A NEW MOTION-ESTIMATION ALGORITHM BASED ON MORPHOLOGICAL CORRELATION

A.C.P. Loui , A.N. Venetsanopoulos, and K.C. Smith*

Department of Electrical Engineering
University of Toronto
Toronto, Ontario
Canada M5S 1A4

ABSTRACT

A novel algorithm is proposed for efficient estimation of two-dimensional planar motion. This approach is based on the idea of morphological correlation first introduced for shape representation [1]. The advantage of this algorithm is that it can be implemented in real-time, using high-speed morphological filters [2] and memory. Most importantly, this technique is capable of estimating the speed of relatively fast-moving objects, the maximum speed being dependent on the size of the object. Promising results have been obtained through simulated experiments.

1. INTRODUCTION

The processing of image sequences embodying motion has many important applications. Industrial applications include dynamic monitoring of assembly processes, dynamic robot vision, navigation and inspection. Commercial applications include bandwidth compression of picture-phone video signals and HDTV signals. Others include medical, remote sensing, and military applications.

It is well known that the problem of analyzing the motion of natural moving objects in a general dynamic scene is a very difficult one. Correspondingly, there have been limited successes in developing general motion-estimation schemes. Basically, two main approaches to the problem of motion detection and estimation have evolved. They are the optic-flow approach [3], [4], [5], and the feature-based approach [6], [7].

In the optic-flow approach, the instantaneous changes in brightness values in the image are analyzed to yield a velocity map called the optic flow. This technique is based on solving a set of constrained equations relating local spatial and temporal derivatives of a moving object. Alternatively, the idea underlying the feature-based approach is to establish correspondence between points, or sets of points, between frames, and then group the sets into objects based upon similarities of motion.

One of the many difficulties associated with existing approaches to motion estimation is the time constraint placed on the process by most of the applications, for examples, in industrial or robotic processes or in video conferencing and picture phone imaging. In this paper, a new approach, which can be easily implemented in real-time, using relatively straightforward, but high-speed, morphological architectures [2], is proposed. In Section 2, some theoretical background and the general assumptions employed in the proposed scheme are outlined. Section 3 presents the basic algorithm which constitutes the core of an efficient integrated structure for moving-object analysis. Finally, in Section 4, the performance of the proposed scheme is analyzed using some experimental examples.

*A.C.P. Loui is now with Bell Communications Research, New Jersey, U.S.A.

2. MORPHOLOGICAL MOTION ESTIMATION

This new approach to motion analysis is based on the idea of a new shape-representation scheme, the Morphological Autocorrelation Transform or MAT [8], which is obtained from the projected motion characteristics of a stationary object. The MAT is comprised of a family of *Geometrical Correlation Functions* or GCFs which are defined [1] as:

$$K_\phi(h) \triangleq \frac{Mes[X \ominus B_h^\phi]}{Mes[Y]} \quad , \quad 0 \le \phi < \pi \tag{1}$$

where B_h^ϕ is a two-point structuring element, X is the image object under consideration, Y is a pre-defined standard shape (e.g., a square of size 200x200 pixels), and $Mes[*]$ is the Lebesgue measure.

There are basically two parts to the analysis: 1) recognition, and 2) estimation of motion-parameters. For recognition, a single frame from the image sequence can be used as input to the recognition system [1]. As far as the motion-estimation is concerned, the required parameters can be obtained quite conveniently from the MAT representation. In fact, the MAT representation demonstrates the duality between the processes of recognition and motion estimation which is implicit in this approach, but never before properly utilized. The consequence is a very efficient scheme for moving-object analysis, in which both the recognition and the motion-estimation processes are carried out concurrently.

The proposed motion-estimation scheme can be described by an integrated structure (see Section 3) which uses information from the correlation between consecutive frames of the object sequence, as well as that from the pre-computed MAT of the corresponding object. When certain quite reasonable assumptions are satisfied, this approach provides a very efficient way of estimating motion parameters. The major assumptions of this approach are outlined as follows:

1. Sampled GCFs in different directions (ϕ) for each of the M objects are available in a database.
2. The speed of an object visible in a minimum of three object frames is reasonably constant.
3. The inter-frame intervals are known.
4. The motion is reasonably smooth and regular. In particular no composite motion including rotation can be involved. Note that this assumption is a usual one, that for natural dynamic scenes, the direction of motion of an object at any time does not change abruptly [9], [10].

CH 3006-4/91/0000 – 0682 $1.00 © IEEE

5. In situations where the image sequence is obtained from a TV camera, it is assumed that the angle between the optical axis of the TV camera (the line of sight) and the normal to the plane on which motion takes place, is small.

Note that one of the advantages of this approach is that it does not impose any restriction on the shape of the viewed object. In particular, the shape of the unknown object may be convex or nonconvex (that is, containing holes); as well it may be polygonal, curvilinear or circular. However, the algorithm does assume that the moving object under analysis is rigid, that is, it should not incorporate flexible components, which may be moving independently.

Unlike the optical-flow approach, where estimation is carried out locally for every pixel in the scene, the present technique deals with the moving object as a whole. In other words, it can be considered to provide a global estimation. However, it cannot be considered simply as a feature-based technique, one which is based on the establishment of frame-to-frame correspondence through matching of features. Rather, in this technique, the interframe or temporal correlation of the image sequence is exploited more globally.

3. THE BASIC ALGORITHM

In this section, the basic algorithm of the proposed scheme is presented. Figure 1 depicts the block diagram of the proposed integrated moving-object-analysis system based on morphological correlation. There are basically three main components in this system - for preprocessing, segmentation and image understanding. In fact, this system can be considered as an enhancement of the system proposed in [1]. For both systems, the reference GCFs of the objects or shapes are assumed to be stored in the data base during the training phase.

According to Figure 1, the input frame will first pass through the preprocessing module, where the image is enhanced and noise is suppressed. Then the resulting image is passed on to the segmentation module. Segmentation is particularly important for scenes containing multiple moving objects. This can be achieved using techniques like thresholding, region growing, and region splitting and merging. After segmentation, different areas of interest are registered as elements on which analysis may be carried out individually. If this is the first frame, it will be provided to the morphological correlator, where the GCFs are computed. However, if this is not the first frame of a new sequence, the morphological correlator may be by-passed. The reason for this is that for classification, and for the extraction of geometrical information such as perimeter and orientation, a single frame is sufficient.

After the MAT has been computed, there are several options available. If it is necessary to know the identity of an object in the sequence, the path to the classifier can be activated. If knowledge of geometrical properties is needed for subsequent actions, the paths to the area, perimeter, or orientation estimators can be activated. This information, together with that from the motion estimator, are then passed to the data-fusion unit where it will be combined in a pre-defined manner. Finally, the outcome of the data-fusion unit is used to determine the subsequent action.

The procedure for the classification step is the same as presented in [1]. That is, a sub-family of the GCFs is derived from the morphological correlator, and classification is carried out based on this smaller subset of the original family of GCFs. However, for motion estimation, all the GCFs from the morphological correlator are needed. It should also be noted that the accuracy of the motion estimator depends basically on the accuracy of the GCF representation. Therefore, to ensure that the resulting speed and motion-direction estimates are accurate, it is necessary to have an accurate representation for each GCF. This is particularly true for those GCFs representing directions other than 0° or 90° in a rectangular-grid configuration. In the proposed structure, this step will be achieved by means of the interpolation module, which is entered immediately before the motion estimator of Figure 1.

The proposed technique can be used effectively when the basic assumptions described in Section 2 are met. With them, the major steps involved in the motion estimator are as follows:

Step 1) Read in frame #1, and denote it F_1.

Step 2) Using frame data in F_1, obtain the GCF family of the unknown object in F_1. Identify the object if necessary.

Step 3) Read in frame #2 and denote it F_2. Calculate the value of the GCF, $K(h_1)$, based on the intersection between the objects in F_1 and F_2, i.e.,

$$K(h_1) = \frac{Mes\,[F_1 \cap F_2]}{Mes\,[F_1]} \tag{2}$$

Step 4) Look up (from the GCF curves of step 2) the spatial shift h_{1i} for each direction, $i = 0, \cdots, m-1$, where m is the number of discrete directions for which data is stored. Since the inter-frame interval, t_1, between frame #1 and frame #2 is known, the corresponding speed for each direction can be calculated as

$$v_i = \frac{h_{1i}}{t_1} \tag{3}$$

The procedure is illustrated in Figure 2 (a) for the case of m=3.

Step 5) Read in frame #3 and denote it F_3. Calculate the value of the GCF, $K(h_2)$, between frame #1 and frame #3, i.e.,

$$K(h_2) = \frac{Mes\,[F_1 \cap F_3]}{Mes\,[F_1]} \tag{4}$$

Then using the knowledge of the inter-frame interval, t_2, based on the intersection between the objects in frame #1 and frame #3, a set of new spatial shifts, h_{2i}, can be found for the respective directions, i.e.,

$$h_{2i} = v_i \times t_2 \ , \quad i = 0, \cdots, m-1 \tag{5}$$

Step 6) Finally, look up the corresponding GCF for each of these h_{2i}. The one that gives a covariance value closest to $K(h_2)$ is the correct one. The corresponding motion parameters, such as speed and motion direction, are readily available. The details of this process are illustrated in Figure 2 (b) for m=3.

Step 7) From the direction estimates and the slope characteristics, determine the number of ambiguities that arise, then shift frame #1 in each of these directions by the amount h_{2i}. Determine the area of intersection between shifted frame #1 and frame #3, for each of these shifts. The one which has the maximum area is in the correct direction.

The last step of the algorithm is needed to resolve the ambiguity associated with the speed and direction estimates. This is due to the fact that the values of the GCF are the same at angles ϕ and $\phi \pm \pi$. In fact, the degree of this ambiguity is directly related to the symmetry of the shape under study; in particular, the more symmetric is the shape, the higher is the degree of ambiguity, and hence the number of iterations. However, the maximum number of iterations is limited; it is equal to the number of GCFs used in the system. Nevertheless, this step is essential in the sense that it increases the confidence level of the resulting estimation of motion direction. Since the operations required by this step are basic shifting and logical operations, the overhead involved is not very significant. The degradation on speed performance is also not very severe since shifting and logical operations can be easily implemented in real-time. If additional frame buffers are available, this step can be implemented in parallel, allowing the operation time to be greatly reduced.

4. PERFORMANCE EVALUATION

In the previous section, the basic idea of motion estimation based on morphological correlation, was presented. It was shown that some steps of the proposed algorithm can be realized using table-lookups with some associated arithmetic and logical operations. This is the case especially for steps 4) and 6). Note that the lookup table of step 4) and step 6) are inverses of one other. However, due to the discrete representation associated with the MAT and its coordinates, some interpolation may be required to achieve an acceptable accuracy in the estimation processes. In this section, we will examine the performance of the proposed motion-estimation scheme for various shapes and different interpolation schemes.

Two interpolation schemes are considered here. The first scheme interpolates correlation samples by taking the average between samples preceding and following the interpolated one. The second scheme is based on a linear interpolation between samples. That is, the equation of a line segment between two samples is first determined. Then intervening samples are interpolated based on this equation. In all experiments, the test shapes are rotated by 139° and are moving at 24.17 ppf (pixels per frame) in the direction 335.56°. To illustrate the need for the last step, the algorithm was first tested using only the first 6 steps described in the previous section. In this case, it was found that the resulting speed and motion-direction estimation is not acceptable using either interpolation scheme.

We next examined the effect of using the algorithm in which an error feedback and verification procedure (that is, step 7) is incorporated. The results produced using this algorithm with simple interpolation are shown in Table 1. There is a significant improvement to be seen in this case, ten out of the twelve estimates being very close to the correct velocity. The slight discrepancies are due to the finite resolution of the MAT used, (employing only eight directions). Note that the number of verification steps is fairly large in some cases (for example, 8 for the annulus). This might not be desirable, even though the associated overhead is not very large.

As indicated experimently, the performance may be further improved by using the linear interpolation scheme. Table 2 shows the estimates which result using linear interpolation with error feedback and verification. It can be seen that all the estimates in this case are very close to the correct value. In addition, the average number of verification steps is fewer than that of the simple interpolation case. This is attributed to the more precise representation available in the linear interpolation scheme.

In conclusion, it has been found that the verification step is very crucial in ensuring the correctness of the resulting estimates. Furthermore it is apparent that the linear interpolation scheme gives a more accurate estimate of motion parameters than does simple interpolation.

5. SUMMARY

In summary, this paper presents an efficient algorithm for estimating motion parameters of a moving-object. It is shown that when certain reasonable assumptions are satisfied, morphological correlation can provide a very efficient means of characterizing two-dimensional planar motion. Most importantly, this technique is capable of estimating the speed of relatively fast-moving objects, the maximum speed being dependent on the size of the object. This approach is inherently more efficient than other feature-based approaches if frame-to-frame correlation is maintained. Fortunately, this is the situation in most cases except for extremely fast-moving objects. Finally, the performance of the algorithm was evaluated through simulated experiments, with the results obtained being very encouraging. Extension of the algorithm to gray-scale images and non-rigid objects is currently underway.

References

1. A.C.P Loui, A.N. Venetsanopoulos, and K.C. Smith, "Two-dimensional shape representation using morphological correlation functions," *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pp. 2165-2168, Albuquerque, New Mexico, April 3-6, 1990.
2. A.C.P. Loui, A.N. Venetsanopoulos, K.C. Smith, and B. Benhabib, "Non-linear pipeline architectures for morphological signal processing," *Proc. IEEE Int. Symp. on Circuits and Systems*, pp. 1438-1441, New Orleans, Louisiana, May 1-3, 1990.
3. J.O. Limb and J.A. Murphy, "Measuring the speed of moving objects from television signals," *IEEE Trans. on Communications*, vol. COM-23, no. 4, pp. 474-478, April 1975.
4. C. Cafforio and F. Rocca, "Methods for measuring small displacement of television images," *IEEE Trans. on Information Theory*, vol. IT-22, no. 5, pp. 573-579, Sept. 1976.
5. W.B. Thompson and S.T. Barnard, "Lower-level estimation and interpretation of visual motion," *IEEE Computer*, vol. 14, no. 8, pp. 20-28, August 1981.
6. S. Ullman, "Analysis of visual motion by biological and computer systems," *IEEE Computer*, vol. 14, no. 8, pp. 57-69, August 1981.
7. J.K. Aggarwal and N. Nandhakumar, "On the computation of motion from sequences of images - a review," *Proc. of the IEEE*, vol. 76, no. 8, pp. 917-935, August 1988.
8. A.C.P. Loui, A.N. Venetsanopoulos, and K.C. Smith, "Morphological autocorrelation transform: A new representation scheme for two-dimensional images," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Sept 1990. under review
9. M. Jekin, "Tracking three-dimensional moving light displays," *Proc. Workshop on Motion Representation*, pp. 66-70, Toronto, Canada, 1983.
10. I.K. Sethi and R. Jain, "Finding trajectories of feature points in a monocular image sequence," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 56-73, Jan. 1987.

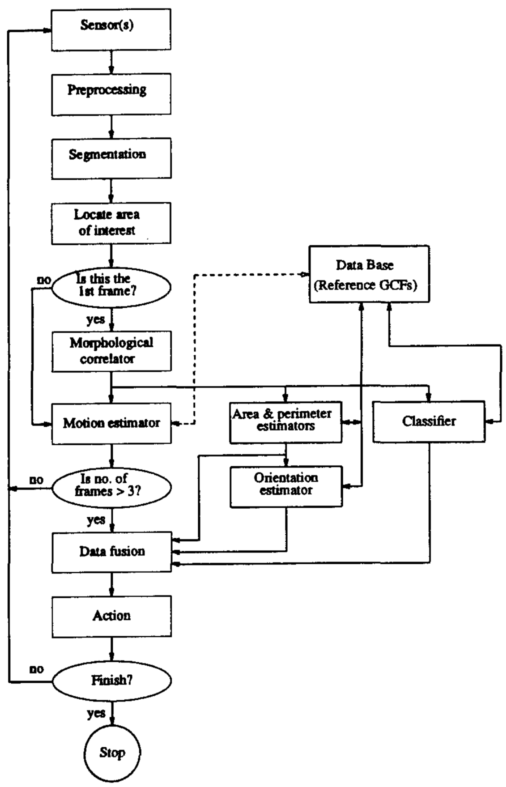

Figure 1 Block diagram of the integrated moving-object-analysis system.

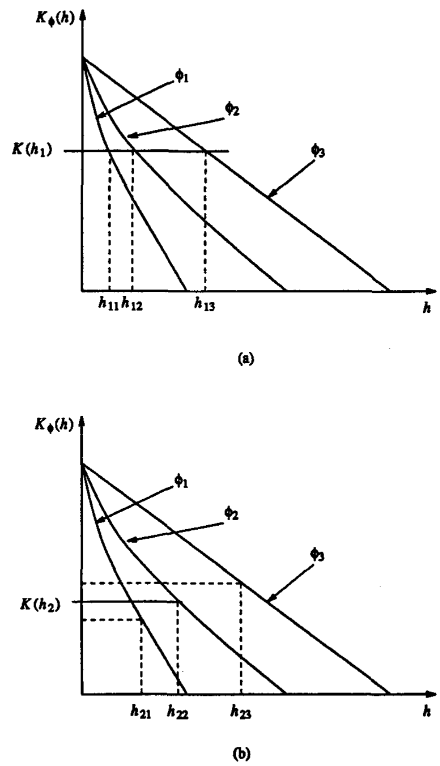

Figure 2 (a) Looking up possible speeds based on $K(h_1)$ (step 4).
(b) Looking up corresponding $K(h_{2i})$ for each h_{2i} (step 6).

shape	simple interpolation		% error	abs. error	number of verification steps
	speed	motion direction	(speed)	(motion direction)	
disc	24.16 ppf	0°	-0.03%	24.44°	2
annulus	23.61 ppf	333.44°	-2.28%	2.12°	8
socket	23.69 ppf	333.44°	-1.99%	2.12°	1
square	24.01 ppf	333.44°	-0.65%	2.12°	4
nut	23.89 ppf	333.44°	-1.16%	2.12°	2
frame	23.99 ppf	333.44°	-0.74%	2.12°	2
ellipse	28.47 ppf	0°	17.79 %	24.44°	2
rectangle	23.61 ppf	333.44°	-2.30%	2.12°	1
triangle	23.40 ppf	333.44°	-3.19%	2.12°	5
tee	23.93 ppf	333.44°	-0.96%	2.12°	1
angle	23.94 ppf	333.44°	-0.94%	2.12°	1
E	26.06 ppf	333.44°	7.83%	2.12°	3

Table 1 Velocity-estimation errors for different shapes (all shapes rotated by 139° and moving at 24.17 ppf in the direction 335.56°), using simple interpolation.

shape	linear interpolation		% error	abs. error	number of verification steps
	speed	motion direction	(speed)	(motion direction)	
disc	24.16 ppf	333.44°	-0.03%	2.12°	1
annulus	24.17 ppf	333.44°	0.00%	2.12°	2
socket	24.25 ppf	333.44°	0.33%	2.12°	1
square	24.61 ppf	333.44°	1.83%	2.12°	1
nut	24.47 ppf	333.44°	1.26%	2.12°	1
frame	24.58 ppf	333.44°	1.73%	2.12°	1
ellipse	23.98 ppf	333.44°	-0.78 %	2.12°	2
rectangle	24.16 ppf	333.44°	-0.01%	2.12°	0
triangle	23.92 ppf	333.44°	-1.00%	2.12°	5
tee	24.52 ppf	333.44°	1.40%	2.12°	0
angle	24.53 ppf	333.44°	1.5%	2.12°	0
E	26.88 ppf	333.44°	11.18%	2.12°	1

Table 2 Velocity-estimation errors for different shapes (all shapes rotated by 139° and moving at 24.17 ppf in the direction 335.56°), using linear interpolation.