t

SEGMENTATION OF NON-RIGID BODIES IN AFFINE MOTION : A NEW FRAMEWORK

Sangeeta Narang*, Naresh Kumar Narang** and Prof. K.K.Biswas***

*IBM Solutions Research Center, New Delhi, India (snarng@in.ibm.com)
**Astt. Director(Engg.) All India Radio, New Delhi, India
***Dept. of Computer Science and Eng., Indian Institute Of Technology, Hauz Khas, New Delhi, India.

ABSTRACT

In this paper a new technique for partitioning a non-rigid body (human body), when it is in affine motion is presented. The technique is based on finding the contour in the "*modified difference frame*". These contours can then be used in identifying the regions of interest. Results show that the technique eliminates a lot of stationary regions and thus can reduce the amount of computational time required for the processing of the image. It reduces significantly the total bit rate required for transmission. It also provides an algorithm for contour tracing using extended boundary concept and contour compression by contour merging.

1. INTRODUCTION

Segmentation-based motion-compensated techniques are considered to be very promising for video compression. These techniques start by segmenting an image into homogeneous regions then performing motion estimation and compensation on the basis of region. The contours in the segmentation and the motion parameters, if any, have to be transmitted to the receiver, producing fewer motion parameters, and no block artifacts are created in reconstructed images. However, significant number of bits are required to encode contours of segmentation [13]. Present technique tries to remove these constraints.

1.1. Image Segmentation

Segmentation is a process to partition the picture into meaningful parts. It is always desirable to develop segmentation algorithms that do not require apriori information. Motion can reveal the information that can not be obtained solely from the analysis of one static scene. Intuitively two pictures of the same scene, taken under different conditions, give more information. The main goal of this process is to divide an image into parts that have a strong correlation with objects or areas of the real world contained in the image. Complete segmentation results in a set of disjoint regions uniquely corresponding with objects in the input image, whereas, partial segmentation results in he regions which do not correspond directly with image objects. Contour tracing is one of the fundamental methods used in shape analysis [4]. In literature the contour of a region R is usually referred to the set of pixels in R that have at least one neighbor which does not belong to R. Thus contour tracing is a technique used to determine the sequence order among boundary pixels based on a clockwise or counterclockwise traversal.

Sangeeta Narang is with IBM Solutions Research Center, New Delhi, IndiIa. She worked on this paper while with Delhi University

1.2. Segmentation using motion information

If the stationary foreground and background regions are not of interest, the stationary regions of the scene image need not be analyzed and attention can be focused in areas where motion is detected [4]. This results in reduction in the computational time. Moreover, the analysis of a sequence of pictures makes it possible to accumulate information (history) which aids in the interpretation process. Many techniques have been proposed to segment the moving objects based on motion information. Human body segmentation is the process to segment a human body from the stationary background. The inputs are the previous and current frame of a sequence of pictures. The output is a segmented picture with most of the stationary background eliminated [4].

In the present paper, *a new method to segment an image when it is in affine motion* is presented. This algorithm is based on finding the contour in the "*modified difference frame*". These segmented frames help transform a sequence of images into a form which requires limited information for storage, transmission and to reconstruct the images at the other end of the multimedia network. The present paper focusses on the reduction of number of regions, contour coding and DFD compression.

2. NEW APPROACH FOR IMAGE SEGMENTATION

Video transmission requires large bandwidth, so it is desirable to send limited information. To obtain the limited information the segmentation of the images has to be carried out to reduce the amount of information to be transmitted. We have used "*modified difference image* *technique*" to get motion information from the video frames and estimate the motion partially on the region where motion feature is detected. If the human body does not move, the number of bits required for transmission reduces significantly. Total number of bits required by our algorithm is given by :

$$N = (bits(DFD(\overrightarrow{v_1}), Q_1) + bits(\overrightarrow{v_1}) + bits(DFD(\overrightarrow{v_2}), Q_2) + bits(\triangle\overrightarrow{v_2}) + ... + bits((DFD(\overrightarrow{v_n}), Q_n) + bits(\triangle\overrightarrow{v_n})$$

where $\overrightarrow{v_i}$ is the motion vector of frame I, $\overrightarrow{\triangle v_i} = \overrightarrow{v_i} - \overrightarrow{v_{i-1}}$, $DFD_i(\overrightarrow{v_i})$ represents the DFD of block i, and $bits(DFD_i(\overrightarrow{v_i}), Q_i)$ is the number of bits required for this frame difference. When the DFD grows, $\overrightarrow{\triangle v_i}$ also grows up since $bits(DFD_i(\overrightarrow{v_i}), Q_i)$ and $bits(\triangle\overrightarrow{v_i})$ also grows. We assume that the object in view is undergoing affine(non-rigid) motion. In what follows we define our new approach for segmenting objects in affine motion. This can be applied to detect the motion of human body parts like the arm movement, the head movement, and the body movement as a whole.

2.1 Motion-preserving simplification

Generally, filters degrade and sometimes eliminate moving components since moving components appear as "small components" on the time axis [13]. Motion-preserving simplification detects and restores moving components while removing still components. We have used thresholding so as to remove very low and very high frequencies.

2.2. Image segmentation

The image segmentation algorithm presented in the present paper assumes that there is a small motion. It involves applying the following steps:

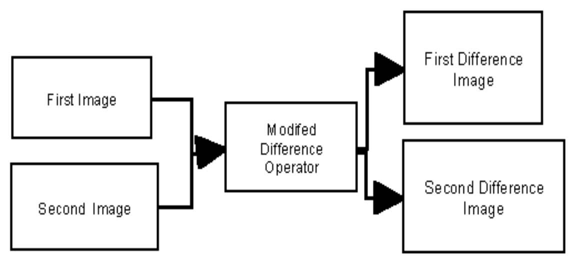
- modified difference operator
- contour tracing
- region/contour merging

2.2.1. *Extraction of motion information with modified difference operator*

Let f1(r,t-1) and f1(r,t) denote the two successive frames at time t-1 and t in an image sequence. A part of a moving component can be obtained by |f1(r,t)-f1(r,t-1)|>0 if there is no noise. In 2-D space this part, usually appears as a linear structure and belongs to the edge of the moving component. From this part, the entire moving component can be recovered. In practice, the natural images are always corrupted by noise and their average gray values may change from one frame to another because of luminous disturbance. Hence |f1(r,t)-f1(r,t-1)| is not always equal to zero in the regions of still components. This can be reduced by thresholding the difference. The luminous disturbance is eliminated by thresholding f(x,t) at threshold T. To recover the entire moving component, we define two images : first difference image $f_1^{'}(r,t)$, which contains those gray values of the previous image which are above threshold except those coordinates where either the gray values in the two input frames are within a specified limit or the gray values in the current frame are a part of the information $(f_1(r,t) \cap f_2(r,t))$. The second image $f_2^{'}(r,t)$ contains all the information of the *current frame* which might have changed from the previous frame, and is required during the reconstruction process except $f_1(r,t) \cap f_2(r,t)$[7]. Fig.1 shows the process.

Fig. 1 : The "modified difference Operator"

These can be represented as :



$$f_1^{'}(r,t) = f_1(r,t) - (f_1(r,t) \cap f_2(r,t)), if\, f_1(r,t) > T$$
$$= f_{1min}, \qquad otherwise$$
$$f_2^{'}(r,t) = f_2(r,t) - (f_1(r,t) \cap f_2(r,t)), if\, f_2(r,t) > T$$
$$= f_{2min}, \qquad otherwise$$

where f1min and f2min are, respectively, the minimum of gray levels for f1 and f2. This recovers all the moving components as well as the still components which are darker than their surroundings. It removes all the still components which are brighter than their surroundings.

2.2.2. *Contour tracing*

The present algorithm is based on a modified/extended contour concept. The extended contour of a region R consists of pixels that belong to R along some directions of traversal (see figure 3). It is a form to represent the shape of the inter-pixel boundary on the original image grid. Determination of the sequence order is not straightforward. A traditional contour tracer will search through its eight neighboring pixels based on some order associated with the direction of its previous step in locating the direction of its next step. In the extended contour, the search order does not exist and a pixel may be surrounded by 8 other pixels that are also elements of the same contour. We will use P0(P), P1(P), ... P7(P) respectively to represent the eight neighbors of a point P at (i,j), as in Fig.2.

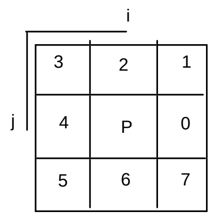
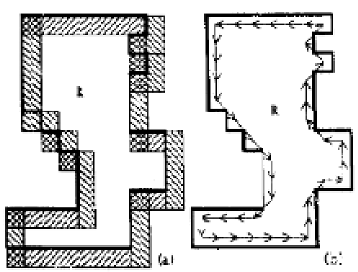


Fig 2: The direction codes used to represent 8 neighbours

Fig.3(a) shows the extended boundary for a simple example where the bold line is the inter-pixel boundary, the hatching \\\ denotes the contribution of Left and Right points, and the hatching /// denotes the contribution of



Upper and Lower points. Fig.5(b) is the corresponding traced contour by a traditional contour tracer

The following boundary definitions are used to derive the extended boundary for R:

1. Left point if $P \in R \text{ and } P_4(P) \in Q \text{ and } Q \neq R$
2. Right point if $P \in R \text{and } P_0(P) \in Q \text{ and } Q \neq R$
3. Upper point if $P \in R \text{and } P_2(P) \in Q \text{ and } Q \neq R$
4. Lower point if $P \in R \text{ and } P_6(P) \in Q \text{ and } Q \neq R$

Now, the extended boundary of R is defined as the union of the following sets :

1. $\{P \mid P \in Left(R) \text{ or } Upper(R)\}$
2. $\{P_6(P) \mid P \in Lower(R) \text{ or } Left(R)\}$
3. $\{P_0(P), P_7(P) \mid P \in Right(R)\}$

The search sequence used in a usual contour tracer does not exist in the extended boundary, so a method/rule is needed to determine the pixel sequence between starting and ending points of a boundary segment. We trace regions in a counterclockwise direction to get the contour sequence, using the direction codes as in fig. 2 and the defintion of the extended boundary. For this algorithm, given the first pixel in the extended boundary, it will traverse the rest of the extended boundary automatically.

2.2.3. Contour merging

The processes of contour tracing produces an over segmentation of the difference image. The step of contour merging aims to merge the homogeneous contours into one contour by boundary elimination. If the average pixel-distance across the boundary between two contours is lower than T/2, then boundary is eliminated and both the contours are merged as one region. Contour merging is carried out by successively removing the boundary with the lowest distance between contours. After contour merging the number of contours decrease by approximately half.

The *chain code technique* has been used for contour coding. It describes the contours by relative movements : turn right, turn left, go straight on a hexagonal lattice of 3x3 / 5x5. This is used to locate the starting point of the new contour. The resulting chain code can be further compressed by arithmetic coding, say of order two or using segmentation sampling to reduce the contour length.

We conducted the experiments on the real video frames. The results show that our method using the current segmentation technique can keep tracking the motion of the human body.

The images used for this project were captured using the video recording system and multimedia systems video capture program at IIT Delhi. Video sequences were first recorded on a VHS tape in the studio. Then images were captured frame by frame one at a time.

3. RESULTS

The algorithm as described was applied to both the rigid and non-rigid bodies with encouraging and consistent results. Figure 4 shows the original images, their difference images and contour traced images for a particular hand motion sequence of frames. Many different sequences were analyzed giving the average compression ratios for Non-rigid Bodies between 3-10 (ie., a 66%-90% compression) and for *Rigid Body* the

Fig. 4 : Results of the segmentation process for a hand motion sequence. The figure shows respectively binary first frame, binary second frame, first difference image, second difference image and the border traced image

value is approximately 3 (i.e., 66% compression), where *Compression ratio* is defined as:

$$\left(\frac{\textit{Number of pixels of information required to be sent}}{\textit{Number of pixels of information actually sent}}\right).$$

4. SUMMARY AND CONCLUSIONS

In this paper a technique has been developed to segment non-rigid bodies in motion. It is interesting in the sense that it is not restricted to one type of motion or posture as is the case in [1]. The results show that use of *two difference image technique* removes the need of the original image at the receiving end except for the first frame of the sequence, without loosing any motion information. The algorithm for image segmentation are inexpensive and accurate. Also the contour tracing using contour merging and motion preserving simplifications reduces the number of contours in the image significantly, thus reducing the number of bits required for transmission to some extent with only a small segmentation compensation errors. The results also show that using the above segmentation technique, as we increase the gray value tolerance slightly for the difference image the compression ratio increases upto a maximum and after that there is no change but there is a slight degradation of the reconstructed image. The algorithm can detect the motion of the background by the moving background also and can be used for differentiating between different objects moving with different velocities. The number of objects moving in the scene is not assumed since we have considered only the motion between frames and not which object has moved. The method is practical since it does not require computation of point correspondences, velocity field and optical flow etc.

5. REFERENCES

[1] Bruckstein A.M., Yanowitz S.D., “A new method for image segmentation”, IEEE Trans. on CVGIP, vol. 46, 1989.

[2] Leung M.K. and Yang Y.H., "Human body motion segmentation in a complex scene." Pattern recognition, Vol. 20 No. 1 pp 55-64 1987.

[3] Liow Yuh-Tay, "A Contour Tracing Algorithm that Preserves Common boundaries between regions.", CVGIP: Image Understanding, Vol. 53, No. 3, May, pp. 313-321, 1991.

[4] Mardia Kanti V., Qian Wei, Kevin M.A. de Souza, “Deformable Template Recognition of Multiple Occluded Objects”, IEEE Trans. on Pattern Anal. and Mach. Intellligence, vol 19, No. 7, Sept. 1997.

[5] Narang N. K., “Tracking non-rigid objects in Image sequences”,M.Tech. thesis, August 1995.

[6] Narang Sangeeta, Narang N. K., Biswas K.K., “Tracking non-rigid objects real world human motion”, submitted to 6th Int. Conf. on Advanced Computing to be held at Pune on Dec. 14-16, 1998.

[7] Rangarajan K. and Shah Mubarak , "Establishing motion correspondence", CVGIP:Image Understanding, Vol. 54 No. 1 July, pp 56-73, 1991.

[8] Sabata Bikash and Aggarwal J. K., " Estimation of Motion from a pair of Range Images: A review.", CVGIP: Image Understanding, Vol. 54, No. 3, November, pp. 309-324, 1991.

[9] Sinclair D., Blake A., Beardsley P., Murray David, "A novel approach to motion segmentation.", Proc, of the British Machine Vision Conference, Glasgow, 1991, pp. 375-378 (Springer Verlag, London 1991).

[10] Wang Demin, Labit Claude, Ronsin Jseph, “Segmentation Based Motion Compensated Video Coding Using Morphological Filters”, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 7, No. 3, June 1997.

[11] Wren Christopher Richard, Azarbayejani Ali, Darrell Trevor, Pentland Alex Paul, “Pfinder : Real-Time Tracking of Human Body”, IEEE Trans. on PAMI, Vol. 19, No. 7, July 1997.

[12] Yen-Kuang Hen, Kung S.Y., “Rate optimization by true motion estimation”, IEEE Workshop on Multimedia Signal Processing, Princeton, NJ, Jun.1997.

[13] Yoshinari Kameda, Michihiko Minoh, Katsuo Ikeda, “Three Dimensional Motion estimation of a Human Body Using a Difference Image”, Asian Conference on Computer Vision, Vol II, 1995.