MULTIPLE PARAMETRIC MOTION MODEL ESTIMATION AND SEGMENTATION

R. Montoliu, F.Pla

Computer Vision Group (www.vision.uji.es)
University Jaume I
Campus Riu Sec, 12071 Castellón (SPAIN)

ABSTRACT

This paper presents a motion estimation and segmentation algorithm based on multiple parametric model estimation that determines the a priori unknown number of motion models present in the data. The algorithm applies a quasi- simultaneously parametric model fitting method based on a general Least Square fitting. Some experiments are showed in order to demonstrate the results obtained using the proposed algorithm.

1. INTRODUCTION

Detecting and segmenting moving objects in a image sequences have been given a large attention in the literature of image sequence analysis. Some of the applications where can be applied motion segmentation techniques are image registration, moving object detection, surveillance, guidance of autonomous vehicles, efficient video compression, traffic monitoring and tracking.

Most of the early works have been concentrated on the estimation of the optical flow computed between image pairs, [1]. These works can be used to detect multiple motions, however, these methods are very sensitive to the quality of the optical flow, that it is usually very inaccurate in the motion boundaries.

Another approach of motion segmentation is the use of parametric motion estimators, which describe the motion over a larger spatial region in terms of a parametric model [2]. These methods are very accurate since they only have to estimate a small number of parameters (6 for the affine motion model) given a large set of constraints, one constraint for each pixel used to estimate the parameters in a region. The problem with the model-based method is that large image regions are often not well modeled by a single parametric model due to the complexity of the motion or the presence of multiple moving objects. When the region of analysis contains multiple moving objects some method that can recover simultaneously the motion parameters and the points (pixels) that support these parameters is needed.

This work has been partially supported by projects GV97-TI-05-27 from the *Conselleria de Educació Cultura i Ciència*, Generalitat Valenciana, and CICYT TIC98-0677-C02-01 from the *Ministerio de Educación y Cultura*.

In [3], Bober and Kittler present a Hough transform based hierarchical algorithm for simultaneously motion estimation and segmentation using the affine motion model. They use robust estimation techniques in order to reduce the influence of the outliers. The Bober and Kittler algorithm estimate first the parameters of the dominant motion present in the entire image. The pixels that support that model are removed, then the algorithm is applied recursively to the remaining pixels until extract all the motion models presents.

In this paper a multiple parametric model fitting algorithm is used in order to simultaneously estimate and segmentate multiple motion models present in a sequence of two images. A generalized Least Squares fitting method is used in order to estimate the motion parameters present in a region. The main difference with dominant motion methods is that our approach can extract quasi- simultaneously all the motion models present in the image sequence. The influence of the outliers is completely removed using two statistical tests that classify the points in supporting model points (inliers) and not-supporting model points (outliers).

In the next section, the process of detecting a single motion model in a sequence of two images is presented, with the explanation of the motion estimation method and the outliers/inliers detection process. The Section 3 explains the multiple parametric motion model algorithm. In Section 4 some experiments are presented in order to show the results obtained using the proposed approach.

2. SINGLE PARAMETRIC MODEL EXTRACTION

In [4], Danuser and Stricker presented a framework for parametric model fitting applied to lines and planes. In this paper an adaptation of this framework has been used in order to be applied it in motion segmentation problems.

The framework has four steps: a generalized Least Square fitting method based on the algorithm proposed by [5], a statistical test for outliers detection which is based on a data snooping procedure, a similar statistical test to collect news points to support the model (inliers) and two statistical tests

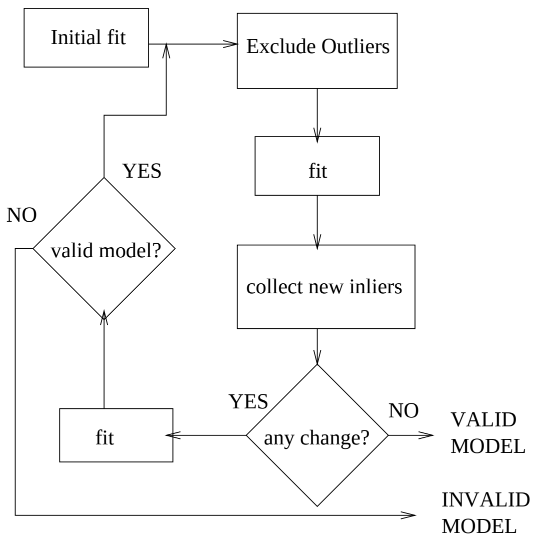

Fig. 1. SPME scheme.

to determine if the model is valid.

The single parametric model extraction (SPME) requires an initial model as input. The initial model consists of a coarse estimate of the model parameters and a list of points that may belong to the model.

The extraction process is showed in the Figure 1. The reader is addressed to [4] for a comprehensive explanation of the method.

2.1. Motion estimation

A Generalized Least Squares-based method is used to estimate the parameters of the model. This method fits a set of functions f_i, each one associated to an observations vector l_i. All functions f_i have a common vector of parameter x,

$$f_i(x, l_i) = 0, i : 1..r$$

Each function f_i has f components, x is a vector of u model parameters and l_i is a vector of n measurements.

The estimation process is an iterative process which starts with an initial guess of the model parameters $\hat{x}_o$. The estimation process is stopped when the increment $\widehat{\Delta x}_t = \hat{x}_{t+1} - \hat{x}_t$ of the estimation $\widehat{x}$ at the iteration t is smaller than a threshold. The increments of the model parameters are obtained from the equation:

$$\begin{aligned}\widehat{\Delta x} &= \left(A^t Q_{ww}^{-1} A\right)^{-1} A^t Q_{ww}^{-1} w \\ \widehat{x}_{t+1} &= \widehat{x}_t + \widehat{\Delta x}\end{aligned}$$

where $A = \frac{\partial f}{\partial x}$ is a $r \times u$ matrix, $B = \frac{\partial f}{\partial l}$ is a diagonal-block matrix (r blocks) where each block is a $f \times n$ matrix. $Q_{ww} = B Q_{ll} B^t$ was introduced to simplify the notation. Finally Q_{ll} is a cofactor matrix which contains all the relations among the various units and precision levels of the observations.

The assumption that the gray level of all the pixels remains constant between two images in a sequence is used as function f_i,

$$f_i = I_1(x'_i, y'_i) - I_2(x_i, y_i) = 0$$

where x_i, y_i are the column and the row of pixel i. $I_1(x'_i, y'_i)$, $I_2(x_i, y_i)$ are the gray level of the first image, at pixel x'_i, y'_i, and the gray level of second image, at pixel x_i, y_i, and $x'_i = p(x_i, y_i)$, $y'_i = q(x_i, y_i)$. The functions p and q modelize the motion model.

Each vector of measurements l_i has $n = 4$ elements: x_i, y_i, $I_1(x'_i, y'_i)$ and $I_2(x_i, y_i)$. The vector of parameters depends on the motion model used in the functions p and q.

For each measurement l_i a matrix B_i and a matrix A_i are defined. The matrices B_i and A_i and the global matrices B and A have the following forms:

$$\begin{aligned}
B_i &= \left(\frac{\partial f_i}{\partial x_i}, \frac{\partial f_i}{\partial y_i}, \frac{\partial f_i}{\partial I_1(x'_i, y'_i)}, \frac{\partial f_i}{\partial I_2(x_i, y_i)}\right)_{(1\times 4)} \\
A_i &= \left(\frac{\partial f_i}{\partial x^1}, \frac{\partial f_i}{\partial x^2}, ..., \frac{\partial f_i}{\partial x^u}\right)_{(1\times u)} \\
B &= \begin{pmatrix} B_1 & 0 & 0 & 0 \\ 0 & B_2 & 0 & 0 \\ ... & ... & ... & ... \\ 0 & 0 & 0 & B_r \end{pmatrix}_{(r\times(r\times 4))} \\
A &= \begin{pmatrix} A_1 \\ A_2 \\ ... \\ A_r \end{pmatrix}_{(r\times u)}
\end{aligned} \tag{1}$$

with r being the number of pixels.

The motion model used in the present approach is the affine motion model that is one of the most used in the literature. The affine transformation corresponds to the orthographic projection of a 3D rigid motion of a planar surface. The affine model can handle translations, scaling, rotations and shear of images. The affine motion model has six parameters, ($u = 6$): $a_1, b_1, c_1, a_2, b_2, c_2$

$$\begin{aligned} x'_i &= p(x_i, y_i) = a_1 x_i + b_1 y_i + c_1 \\ y'_i &= q(x_i, y_i) = a_2 x_i + b_2 y_i + c_2 \end{aligned}$$

Therefore, the Jacobian A_i and B_i introduced above become:

$$
\begin{aligned}
A_i &= \left(\frac{\partial f_i}{\partial a_1}, \frac{\partial f_i}{\partial b_1}, \frac{\partial f_i}{\partial c_1}, \frac{\partial f_i}{\partial a_2}, \frac{\partial f_i}{\partial b_2}, \frac{\partial f_i}{\partial c_2} \right)_{(1\times 6)} \\
&= (x_i I_x^1(x_i', y_i'), y_i I_x^1(x_i', y_i'), I_x^1(x_i', y_i'), \\
&\quad\; x_i I_y^1(x_i', y_i'), y_i I_y^1(x_i', y_i'), I_y^1(x_i', y_i')) \\
B_i &= \left(\frac{\partial f_i}{\partial x_i}, \frac{\partial f_i}{\partial y_i}, \frac{\partial f_i}{\partial I_1}, \frac{\partial f_i}{\partial I_2} \right)_{(1\times 4)} \\
&= (a_1 I_x^1(x_i', y_i') + a_2 I_y^1(x_i', y_i') - I_x^2(x_i, y_i), \\
&\quad\; b_1 I_x^1(x_i', y_i') + b_2 I_y^1(x_i', y_i') - I_y^2(x_i, y_i), 1.0, -1.0)
\end{aligned}
$$

Where $I_x^1(x_i', y_i')$, $I_y^1(x_i', y_i')$, are the gradient of first image at the pixel (x_i', y_i') in x and y direction, and $I_x^2(x_i, y_i)$, $I_y^2(x_i, y_i)$, are the gradient of second image at the pixel (x_i, y_i) in x and y direction.

Jacobian Matrices A_i and B_i can also be calculated for other motion models, e.g projective (8-parameters) motion model, as expressed in (1).

The outliers detection is based in data snooping which relies on a statistical test of the residuals of each point. For each point j belonging to the model, a statistical test is performed, if the point does not support the model, then it is excluded from the model and added to a not yet classified set.

A similar statistical test is used to collect new points to support the model. For each not yet classified point k, a statistical test is performed, if the point support the model, then it is excluded from the not yet classified set and added to the model.

A new iteration loop starts with two statistical tests that verify the validity of the current model. The first test compares the current standard deviation with the standard deviation of previous iteration, the second test is applied to check for deficiencies in the stochastic model, as well as in the deterministic model. A model become invalid if one or both tests fail.

3. MULTIPLE PARAMETRIC MODELS EXTRACTION (MPME)

The Multiple models extraction procedure can extract quasi simultaneously all the models present in the data without a priori knowledge of the number of models. Figure 2 shows a scheme of the algorithm.

The procedure starts with a successful initialization of a model. Then the SPME scheme is applied in the initial model in order to obtain a valid or an invalid model. If a valid model M_n has been extracted, then a data exchange procedure is started. The goal of this procedure is to reclassify points that have been classified in an early model M_m, where $m < n$, but they are closer to the new model. The

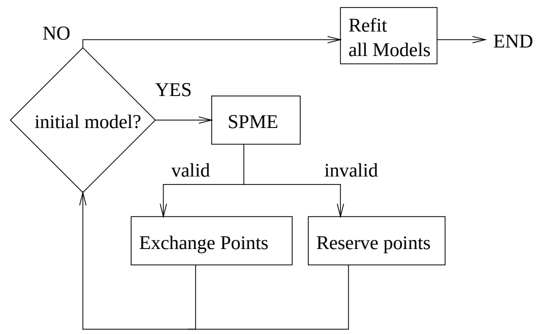

Fig. 2. MPME scheme.

points of invalid models are reserved in order to exclude them from the next model initialization. However, they are still available to be recaptured in the inliers test. If no further model can be initialized, then all the valid models are refitted based on theirs final inliers data sets. The initialization of new models is stopped if all the points were classified or a number of trials were unsuccessful.

At the end, a list of valid models are obtained, each one with a list of points that support the model together with the estimated motion parameters.

Using this algorithm, although the single models are extracted sequentially, the final result is practically [4] independent of the order in which they have been extracted.

4. EXPERIMENTS AND RESULTS

In the first experiment two synthetic image sequences have been used in order to test the proposed SPME algorithm. In the first image sequence a change of scale plus a translation have been performed. In the second image sequence a rotation have been performed. Both image sequences has been generated using the ideal interpolation function. Figure 3 shows the two images of each sequence and the results obtained assuming only one motion model present in the sequence. The white points are the inliers, and the black points are the outliers.

The outliers points in both sequences mainly correspond to points that the ideal interpolation function could not calculate a valid gray level since it had not enough information.

The motion parameters estimated for both image sequences of the motion model found (white points) are showed in Table 1.

The second experiment is directed to show the results obtained with MPME algorithm in a motion segmentation problem. Preliminary tests have been made using a translational motion model, that is an instance of the affine motion model.

In Figure 1 the results of the motion segmentation are

	Scale/translation	*Rotation*
$\hat{a}_1$	1.080020 (1.08)	0.987613 (0.9876)
$\hat{b}_1$	-0.000038 (0.0)	-0.156461 (-0.1564)
$\hat{c}_1$	-0.004885 (0.0)	0.001782 (0.0)
$\hat{a}_2$	0.000165 (0.0)	0.156394 (0.1564)
$\hat{b}_2$	1.079699 (1.08)	0.987720 (0.9876)
$\hat{c}_2$	4.006840 (4.0)	0.002490 (0.0)

Table 1. Affine parameters estimated using SPME. In parenthesis true values.

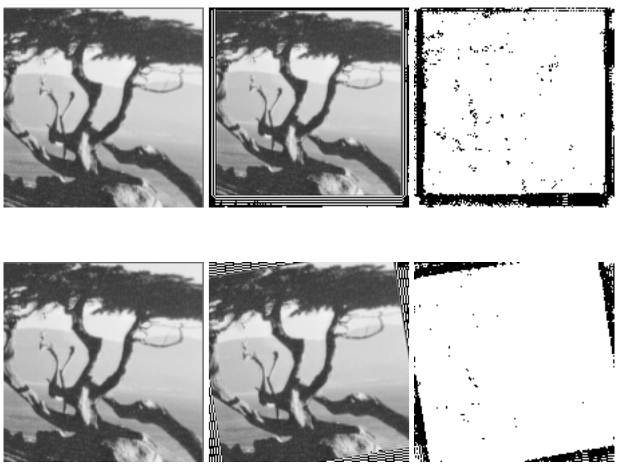

Fig. 3. Image sequences of experiment 1 and results.

showed, the pixels corresponding at each model are labeled using different gray levels, the non-classified points are labeled with white color. The parameters and number of supporting points of the three meaningful models found are showed in Table 3.

In Figure 1 some points that a priori belong to model 2 and 3, the MPME procedure classify them in the model 1 (no motion), since these points are very-low textured and certainly could be also classified into model 2 and 3, respectively.

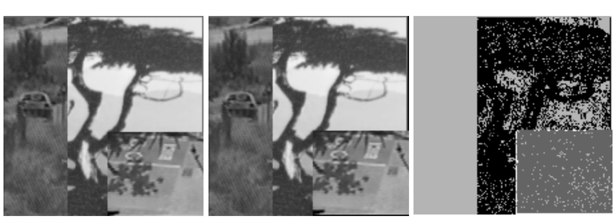

Fig. 4. Three translational motion models sequence.

Model	c_1 (true)	c_2 (true)	No. of points
1	-0.0022 (0.0)	0.0031 (0.0)	9548
2	-0.0043 (0.0)	0.9970 (1.0)	4947
3	0.9960 (1.0)	0.0040 (0.0)	7773
Outliers			232

Table 2. Results of the Three translational motion models sequence, ordered by gray level

5. CONCLUSION

A multiple parametric model motion estimation and segmentation algorithm has been presented in this paper. The results of the algorithm are a set of models, each one containing the parameters estimated for the model and a list of points that support the motion parameters estimated. An affine motion model has been used in the first experiment to test the single parametric motion model extraction algorithm and a translational motion model has been used in the second experiment to test the multiple parametric motion model algorithm.

The results obtained show that the proposed approach improve the parameter estimation , discarding outliers, when only one affine motion model is presented in the region of analysis, and can extracts multiple translational motion models without using neighborhood information.

Further work is directed to test the proposed algorithm with more complex affine sequences using other real scenes applying affine motion estimation/segmentation and adding multi-resolution and spatial constrains to improve the final results.

6. REFERENCES

[1] B.K.P. Horn and B.G. Schunck, "Determinig optical flow," in *MIT AI*, 1980, A.I. Memo No 572.

[2] S. Ayer, P. Schroeter, and J. Bigun, "Segmentation of moving objects by robust motion parameter estimation over multiple frames," in *ECCV94*, 1994, pp. B:316–327.

[3] M. Bober and J. Kittler, "A hough transform based hierarchical algorithm for motion segmentation and estimation," in *4th International workshop on Time-Varying image Processing and Moving Object Recognition*, V. Cappellini, Ed., 1993, pp. 335–342.

[4] G. Danuser and M. Stricker, "Parametric model-fitting: From inlier characterization to outlier detection," *PAMI*, vol. 20, no. 3, pp. 263–280, March 1998.

[5] H.I. Britt and R.H. Luecke., "The estimation of parameters in nonlinear implicit models.," *Technometrics*, vol. 15, no. 2, pp. 233–247, 1973.