PARAMETER ESTIMATION OF NON - TRANSLATIONAL MOTION FIELDS

Constantinos Dimou   *Ioannis Pitas*
Dept. of Informatics, University of Thessaloniki
P.O. BOX 451 Thessaloniki, GREECE
Tel.: +30-31-996304, FAX: +30-31-996304
e-mail: `dinos, pitas@zeus.csd.auth.gr`

ABSTRACT

Motion estimation is a very important topic in computer vision and image sequence compression. However, most commonly used motion estimation algorithms do not take into consideration any motion invariances that a certain local motion might possess. In this paper, a technique for estimating the invariant motion parameters of non-translational motion fields is proposed, which leads to more efficient estimation, smoothing or coding of the motion field. It is shown that the algorithm performs well, even in high noise levels, i.e., in the case of noisy output of the motion estimator.

1 INTRODUCTION

A number of different approaches have been proposed for extracting motion information from image sequences, such as the commonly used block matching algorithms and their variations [1], recursive algorithms or newly proposed cluster matching techniques [2]. However, these algorithms do not take into consideration any motion invariances that a local motion might possess. Even when applied to image sequences which possess certain motion invariant properties, they tend to give a translational estimation of the motion field. It is obvious that if such motion invariants of a certain motion type were known, then all vectors which belong to that motion type could be easily described, which would lead to a more efficient way of representation of the entire motion field.

Previous work in identification and recovery of non-translational flow patterns was done in [3], where a classification scheme based on local properties of the motion field was developed with good results. However, this scheme is applied in cases when there is only one such pattern in the scene. Very little work has been done in estimating motion invariant parameters from scenes which contain different types of motions.

In this paper a technique for simultaneous parameter estimation and detection of non-translational motion fields is presented. The present work is concerned only with the estimation of rotational and divergent motion fields as they are the most commonly appearing non-translational motion types, but it can also be expanded to other types of motions possessing invariant properties as well. The estimator works in two steps. Initially, the estimation of the motion centers is performed. Then the motion velocity estimation and boundary detection is performed based on the results of the center estimation. It will be shown that this technique allows the simultaneous detection of a number of such motion fields, with very little a priori knowledge about the scene.

2 INVARIANT PROPERTIES OF NON-TRANSLATIONAL MOTION FIELDS

The velocity field $\dot{x} = u(x, y), \dot{y} = v(x, y)$ induced on an image plane with equation $Z = pX + qY + r$ moving with translational velocity (a, b, c) at $(0, 0, r)$ and rotational velocity $(\omega_1, \omega_2, \omega_3)$ is given by [4]:

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{1}$$

where $A = p\omega_2$, $B = q\omega_2 - \omega_3$, $C = -p\omega_1 + \omega_3$, $D = -q\omega_1$. From these flow parameters we obtain the invariants:

$$\begin{aligned} V &= a + ib \\ T &= A + D \\ R &= C - B \end{aligned} \tag{2}$$

which correspond to translational, divergent and rotational flow types, respectively. A pure rotational flow in particular corresponds to all parameters but R zero, and a pure divergent flow corresponds to all parameters but T zero.

The invariant R is called the *rotation*, and corresponds to the rotational velocity:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{R}{2} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \tag{3}$$

where (x_0, y_0) is a spatial displacement of the center.

In an equivalent manner, the invariant T is called the *divergence*, and corresponds to the divergent velocity:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{T}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \tag{4}$$

In the case where a motion estimation algorithm is applied to an image sequence containing a rotational or a divergent field, the estimated optical flow field will contain errors produced by the algorithm. These errors can be modeled as additive noise applied to the original motion field, in the form of noise vectors $\underline{\mathbf{n}} = [n_x, n_y]^T$, where n_x and n_y are considered *iid* random variables. The following analysis will focus on rotational motion fields only. The extension in the case of divergent fields is straightforward and therefore is omitted.

3 MOTION CENTER ESTIMATION

We consider two motion vectors, namely $\underline{\mathbf{V}}_{\mathbf{1}} = [V_{x_1}, V_{y_1}]^T$ and $\underline{\mathbf{V}}_{\mathbf{2}} = [V_{x_2}, V_{y_2}]^T$ belonging to the same rotational field, with starting points (x_1, y_1) and (x_2, y_2), respectively. The corresponding noise vectors are $\underline{\mathbf{N}}_{\mathbf{1}} = [n_{x_1}, n_{y_1}]^T$ and $\underline{\mathbf{N}}_{\mathbf{2}} = [n_{x_2}, n_{y_2}]^T$, respectively. The noisy motion vectors, namely $\underline{\mathbf{S}}_{\mathbf{1}}$ and $\underline{\mathbf{S}}_{\mathbf{2}}$, as estimated from the motion estimator are:

$$\begin{aligned} \underline{\mathbf{S}}_{\mathbf{1}} &= \underline{\mathbf{V}}_{\mathbf{1}} + \underline{\mathbf{N}}_{\mathbf{1}} \\ \underline{\mathbf{S}}_{\mathbf{2}} &= \underline{\mathbf{V}}_{\mathbf{2}} + \underline{\mathbf{N}}_{\mathbf{2}} \end{aligned} \quad (5)$$

We base our analysis on two observations, coming from local geometrical characteristics. First, in a pure rotational field all perpendicular lines passing through each vector starting point intersect in the center of the rotational field. Similarly, in a pure divergent field all vector directional lines pass through the center of the divergent field. These lines will be referred to as *radial lines* in the rest of the paper. In the case of a noisy motion field the intersection of all such radial lines will form a spatial point concentration around the true motion center. In the case of a scene containing two or more rotational motions, an equal number of such concentrations is expected.

The coordinates x_s and y_s of the intersection point of the radial lines corresponding to vectors $\underline{\mathbf{S}}_{\mathbf{1}}$ and $\underline{\mathbf{S}}_{\mathbf{2}}$ are computed as functions of the four *iid* random variables n_{1_x}, n_{2_x}, n_{1_y}, n_{2_y}:

$$\begin{aligned} x_s &= f(n_{1_x}, n_{2_x}, n_{1_y}, n_{2_y}) \\ y_s &= g(n_{1_x}, n_{2_x}, n_{1_y}, n_{2_y}) \end{aligned} \quad (6)$$

The model distribution for the noise variables is assumed to be zero-mean Gaussian, with density function [5]:

$$\varphi_{n_x}(n_x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{n_x^2}{2\sigma^2}} \quad (7)$$

After some algebraic manipulation, the distribution of the intersection point (x_s, y_s) is computed as:

$$f_{x_s,y_s}(x_s, y_s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sigma^4 4\pi^2}\frac{1}{|J|} e^{-\frac{w^2+v^2}{2\sigma^2}} e^{-\frac{n_{y1i}^2+n_{y2i}^2}{2\sigma^2}} \, dwdv \quad (8)$$

where w, v are two dummy variables required for the integration, n_{y1i}, n_{y2i} are functions of x_s, y_s, w, v, and J is the Jacobian determinant of f and g.

It is difficult to obtain a closed form solution from (8). Therefore, the pdf values have been calculated numerically in a rectangular grid. Clearly (8) does not represent a normal distribution.

Because the exact mechanisms that produce errors in block matching algorithms are not known exactly and can not be described by a single distribution model $F(x)$, it is natural to have outlying observations from the assumed model distribution. This means that, even if the density function of the intersection points could be derived in a closed form from (8), we would still have some observations which would not belong to the model. Therefore, the location estimator used to estimate the center of the motion field must have good robustness properties in order to reject the outlying observations.

An estimator with very good robustness properties is the well known *sample median* of the observations [7]. The median possesses a finite gross-error sensitivity $\gamma*$, which makes it B-robust. In particular:

$$\gamma * (\psi, F) \geq \frac{1}{2f(0)} \quad (9)$$

Thus, the median is the *most* B-robust estimator. However, it should be noted that the median does *not* reject outliers, since its rejection point $\rho*$ is infinite. In order to increase the robustness of the estimation, we use a variation of the sample median estimator, called the *skipped median* estimator [8]. The skipped median belongs to the class of the so called *redescending M-estimators*, which are able to reject outliers entirely.

The calculation of all intersection points on the parameter space is obtained by using a technique principally similar to that of the Hough transform. The parameter space is the plane defined by (x_s, y_s). Initially the area of this plane is divided into accumulator cells having variable size. It is obvious that, as in the standard Hough algorithm, the size of the accumulator cells controls the accuracy of the algorithm. Having set the accumulator cells, the radial line parameters for each vector in the field are computed. Finally, from these parameters the intersection points of two or more radial lines are computed and the corresponding cell is updated.

In the case of a motion field containing several non-translational motions, a corresponding number of concentrations of intersection points is expected. Therefore, a clustering scheme able to locate the motion centers from the intersection point concentrations is required in this case. Moreover, the clustering procedure should correspond to the statistical properties of the intersection point data set.

Such a clustering scheme is the widely used *Learning Vector Quantizer* (LVQ) neural network. LVQ is an autoassociative nearest-neighbor classifier which classifies arbitrary patterns into classes using an encoding procedure related to competitive learning. An interesting variation of LVQ, called marginal median LVQ (MMLVQ), which uses the median estimator in its classification criterion, was proposed in [6]. Because it is based on order statistics, marginal median LVQ has superior performance for non-Gaussian distributions. Moreover, we have used a variation of MMLVQ based on the skipped median estimator presented previously. The resulting scheme is a *Skipped Median LVQ* (SMLVQ). SMLVQ updates the reference vectors at a time t by using the skipped median of the set of vector-valued observations that have been already assigned to that class until time t. Only approximate knowledge of the number of non-translational motions is required for the proper initialization of the LVQ reference vectors.

Based on these characteristics, SMLVQ ensures more robustness in the classification procedure. Therefore, it was used as a location estimator of the center of the concentration of intersection points.

4 MOTION VELOCITY ESTIMATION AND BORDER DETECTION

The motion invariance estimation is based on the result of the center estimator. One problem is still unresolved, namely the detection of the border of the motion field, so that all the velocity observations are assigned only to the field where they belong to. A simultaneous technique for border detection and parameter estimation was developed in this case, which can be used to estimate the angular velocity of the motion field at the same time.

Considering a vector $\mathbf{S} = [S_x, S_y]^T$ of a rotational motion corrupted by noise, we conclude from (1) and (5) that:

$$\begin{aligned} S_x &= -\frac{R}{2}(y_s - y_0) \\ S_y &= \frac{R}{2}(x_s - x_0) \end{aligned} \qquad (10)$$

Therefore, the observed angular velocity samples $R_{(x)}$, $R_{(y)}$ from the noisy motion field are:

$$\begin{aligned} R_{(x)} &= 2\frac{V_x + n_x}{-d_y} = R + \frac{1}{-d_y} n_x \\ R_{(y)} &= 2\frac{V_x + n_x}{d_x} = R + \frac{1}{d_x} n_y \end{aligned} \qquad (11)$$

where d_x, d_y denote the marginal vector distances from the rotation center. In a completely equivalent manner the equations for divergent motion fields are derived. The parameters d_x, d_y act as scale parameters in (11), so that vectors near the motion center are expected to have larger noise variance than the rest. In order to compensate for this and simultaneously to ensure robustness into the estimation, we apply an adaptive a-trimmed mean estimator [9], with a being inversely proportional to d_x, d_y:

$$\begin{aligned} \alpha &= \frac{0.5}{d_x} \quad for \ R_{(x)} \\ \alpha &= \frac{0.5}{d_y} \quad for \ R_{(y)} \end{aligned} \qquad (12)$$

As d_x, d_y take increasing discrete values starting from 1, the adaptive α-trimmed mean estimator changes from the sample median to the sample arithmetic mean, which produces smoother output in small distances from the motion center.

The estimation is performed in concentric discrete patterns (circles or squares) around the motion center, and at each step velocity samples inside the current circle are included in the sample data set. The first order derivative of the estimator output is computed at each step and compared with a predefined threshold. The iterations stop when the derivative falls below the threshold, indicating that the motion border was reached.

5 SIMULATION RESULTS

The overall estimation technique can be applied for the simultaneous detection of several rotational or divergent field parameters. The estimation is performed in two steps, as mentioned earlier. In the first step we apply SMLVQ to the intersection point concentrations. After locating the motion centers the angular velocity estimation and border detection are performed separately for each motion in the second step.

Extensive simulation experiments have been performed. A synthesized motion field representing a moving car (Fig.1) was constructed in order to test the performance of the algorithm. This motion field comprised of a cluster having only translational properties (car body), and two rotational fields representing the two wheels of the car. Gaussian noise was added to the field in order to simulate the errors produced by the motion estimator. The intersection points appear in two concentrations around the two centers of the wheels. SMLVQ was applied to these concentrations in order to locate the centers of the two rotations.

In the second step the rotational velocity was estimated based on the center estimation. In Fig.2 the velocity estimator output vs. distance from the estimated motion center is shown. The iterations stop when the gradient change exeeds a predefined threshold.

From the overall obtained results (Fig.3) it is shown that the algorithm performs very well, even at high noise levels.

6 CONCLUSIONS

In this paper a technique for estimating the parameters of non-translational motion fields and for segmenting two or more such motions was presented. A technique similar to that of the standard Hough transform combined with skipped median LVQ was used in order to locate the motion centers, and a combined robust algorithm was applied to estimate the non-translational motion velocity and motion boundaries. The estimation of the above parameters allow the clustering of the different motion fields, which in turn leads to better clustering and coding of the entire motion field. Our experimental results have shown that the proposed technique is relatively insensitive to noise.

References

[1] A. N. Netravali and B. G. Haskell, "Digital Pictures, Representation and Compression", Plenum Press, 1989, 5.2.3g, pp. 335-340.

[2] Dane P. Kottke and Ying Sun, "Motion Estimation Via Cluster Matching", *IEEE Trans. On Pattern Analysis And Machine Intel. (PAMI)*, Vol. 16, No. 11, Nov. 1994, pp. 1128-1132.

[3] Chiao-Fe Shu and Ramesh C. Jain, "Vector Field Analysis for Oriented Patterns", *IEEE Trans. On Pattern Analysis And Machine Intel. (PAMI)*, Vol. 16, No. 9, Sept. 1994, pp. 946-950.

[4] I. Kanatani, "Group-Theoretical Methods for Image Understanding", pp. 35-40.

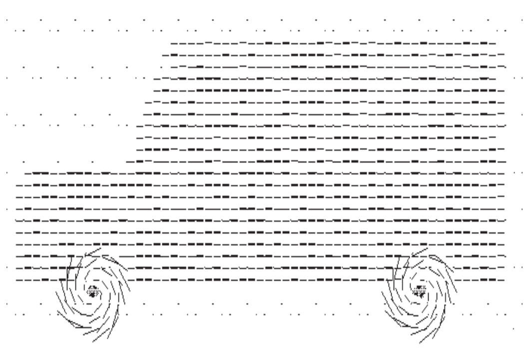
[5] A. Papoulis, *Probability Random Variables and Stochastic Processes*, Mc Graw-Hill, 1991.

[6] I. Pitas, C. Kotropoulos, N. Nikolaidis, R. Yang and M. Gabbouj, "A Class of Order Statistics Learning Vector Quantizers", *IEEE Int. Symposium on Circuits and Systems (ISCAS '94)*, vol. 6, pp. 387-390.

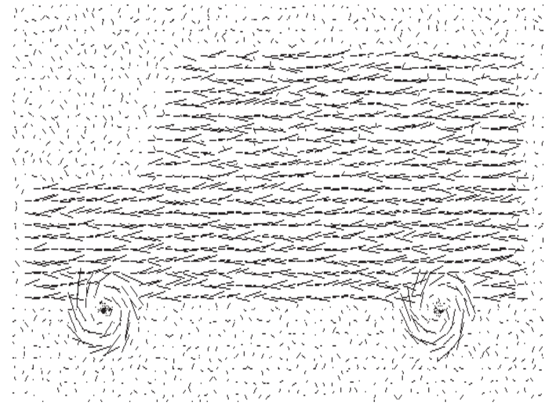
[7] I. Pitas, A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Kluwer Academic Publishers, 1990.

[8] F. Hampel, E. Ronchetti, P. Rousseauw, W. Stahel, *Robust Statistics*, John Wiley, 1986.

[9] A. Restrepo and A.C. Bovic, "Adaptive Trimmed Mean Filters for Image Restoration", *IEEE Trans. On Acoustics, Speech and Signal Processing*, Vol. 36, No. 8, Aug. 1988, pp. 1326-1337.



(a)



(b)

Figure 1: Concentration of intersection points around the centers obtained for a synthetic motion field corrupted with additive Gaussian noise with standard deviation 1 (a) and 5 (b). The computed intersection points are shown with bold dots.

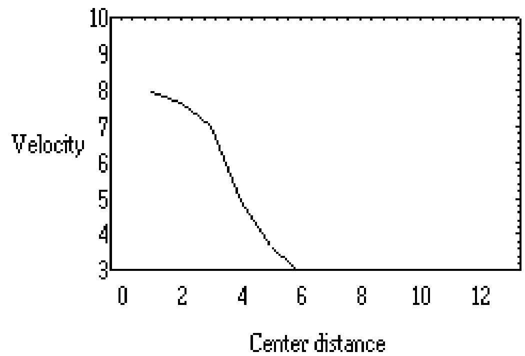


Figure 2: Estimator output values vs. distance from the motion center using the adaptive a-trimmed mean algorithm for noise standard deviation of 5.

	Front Wheel		
	Invariance	Radius	Center Estim.
original	6.66	3	$(120, 280)$
$\sigma = 1$	6.71	3	$(120, 279)$
$\sigma = 3$	6.84	3	$(120, 278)$
$\sigma = 5$	6.96	3	$(121, 279)$
	Back Wheel		
	Invariance	Radius	Center Estim.
original	6.66	3	$(500, 280)$
$\sigma = 1$	6.72	3	$(500, 279)$
$\sigma = 3$	6.85	3	$(499, 279)$
$\sigma = 5$	7.03	3	$(502, 280)$

Figure 3: Estimated parameters of the front and back wheel of the simulated car.