Lausanne, Switzerland March. 93 PCS-93 pp. 10.1 - 10.2

Global Motion Estimation in Image-Sequences of 3-D Scenes for Coding Applications [1]

Amichay Amitay and David Malah

Department of Electrical Engineering
Technion - Israel Institute for Technology
Technion city, Haifa 32000, Israel

Techniques for *global* motion estimation in image sequences are of great interest for image-sequence coding. The better the motion compensation, the better is the prediction of the picture to be coded, thus bringing about a reduction in bit rate and/or improved quality. This reduction is due to the smaller prediction error as well as to the reduced motion-vector information. Current coding standards apply only *local* motion compensation, assuming that blocks in the image have translatory motion only. However, in 3-D scenes this is usually an inadequate assumption, even when the camera is just panned, since the motion magnitude may vary gradually from pixel to pixel in the block depending on the distance to the camera. If the camera is zoomed this is the case for 2D scenes as well.

Most reported methods attempt to find a single set of parameters describing the global motion [1]. This gives a good description when 2-D scenes are involved. For 3-D scenes, a single set of motion parameters may not adequately describe the global motion if pixels in the picture are not at the same distance from the camera. Furthermore, large moving objects reduce the accuracy of the estimated parameters. Hough-based [1] methods are more robust when there is more than one dominant motion in the picture. Still, in [1], only the motion of the largest moving region is found and the full description of the motion in the picture is not always obtained.

For general 3-D scenes segmentation-based methods are therefore needed. In [2], pixel-based segmentation is used. The segmentation is done according to the spatial and temporal gradients combined with contour and texture information. This approach is computationally costly and the overhead for transmitting the image segmentation is large. Therefore, in our work block-based segmentation is used utilizing the blocks' motion vectors.

The proposed method in this work consists of *simultaneous* block-segmentation and estimation of the parameters describing the motion of each segment . The segmentation is initialized using the Hough transform and the motion vectors obtained by Block-Matching (BM). It is further improved by applying a Gibbs-Markov model. Finally, spliting/merging of segments is applied and a detailed 8-parameter motion description is found for each segment. Some more details of these steps are given below.

Initial Hough-Based Segmentation

Least-Squares (LS) provides a good estimation of the motion parameters if it uses motion vectors of an homogeneously moving area, i.e., of a region moving according to an expected motion model. However, when the picture contains several regions having different motion parameters, LS finds an "average" parameter set, which really doesn't describe accurately any of the regions' motion. Therefore, segmentation should be done first, and only after the segmentation defines homogeneously moving regions, LS can be applied to each segment separately.

A robust method for segmentation is the Hough Transform [1]. In this method, the BM motion vectors are used for voting for parameter sets in the parameter space, representing the motion of the different segments. The Hough Transform used here is based on a 3-parameter model (zoom pan and tilt), because the amount of computations involved is exponentially increasing with the number of parameters, so that applying it to a higher order model is usually computationally prohibitive. Every set of parameters, which carries enough support in the voting process, corresponds to a distinct segment. The blocks supporting it define the segment.

Gibbs-Markov Model-Based Segmentation

The Hough Transform provides the initial segmentation, which could still be unsatisfactory in some ways. First, some blocks in a region may be matched to the wrong segment, as the result of

1 *Supported in part by the Samuel Neaman Institute for Advanced studies in Sceience and Technology of the Technion*

stray motion vectors, or because a vector may vote for more than one set of parameters. Another problem is inaccuracy, especially at segment boundaries. Both of these effects are corrected in the proposed algorithm by using a Gibbs-Markov model-based segmentation.

The problem we wish to solve is : Having two successive pictures (A , B) of a sequence, we look for the best segmentation field (S) (and the parameters associated with it). For this purpose we want to find S which will maximize the conditional probability density function $p(S/A,B)$.

A Gibbs-Markov model-based cost function is defined. The cost function is constructed to support small prediction errors, homogeneous segmentation field, and consistency of the segmentation in the previous and in consecutive frames.

The optimization is done for each block separately, thus finding a *local* minimum of the cost function. This is done using an iterative algorithm known as the Iterated Conditional Modes (ICM) algorithm [3], and requires a reasonable amount of computation while providing good results.

Joint Spliting/Merging Decision and Detailed Motion Description

The sets of parameters found by the Hough Transform correspond to a 3-parameter motion-model. The next step is the estimation of the parameters of an 8-parameter projective transformation model for each of the segments. The estimation is coupled with splitting/merging of segments, to better match the description by 8-parameter models. Splitting and merging may be needed because regions which have different 3-parameters sets may have the same 8-parametrs set, and vice versa.

In the literature, pixel-difference-based spliting/merging decisions are mostly used. In our work the decision is based on the local motion-vectors of blocks, which requires less computations. It is done using a hypothesis testing technique by which a decision is made between the hypotheses that either two regions belong to the same segment or to two different segments. To make this decision the Maximum-Likelihood approach is applied. The probability distribution used is that of the error between the expected motion vector according to the model and the one found by BM.

Simulation Results

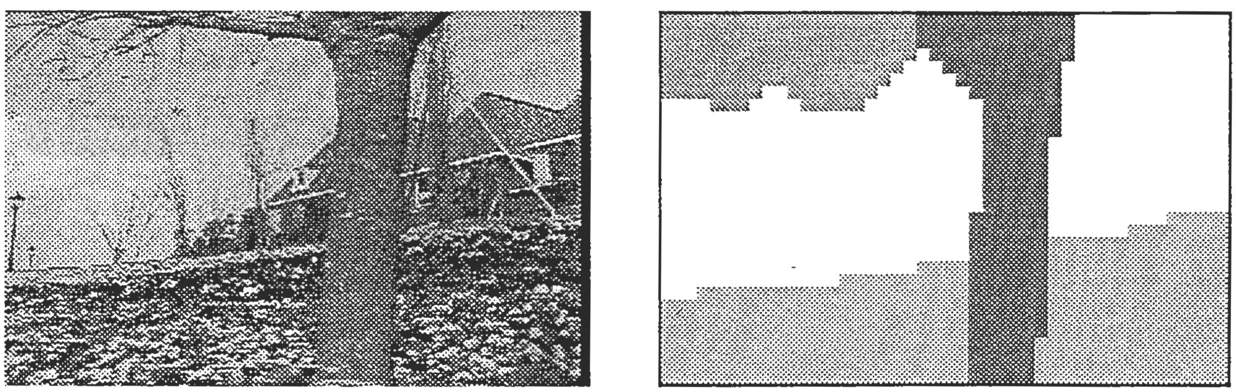
A global motion compensation scheme based on the described approach for global motion estimation was incorporated into a RM8-based image-sequence coder. In comparison to RM8, a 30% reduction in bit-rate was obtained in coding the ISO test sequence "Garden-Flower", for a wide range of rates. This is attributed to the reduced prediction error obtained by the improved motion compensation technique proposed. The segmentation of a typical frame of the sequence is shown in Fig 1.

<u>References</u>

[1] Z. Eisips and D. Malah, "Global motion estimation for image sequence coding applications", Proc. 17th convention of Electrical and Electronics Engineering in Israel ,May 1991, pp. 186-189.

[2] Nobert Diehl, "Object-oriented motion estimation and segmentation in image sequences", Signal Processing: Image Communication 3, 1991, pp. 23-56.

[3] Julian Besag, "On the statistical analysis of dirty pictures", J.R.Statist.Soc.B, 1986, pp 259-302.



(a) (b)

Fig 1 (a) Original picture from the Garden-Flower sequence. (b) Final segmentation results by proposed algorithm.