Identification and Velocity Computation of Multiple Moving Objects in Images

SABRI A. MAHMOUD, Senior Member, IEEE

MOSTAFA S. AFIFI, Senior Member, IEEE
King Saud University
Saudi Arabia

ROGER J. GREEN, Senior Member, IEEE
University of Bradford
United Kingdom

Analysis for the identification and velocity computations of multiple moving objects in sequences of images is presented. Previously published analysis of the authors addressed sequences with one or two moving objects. Here the analysis is extended to sequences with multiple moving objects.

Simulation is performed to demonstrate the applicability of this technique. Multiple moving objects appear in the spectral domain as groups of individual peaks which identify the objects and is used to compute their velocities. Algorithms for evaluation of the spectral peaks of multiple moving objects, which completely identify the velocities of these objects, are also presented.

Manuscript received July 20, 1988.

IEEE Log No. 36731.

Authors' address: S. A. Mahmoud and M. S. Afifi, Computer Engineering Dep't., King Saud University, P.O. Box 51405, Riyadh 11543, Saudi Arabia; R. J. Green, Electrical Engineering Dep't., University of Bradford, Bradford WY BD7 1DP, UK.

0018-9251/90/0700-0586 $1.00 © 1990 IEEE

I. INTRODUCTION

Recognition of moving objects and extraction of their vector velocity is useful in many applications. Examples are in industrial and military applications, where a controlled camera can be mounted for identification of moving objects, in biomedical cell motion analysis, in tracking of multitargets from video data, in tracking dust storms and clouds for weather forecasts, and in highway traffic analysis. A method is presented for the detection of objects and velocity computation of multiple targets in sequences of images.

This work is an extension to previous work [8, 9] by the addition of the following: 1) analytical formulation to handle multiple moving objects in sequences of images, 2) study of the effects of the moving objects having distinct or similar velocities and entering or leaving the sequence at similar or different times and locations, and 3) presenting algorithms for discriminating between moving objects with similar velocities.

Researchers have used segment and match techniques to acquire velocity information [1, 10, 12]. Static images are segmented and then feature points are matched to establish correspondence between objects in successive frames. This technique is sensitive to segmentation errors and the success of the algorithm is based on accurate segmentation of static frames which is rarely satisfied in real world scenes, especially in noisy environments.

Researchers in [4–6] used differencing techniques to extract moving objects in a sequence. In their technique two frames are subtracted and the resulting frame gives the change between the two frames. This has limitations as it requires the images to be exactly registered, illumination to be invariant, and the moving objects to be totally displaced. In addition, if the sequence is noisy and the moving object grey level is low, the result of differencing is ambiguous whether it is showing motion changes or noise.

References [2, 3] applied temporal-spatial gradient techniques which utilize low level estimation of velocities of individual pixels from the analysis of a pair of frames. These methods give best results when the moving objects have smooth edges and the surfaces contain prominent texture. Their algorithm requires registered image sequences, and the velocity estimate is approximate (it is close to accurate velocity at the boundaries of the objects, while this is not the case at the surfaces of the objects).

The above three techniques have other limitations when multiple moving objects are in the sequence. The correspondence and occlusion problems are more severe.

Most researchers of time-varying images use only two or three frames of a sequence. The analysis based on few frames misses complete information about the motion of objects. References [11, 14] show that the human visual system requires an extended frame sequence to recover the structure of moving patterns. A longer sequence of frames allows the use of velocity information in analyzing the problem.

A sequence of frames is used in [13] for finding trajectories of feature points in an image sequence and tokens are used to solve the correspondence problem, assuming smoothness of velocity. Selecting interesting features, however, requires human interaction, moreover stationary objects also have interesting features.

Here the fast Fourier transform is applied to the time sequence. Then, peaks are detected in the spectrum. The temporal frequencies at which the peaks are detected are used to compute the velocities of the moving objects, by dividing the temporal frequencies by the corresponding spatial frequency. Hence, the velocities $V_1, V_2, V_3, \ldots$,etc. are computed. All stationary objects and time-invariant background appear as zero frequency component in the transform domain, and hence do not affect the velocity estimation.

The analytical model and formulations for multiple moving objects, in a time sequence, are presented in Section II. Then, the discrete Fourier transform is applied to the sequence, followed by velocity computation.

Section III presents an algorithm to discriminate between moving objects having the same velocity. These algorithms take advantage of the moving objects entering and/or leaving the sequence at different times and locations.

Section IV covers the simulation results of this method. Conclusions are presented in Section V. The detailed analytical derivations appear in Appendix I.

II. ANALYTICAL FORMULATION

To simplify the processing of the sequence, the transformation of the three-dimensional sequence to two-dimensional x- and y-sequences, which was presented in [8], is assumed here. The processing is applied to the generated two-dimensional sequences separately, where the velocities in the x- and y-directions are V_x and V_y, from which the velocity in the $x - y$ plane is found

$$V = \begin{bmatrix} V_x \\ V_y \end{bmatrix}. \tag{1}$$

A. The Case of Three Moving Objects

Multiple moving objects in a sequence can be modeled as a light intensity function $o[n,m]$ at a point n of frame m, and o is proportional to the brightness (or gray level) of the object at that point. The sequence consists of two contributions, one representing the moving objects and the other representing the background. In order to simplify the description of the moving objects, an assumed binary mask $o_c[n,m]$ is used. The model of the sequence is given by

$$g[n,m] = o_c[n,m]o[n,m] + (1 - o_c[n,m])b[n,m] \tag{2}$$

where $g[n,m]$ is the recorded sequence, $o[n,m]$ is the moving objects function in the time sequence, $b[n,m]$ is the time-varying background, and $o_c[n,m]$ is equal to 1 for all pixels corresponding to the moving objects and is equal to 0 otherwise.

Note that the moving objects, in the above model, occlude the background in the sequence.

In this paper we consider the multiple moving objects term $o[n,m]$. The effects of the other terms (i.e., the background term, whether it be stationary, time-invariant, or time-varying) was addressed in [9].

The case of three moving objects in a sequence is analyzed first, then the generalization to any number of moving objects is presented.

A model for three moving objects, one pixel in size each, in a sequence is given by

$$\begin{aligned} o[n,m] = {} & A_1\delta[n - L_{01} - mV_1] + A_2\delta[n - L_{02} - mV_2] \\ & + A_3\delta[n - L_{03} - mV_3] \\ & - A_1\delta[n - L_{01} - mV_1]\delta[n - L_{02} - mV_2] \\ & - A_1\delta[n - L_{01} - mV_1]\delta[n - L_{03} - mV_3] \\ & - A_2\delta[n - L_{02} - mV_2]\delta[n - L_{03} - mV_3] \\ & + A_1\delta[n - L_{01} - mV_1]\delta[n - L_{02} - mV_2] \\ & \times \delta[n - L_{03} - mV_3]. \end{aligned} \tag{3}$$

where n,m are the pixel coordinates at data point n and frame m; A_1, A_2, A_3 are the grey level values for the first, second, and third moving objects, respectively; L_{01}, L_{02}, L_{03} are the initial distance positions (from the first data point of the first frame) of the first, second, and third moving objects, respectively; V_1, V_2, V_3 are the velocities of the first, second, and third moving objects, respectively; and $\delta[\ ]$ is the dirac delta function.

In this model we arbitrarily assume that the higher indexed moving object occludes the lower indexed moving object when they cross in the sequence (at the moment when one object is occluding the other). This occurs when the moving objects are at the same location at the same frame. (When other rules are applied, only the amplitude of the occlusion term is changed accordingly).

The above model consists of the following three groups of terms. 1) The first three terms represent the path of the first, the second, and the third moving objects, respectively. 2) The second three terms represent the occlusion terms when one moving object occludes another in the sequence. 3) The last term represents the occlusion when one moving object occludes the other two moving objects.

The discrete Fourier transform of (3) is given by

$$o[k,f] = \sum_{m=0}^{m=M-1} \sum_{n=0}^{n=N-1} o[n,m] e^{-j2\pi fm/M} e^{-j2\pi kn/N} \tag{4}$$

where N,M are the total number of data points and frames, respectively. Expanding (4), as shown in Appendix I, we get,

2) The occlusion terms are functions of all frequencies, and not limited to particular ones. A particular term is zero when there is no actual occlusion in the sequence.

3) Each of the first three terms of (5) is nonzero only at $(f/M) + (f_i/N) = 0$ or $f = -f_i * M/N$.

Assume that $N = M$ (if this is not the case, the image can be appended with zeros to make $N = M$). Hence, each of the first three terms is nonzero at $f = -f_i$.

$$\begin{aligned} o[k,f] = {} & \frac{A_1 M_1 \operatorname{sinc}[(f/M + f_1/N)M_1] e^{-j2\pi k L_{01}/N} e^{-j2\pi(f/M + f_1/N)(2m_1 + M_1 - 1)}}{\operatorname{sinc}[f/M + f_1/N]} \\ & + \frac{A_2 M_2 \operatorname{sinc}[(f/M + f_2/N)M_2] e^{-j2\pi k L_{02}/N} e^{-j2\pi(f/M + f_2/N)(2m_2 + M_2 - 1)}}{\operatorname{sinc}[f/M + f_2/N]} \\ & + \frac{A_3 M_3 \operatorname{sinc}[(f/M + f_3/N)M_3] e^{-j2\pi k L_{03}/N} e^{-j2\pi(f/M + f_3/N)(2m_3 + M_3 - 1)}}{\operatorname{sinc}[f/M + f_3/N]} \\ & - A_1 e^{-j2\pi k L_{01}/N} e^{-j2\pi k(f/M + f_1/N)(L_{01} - L_{02})/(f_2 - f_1)} \in [k(L_{01} - L_{02})/(f_2 - f_1), m_2, M_{12}] \\ & - A_1 e^{-j2\pi k L_{01}/N} e^{-j2\pi k(f/M + f_1/N)(L_{01} - L_{03})/(f_3 - f_1)} \in [k(L_{01} - L_{03})/(f_3 - f_1), m_3, M_{13}] \\ & - A_1 e^{-j2\pi k L_{02}/N} e^{-j2\pi k(f/M + f_2/N)(L_{02} - L_{03})/(f_3 - f_2)} \in [k(L_{02} - L_{03})/(f_3 - f_2), m_3, M_{23}] \\ & + A_1 e^{-j2\pi k L_{01}/N} \delta[L_{01} - L_{03} + (f_1 - f_3)(L_{01} - L_{02})/(f_2 - f_1)] \\ & \times e^{-j2\pi k(f/M + f_1/N)(L_{01} - L_{02})/(f_2 - f_1)} \in [k(L_{01} - L_{02})/(f_2 - f_1), m_3, M_m] \end{aligned} \tag{5}$$

where $\operatorname{sinc}[Z] = \sin[\pi Z]/\pi Z$ and $\in$ is defined by (A8) in the Appendix.

The following 3 points are made clear by (5).

1) The contribution of a moving object at any frequency is proportional to the amplitude of the object (A_i) and to the number of frames (M_i) that the moving object is in the sequence. The relation between the number of frames that a moving object stays in the sequence (M_i), the velocity of the moving object V_i, and the initial position L_i is given by

$$M_i = (X_i - L_i)/V_i$$

where X_i is the final distance position (from the first data point) of moving object (i) in the image sequence (i.e., the position of the moving object before leaving the image sequence), and L_i is the initial distance position (from the first data point) of moving object (i) at the time this moving object enters the sequence. Hence, the number of frames is inversely proportional to the initial position (L_i) and to its velocity (V_i). (Note that L_i is a function of L_{oi} as given by (A2).)

Detecting a peak in the spectrum at a frequency f_p indicates that, at least, one of the first three terms is nonzero (as the contribution of the first three terms is more than the other terms at a peak). This indicates that $f_p = -f_i$. Substituting the value of f_i from (A7), we have

$$f_p = -kV_i$$

or

$$V_i = -f_p/k.$$

Hence, the velocities of the moving objects are found by detecting peaks in the spectrum given by (5), and finding the temporal frequencies at which the peaks exist. Let the temporal frequencies be at f_{1p}, f_{2p}, f_{3p} (for a specified value of $k = k_s$). Then, the velocities of the moving objects are

$$\begin{aligned} V_1 &= -f_{1p}/k_s; \\ V_2 &= -f_{2p}/k_s, \\ V_3 &= -f_{3p}/k_s. \end{aligned} \tag{6}$$

B. The Case of Multiple Moving Objects

A model for multiple moving objects in a time sequence is given by

$$
\begin{aligned}
o[n,m] &= \sum_{i=1}^{i=r} A_i \delta[n - L_{0i} - mV_i] \\
&\quad - \sum_{i=1}^{i=r-1} \sum_{j=i+1}^{j=r} A_i \delta[n - L_{0i} - mV_i] \\
&\quad \times \delta[n - L_{0j} - mV_j] + \sum_{i=1}^{i=r-2} \sum_{j=i+1}^{j=r-1} \sum_{c=i+2}^{c=r} \\
&\quad \times A_i \delta[n - L_{0i} - mV_i] \delta[n - L_{0j} - mV_j] \\
&\quad \times \delta[n - L_{0c} - mV_c] - \sum_{i=1}^{i=r-3} \sum_{j=i+1}^{j=r-2} \sum_{c=i+2}^{c=r-1} \sum_{d=i+3}^{d=r} \\
&\quad \times A_i \delta[n - L_{0i} - mV_i] \delta[n - L_{0j} - mV_j] \\
&\quad \times \delta[n - L_{0c} - mV_c] \delta[n - L_{0d} - mV_d] \\
&\quad + \cdots
\end{aligned} \tag{7}
$$

where r is the number of moving objects in the time sequence, and other variables are similar to the three moving objects case.

In this model, as in the previous case, we arbitrarily assumed that the higher indexed moving object occludes the lower indexed moving object when they cross in the sequence.

The above model summations are as follows.

1) The first summation represents the motion of the individual moving objects.

2) The second summation represents the occlusion terms when one moving object occludes another in the sequence.

3) The third summation represents the occlusion terms when one moving object occludes two moving objects.

4) The fourth summation represents the occlusion terms when one moving object occludes three moving objects.

More summation terms can be added in the same way for moving objects occluding a larger number of moving objects.

It can be deduced from (5), that the effect of the occlusion terms is very small, such that they have no appreciable effect on the velocity computations. Moreover, in real applications, it is very rare that three moving objects be at the same time at the same location (i.e., the probability that one moving object occluding another two objects is very small). The occlusion term, in fact, tends to reduce the spectral peaks by a negligible value. Hence, we can ignore the occlusion terms where more than two moving objects occlude one another. The new model is as follows:

$$
\begin{aligned}
o[n,m] &= \sum_{i=1}^{i=r} A_i \delta[n - L_{0i} - mV_i] \\
&\quad - \sum_{i=1}^{i=r-1} \sum_{j=i+1}^{j=r} A_i \delta[n - L_{0i} - mV_i] \delta[n - L_{0j} - mV_j].
\end{aligned} \tag{8}
$$

The result of applying the discrete Fourier transform to (8) can be deduced from the three moving objects case, and is given by

$$
\begin{aligned}
o[k,f] &= \sum_{i=1}^{i=r} \frac{A_i M_i \operatorname{sinc}[(f/M + f_i/N)M_i] e^{-j2\pi k L_{0i}/N} e^{-j2\pi(f/M + f_i/N)(2m_i + M_i - 1)}}{\operatorname{sinc}[f/M + f_i/N]} \\
&\quad - \sum_{i=1}^{i=r-1} \sum_{j=i+1}^{j=r} A_i e^{-j2\pi k L_{0i}/N} e^{-j2\pi k(f/M + f_i/N)(L_{0i} - L_{0j})/(f_j - f_i)} \\
&\in [k(L_{0i} - L_{0j})/(f_j - f_i), m_j, M_{ij}].
\end{aligned} \tag{9}
$$

The above analysis was carried out for objects of one pixel in size. It can be easily extended to multiple large moving objects by using the above results and those of reference [8].

III. ALGORITHMS FOR DISCRIMINATION BETWEEN MOVING OBJECTS OF SAME VELOCITY

The above analysis and formulation for velocity computation of multiple moving objects in sequences of images, indicates an ambiguity when moving objects have the same velocity. In this case, one peak in the spectrum corresponds to that velocity. This peak is a result of the sum of the spectral peaks of all the moving objects with that velocity as given by (5) and (6). This peak, although giving accurate velocity estimation, is wrongly associated with only one moving object. Hence, the number of moving objects in that situation is inaccurate.

The number of moving objects detected, for a sequence with multiple moving objects of different velocities, is not ambiguous. This is not true in the case of moving objects with similar velocities. In real applications, the moving objects normally start at different locations and enter and/or leave the sequence at different times. Generating sequences with subsets of the original sequence and applying the velocity technique to these sequences produces spectrum for subsets of the moving objects. The number of moving objects can be estimated by comparing the values of the spectrum peaks of the generated sequences.

A. Algorithm Based on Time Discrimination

Assuming that the moving objects enter and/or leave the sequence at different times, the following algorithm is applied.

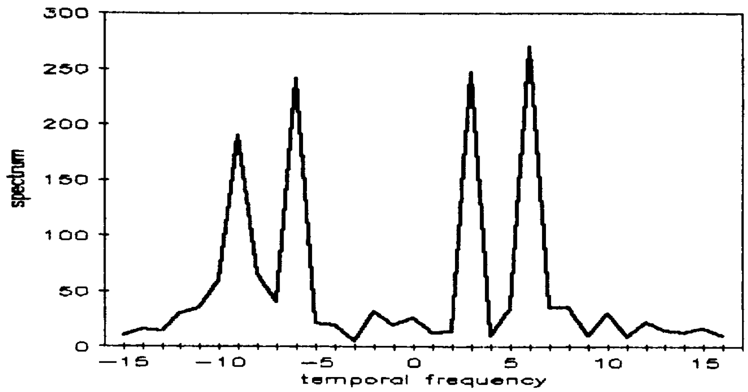


Fig. 1. Spectrum of sequence of 4 moving objects with distinct velocities at spatial frequency of 6.

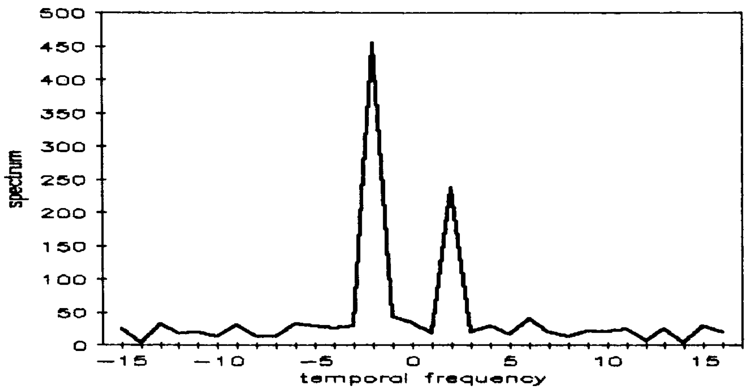


Fig. 2. Spectrum of sequence of 4 moving objects, three with similar velocities, at spatial frequency of 2.

1) Image sequences with contiguous subsets of the total number of frames are generated from the original image sequence. These sequences have regions of overlap and span all the frames of the original sequence.

2) The velocity computation procedure is then applied to each of the sequences.

3) The values of the spectral peaks are compared with other sequences.

B. Algorithm Based on Location Discrimination

Assuming that the moving objects start at different locations, the following algorithm is applied.

1) Image sequences are generated, from the original image sequence, in similar manner as in the previous algorithm (each generated sequence has a contiguous subset of the total number of data points of the original sequence). The generated sequences span all the image sequence data and adjacent regions overlap.

2) The velocity computation is applied to these generated sequences.

3) The values of the spectral peaks are compared. The comparison between the peak values of different generated sequences indicates whether a certain peak is a result of one or more moving objects, and hence the number of moving objects can be estimated.

IV. SIMULATION RESULTS

Simulations were performed to show the effects of having multiple moving objects in a sequence.

1) *Moving Objects with Distinct Velocities*: Fig. 1 shows the spectrum of a sequence with four moving objects at a spatial frequency of six. It is clear from this figure that there are four spectral peaks at −9, −6, 3, and 6 temporal frequencies. Hence, the velocities of the moving objects are $-(-9/6) = 1.5$ pixels per frame (ppf), $-(-6/6) = 1$ ppf, $-(3/6) = -0.5$ ppf, and $-(6/6) = -1$ ppf, as given by (6). It is clear that there are four moving objects. In this case the number of the moving objects in a sequence is accurately found.

2) *Moving Objects with Similar Velocities*: Fig. 2 shows the spectrum, at a spatial frequency of two, for a sequence of four moving objects where three with same velocity and the third is at different velocity. The

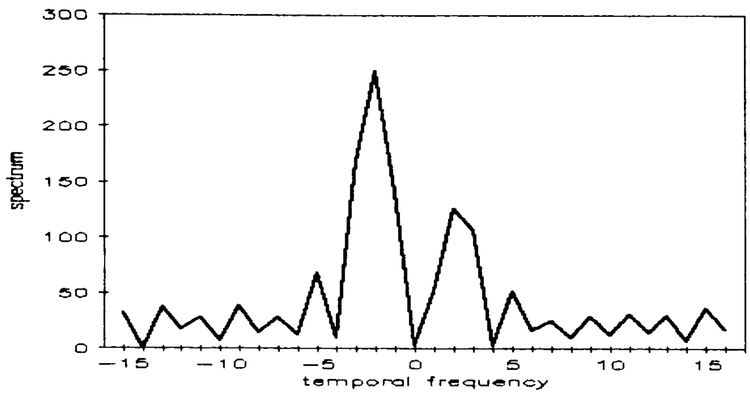


Fig. 3. Spectrum of first 16 frames of sequence of Fig. 2 at spatial frequency of 2.

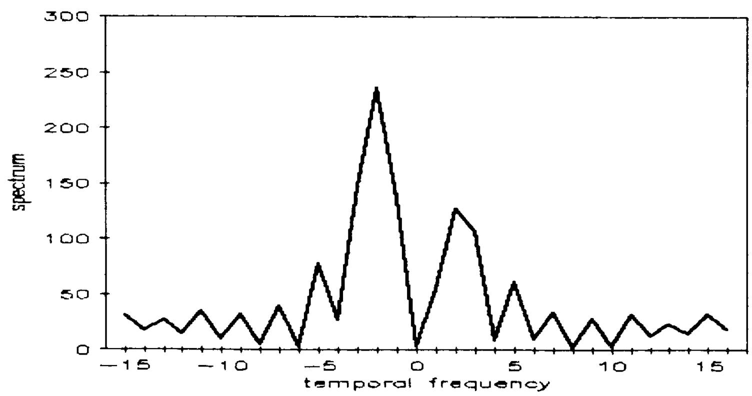


Fig. 4. Spectrum of frames 9–28 of sequence of Fig. 2 at spatial frequency of 2.

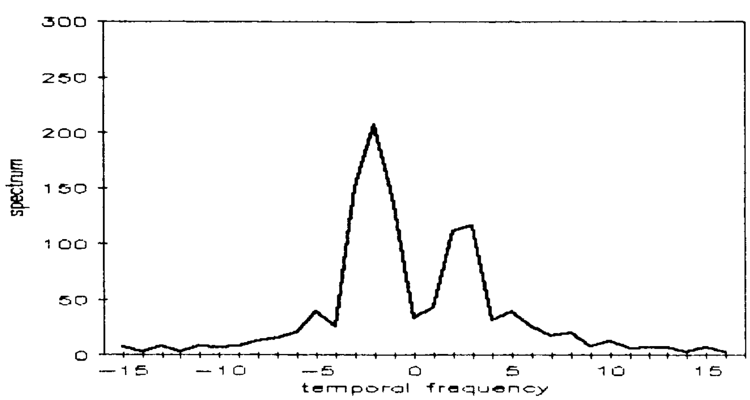


Fig. 5. Spectrum of last 16 frames of sequence of Fig. 2 at spatial frequency of 2.

figure shows two peaks corresponding to temporal frequencies of -2 and 2, hence the velocities are 1 and -1 ppf, respectively. Although we have four moving objects, it is not clear from the figure whether we have two or more moving objects.

a) *Sequences generated from subsets of the image sequence frames*: To find the correct number of moving objects in a sequence, subsequences consisting of sections of the original sequence are generated then the above procedure is applied to these subsequences.

Figs. 3, 4, and 5 show the spectrum, at a spatial frequency of two, for generated sequences consisting of frames 1 to 16, 9 to 24, and 17 to 32 of the original sequence, respectively. The three figures show that the moving objects are at velocities of 1 and -1 ppf. The peak values for the -1 ppf velocity is nearly constant

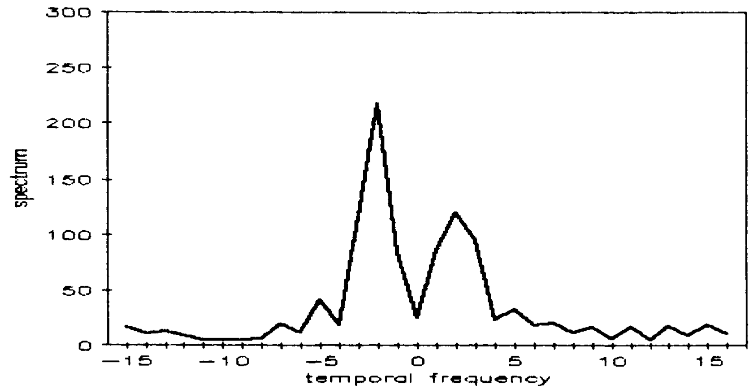


Fig. 6. Spectrum of first 16 data points of all frames of sequence of Fig. 2 at spatial frequency of 2.

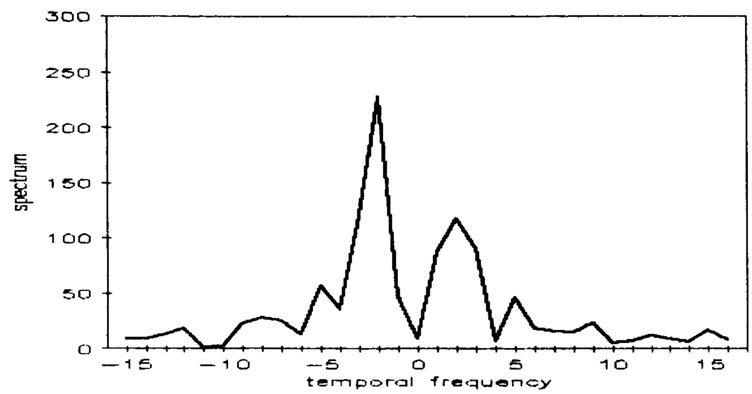


Fig. 7. Spectrum of middle 16 data points of all frames of sequence of Fig. 2 at spatial frequency of 2.

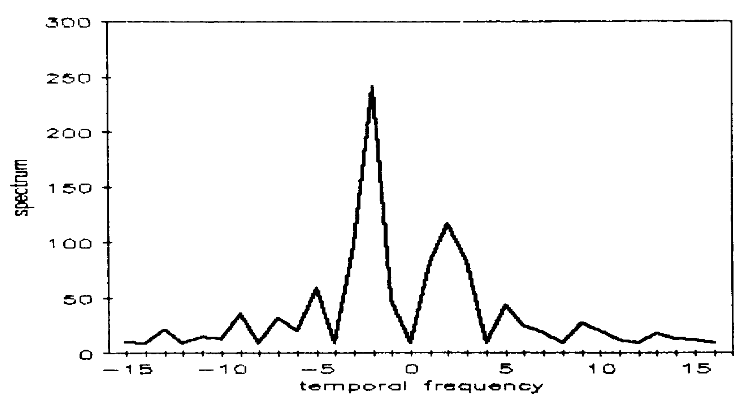


Fig. 8. Spectrum of last 16 data points of all frames of sequence of Fig. 2 at spatial frequency of 2.

indicating that this peak is a result of one moving object. The spectral peaks corresponding to a velocity of 1 ppf are changing (249.4, 236.7, and 208 in Figs. 3, 4, and 5, respectively). The decrease in the spectral peak in Figs. 4 and 5 from that of Fig. 3 indicates that, most probably, one moving object left the sequence. Hence, the number of moving objects in the sequence is more than two.

b) *Sequences generated from subsets of the image sequence data points*: Figs. 6, 7, and 8 show the spectrum, at a spatial frequency of two, for generated sequences consisting of data points 1 to 16, 9 to 24, and 17 to 32 of the original sequence, respectively. The three figures show that we have moving objects at velocities of 1 and −1 ppf. The peak values for the −1 ppf velocity are nearly constant, indicating that this peak is a result of one moving object. The spectral peaks corresponding to a velocity of 1 ppf are changing (217.1, 228.1, and 241.2 in Figs. 6, 7, and 8, respectively). The increase in the spectral peak in Figs. 7 and 8 compared with Fig. 6 indicates that, most probably, one moving object entered the sequence. Hence, the number of moving objects in the sequence is more than two.

The values of the spectral peaks for the generated sequences are used to gain more insight into the sequence, and can be used to get the actual number of moving objects in the sequence. It may be necessary, in some cases, to generate more sequences with less number of frames from the original sequence, or less data points (to refine the sensing range of detection of the objects). Combining the results of the different sequences (whether generated from subsets of the sequence frames, or data points) gives accurate estimation and confidence in the results, and clarifies any ambiguities.

V. CONCLUSIONS

This algorithm leads to better velocity estimation as it uses the whole image sequence, unlike most other techniques which use only two or three frames.

The velocity computation is done in the transformation domain using one-dimensional Fourier transform, hence it is efficient. No segmentation, matching or correspondence between frames are needed. It is less sensitive to occlusion and noise than most other methods. The spectrum component of the stationary objects is at zero temporal frequency which is not considered for velocity computation, hence correct velocities are estimated in the presence of stationary objects.

In the case of moving objects with similar velocities, algorithms which apply this technique to the generated subsequences are introduced. The values of the spectrum peaks are compared to trace the moving objects entering or leaving the sequence. Hence the number of moving objects in the sequence is found, thus eliminating the ambiguity problem.

This method, however, requires many frames, which are hard to obtain in some applications. At this time the method does not attempt to recognize the moving objects in the time domain.

Research is in progress to expand the algorithm, add the time-domain recognition, and incorporate previous experience and history, for systematic determination of the moving objects. In addition, the above technique will be applied to real images.

APPENDIX I

Derivation of the discrete Fourier transform for three moving objects in a time sequence:

A model for three moving objects in time sequence, each of one pixel in size, is given by

$$\begin{aligned} o[n,m] &= A_1\delta[n-L_1-(m-m_1)V_1] \\ &\quad + A_2\delta[n-L_2-(m-m_2)V_2] \\ &\quad + A_3\delta[n-L_3-(m-m_3)V_3] \\ &\quad - A_1\delta[n-L_1-(m-m_1)V_1]\delta[n-L_2-(m-m_2)V_2] \\ &\quad - A_1\delta[n-L_1-(m-m_1)V_1]\delta[n-L_3-(m-m_3)V_3] \\ &\quad - A_2\delta[n-L_2-(m-m_2)V_2]\delta[n-L_3-(m-m_3)V_3] \\ &\quad + A_1\delta[n-L_1-(m-m_1)V_1]\delta[n-L_2-(m-m_2)V_2] \\ &\quad \times \delta[n-L_3-[m-m_3]V_3]. \end{aligned} \tag{A1}$$

where n,m are the pixel coordinates at data point n and frame m; A_1,A_2,A_3 are the grey level values for the first, second, and third moving objects, respectively; L_1,L_2,L_3 are the initial distance positions (from the first data point) of the first, second, and third moving objects, respectively, at the time each object appears in the sequence; m_1,m_2,m_3 are the frame numbers at which the moving objects one, two, and three appear in the sequence, respectively; V_1,V_2,V_3 are the velocities of the first, second, and third moving objects, respectively; $\delta[\ ]$ is the dirac delta function. In this model, we arbitrarily assumed that the higher indexed moving object occludes the lower indexed moving object when they cross in the sequence. This occurs when the moving objects are at the same location at the same frame. (When other rules are applied, only the amplitude of the occlusion term is changed accordingly.) Let $m_1 \le m_2 \le m_3$, the low-indexed moving object appears before, or at the same time as the high-indexed moving object appears in the sequence (this is done by labeling the moving objects according to the time each appears in the sequence); and $L_1,L_2,L_3,m_1V_1,m_2V_2,m_3V_3$ are integers.

The above model consists of the following three groups of terms. 1) The first three terms represent the motion of the first, second, and third moving objects, respectively. 2) The next three terms represent the occlusion terms, when one moving object occludes another in the sequence. 3) The last term represents the occlusion, when one moving object occludes the other two moving objects.

Let

$$\begin{aligned} L_{01} &= L_1 - m_1V_1 \\ L_{02} &= L_2 - m_2V_2 \\ L_{03} &= L_3 - m_3V_3 \end{aligned} \tag{A2}$$

where L_{01}, L_{02}, and L_{03} are the initial distance positions (from the first data point of the first frame) of the first, second, and third moving objects, respectively.

Substituting (A2) in (A1), we have

$$\begin{aligned} o[n,m] = {} & A_1\delta[n-L_{01}-mV_1] + A_2\delta[n-L_{02}-mV_2] \\ & + A_3\delta[n-L_{03}-mV_3] \\ & - A_1\delta[n-L_{01}-mV_1]\delta[n-L_{02}-mV_2] \\ & - A_1\delta[n-L_{01}-mV_1]\delta[n-L_{03}-mV_3] \\ & - A_2\delta[n-L_{02}-mV_2]\delta[n-L_{03}-mV_3] \\ & + A_1\delta[n-L_{01}-mV_1]\delta[n-L_{02}-mV_2] \\ & \times \delta[n-L_{03}-mV_3]. \end{aligned} \tag{A3}$$

The discrete Fourier transform of equation (A3) is given by

$$o[k,f] = \sum_{m=0}^{m=M-1} \sum_{n=0}^{n=N-1} o[n,m] e^{-j2\pi f m/M} e^{-j2\pi kn/N} \tag{A4}$$

where N, M are the total number of data points and frames, respectively.

The moving objects appear in the sequence in the following frames:

from frames m_1 to $m_1 + M_1 - 1$ for the first moving object;

from frames m_2 to $m_2 + M_2 - 1$ for the second moving object;

from frames m_3 to $m_3 + M_3 - 1$ for the third moving object;

from frames m_2 to $M_{12} - 1$ for the first occlusion term;

from frames m_3 to $M_{13} - 1$ for the second occlusion term;

from frames m_3 to $M_{23} - 1$ for the third occlusion term;

from frames m_3 to $M_m - 1$ for the last occlusion term;

where M_1, M_2, and M_3 are the number of frames the first, second, and third moving objects are in the sequence, respectively; $M_{12} = \min[(m_1 + M_1), (m_2 + M_2)]$, is the last frame in which the first and second moving objects are simultaneously in the sequence; $M_{13} = \min[(m_1 + M_1), (m_3 + M_3)]$, is the last frame in which the first and third moving objects are simultaneously in the sequence; $M_{23} = \min[(m_2 + M_2), (m_3 + M_3)]$, is the last frame in which the second and third moving objects are simultaneously in the sequence; $M_m = \min[M_{12}, M_{13}, M_{23}]$, is the last frame in which the three moving objects are simultaneously in the sequence.

Applying the limits in which each moving object appears in the sequence:

$$\begin{aligned} o[k,f] = {} & \sum_{m=m_1}^{m=m_1+M_1-1} \sum_{n=0}^{n=N-1} A_1\delta[n-L_{01}-mV_1] e^{-j2\pi f m/M} e^{-j2\pi kn/N} \\ & + \sum_{m=m_2}^{m=m_2+M_2-1} \sum_{n=0}^{n=N-1} A_2\delta[n-L_{02}-mV_2] e^{-j2\pi f m/M} e^{-j2\pi kn/N} \\ & + \sum_{m=m_3}^{m=m_3+M_3-1} \sum_{n=0}^{n=N-1} A_3\delta[n-L_{03}-mV_3] e^{-j2\pi f m/M} e^{-j2\pi kn/N} \\ & - \sum_{m=m_2}^{m=M_{12}-1} \sum_{n=0}^{n=N-1} A_1\delta[n-L_{01}-mV_1]\delta[n-L_{02}-mV_2] e^{-j2\pi f m/M} e^{-j2\pi kn/N} \\ & - \sum_{m=m_3}^{m=M_{13}-1} \sum_{n=0}^{n=N-1} A_1\delta[n-L_{01}-mV_1]\delta[n-L_{03}-mV_3] e^{-j2\pi f m/M} e^{-j2\pi kn/N} \\ & - \sum_{m=m_3}^{m=M_{23}-1} \sum_{n=0}^{n=N-1} A_2\delta[n-L_{02}-mV_2]\delta[n-L_{03}-mV_3] e^{-j2\pi f m/M} e^{-j2\pi kn/N} \\ & + \sum_{m=m_3}^{m=M_m-1} \sum_{n=0}^{n=N-1} A_1\delta[n-L_{01}-mV_1]\delta[n-L_{02}-mV_2]\delta[n-L_{03}-mV_3] e^{-j2\pi f m/M} e^{-j2\pi kn/N}. \end{aligned} \tag{A5}$$

Summing with respect to n, all the terms are zero except at n between 0 and $N-1$ and the following conditions:

the first term is zero except at $n = L_{01} + mV_1$;

the second term is zero except at $n = L_{02} + mV_2$;

the third term is zero except at $n = L_{03} + mV_3$;
the fourth term is zero except at $n = L_{01} + mV_1$ and $n = L_{02} + mV_2$;
the fifth term is zero except at $n = L_{01} + mV_1$ and $n = L_{03} + mV_3$;
the sixth term is zero except at $n = L_{02} + mV_2$ and $n = L_{03} + mV_3$;
the last term is zero except at $n = L_{01} + mV_1$, $n = L_{02} + mV_2$, and $n = L_{03} + mV_3$.

Summing with respect to m, the occlusion terms are zero except at m between the limits of the summation and the following conditions:

the first occlusion term is zero except at $m = k(L_{01} - L_{02})/(f2 - f1)$;
the second occlusion term is zero except at $m = k(L_{01} - L_{03})/(f3 - f1)$;
the third occlusion term is zero except at $m = k(L_{02} - L_{03})/(f3 - f2)$;

$$\begin{aligned} o[k,f] = {} & A_1 e^{-j2\pi k L_{01}/N} \sum_{m=m_1}^{m=m_1+M_1-1} e^{-j2\pi m(f/M + kV_1/N)} \\ & + A_2 e^{-j2\pi k L_{02}/N} \sum_{m=m_2}^{m=m_2+M_2-1} e^{-j2\pi m(f/M + kV_2/N)} + A_3 e^{-j2\pi k L_{03}/N} \sum_{m=m_3}^{m=m_3+M_3-1} e^{-j2\pi m(f/M + kV_3/N)} \\ & - A_1 e^{-j2\pi k L_{01}/N} \sum_{m=m_2}^{m=M_{12}-1} \delta[L_{01} - L_{02} + m(V_1 - V_2)] e^{-j2\pi m(f/M + kV_1/N)} \\ & - A_1 e^{-j2\pi k L_{01}/N} \sum_{m=m_3}^{m=M_{13}-1} \delta[L_{01} - L_{03} + m(V_1 - V_3)] e^{-j2\pi m(f/M + kV_1/N)} \\ & - A_2 e^{-j2\pi k L_{02}/N} \sum_{m=m_3}^{m=M_{23}-1} \delta[L_{02} - L_{02} + m(V_2 - V_3)] e^{-j2\pi m(f/M + kV_2/N)} \\ & + A_1 e^{-j2\pi k L_{01}/N} \sum_{m=m_3}^{m=M_m-1} \delta[L_{01} - L_{02} + m(V_1 - V_2)] \delta[L_{01} - L_{03} + m(V_1 - V_3)] e^{-j2\pi m(f/M + kV_1/N)}. \end{aligned} \quad \text{(A6)}$$

Let

$$kV_1 = f_1; \quad kV_2 = f_2; \quad kV_3 = f_3 \quad \text{(A7a)}$$

$$V_1 = f_1/k; \quad V_2 = f_2/k; \quad V_3 = f_3/k. \quad \text{(A7b)}$$

Substituting equation (A7a) in equation (A6), we have

the last occlusion term is zero except at $m = k(L_{01} - L_{02})/(f2 - f1)$ and $m = k(L_{02} - L_{03})/(f3 - f2)$.

Let ϵ be a function such that

$$\begin{aligned} \epsilon[x, x_1, x_2] &= 1, \quad \text{for} \quad x_1 \le x < x_2 \\ &= 0, \quad \text{otherwise} \end{aligned} \quad \text{(A8)}$$

$$\begin{aligned} o[k,f] = {} & A_1 e^{-j2\pi k L_{01}/N} \sum_{m=m_1}^{m=m_1+M_1-1} e^{-j2\pi m(f/M + f_1/N)} \\ & + A_2 e^{-j2\pi k L_{02}/N} \sum_{m=m_2}^{m=m_2+M_2-1} e^{-j2\pi m(f/M + f_2/N)} + A_3 e^{-j2\pi k L_{03}/N} \sum_{m=m_3}^{m=m_3+M_3-1} e^{-j2\pi m(f/M + f_3/N)} \\ & - A_1 e^{-j2\pi k L_{01}/N} \sum_{m=m_2}^{m=M_{12}-1} \delta[L_{01} - L_{02} + m(f_1 - f_2)/k] e^{-j2\pi m(f/M + f_1/N)} \\ & - A_1 e^{-j2\pi k L_{01}/N} \sum_{m=m_3}^{m=M_{13}-1} \delta[L_{01} - L_{03} + m(f_1 - f_3)/k] e^{-j2\pi m(f/M + f_1/N)} \\ & - A_2 e^{-j2\pi k L_{02}/N} \sum_{m=m_3}^{m=M_{23}-1} \delta[L_{02} - L_{03} + m(f_2 - f_3)/k] e^{-j2\pi m(f/M + f_2/N)} \\ & + A_1 e^{-j2\pi k L_{01}/N} \sum_{m=m_3}^{m=M_m-1} \delta[L_{01} - L_{02} + m(f_1 - f_2)] \delta[L_{01} - L_{03} + m(f_1 - f_3)/k] e^{-j2\pi m(f/M + f_1/N)}. \end{aligned}$$

or

$$\epsilon[x,x_1,x_2] = u[x-x_1] - u[x-x_2]$$

where $u[x]$ is the unit step function.

The conditions for the occlusion terms to be nonzero can be defined in terms (A8). The first occlusion term function $\epsilon[k(L_{01}-L_{02})/(f_2-f_1),m_2,M_{12}]$, the second occlusion term function $\epsilon[k(L_{01}-L_{03})/(f_3-f_1),m_3,M_{13}]$, the third occlusion term function $\epsilon[k(L_{02}-L_{03})/(f_3-f_2),m_3,M_{23}]$, the last occlusion term function $\epsilon[k(L_{01}-L_{02})/(f_2-f_1),m_3,M_m]$,

$$\begin{aligned}
o[k,f] = {} & \frac{A_1 e^{-j2\pi k L_{01}/N}\{(1-e^{-j2\pi(f/M+f_1/N)(m_1+M_1)}) - (1-e^{-j2\pi(f/M+f_1/N)m_1})\}}{1-e^{-j2\pi(f/M+f_1/N)}} \\
& + \frac{A_2 e^{-j2\pi k L_{02}/N}\{(1-e^{-j2\pi(f/M+f_2/N)(m_2+M_2)}) - (1-e^{-j2\pi(f/M+f_2/N)m_2})\}}{1-e^{-j2\pi(f/M+f_2/N)}} \\
& + \frac{A_3 e^{-j2\pi k L_{03}/N}\{(1-e^{-j2\pi(f/M+f_3/N)(m_3+M_3)}) - (1-e^{-j2\pi(f/M+f_3/N)m_3})\}}{1-e^{-j2\pi(f/M+f_3/N)}} \\
& - A_1 e^{-j2\pi k L_{01}/N} e^{-j2\pi k(f/M+f_1/N)(L_{01}-L_{02})/(f_2-f_1)} \in [k(L_{01}-L_{02})/(f_2-f_1),m_2,M_{12}] \\
& - A_1 e^{-j2\pi k L_{01}/N} e^{-j2\pi k(f/M+f_1/N)(L_{01}-L_{03})/(f_3-f_1)} \in [k(L_{01}-L_{03})/(f_3-f_1),m_3,M_{13}] \\
& - A_1 e^{-j2\pi k L_{02}/N} e^{-j2\pi k(f/M+f_2/N)(L_{02}-L_{03})/(f_3-f_2)} \in [k(L_{02}-L_{03})/(f_3-f_2),m_3,M_{23}] \\
& + A_1 e^{-j2\pi k L_{01}/N} \delta[L_{01}-L_{03}+(f_1-f_3)(L_{01}-L_{02})/(f_2-f_1)] \\
& \times e^{-j2\pi k(f/M+f_1/N)(L_{01}-L_{02})/(f_2-f_1)} \in [k(L_{01}-L_{02})/(f_2-f_1),m_3,M_m]
\end{aligned} \tag{A9}$$

Rearranging terms in (A9), we have,

$$\begin{aligned}
o[k,f] = {} & \frac{A_1 M_1 \,\mathrm{sinc}[(f/M+f_1/N)M_1] e^{-j2\pi k L_{01}/N} e^{-j2\pi(f/M+f_1/N)(2m_1+M_1-1)}}{\mathrm{sinc}[f/M+f_1/N]} \\
& + \frac{A_2 M_2 \,\mathrm{sinc}[(f/M+f_2/N)M_2] e^{-j2\pi k L_{02}/N} e^{-j2\pi(f/M+f_2/N)(2m_2+M_2-1)}}{\mathrm{sinc}[f/M+f_2/N]} \\
& + \frac{A_3 M_3 \,\mathrm{sinc}[(f/M+f_3/N)M_3] e^{-j2\pi k L_{03}/N} e^{-j2\pi(f/M+f_3/N)(2m_3+M_3-1)}}{\mathrm{sinc}[f/M+f_3/N]} \\
& - A_1 e^{-j2\pi k L_{01}/N} e^{-j2\pi k(f/M+f_1/N)(L_{01}-L_{02})/(f_2-f_1)} \in [k(L_{01}-L_{02})/(f_2-f_1),m_2,M_{12}] \\
& - A_1 e^{-j2\pi k L_{01}/N} e^{-j2\pi k(f/M+f_1/N)(L_{01}-L_{03})/(f_3-f_1)} \in [k(L_{01}-L_{03})/(f_3-f_1),m_3,M_{13}] \\
& - A_1 e^{-j2\pi k L_{02}/N} e^{-j2\pi k(f/M+f_2/N)(L_{02}-L_{03})/(f_3-f_2)} \in [k(L_{02}-L_{03})/(f_3-f_2),m_3,M_{23}] \\
& + A_1 e^{-j2\pi k L_{01}/N} \delta[L_{01}-L_{03}+(f_1-f_3)(L_{01}-L_{02})/(f_2-f_1)] \\
& \times e^{-j2\pi k(f/M+f_1/N)(L_{01}-L_{02})/(f_2-f_1)} \in [k(L_{01}-L_{02})/(f_2-f_1),m_3,M_m]
\end{aligned} \tag{A10}$$

where $\mathrm{sinc}[X] = \sin[\pi X]/\pi X$.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the referees for their time and effort in carefully reviewing the preliminary manuscript and for their constructive criticism. The incorporation of their suggestions and comments significantly improved the clarity of the final manuscript.

REFERENCES

[1] Aggrawal, J. K., and Duda, R. O. (1975)
Computer analysis of moving polygonal images.
IEEE Transactions on Computers, **C-24**, 10 (Oct. 1975), 966–975.

[2] Cafforio, C., and Fabio, R. (1976)
Methods for measuring small displacements of TV images.
IEEE Transactions on Information Theory, **IT-22**, 5 (Sept. 1976), 573–579.

[3] Fennema, C. L., and Thompson, W. B. (1979)
Velocity determination in scenes containing several moving objects.
Computer Graphics and Image Processing, **9** (Apr. 1979), 301–315.

[4] Jain, R., and Nagel, H. H. (1979)
On the analysis of accumulative difference pictures from image sequences of real world scenes.
IEEE Transactions on Pattern Analysis and Machine Intelligence, **PAMI-1**, 2 (Apr. 1979), 206–214.

[5] Jain, R., Martin, W. N., and Aggarwal, J. K. (1979)
Segmentation through the detection of changes due to motion.
Computer Graphics and Image Processing, **11** (Sept. 1979), 13–34.

[6] Jain, R. (1981)
Extraction of motion information from peripheral processing.
IEEE Transactions on Pattern Analysis and Machine Intelligence, **PAMI-3**, 5 (1981), 489–503.

[7] Mahmoud, S. A., Afifi, M. S., and Green, R. J. (1984)
Application of the exponential area transform for detecting the velocities of moving objects.
Presented at the Arabian Productivity Advancement using Computers/Graphics Conference, Riyadh, Saudi Arabia, Feb. 20–23, 1984.

[8] Mahmoud, S. A., Afifi, M. S., and Green, R. J. (1988)
Recognition of large moving objects in images.
IEEE Transactions on Acoustics, Speech, and Signal Processing, Nov. 1988, 1790–1791.

[9] Mahmoud, S. A., Afifi, M. S., and Green, R. J. (1990)
The effects of background on velocity computation of moving objects in a sequence of images.
Journal of Microcomputer Applications, UK, Jan. 1990, 89–96.

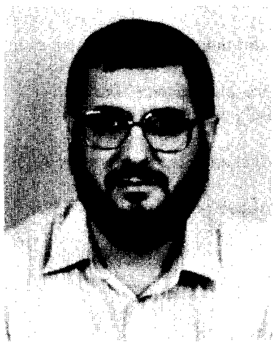
[10] Martin, W. N., and Aggrawal, J. K. (1979)
Computer analysis of dynamic scenes containing curvilinear figures.
Pattern Recognition, **11** (1979), 169–178.

[11] Ramachandran, V. S., and Antis, S. M. (1984)
Extrapolation of motion path in human visual perception.
Vision Research, **23** (1984), 83–85.

[12] Roach, J. W., and Aggrawal, J. K. (1979)
Computer tracking of objects moving in space.
IEEE Transactions on Pattern Analysis and Machine Intelligence, **PAMI-1**, 2 (Apr. 1979), 127–135.

[13] Sethi, I. K., and Jain, R. (1987)
Finding trajectories of feature points in a monocular image sequence.
IEEE Transactions on Pattern Analysis and Machine Intelligence, **PAMI-9**, 1 (1987), 56–73.

[14] Todd, J. T. (1982)
Visual information about rigid and nonrigid motion: A geometric analysis.
Journal of Experimental Psychology, **8** (1982), 238–252.



Sabri A. Mahmoud (M'89) received his M.S. in computer sciences from Stevens Institute of Technology, Hoboken, NJ, in 1980. He received his Ph.D. in information systems engineering from the University of Bradford, United Kingdom, in 1987.

He worked with Royal Saudi Naval Forces as Senior Systems Engineer for five years and is currently an Assistant Professor at the Computer Engineering Department at King Saud University, Saudi Arabia. His research interests include image processing, pattern recognition, parallel processing, and the use of fuzzy concepts and petri nets in the above applications.

Dr. Mahmoud is a member of IEEE Acoustic, Speech, and Signal Processing Society.

Mostafa S. Afifi (M'60—SM'78) was born in Egypt in 1936. He received his D.Sc. in 1967 from Delft University of Technology, The Netherlands, in microwaves and antennas.

From 1963 to 1970 he was with Delft University, from 1970–1973 with Northrop Page Communications Engineers, from 1973–1974 with Fairchild Space and Electronics Company, and from 1974–1982 with General Electric Space Center. Since 1982 he has been with the Electrical Engineering Department, King Saud University, Saudi Arabia, where he is currently a Professor. His research interests include telecommunications, imaging and signal processing systems.

He is senior member of the IEEE, was a Chairman of the AP/MTT Philadelphia section in 1981 and elected treasurer in 1982.

Roger J. Green (M'86—SM'88) obtained his B.Sc. in electronics from Manchester University, U.K., in 1973 and his Ph.D. in communications from Bradford University, U.K., in 1976.

After several years in industry, he worked at Leads University as a lecturer, and moved to Bradford University as a full lecturer in 1981. His research interests include imaging systems, electro-optics, and optoelectronics.

Dr. Green is an IEE member, and a Senior Member of the IEEE. He is the holder of a patent on low noise optical detection, and has published widely.