Tracking of Moving Objects Using Morphological Segmentation, Statistical Moments and Hough Transform

Muhammad Bilal Ahmad*, Min Hyuk Chang** and Park JongAn**

* Kwangju Institute of Science and Technology, Kwangju, Korea
(Tel : +82-62-970-2419; E-mail: bilal@kjist.ac.kr)
** Electronics Information & Communications Engineering, Chosun University, Kwangju, Korea
(Tel : +82-62-230-7064; E-mail: japark@chosun.ac.kr)

Abstract: This paper describes real time object tracking of 3D objects in 2D image sequences. The moving objects are segmented from the image sequence using morphological operations. The moving objects are segmented by the method of differential image followed by the process of morphological dilation. The moving objects are recognized and tracked using statistical moments. The direction of moving objects are determined by the Hough transform. The straight lines in the moving objects are found with the help of Hough transform. The direction of the moving object is calculated from the orientation of the straight lines in the direction of the principal axes of the moving objects. The direction of the moving object and the displacement of the object in the image sequence is used to calculate the velocity of the moving objects. The simulation results of the proposed method are promising on the test images.

Keywords: Object tracking, Morphological Segmentation, Statistical descriptors, Hough transform

1. INTRODUCTION

Segmentation and tracking semantic objects in video are essential tasks for the most content-based digital video applications. Object tracking is an important problem in the field of content-based video processing. When a physical object appears in several consecutive frames, it is necessary to identify its appearances in different frames for purposes of processing. Object tracking attempts to locate, in successive frames, all objects that appear in the current frame. The most straightforward approach to this task is to consider objects as rectangular blocks and use traditional block matching algorithms [1]. However, since objects may have irregular shapes and deformations in different frames, video spatial segmentation and object temporal tracking can be combined [2]-[3].

In object tracking, pattern recognition is to deal with the geophysical data based on the information contained in the image sequences. An automatic interpretation or recognition of geophysical data is very difficult from the image sequences. A lot of efforts have been found in the literature, and still a lot of research is needed for automatic recognition of moving objects in the image sequences. Most methods of object tracking such as optical flow, block matching , etc are highly computational and hence difficult to apply in the run time applications. For object tracking, we are mostly interested in the rigid body motion. Every moving object, say, vehicles, walking human being, etc are all consists of rigid body motion. We can exploit this rigid body motion for tracking the objects.

In this paper, we have proposed an effective moving objects tracking based on the orientation of the moving objects. Moving objects locations are found in the image sequence by the method of differential edge image followed by morphological dilation [6,11]. After locating the moving objects in the image sequences, we extract different high-level features directly from the regions of pixels in the images, and describe them by various statistical measures. Such measures are usually represented by a single value. Measurements of area, length, perimeter, elongation, compactness, moments of inertia are usually called statistical geometrical descriptions [3,4]. We use the statistical geometrical descriptions to recognize the moving objects in the image sequences. The principal axes of inertia for the moving objects in the image sequences is used for extracting the direction of the moving objects. The straight lines in the moving objects are determined by the Hough transform. The straight lines in the moving objects that are almost aligned with the principal axes are averaged to find the direction of the moving objects. We track the moving objects using probabilistic approaches. We assumed that the different moving objects do not have common statistical descriptors. We also assumed that the velocity of the moving objects is not too high, and we restrict the search area for tracking of the individual moving objects within the most probable range. The former assumption is useful for tracking dissimilar statistical or geometrical moving objects. From the later assumption, we can also track similar statistical or geometrical objects, such as many similar moving balls in the image sequence.

The simulation results of our method of tracking and recognizing the moving objects using statistical descriptors, and Hough transform are very promising on the test images. The execution speed of our method is very fast and the results are very good.

This paper is organized as follows. Section 2 describes the segmentation of the moving objects using differential edge images followed by the process of morphological dilation. Section 3 describes the different statistical descriptors that will be utilized for tracking and recognizing the objects. Section 4 explains the Hough transform to find the direction of the moving objects. Simulation results are shown in section 5. At the end, we will conclude our paper with few final remarks.

2. SEGMENTATION OF MOVING OBJECTS

We first segment the moving objects in the input image sequence. Sobel edge detector is applied on the two input image sequence. For removing the background (still part) in the images, we find the binary difference image from the resulting two input edge maps, as:

$$D(x, y) = ABS(E2(x,y) - E1(x,y)), \quad (1)$$

where E2(x,y), E1(x,y) are the two binary edge maps of the input image sequence, and D(x,y) is the resulting binary difference image. The resulting binary difference image D(x,y) gives us the possible location of moving objects. To find the areas of moving objects, we binary dilate the difference image D(x,y) as:

$$DL = dilate(D), \quad (2)$$

where DL is the dilated image of the difference binary image D. The dilated image DL detects the areas of moving objects in the image sequence. The dilation of the difference image should be enough such that the areas of the moving objects detected in the dilated image must be greater than the actual areas of the moving objects. In the dilated image DL, all possible moving objects (both real and erroneous moving objects) are detected. The erroneous moving objects are detected due to the presence of noise in the images. We applied a thresholding method to extract the real moving objects from the dilated image DL. We first label the moving objects in the dilated image DL and then calculate the binary areas of each of the moving objects. We threshold the real moving objects that have considerable area in the dilated image as:

if $A[DL(j)] > T_{area}$
    Real Moving Object (keep it)
else
    Erroneous Moving Object (discard it) (3)

where A[DL(j)] calculates the binary area (count the number of 1s of the labeled object) of j^{th} labeled object in DL, and T_{area} is the threshold, the value of which depends on the size of input images, and the distance of camera from the scene. We discard the erroneous moving objects by replacing 1s with 0s in that area. Finally, we get the image, which contains only real moving objects in the image sequence. We then calculate the statistical descriptors in those actual moving areas.

3. OBJECT TRACKING USING STATISTICAL DESCRIPTORS

After segmenting the moving objects from the input image sequence, a matching algorithm is needed between the regions in the two consecutive images for tracking and recognizing the moving objects. A region matching or similarity is obtained by comparing the statistical descriptors of the two regions. Since the images may have translational, rotational, and scaling differences (objects may move further or closer to the camera), the region or shape measures should be invariant with respect to translational, rotation and scaling. One kind of such invariants belongs to statistical moments, called statistical invariant descriptors. We use statistical invariant descriptors for tracking of the moving objects. We will first briefly explain the statistical moments and invariant descriptors, and then explain the utilization of these statistical descriptors for tracking the moving objects.

3.1 Statistical Moments and Invariant Descriptors

The moment invariants are moment-based descriptors of planar shapes, which are invariant under general translation, rotational and scaling transformations. Such statistical moments work directly with regions of pixels in the image using statistical measures. Such measures are usually represented by a single value. These can be calculated as a simple by-product of the segmentation procedures. Such statistical descriptors usually find area, length, perimeter, elongation, Moments of Inertia, etc. The moments of a binary image b(x, y) are calculated as:

$$\mu_{pq} = \sum_x \sum_y b(x,y) x^p y^q, \quad (4)$$

where p and q define the order of moment. Where *b(x,y)* can be omitted as it has only 1 and 0 values, so sums are only taken where *b(x,y)* has values 1. The center of gravity of the object can be found from moments as:

$$\bar{x} = \frac{\mu_{10}}{\mu_{00}}, \quad \bar{y} = \frac{\mu_{01}}{\mu_{00}}, \quad (5)$$

where ($\bar{x}$, $\bar{y}$) are the coordinates of the center of gravity. The pq^{th} discrete central moment m_{pq} of a region is defined by

$$m_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q \quad (6)$$

where the sums are taken over all points (x,y). Hu [6] proposed seven new moments from the central moments that are invariant to changes of position, scale and orientation of the object represented by the region using central moments of lower orders. All the seven moments are translational, rotational and scale invariants. These invariants will help us in the object tracking of the moving objects.

The principal axes of inertia define a natural coordinate system for a region. Let θ be the angle that the x-axis of the natural coordinate system (the pricicpal axes) makes with the x-axis of the reference coordinate system. Then θ is given by

$$\theta = \frac{1}{2} \tan^{-1}\left[\frac{2m_{11}}{m_{20} - m_{02}}\right] \quad (7)$$

From the principal axes of inertia, we can find the direction of the moving objects.

3.2 Tracking of Moving Objects

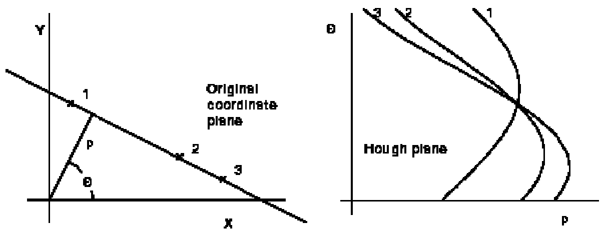
For tracking of moving objects, the seven statistical descriptors are calculated for the detected moving regions of the input image sequence. There are translation and rotation of moving objects due to motion from one image frame to another image frame, and also the object can move far or closeer from the camera, which results in the different size of the object in terms of pixels for the fixed camera position. The next step is the comparison of the statistical descriptors in the two images. Here we have assumed that either the motion of the objects are very small, or the frame rate is very high, so that we can restrict the search area for tracking of the individual moving objects within the most probable range. With the help of the statistical descriptors, we recognize and track different kind of moving objects. We found the statistical invariant descriptors for every detected moving region in the two images, and then track the moving objects within the search region by comparing the statistical descriptors.

4. FINDING VELOCITY VECTORS FOR THE MOVING OBJECTS USING HOUGH TRANSFORM

After tracking the moving objects in the input image sequence, we determine the principle axes using (7) for each of the segmented moving objects. The principal axes do not give the true direction of the moving object, because of 2D image representation of 3D objects. However, the principal axes give the rough estimate of the direction of the moving objects. To find the true direction, we need to determine the straight lines in the object. The Hough transform is used to find the straight lines in the moving objects.

4.1 Straight lines using the Hough transform

Hough transform can be efficiently used to search the straight lines in the images. Let us consider the line in an image in its parametric form $\rho = x\cos\theta + y\sin\theta$, where ρ is the perpendicular distance from the origin and θ the angle with the normal in figure 1 (a). Collinear points (x_i,y_i), with i = 1,…,N, are transformed into N sinusoidal curves $\rho = x_i \cos\theta + y_i \sin\theta$ in the (ρ,θ) plane, which intersect in the point (ρ,θ) in figure 1 (b).



(a)

(b)

Figure 1. (a) A straight line in xy-plane (b) the transformed (ρ,θ) domain.

In an image analysis context, the coordinates of the point(s) of edge segments i.e. (x_i,y_i) in the image are known and therefore serve as constants in the parametric line equation, while ρ and θ are the unknown variables we seek. If we plot the possible (ρ,θ) values defined by each (x_i,y_i), points in cartesian image space map to curves (i.e. sinusoids) in the polar Hough parameter space. This point-to-curve transformation is the Hough transformation for straight lines. When viewed in Hough parameter space, points which are collinear in the cartesian image space become readily apparent as they yield curves which intersect at a common (ρ,θ) point.

The transform is implemented by quantizing the Hough parameter space into finite intervals or accumulator cells. (i.e. a multidimensional array). As the algorithm runs, each (x_i,y_i) is transformed into a discretized (ρ,θ) curve and the accumulator cells which lie along this curve are incremented. Peaks in the accumulator array represent strong evidence that a corresponding straight line exists in the image. All pixels which lie on the same line in (x,y) space are represented by lines which all pass through a single point in (ρ, θ) space.

The single point through which they all pass gives the values of ρ and θ in the equation of the line $\rho = x\cos\theta + y\sin\theta$.

To detect straight lines in an image, we do the following procedure:

- Edge image of the input image is found.
- Quantise (ρ,θ) space into a two-dimensional array A for appropriate steps of ρ and θ.
- Initialise all elements of A(ρ, θ) to zero.
- For each edge pixel (x',y') which lies on some edge in the image, we add 1 to all elements of A(ρ, θ) whose indices ρ and θ satisfy $\rho = x'\cos\theta + y'\sin\theta$.
- Search for elements of A(ρ, θ) which have large values. Each one found corresponds to a line in the original image.
- These large values are transformed back using inverse Hough transform. Now the parameter space (ρ, θ) is converted into spatial space (x,y).

4.2 Object Orientation

We determined all the straight lines using the Hough transform for the every tracked object in the image sequence. The orientation of the moving object is determined from the straight lines and the principal axes of the object as shown in figure 2. The x-axis of the principal axes is selected as the reference axis. The straight lines that make a greater angle than the threshold angle are discarded. The angles that the remaining straight lines in the object make with the principal axes are averaged. The average angle thus determined is the true orientation of the 3D moving objects.

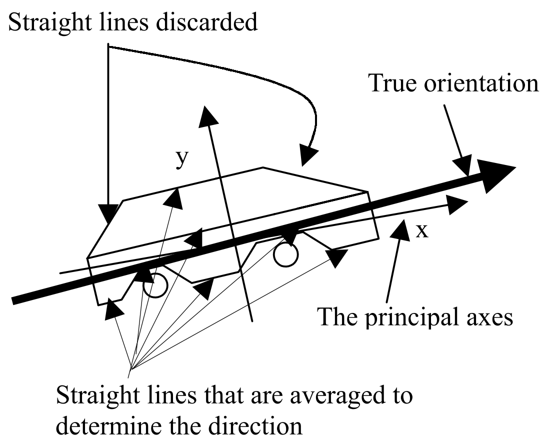


Figure 2. Determining the object direction

The direction of moving object is found from the law of cosines from the orientation angles of the individual moving object in the two consecutive images. From Figure 3, we can find the direction of the moving object. In Figure 3, let L1 and L2 be the two lines making angle θ_1 and θ_2 with respect to x-axis of the reference frame, respectively. L1 and L2 correspond to the true orientation of the moving object as determined in figure 2. The mathematical derivation for the moving object direction θ_3 with respect to x can be derived as:

For L1 : $y = m_1 x + c_1$

For L2 : $y = m_2 x + c_2$

By solving the above equations, the intersection point of L1 and L2 can be found as:

$$x_{\text{int}} = \frac{c_2 - c_1}{m_1 - m_2},$$

$$y_{\text{int}} = m_1 x_{\text{int}} + c_1$$

The origin in Figure 3 is the center of gravity of the object in image 1. From law of cosines

$$l_3^2 = l_1^2 + l_2^2 - 2 l_1 l_2 \cos(\pi + \theta_1 - \theta_2)$$

and

$$\cos(\theta_3 - \theta_1) = \frac{l_1^2 + l_3^2 - l_2^2}{2 l_1^2 l_3^2}$$

The angle θ_3 gives the direction of the moving object. The small l_1, l_2, l_3, are the magnitudes of L1, L2 and L3 lines. For calculating the magnitude of the velocity vector, the Euclidean distance of the two centers of gravity (centers of gravity marked as stars in figure 2) is measured. From the angle θ_3, and Euclidean distance of the centers of gravity, we calculate the velocity vectors of the moving objects. Same method is applied for extracting the velocity vectors of each individual moving objects.

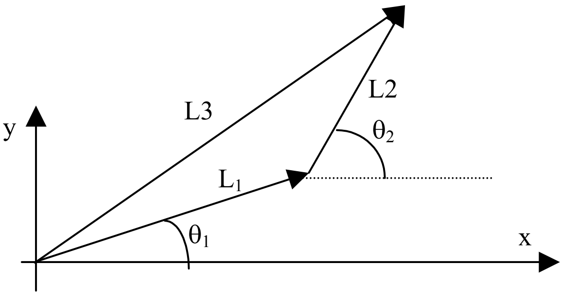


Fig. 3 Determining the direction of the moving object

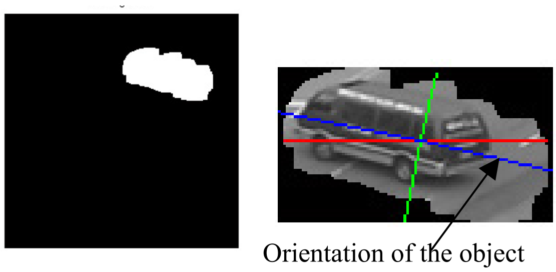
5. SIMULATION RESULTS

For simulation, 256 x 256 gray-level image sequences are used. One test sequence is shown in figure 4. First we segment the moving objects from the input image sequence using the proposed differential edge algorithm. The mask obtained in the figure 5(a) is used to segment the moving objects from the input image sequence as shown in figure 5(b). The statistical descriptors are calculated for the segmented moving regions only. The moving objects are recognized using the similarity of statistical descriptors. The direction of the moving object is determined using the Hough transform and the principal axes. The principal axes as shown in figure 5(b) doesn't give the right direction of the 3D object, whereas the direction obtained by using the Hough transform represents more accurate direction of the moving object. Figure (6) shows the tracking result of three test image sequences. The three test moving objects are accurately tracked in the image sequences.

6. CONCLUSIONS

In this paper, a new algorithm is proposed for segmenting, recognizing, tracking and finding the velocity vectors for moving objects in a video stream. There are many popular techniques for finding velocity vectors, such as optical flow, and block matching algorithm, but they are time-consuming algorithms. Our method is computationally fast and gives compact information about the moving objects. From the input video stream, we segment the moving objects using the edge differential algorithm. For tracking of the moving objects, we proposed method based on the statistical invariant moments or descriptors, which are invariant to translation, rotation and scaling transformation.

Figure 4. A test sequence



(a) (b)

Figure 5. (a) The mask for segmenting the moving objects (b) the direction and the principal axes of the object.

Figure (6) Object tracking on three test image sequences

We use seven different descriptors based on the lower order moments for tracking of the moving objects. After tracking, we found the orientation of the moving objects using the principal axes of inertia and the Hough transform. From the knowledge of the orientation of the moving object in the consecutive image frames, we found the direction of the moving objects. From the displacement of the center of gravity, we found the Euclidean distance of the moving objects. The final velocity vector for a moving object is calculated from the orientation angles, and the Euclidean distance of the centers of gravity of the object. The simulation results show that the proposed gives accurate results. The process of edge detection and segmentation accurately find the location and areas of the real moving objects, and hence the extractions of moving information are very easy and accurate. The statistical invariant descriptors are very accurate for tracking different kind of objects. The orientation of the objects is more accurately determined from the Hough transform. The idea of extracting velocity vectors using the orientation and distance information works well for the simultaneous translation and rotation transformations.

REFERENCES

[1] A.M. Tekalp, Digital Video Processing, Parentice Hall, 1995.

[2] R.C. Gonzalez and R. E. Woods, Digital Image Processing, Addison-Wesley, 1993Allen

[3] Berthold Klaus Paul Horn "Robot Vision" The MIT Electrical Engineering and Computer Science Series.

[4] Robert M. Haralick, Linda G. Shapiro "Computer and Robot Vision", vol. 1.

[5] B. K. P. Horn and B. G. Schunck, 1981, "Determining optical flow", Artif. Intell. 17, pp.185~203.

[6] M. K. Hu, "Visual pattern recognition by moment invariants", IEEE Trans. Inform. Theory, vol. IT-8, no. 2, pp. 179-187, 1962

[7] N. Diehl, "Object Oriented Motion Estimation and Segmentation in Image Sequences", Signal Processing: Image Communication, Vol. 3, No. 1, pp. 23-56, Feb. 1991.

[8] P. J. Burt, et al., 1989, "Object tracking with a moving camera", Proc. IEEE Workshop on Visual Motion, Irvine, California, U.S.A., pp.2~12

[9] C. Cafforio and F. Rocca, 1979, "Tracking moving objects in television images", Signal Process. 1, pp.133~140

[10] W. B. Thompson, 1980, "Combining motion and contrast for segmentation", IEEE Trans. PAMI, 2(6), pp.543~549

[11] M. Etoh et. al., 1993, "Segmentation and 2D motion estimate by region fragments", Proc. 4th Int. Conf. Computer Vision, pp.192~199

[12] Yvan Petillot, Ioseba Tena Ruiz, and David M. Lane, "Underwater vehicle obstacle avoidance and path planning using a multi-beam forward looking sonar", IEEE Journal of Oceanic Engineering, Vol. 26, No. 2, April 2001.

[13] Alexander Kirchner, and Christian Ameling, "Integrated obstacle and road tracking using a laser scanner", Proc. of the IEEE Intelligent Vehicles Symposium 2000, Dearborn (MI), USA October 3-5, 2000.

[14] Chao He, Yuan F. Zheng, and Stanley C. Ahalt, "Object tracking using the Gabor wavelet transform and the golden section algorithm", IEEE transactions on multimedia, vol. 4, No. 4, December 2002.

[15] R. C. Baker and B. Charlie, "Nonlinear unstable systems," International Journal of Control, Vol. 23, No. 4, pp. 123-145, 1989.

[16] G.-D. Hong, "Linear controllable systems," Nature, Vol. 135, pp. 18-27, 1990.

[17] K. S. Hong and C. S. Kim, "Linear stable systems," IEEE Trans. on Automatic Control, Vol. 33, No. 3, pp. 1234-1245, 1993.

[18] Z. Shiler, S. Filter, and S. Dubowski, "Time optimal paths and acceleration lines of robotic manipulators," Proc. of the 26th Conf. on Decision and Control, pp. 98-99, 1987.

[19] M. Young, The Technical Writer's Handbook, Mill Valley, Seoul, 1989.