In Proceedings of IEEE International Conference on Image Processing (ICIP'01), Thessaloniki, Greece, October 7–10, 2001, vol. **II**, pp. 917–920.

ANALYTIC SOLUTIONS FOR MULTIPLE MOTIONS

Cicero Mota, Ingo Stuke, and Erhardt Barth

Institute for Signal Processing, University of Lübeck
Ratzeburger Allee 160, 23538 Lübeck, Germany

ABSTRACT

A novel framework for single and multiple motion estimation is presented. It is based on a generalized structure tensor that contains blurred products of directional derivatives. The order of differentiation increases with the number of motions but more general linear filters can be used instead of derivatives. From the general framework, a hierarchical algorithm for motion estimation is derived and its performance is demonstrated on a synthetic sequence.

1. INTRODUCTION

Multiple motions can occur in computer-vision applications, e.g., in case of semi-transparencies, and also in medical imaging, when different layers of tissue move differently.

We consider image sequences defined by intensity $f(x, y, t)$. The classical constant brightness constraint for the motion vector $\mathbf{v} = (v_x, v_y)$ is

$$v_x f_x + v_y f_y + f_t = 0 \tag{1}$$

Under the hypothesis that $\mathbf{v}$ is approximately constant, the method of total least squares applied to Eq. 1 allows for computing the motion vector by an eigenvalue analysis of the *structure tensor* $\mathbf{J} = \omega * (\nabla f^T \nabla f)$, where ∇f is the gradient and ω a convolution kernel - see [1] for a review.

However, it has been shown [2] that the eigenvalue analysis can be replaced by estimates of $\mathbf{v}$ based on the minors of $\mathbf{J}$, i.e., the following vectors

$$\mathbf{v}_i = (M_{3i}, -M_{2i})/M_{1i} \tag{2}$$

$i = 1, 2, 3$, are all equal to the motion vector $\mathbf{v}$. The numbers M_{ij} are the determinants of the matrices obtained from $\mathbf{J}$ by eliminating the row $4 - i$ and the column $4 - j$, e.g., $M_{11} = (\omega * {f_x}^2)(\omega * {f_y}^2) - (\omega * (f_x f_y))^2$, i.e., M_{ij} are the minors of $\mathbf{J}$.

Here we extend the methods based on the minors of $\mathbf{J}$ to the case of multiple motions. An overview of the problem of multiple motions has been given in [3] and robust methods for multiple motions have been proposed. To our knowledge, the problem of two motions has been first solved in [4, 5] by the use of spatio-temporal Gabor filters, implemented in the frequency domain, and fourth-order moments derived from these filters. The resulting six-dimensional eigensystem has been used to estimate the *mixed motion parameters* [4]. A three-dimensional eigensystem was then used to separate the motion vectors. The main difference to our approach is that we do not need to solve such eigensystems. A further consequence is that we can extend our approach to more than two motions and that we obtain a higher resolution. A recent analysis of the spectral properties of multiple motions can be found in [6]. Others have introduced the useful and intuitive notions of 'nulling filters' and 'layers' [7, 8]. Their approach is more general in that it treats the separation of motions into layers, but is also limited to the use of a discrete set of possible motions and a probabilistic procedure for finding the most likely motions out of the set. We can conclude, that simple analytical solutions for multiple motions have not been proposed.

C.M. is affiliated with the University of Amazonas, Brazil and is supported by the DAAD under A/99/22641.

2. EIGENVECTORS AND MINORS

Motion estimation is often treated as an optimization problem, e.g. by using least-squares methods. The optimization problem then leads to an eigenvalue problem, for example in the case of the tensor-based methods mentioned above.

Fact 1. *If a matrix has a single zero eigenvalue, the corresponding eigenvector can be evaluated in terms of the minors of that matrix.*

Proof. Let $\mathbf{A}$ be a matrix of order m and $\mathbf{X}$ an eigenvector of $\mathbf{A}$ with zero eigenvalue, i.e., $\mathbf{A}\mathbf{X}^T = \mathbf{0}$. If we denote the rows of $\mathbf{A}$ by $\mathbf{A}_i$, we can write $\mathbf{A}_i\mathbf{X}^T = 0$, which means that the rows of $\mathbf{A}$ are orthogonal to $\mathbf{X}$. The rows of $\mathbf{A}$ are linearly dependent since $\det(\mathbf{A}) = 0$, in consequence, a vector orthogonal to $m - 1$ independent rows will also be orthogonal to the remaining one. Now we note that $\det(\mathbf{A}_1, \ldots, \mathbf{A}_{m-1}, \mathbf{A}_1) = 0$, which is the same as $\mathbf{A}_1\mathbf{M}_1^T = 0$, where the components M_{1j} of $\mathbf{M}_1$ are the minors of $\mathbf{A}$ computed by skipping the last row. Therefore,

$\mathbf{X}$ and $\mathbf{M}_1$ are aligned to each other. We can now skip the other rows $m+1-i$ to show that $\mathbf{X} \propto \mathbf{M}_i$ for $i = 1, \dots, m$, which concludes the proof. □

Due to the above proof, we can easily relate the result in Eq. 2 to previous methods that use the eigenvectors of $\mathbf{J}$ [1]. However, for single motions, the method based on the minors has been shown to be faster and more accurate [2].

As we will show, the use of minors for the estimation of motion parameters combined with the algebraic separation of motion vectors provides finite low-complexity algorithms for the estimation of up to *four* motions. This is impossible with eigenvalue analysis because a finite algorithm for eigenvalue analysis is unknown (the estimation of only two motions is equivalent to finding the roots of polynomials of degree *six*).

3. THEORY OF SINGLE AND MULTIPLE MOTIONS

The optical flow can be estimated from the output of linear filters by convolving Eq. 1 with a smooth function g to arrive at [9]

$$(\alpha(\mathbf{v})f) * g = f * (\alpha(\mathbf{v})g) = 0 \tag{3}$$

where $\alpha(\mathbf{v}) = v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} + \frac{\partial}{\partial t}$ is the derivative operator along the direction $\mathbf{V} = (v_x, v_y, 1)$.

The right equality of the above equation is the same as

$$v_x f * g_x + v_y f * g_y + f * g_t = 0, \tag{4}$$

which has to be true for any kernel g. We can then look for a pair of such kernels that make the above equation well posed and solve Eq. 4 for $\mathbf{v}$.

So far, f was supposed differentiable. It is possible, however, to show that Eq. 4 is still valid even if f is not differentiable. Note that

$$f * g(\mathbf{x}, t) = \int f(\mathbf{y} - \mathbf{v}\tau) g(\mathbf{x} - \mathbf{y}, t - \tau)\, d\mathbf{y} d\tau$$

is differentiable. With the change of variables $\mathbf{z} = \mathbf{y} - \mathbf{v}\tau,\ s = t - \tau$, we find

$$f * g(\mathbf{x}, t) = \int f(\mathbf{z}) g(\mathbf{x} - t\mathbf{v} - \mathbf{z} + s\mathbf{v}, s)\, d\mathbf{z} ds.$$

By differentiation under the integral sign, it is easy to see that $f * g$ satisfies Eq. 4.

Now suppose that f is the additive superposition of n motions with velocity $\mathbf{v}_i = (v_{ix}, v_{iy})$, i.e., [1]

$$f(x, t) = f_1(\mathbf{x} - \mathbf{v}_1 t) + \cdots + f_n(\mathbf{x} - \mathbf{v}_n t) \tag{5}$$

The following fact extends the above method to the case of multiple motions.

[1]The case of multiplicative motions can be treated in an analogous way by first taking the logarithm of f.

Fact 2. *Let L be a space-invariant linear operator with kernel g that is n-times differentiable. Then $L(f)$ is also n-times differentiable and the concatenated derivative of $L(f)$ is zero:*

$$\alpha(\mathbf{v}_1) \cdots \alpha(\mathbf{v}_n) L(f) = 0 \tag{6}$$

Proof. From the linearity and spatial invariance of L, we can write $L(f) = f_1 * g + f_2 * g + \cdots + f_n * g$. By the previous discussion for one motion, we know that $\alpha(\mathbf{v}_i)(f_i * g) = 0$ and in consequence $\alpha(\mathbf{v}_1) \cdots \alpha(\mathbf{v}_n)(f_i * g) = 0$, for $i = 1, \dots, n$, which concludes the proof. □

The above result is of much interest since it shows that for any space invariant linear operator the motion parameters for f and $L(f)$ are the same. In the particular case, where the linear operator L is the identity and the image f is $n-$times differentiable, Eq. 6 becomes $\alpha(\mathbf{v}_1) \cdots \alpha(\mathbf{v}_n) f = 0$, which was introduced in [5] for the estimation of the mixed motion parameters of f that are the coordinates of $\alpha(\mathbf{v}_1) \cdots \alpha(\mathbf{v}_n)$ in the canonical basis for differential operators.

We will now show how to estimate the multiple-motion parameters. First, we expand Eq. 6 to

$$\sum_I c_I f * g_I = 0 \tag{7}$$

where $I = (i_1, i_2, \dots, i_n)$ are ordered sequences with elements $i_j \in \{x, y, t\}$ and g_I are linear kernels obtained as partial derivatives of g with respect to the elements in I. The mixed motion parameters c_I are homogeneous symmetric functions of the coordinates of the motion vectors.

Eq. 7 is valid for any filter g. Therefore, in order to estimate the motion parameters, we can look for a set of filters $\{g_k\}_{k=1,\dots,\ell}$ that makes the linear system of equations

$$\sum_I c_I f * g_{k,I} = 0 \tag{8}$$

well posed for the variables c_I.

In what follows we will use simple algebraic arguments, but one can arrive at the same conclusions with classical total least square reasoning. Eq. 8 can be rewritten as

$$\mathbf{L}\mathbf{V} = \mathbf{0} \tag{9}$$

where $\mathbf{L} = (f * g_{k,I})$ and $\mathbf{V} = (c_I)^T$. After multiplying Eq. 9 by $\mathbf{L}^T$ to obtain a $m \times m$ system of equations, we perform a weighted integration along a small neighborhood of the point in question, i.e. a convolution with a kernel $\omega(x)$,

$$\int \mathbf{L}(x)^T \mathbf{L}(x) \mathbf{V}(x) \omega(x)\, dx = \mathbf{0} \tag{10}$$

to make the system well posed. Since we are supposing that the motion vectors are locally constant, we can take $\mathbf{V}$ out of the integral and obtain

$$\mathbf{J}_n \mathbf{V} = 0 \tag{11}$$

where

$$\mathbf{J}_n = \int \mathbf{L}(x)^T \mathbf{L}(x) \omega(x)\, dx \tag{12}$$

We call $\mathbf{J}_n$ the *generalized structure tensor for* n *motions.* Eq. 11 shows that the mixed motion parameters in $\mathbf{V}$ will form an eigenvector related to the zero eigenvalue of $\mathbf{J}_n$ and therefore can be computed from its minors as shown in Section 2. More precisely, we have up to $m = \text{ord}(\mathbf{J}_n)$ different estimates for the mixed motion parameters given by

$$\mathbf{V}_i \propto (M_{im}, -M_{im-1}, \ldots, (-1)^m M_{i1}), \tag{13}$$

where $M_{ij}, i = 1, \ldots, m$ are the minors of $\mathbf{J}_n$.

3.1. Separation of the motion vectors

Now, we show how to recover the motion vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ from their mixed coefficients c_I in $\mathbf{V}$. In order to accomplish this, we remember that c_I are homogeneous symmetric functions of degree less than n of the coordinates of the motion vectors . This observation is the basis for our solution, that we obtain by interpreting $\mathbf{v}_i$ as complex numbers, that is $\mathbf{v}_i = v_{ix} + jv_{iy}$, where $j^2 = -1$. In this case, the motion vectors will be the roots of a complex polynomial $Q_n(z)$ whose coefficients are functions of (c_I):

$$Q_n(z) = z^n - A_{n-1} z^{n-1} + \cdots + (-1)^n A_0 \tag{14}$$

To compute the coefficients, we just note that A_i are homogeneous symmetric functions of degree $n - i$ of $\mathbf{v}_1, ..., \mathbf{v}_n$. For example, the coefficients of $Q_n(z)$ for two and three motions are:

- Two motions: $A_1 = c_{xt} + jc_{yt}$ and $A_0 = c_{xx} - c_{yy} + jc_{xy}$.
- Three motions: $A_2 = c_{xtt} + jc_{ytt}$, $A_1 = c_{xxt} - c_{yyt} + jc_{xyt}$ and $A_0 = c_{xxx} - c_{xyy} + j(c_{xxy} - c_{yyy})$

For more motions, the coefficients of $Q_n(z)$ can be evaluated in analogy.

3.2. Confidence measures

We have shown how to estimate multiple additive motions and now we consider the problem of detecting multiple motions, i.e., we want to quantify the confidence in the assumptions that we made.

In the case of one motion, the confidence is high if one eigenvalue of $\mathbf{J}$ is small and the other two are significant, i.e., rank$(\mathbf{J}) = 2$ [1]. This case excludes regions with aperture problems (two small eigenvalues) and occlusions etc. (three significant eigenvalues). With n motions the confidence is still high if the rank$(\mathbf{J}_n) = m - 1$, where $m = \text{ord}(\mathbf{J}_n)$. Since with our new method we do not compute the eigenvalues, we will in the following define confidence measures that do not need them.

$\mathbf{J}_n$ is a symmetric, positive semidefinite matrix. Since the characteristic polynomial $p(\lambda) = \det(\mathbf{J}_n - \lambda \mathbf{I})$ does not depend on a particular representation of this matrix, the same is true for the following numbers, i.e., they are invariants of $\mathbf{J}_n$:

$$\begin{aligned} K &= \det(\mathbf{J}_n) = \lambda_1 \lambda_2 \cdots \lambda_m \\ S &= \frac{1}{m}(M_{11} + M_{22} + \cdots + M_{mm}) \\ &= \frac{1}{m} \sum_i \lambda_1 \cdots \hat{\lambda_i} \cdots \lambda_m \\ H &= \frac{1}{m}\text{trace}(\mathbf{J}_n) = \frac{1}{m}(\lambda_1 + \lambda_2 + \cdots + \lambda_m) \end{aligned} \tag{15}$$

where $\hat{\lambda_i}$ indicates to skip λ_i. [2]

With the above measures, the confidence criterion translates to $K = 0$ and $S \neq 0$. Before we can compare K with S, we need to know how these numbers scale relative to each other. We found that $K^{1/m} \leq S^{1/m-1} \leq H$. This means that the confidence criterion ($K = 0$ and $S \neq 0$) becomes $K^{1/m} \ll S^{1/m-1}$ or, equivalently, $K^{1/m} < \epsilon S^{1/m-1}$.

4. LOW-COMPLEXITY ALGORITHMS FOR MULTIPLE MOTIONS

Algorithm 1 Single or multiple motion estimation

1: compute $\mathbf{J}_n$ according to Eq. 12
2: **if** $K^{1/m} < \epsilon S^{1/m-1}$ (high confidence) **then**
3: compute $\mathbf{V}_1, \ldots, \mathbf{V}_m$ from the minors of $\mathbf{J}_n$ (Eq. 13)
4: compute the mixed motion parameters $\mathbf{V} = \alpha_1 \mathbf{V}_1 + \cdots + \alpha_m \mathbf{V}_m$
5: **if** n = 1 **then**
6: $\mathbf{v} = (V_x, V_y)$
7: **else**
8: $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are the roots of $Q_n(z)$ in Eq. 14

Our hierarchical algorithm first evaluates the confidence in one motion and estimates that one motion in case of high confidence. Otherwise, the confidence for two motions is evaluated and two motions are estimated. This procedure can be iterated for up to n motions - see Alg. 1.

5. RESULTS

Simulation results are presented in Fig. 1 for a synthetic sequence with one, zero, two, and three motions in the four

[2] We need not compute the eigenvalues; they are used here only to illustrate the invariance of K, S and H.

quadrants. The motions resulted from additive superposition of spatial noise patterns that move in different directions (see results). We used Alg. 1 and stopped after three motions. The kernels $g_{k,I}$ had $k = 1$ and the convolution was performed in the Fourier domain with a cutoff frequency of 0.3 of the maximum frequency. The convolution kernel ω was Gaussian with sigmas of $(2, 2, 1)$ for (x, y, t) respectively. Further, we used $\alpha_i = M_{1i} / \sum_{i=1,..,m} M_{1i}^2$, and $\epsilon = 0.2, 0.3, 0.6$ for one, two, and three motions respectively.

We obtained the following mean errors e with standard deviations σ_e given below (in pixels per frame for the full frame 16) in the format $(e_{\mathbf{v}_x}/\sigma_{e_{\mathbf{v}_x}}, e_{\mathbf{v}_y}/\sigma_{e_{\mathbf{v}_y}})$ for one motion: (-0.003/ 0.015, -0.004/0.019), two motions: (0/0.004, 0.001/0.004), (0/0.003, -0.001/0.005), three motions: (-0.004/ 0.008, 0/0.006), (0/0.007, -0.004/0.008), (0.008/0.026, 0.008/ 0.021). The precision is high, although we used rather small kernels g and ω. However, the issue of optimizing the kernels still needs to be resolved.

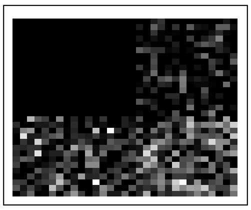
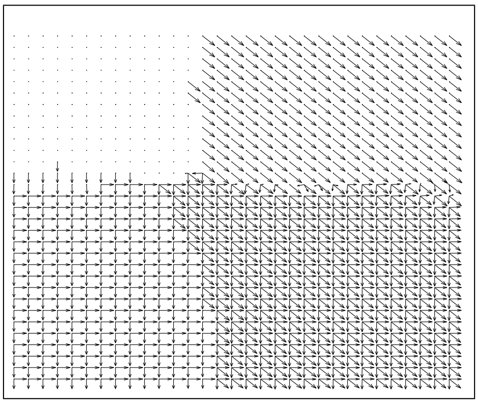


Fig. 1. 30×30 middle section of frame 16 of the $64 \times 64 \times 32$ input sequence (left) and estimated motion vectors (below).



6. SUMMARY AND CONCLUSIONS

We have presented a general framework for estimation of single and multiple motions. The methods we derived are based on derivatives, with an order that increases with the number of motions, but can be generalized to the use of more general linear filters. From the output of these filters we constructed a generalized structure tensor $\mathbf{J}_n$ of size $m = (n + 1)(n + 2)/2$ for n motions and obtained the motion vectors from the minors of $\mathbf{J}_n$. In case of multiple motions, we have shown how to separate the mixed motion parameters by solving for the roots of a complex polynomial. In addition, we have shown how to detect multiple motions, i.e., we have derived confidence measures for the presence of multiple motions.

Our method allows for closed-form solutions for up to four transparent motions, whereas, in case of two and more motions, the eigenvalue problem can only be solved approximately by iterative methods. We should also note that our results were obtained without any additional regularization.

In conclusion, we have presented a novel hybrid method for single and multiple motion estimation.

7. REFERENCES

[1] Horst Haußecker and Hagen Spies, "Motion," in *Handbook of Computer Vision and Applications*, Bernd Jähne, Horst Haußecker, and Peter Geißler, Eds., vol. 2, pp. 309–96. Academic Press, 1999.

[2] E Barth, "The minors of the structure tensor," in *Musterertkennung 2000*, G Sommer, Ed. 2000, pp. 221–228, Springer, Berlin.

[3] Michael J Black and P Anandan, "The robust estimation of multiple motions: parametric and piecewise-smooth flow fields," *Computer Vision and Image Understanding*, vol. 63, no. 1, pp. 75–104, Jan. 1996.

[4] Masahiko Shizawa and Kenji Mase, "Simultaneous multiple optical flow estimation," in *IEEE Conf. Computer Vision and Pattern Recognition*, Atlantic City, NJ, June 1990, vol. I, pp. 274–8, IEEE Computer Press.

[5] Masahiko Shizawa and Kenji Mase, "A unified computational theory for motion transparency and motion boundaries based on eigenenergy analysis," in *IEEE Conf. Computer Vision and Pattern Recognition*, Maui, HI, June 1991, pp. 289–95, IEEE Computer Press.

[6] Weichuan Yu, Kostas Daniilidis, Steven Beauchemin, and Gerald Sommer, "Detection and characterization of multiple motion points," in *IEEE Conf. Computer Vision and Pattern Recognition*, Fort Collins, CO, June 23–25, 1999, vol. I, pp. 171–7, IEEE Computer Press.

[7] E P Simoncelli, "Distributed representation and analysis of visual motion," Tech. Rep. 209, MIT Media Laboratory, Cambridge, MA, 1993.

[8] John Y A Wang and Edward H Adelson, "Representing moving images with layers," *IEEE Transactions on Image Processing*, vol. 3, no. 5, pp. 625–38, 1994.

[9] M V Srinivasan, "Generalized gradient schemes for the measurement of two-dimensional image motion," *Biol Cybernetics*, vol. 63, pp. 421–31, 1990.