Similarly, Fennema and Thompson [6] use the spatial and temporal gradients of intensity to obtain point velocities, and then consider all points with similar velocities to be part of the same object. Thompson [7] develops a grouping scheme

Manuscript received April 30, 1984; revised October 22, 1984. Recommended for acceptance by Ruzena Bajcsy. This work was supported in part by the National Science Foundation under Grant MCS-81-05215, the Air Force Office of Scientific Research under Contract F49620-83-0140, and by Zonta International.

K. M. Mutch is with the Department of Computer Science, Arizona State University, Tempe, AZ 85287.

W. B. Thompson is with the Department of Computer Science, University of Minnesota, Minneapolis, MN 55455.

based upon both intensity and velocity information. Regions of both identical intensity and identical velocity are formed, followed by merging of adjacent regions based upon similarities, or at least lack of conflict, in intensity and velocity. With the exception of Clocksin's work [2], these flow-based techniques are unable to provide any indication of the occluded surface at an edge.

Accretion and deletion are fundamental to motion analysis based on differencing [8], [9]. These techniques subtract one image from another and then use the presence of regions of significant difference to infer properties of object boundaries and motion. The approach is most effective when a reasonably homogeneous object is moving relative to a homogeneous background with different luminence. Covering and uncovering of the background leads to significant differences between frames, allowing boundaries to be located. Analysis of these difference regions over time can often be used to associate the difference region with adjacent, nonchanging areas of the image sequence and thereby identify which side of the boundary is being occluded. This scheme is intensity based, and suffers when intensity contrasts occur that are not related to object structure. A textured object which changes location on the image plane, for example, will produce many regions of intensity difference which are not accretion or deletion regions.

Only limited experimentation has been directed at the role of accretion and deletion in human perception. Kaplan [10] showed that patterns of accretion and deletion in fields of moving random dots provide sufficient information for the judgment of relative depth by human subjects. In his stimuli, a single edge separated two regions of random dots, where each region moved coherently. The edge was implicit, being the line along which accretion and/or deletion occurred, and thus was not visible if all of the dots were stationary. Subjects consistently perceived the more distant surface to be the one which was undergoing accretion or deletion at a greater rate. This was true even when the implicit edge moved with a velocity different than the velocity of points on either surface. In these cases of inconsistent edge motion, there was more ambiguity in the perceptions of subjects, although the statistically significant perception was that the surface with a greater rate of accretion or deletion was more distant. This suggests that both edge velocity and accretion/deletion are important factors in the perception of depth at an edge, but that accretion/deletion information may be dominant.

III. DETECTING ACCRETION/DELETION REGIONS

A motion-based scheme for identifying accretion and deletion regions is developed here. To recover motion on the image plane, corresponding structures in each frame of an image pair are located. The result of this is a *disparity vector field*, where each vector represents the change in image plane location of a structure. (Disparity is the discrete representation of optical flow arising from image sequences that are discretely sampled in time.) This correspondence is accomplished by *token matching*. A token is a distinctive region in the image which is identified by some predefined local operator. A set of tokens is obtained for each image in the pair, and an

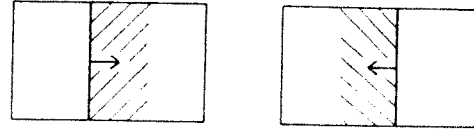


Fig. 1. Location of an accretion/deletion region relative to the boundary indicates the direction of the occluded surface. In both cases shown above, the vertical line represents a boundary and the shaded area represents an accretion or deletion region. The arrow points toward the occluded surface.

organized search is performed to match tokens from the first image to corresponding tokens in the second image using the relaxation labeling technique described in [11]. Possible matches between tokens in the two frames are evaluated based on two criteria: the similarity between properties of the tokens, and a surface smoothness measure that favors matches with disparities similar to neighboring tokens. An important aspect of this particular matching technique is that it can determine that a token in one frame is *unmatchable* if no token in the other frame satisfies the appropriate matching criteria. By basing the analysis on the motion of tokens in the image, many of the intensity contrast problems of a differencing system are circumvented.

Regions of accretion and deletion are identified by analyzing unmatchable tokens in either image. A token may not be matchable either because the token detector failed to find the corresponding structure in the other image of the pair, or because the corresponding token is not visible in the other image. Regions with a high ratio of unmatchable tokens to total tokens are likely to be regions of accretion or deletion. This motion-based, token-matching approach is an implementation of Kaplan's model for detecting such regions [12]. Kaplan argues that accretion and deletion are detected in the human visual system by isolating clusters of elements of optical texture, tracking them over time, and responding when they change in some way that is not topologically permissible. Token identification is equivalent to isolating elements of optical texture; token matching serves the purpose of tracking such elements over time; and analyzing unmatchable tokens is a response to some change which may be due to appearance or disappearance of a region.

IV. IDENTIFYING OCCLUDED SURFACES

Not only can accretion/deletion patterns be used to locate boundaries, they provide information that allows the identification of the side of the boundary being occluded. Such information is beneficial when interpreting dynamic scenes. Several specific approaches are possible, though all are based on associating the accretion or deletion region with a surface on one side of the boundary. That surface is the one being occluded. One approach relies upon the relative location of the accretion/deletion region with respect to the precise position of the image of the boundary. This boundary is the actual point of occlusion, the accretion/deletion region being on the same side of the boundary as the occluded surface. Fig. 1 illustrates this concept. The primary difficulty in this approach is identifying the boundary location relative to the accretion/deletion region. In particular, motion-based edge de-

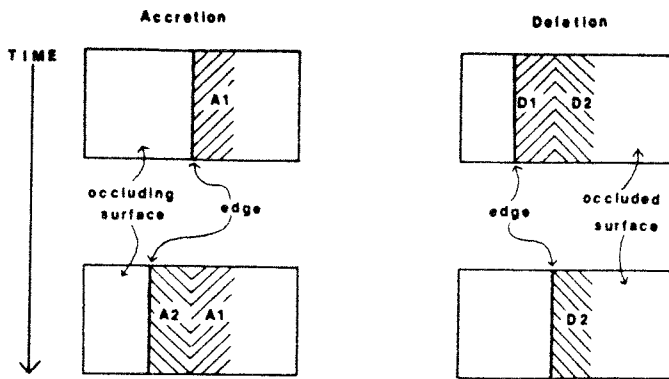


Fig. 2. The location of new accretion/deletion regions relative to previous such regions indicates the direction of the occluded surface. The second accretion region appears to the opposite side of the first accretion region as the occluded surface. The second deletion region occurs to the same side of the first deletion region as the occluded surface.

tection cannot locate the boundary precisely enough without first knowing which surface is occluded. The inadequacies of intensity-based edge detectors for this purpose are well known, particularly when applied to textured surfaces.

An alternative approach involves identifying the location of an accretion or deletion region relative to the location of such a region at a previous instant in time [13]. New accretion regions will appear to the side of previous accretion regions opposite the remainder of the occluded surface. New deletion regions will occur on the same side of previous deletion regions as the occluded surface (see Fig. 2). A disadvantage of this approach is the necessity to track and locate whole accretion/deletion regions over time.

The approach for identifying occluded surfaces from accretion/deletion regions which is developed in this paper requires the recognition of similarities between such regions and one of the two surfaces on either side of the boundary. Since the accretion/deletion region belongs to the occluded surface, it will share certain properties with that surface. The common property could be intensity or texture, although the problems inherent in most intensity-based analyses make these alternatives undesirable. Once again, motion-based properties may be more reliable. One such property is the disparity of tokens on the image plane. Disparity varies slowly over the surface of almost all rigid objects. Accretion or deletion tokens will thus exhibit disparities which are nearly identical to nearby token disparities on the same surface, while token disparities on different surfaces will usually vary.

V. IMPLEMENTATION

The system which was developed to detect occluded surfaces from regions of accretion or deletion uses token matching to obtain disparity vector fields. Unmatched tokens in clusters of high density are classified as accretion or deletion tokens, depending upon whether they have matches in subsequent or previous frame pairs. The disparity of accretion tokens after their appearance, or of deletion tokens prior to their disappearance, is obtained. Nearby tokens which are not accretion or deletion tokens and which have known disparities are identified and are used to identify the surface to which the ac-

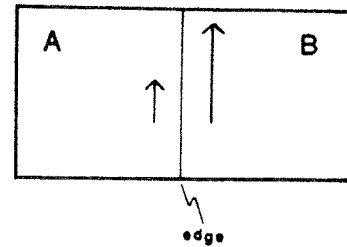


Fig. 3. The optical flow of surfaces *A* and *B* is indicated by the vectors on those surfaces. Since neither surface exhibits any flow perpendicular to the edge, there will be no accretion or deletion regions.

tion or deletion tokens belong. Such tokens with similar disparities to an accretion or deletion point lie on the occluded surface.

Three frames in an image sequence are required. Disparity fields *D1* and *D2* are obtained for frames 1 and 2, and for frames 2 and 3, respectively. Accretion points are not visible in frame 1, but do appear in frames 2 and 3. Tokens first appearing in frame 2, and thus having no associated match in frame 1, are noted. If these tokens have a match in frame 3, and if they are in a region with a high ratio of such tokens to total tokens, they are considered to be points of accretion. The disparity of accretion points is provided by *D2*. For every accretion point, a search is made within a neighborhood about the point location in frame 2. Tokens which are matched in *D2*, but which are not marked as accretion points are found. All of these tokens which have disparities similar to the accretion point are considered as a cluster. A vector pointing towards the center of the cluster is assigned to each accretion point, and indicates the direction from that point to the occluded surface.

Deletion points are visible in frames 1 and 2, but not frame 3. Tokens which are indicated as unmatchable in frame 2 are noted. If these tokens have a match in frame 1 and if they are in a region with a high ratio of such tokens to total tokens, they are considered to be points of deletion. The disparity of deletion points is provided by *D1*. For every deletion point, a search is made within a neighborhood about that point location in frame 1. Tokens which are matched in *D1*, but which are not marked as deletion points, are found. All of these tokens which have disparities similar to the deletion point are considered as a cluster. As before, a vector in the direction of the center of the cluster is assigned to each deletion point and indicates the direction from that point to the occluded surface.

VI. LIMITATIONS

This boundary detection technique requires a moderately dense token set, both to find accretion/deletion regions, and to determine image-plane displacements. This means that the two surfaces adjacent to the edge must be distinctly textured. In addition, there must be some component of optical flow perpendicular to the occlusion boundary, or neither accretion nor deletion will occur. In particular, motion exactly parallel to the boundary will produce no accretion or deletion regions (see Fig. 3). Perspective viewing of translating objects in principle leads to difficulties similar to those associated with rota-

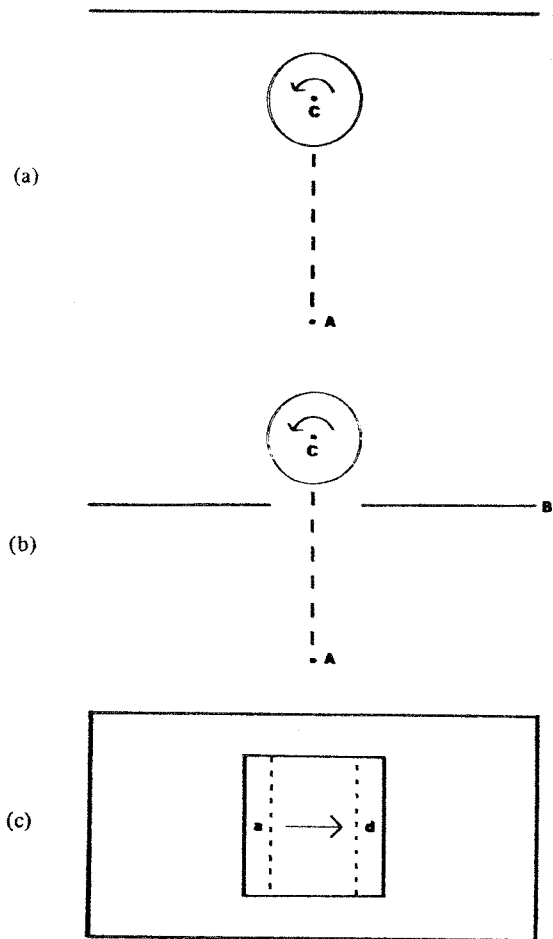


Fig. 4. (a) Overhead view of a cylinder rotating counter-clockwise about an axis at C , in front of a stationary background B . The viewer is at A , and the line of sight is along the dotted line. (b) The rotating cylinder seen through an aperture in surface B which is now in front. (c) When either (a) or (b) are viewed from point A , accretion regions a will occur along the left edge on the cylinder, deletion regions d along the right edge. While the cylinder is correctly identified as the occluded surface, there is insufficient information to determine the relative depth between the cylinder and the surface at B .

tion in depth (see below), as the perspective effects can be locally described as a combination of rotation and scale change. Fortunately, the practical difficulties caused by perspective effects are minimal. When objects are translating in front of a background, the size of accretion/deletion regions due to translation is almost always much greater than accretion/deletion regions that appear due to effective rotation of the object.

Certain rotations lead to potentially confusing situations when analyzing occlusion boundaries. Fig. 4(a) shows an overhead view of a cylinder rotating in depth. Fig. 4(c) shows the accretion/deletion regions that arise if there is no relative motion between the cylinder and the background surface. The analysis above assigns the accretion and deletion regions to the cylinder. Thus, the cylinder, not the background surface, is indicated as the surface being dynamically occluded. This is the correct interpretation, as the rotation in depth causes the cylinder to occlude itself over time. However, while the dynamic occlusion is correctly recognized, no information



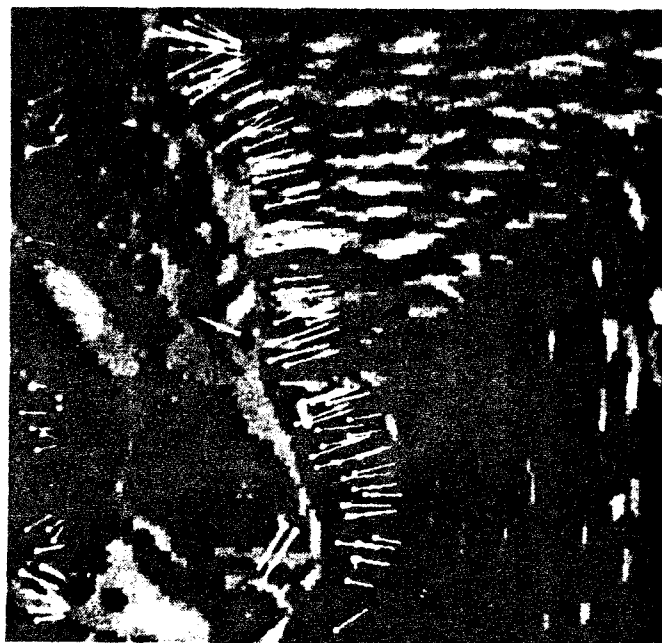
Fig. 5. Image sequence in which the leopard is translating from left to right.

is directly available about the relative depths to the surfaces on either side of the boundary. In fact, it is possible that the surrounding surface in the image is actually in front of the cylinder [Fig. 4(b)], yet generates the same image sequence.

A different complication occurs if the rotating object is moving with respect to the background surface, the cross section of the object is not circular, or the object is not rotating about its axis of symmetry. In all of these situations, accretion and/or deletion will be occurring on both sides of the actual boundary. The method given above is still valid and will identify both sides of the boundary as occluded surfaces. The problem again arises when trying to infer relative depth given an identification of the occluded surface. The determination of relative depth at a dynamic occlusion boundary when rotation is occurring is made possible by combining accretion/deletion analysis as described in this paper with an optical flow based approach [4]. This second technique uses the relationship between the flow of a boundary and the surface flows on either side of the boundary to identify *occluding* surfaces. Accretion/deletion analysis locates *occluded* surfaces. When taken together, both the occlusion of one surface by another and the self-occlusion resulting from rotation in depth can be recognized and appropriately interpreted.

VII. EXAMPLE

The system implementation described above was applied twice to the image sequence shown in Fig. 5, first processing the sequence in the order shown, then in the reverse order. All images had a resolution of 128×128 . There were approximately 1000 tokens identified in each image, and over 800 of these were matched in every image pair. As is usual with token matching systems, the density of tokens (and thus disparity vectors) varied across the image, being higher in areas of fine texture. An 11×11 square neighborhood, centered at the unmatched point, was used for computing the density of unmatched tokens. This size was chosen to be small enough so that most of the neighborhood fell within the accretion/deletion region, yet big enough to contain a reasonable number of tokens (usually 6 to 12). If 80 percent of the tokens in this neighborhood were unmatched in the same way as the point under consideration, then the point was labeled "accretion" or "deletion." This ratio was chosen to be selectively high, and yet to allow for some incorrect matches in the neighborhood, or some extension of the neighborhood out of the accretion/deletion region. A 31×31 window, centered at the accretion/deletion point, was searched to find clusters of similar disparity vectors. This size was chosen to be large enough to include portions of both surfaces outside the accretion/deletion region, yet not so large as to extend beyond these sur-



(a)



(b)

Fig. 6. (a) Results of occluded surface determination based upon accretion/deletion regions for the sequence of three frames in Fig. 5. Square white boxes are locations of accretion or deletion points in frame 2. The line emanating from each box points in the direction of the occluded surface. (b) Results of occluded surface determination when the sequence of three frames in Fig. 5 is processed in the reverse order.

faces. Disparity vectors were considered "similar" if they differed by no more than 2 pixels in each of the x and y components. The actual values of most of these parameters will, in general, depend upon factors such as the resolution of the images, the amount of texture, and the maximum expected disparity.

The results of processing in the forward direction are shown in Fig. 6(a). All of the square white points represent accretion

or deletion frame 2 tokens, which were matched in the second (D2) or first (D1) disparity field. The line which emanates from each box projects toward the surface which the algorithm indicates is being occluded. The set of tokens to the right of the leopard are deletion points. Tokens near the left border of the image are accretion points, which appear as more of the leopard moves into the field of view. Vectors associated with these points indicate that the leopard is being occluded by the surrounding frame. Except for six noise points, all accretion and deletion tokens have an associated vector pointing in the correct direction. The noise points are not in the accretion or deletion regions, but rather occur in or near untextured regions, or on the edge of the accretion/deletion regions. As a result, there are either no other tokens in the vicinity, or else a large number of unmatched tokens in the neighboring accretion/deletion region. These points are thus incorrectly identified as accretion or deletion points.

Fig. 6(b) shows the results when the image sequence of Fig. 5 is processed in the reverse order. The disparity field D1 is now the set of matches for frames 3 and 2, and D2 for frames 2 and 1. Tokens to the right of the leopard are now accretion points, and tokens near the left border of the image are deletion points. Once again, except for nine noise points, all vectors correctly point toward the occluded surface. The noise points are due to the same causes described in the previous paragraph.

REFERENCES

- [1] K. Nakayama and J. M. Loomis, "Optical velocity patterns, velocity sensitive neurons, and space perception: A hypothesis," *Perception*, vol. 3, pp. 63-80, 1974.
- [2] W. F. Clocksin, "Perception of surface slant and edge labels from optical flow: A computational approach," *Perception*, vol. 9, pp. 253-269, 1980.
- [3] W. B. Thompson, K. M. Mutch, and V. A. Berzins, "Edge detection in optical flow fields," presented at 2nd Nat. Conf. Artificial Intell., Aug. 1982.
- [4] —, "Dynamic occlusion analysis in optical flow fields." Comput. Sci. Dep., Univ. Minnesota. Tech. Rep. 84-6, 1984.
- [5] J. L. Potter, "Scene segmentation using motion information," *Comput. Graphics Image Processing*, vol. 6, pp. 558-581, Dec. 1977.
- [6] C. L. Fennema and W. B. Thompson, "Velocity determination in scenes containing several moving objects," *Comput. Graphics Image Processing*, vol. 9, pp. 301-315, Apr. 1979.
- [7] W. B. Thompson, "Combining motion and contrast for segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, pp. 543-549, Nov. 1980.
- [8] R. Jain, W. N. Martin, and J. K. Aggarwal, "Segmentation through the detection of changes due to motion," *Comput. Graphics Image Processing*, vol. 11, pp. 13-34, Sept. 1979.
- [9] R. Jain, "Extraction of motion information from peripheral processes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, no. 5, pp. 489-503, 1981.
- [10] G. A. Kaplan, "Kinetic disruption of optical texture: The perception of depth at an edge," *Perception Psychophys.*, vol. 6, pp. 193-198, 1969.
- [11] S. T. Barnard and W. B. Thompson, "Disparity analysis of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, pp. 333-340, July 1980.
- [12] G. A. Kaplan, "Kinetic disruption of optical texture: The perception of depth at an edge," Ph.D. dissertation, Cornell Univ., Ithaca, NY, 1968.
- [13] W. B. Thompson and S. T. Barnard, "Low-level estimation and interpretation of visual motion," *Comput.*, Aug. 1981.



Kathleen M. Mutch (S'80-M'83) received the M.S. and Ph.D. degrees in computer science from the University of Minnesota, Minneapolis, in 1981 and 1983, respectively.

She is currently an Assistant Professor in the Computer Science Department at Arizona State University, Tempe, AZ. Her research interests include time-varying image analysis and applications of computer vision.

Dr. Mutch is a member of the Association for Computing Machinery, SIGART, SIGCAPH, and the American Association for Artificial Intelligence.



William B. Thompson (S'72-M'75) received the Sc.B. degree in physics from Brown University, Providence, RI, in 1970, and the M.S. and Ph.D. degrees in computer science from the University of Southern California, Los Angeles, in 1972 and 1975, respectively.

He is currently an Associate Professor in the Computer Science Department of the University of Minnesota, Minneapolis, MN. Previously, he was a member of the Image Processing Institute at the University of Southern California. His primary research interest is in the area of computer vision, with an emphasis on the development of techniques for perceiving spatial organization. In addition, he is a Principal in the expert problem solving research group at the University of Minnesota.

Dr. Thompson is a member of the American Association for Artificial Intelligence and the Association for Computing Machinery.