Multiple Constraints for Optical Flow

Massimo Tistarelli

University of Genoa
Department of Communication, Computer and Systems Science
Integrated Laboratory for Advanced Robotics (LIRA - Lab)
Via Opera Pia 11A - 16145 Genoa, Italy

Abstract. The computation of the optical flow field from an image sequences requires the definition of constraints on the temporal change of image features. In general, these constraints limit the motion of the body in space and/or of the features on the image plane.
In this paper the implications in the use of multiple constraints in the computational schema are considered. It is shown that differential constraints correspond to an implicit feature tracking. Consequently, the results strictly depend upon the local gray level structure. The best results (either in terms of measurement accuracy and speed in the computation) are obtained by selecting and applying the constraints which are best "tuned" to the particular image feature under consideration.
Several experiments are presented both from a synthetic scene and from real image sequences.

1 Introduction

In every computational schema it is always desiderable to overconstrain the problem. This practice generally allows to more precisely determine the solution. Moreover, it is also important to use redundant informations to enforce robustness with respect to measurement noise in the input data. There are many examples in the literature, where optical flow is computed by means of several constraint equations applied to many image points [1]. The question is: is always the problem posed in the correct way? What is the best way to overconstrain the computational problem and make it *well posed*?

The aim of this paper is to analyze flow constraints commonly used in differential methods to compute optical flow, and understand the relation with the underlaying intensity pattern. The problem is faced in terms of the geometrical properties of the constraint equations in relation with the distribution of the image brightness. The analysis of the equations in terms of the response of the differential operators to different intensity patterns, allows to understand which is the best combination of constraints at a given image point.

2 Motion and optical flow

Uras et al. [2, 3], among others [4, 5], discovered that the aperture problem is a "false problem". In fact, it can be easily overcomed as soon as enough "structure" is present in the image brightness. They suggested to use several second

Lecture Notes in Computer Science, Vol. 800
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94

order derivative operators to compute the optical flow which best approximates different hypothesis. Within a different framework, Nagel [1] developed a set of linear equations in the first and second order derivatives of the image brightness, to compute the optical flow and its derivatives.

For example, assuming the flow field to be locally constant, the time derivative of the gradient can be used, yelding two equations in the two velocity components:

$$\frac{d}{dt}\nabla E = 0 \tag{1}$$

where $E(x, y, t)$ is the image brightness of the point (x, y) at time t.

Several researchers [6, 7, 1, 8] exploited the integration of multiple constraint equations, to overconstrain the computational problem. In particular Tistarelli and Sandini [9] considered a locally constant flow model (not implying any particular movement of the objects in space), adding the brightness constancy equation [10] to obtain three equations in two unknowns:

$$\frac{d}{dt}E = 0 \qquad \frac{d}{dt}\nabla E = 0 \tag{2}$$

In this case the optical flow is computed by solving the over-determined system of linear equations in the unknown terms $(u, v) = \mathbf{V}$. These equations can be solved in closed form, in the least square sense, for each point on the image plane [8].

3 Using multiple constraints

In this section we consider three methods for computing optical flow, using multiple differential constraints[1]:

1) The direct solution of (1) provides the velocity vector $\mathbf{V}$ for each point on the image, where the Hessian matrix of the image brightness is not singular.
2) A unique solution can be obtained, in closed form, by applying a least squares estimator to (2).
3) Another method is based on integrating measurements from neighbouring points in the image.

These methods deserve advantages as well as disadvantages, which can be analyzed by considering a geometrical explanation of the solution. The vector equation (1) corresponds to the intersection of two straight lines in the (u, v), velocity space. On the other hand equations (2) correspond to three straight lines intersecting in, at least, three points. A geometrical explanation of the three considered methods is shown in figure 1. The three plots represent the constraint equations in the velocity (u, v) space according to each different algorithm. In

[1] Even though we are considering three basic constraints, which have been formulated on the basis of the first and second order partial derivatives of the image brightness, the following analysis and results still apply for any number of constraints, at least two, which are somehow related to the image brightness function.

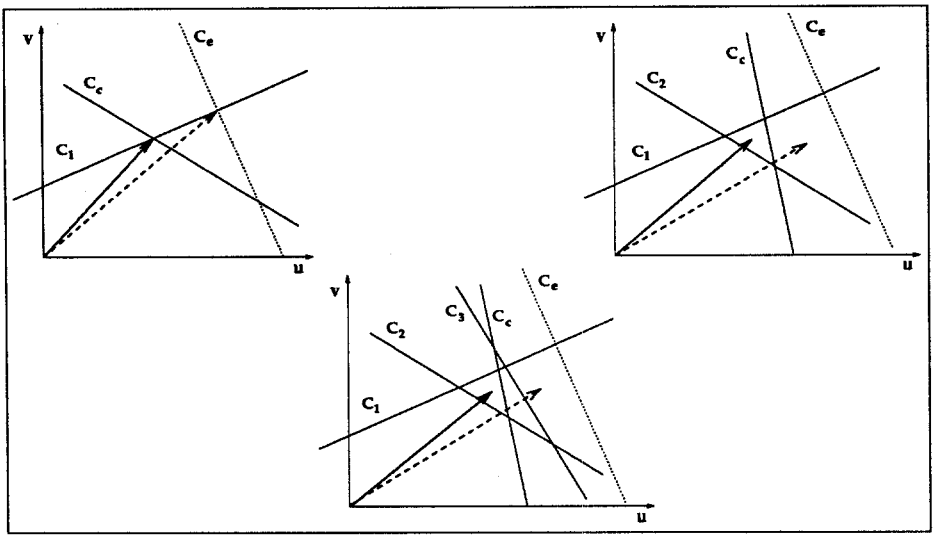

Fig. 1. Effects of the errors in one constraint equation. The wrong constraint is the dotted line C_e while the correct constraint line is C_c. The correct flow vector is bold, while the wrong one is dashed. The geometry of the constraint equations is shown in case of: the constraint equation (1) (upper left); the constraint equation (2) (upper right); multiple data points applying brightness constancy equation (lower).

figure 1 the effect of the errors, represented by a wrong constraint line, is shown. It is interesting to note that in the case of equation (2) it is impossible to determine, at least from the geometry of the system, which one is the wrong constraint line to correct the solution. This is due to the fact that all the lines are considered simultaneously. In the third plot we can notice that the accuracy can be very bad, because the pseudo intersection can be affected by wrong estimates.

3.1 Mapping flow constraints to the image brightness structure

How can we take advantage of the beneficial features of the methods described? Let us consider equation (2). We can rewrite it explicitely:

$$\begin{bmatrix} E_x & E_y \\ E_{xx} & E_{xy} \\ E_{yx} & E_{yy} \end{bmatrix} \cdot \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} E_t \\ E_{xt} \\ E_{yt} \end{bmatrix} \tag{3}$$

Instead of taking all three equations at the same time, it is possible to consider three equation pairs separately and study their stability. This corresponds to consider each intersection in figure 1 separately and determine those which are better conditioned. This is done simply by looking at the matrices[2]:

$$\mathbf{M}_1 = \begin{bmatrix} E_x & E_y \\ E_{xx} & E_{xy} \end{bmatrix} \quad \mathbf{M}_2 = \begin{bmatrix} E_{xx} & E_{xy} \\ E_{yx} & E_{yy} \end{bmatrix} \quad \mathbf{M}_3 = \begin{bmatrix} E_x & E_y \\ E_{yx} & E_{yy} \end{bmatrix} \tag{5}$$

[2] In general we have to consider a set of n functions of the image brightness $\mathbf{F}(E)$

	Algorithm 1: best two equations	Algorithm 2: least squares	Algorithm 3: mean of two intersections
Number of Vectors	33156	31112	26668
Optic Flow timing	15 sec	17.5 sec	15.4 sec

Table 1. Statistics of the three algorithms presented. The equations applied for algorithm 1 are (cfr. equation (5)): M_1: 16960 M_2: 1093 M_3: 15103. The data has been obtained by computing the optical flow on a *Sun* SPARC IPC^{TM} workstation.

A first option is to take as correct solution the intesection corresponding to the equation with the best conditioned matrix (for example with the highest determinant). In this way $\mathbf{V}$ will be given as:

$$\mathbf{V} = \mathbf{M_i}^{-1}\,\mathbf{b}_i \iff \det \mathbf{M}_i > \{\det \mathbf{M_j}\,,\,\det \mathbf{M_k}\}\,, \qquad (6)$$

$$i,j,k \in \{1,2,3\}\,,\quad i \neq j \neq k\,,\quad \mathbf{b}_1 = \begin{bmatrix} E_t \\ E_{xt} \end{bmatrix}\ \mathbf{b}_2 = \begin{bmatrix} E_{xt} \\ E_{yt} \end{bmatrix}\ \mathbf{b}_3 = \begin{bmatrix} E_t \\ E_{yt} \end{bmatrix}$$

In the same way it is possible to solve simultaneously all three equations, from the original set in (2), for each image point. This solution applies only if all three determinants are greater than the given threshold Th. This condition enforces a high reliability for all the three constraint lines.
In figure 2 one image from a sequence of a flat picture approaching the camera is shown. The sequence is composed of 37 images 256×256 pixels with 8 bits of resolution in intensity. In figure 2 (b) the optical flows obtained by applying different methods is shown. As reported in table 1, the vector density of the three flow fields is quite different.

4 *Implicit* feature tracking

In general it is possible to regard any algorthm for the computation of the optical flow field as a solution to the tracking problem. In fact, even though differential methods do not establish any explicit correspondence between image features over time, still the differential operators involved (in any framework) track characteristic intensity patterns over time [11].
This is an interesting aspect which has never been considered in the literature (except for few cases), but dramatically changes the perspective under which differential techniques should be analyzed and applied. Flow and/or motion constraints should not be applied regardless of the image sequence to be processed,

with the constraints:

$$\frac{d}{dt}\mathbf{F}(E) = 0 \qquad \frac{d}{dt}\nabla\mathbf{F}(E) = 0 \qquad (4)$$

where $\nabla\mathbf{F}(E)$ represents the gradient operator applied to each element of the vector of functions $\mathbf{F}(E)$. These constraints result in a set of $3 \times n$ equation pairs.

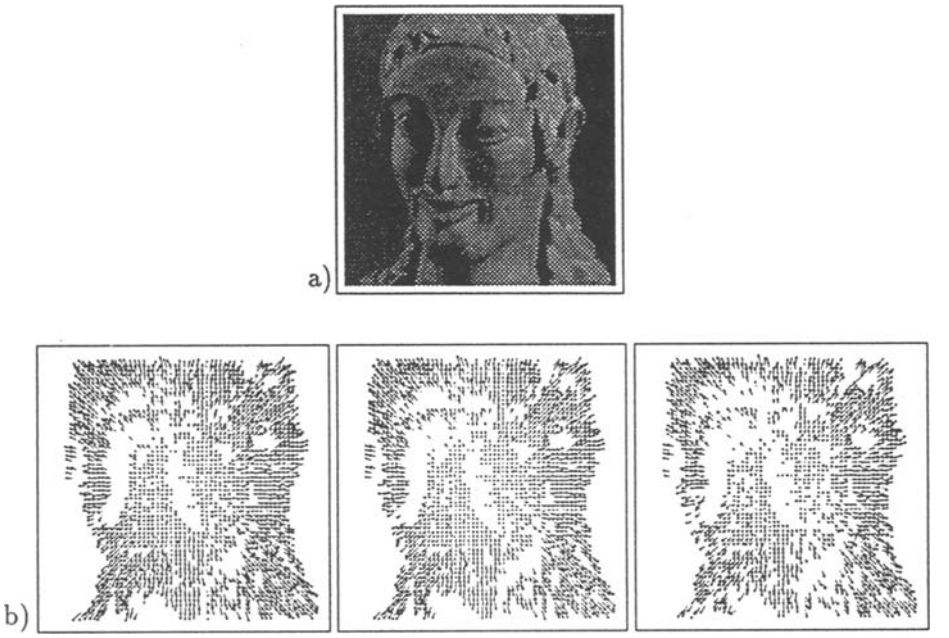

Fig. 2. (a) One image from a set of 37 of a flat picture moving toward the camera along the optical axis. (b) Optical flows computed by: (left) applying the equation pair with the highest determinant; (middle) computing the pseudo-inverse of equations 3; (right) taking the average of the flow vectors computed from the equation pairs with highest determinant.

but rather, purposively selected as a function of the local distribution of the image brightness[3]. In general, this is implicitely performed when tuning thresholds on the parameters of the algorithm or discarding flow vectors with a low confidence (in any way it has been computed in the first place) [11]. It is possible to avoid wong measurements only by tuning the constraints to the underlaying intensity distribution. As shown in the experiments, a side effect of this procedure is the reduction of the processing time.

4.1 Effects of the image intensity distribution

As demonstrated by Nagel [5], the brightness constancy equation and the stationarity of the intensity gradient (SIG) are ill-conditioned when characteristic gray patterns occur on the image. In particular, both are ill-conditioned whenever the gray level variation is linear, while the SIG equations are ill-conditioned

[3] It is worth noting that in the work by Fleet et al. [12] this is performed by applying multiple filtering stages tuned to characteristic frequency bands in the image sequence.

at gray level corners and are best conditioned at maxima of the image brightness. This fact explains why the optical flow, computed by means of the constraint equations expressed in (1), is much less dense than applying all three equations (3)[4]. Therefore, whenever the gray level pattern is not a corner or a maximum, all three equations are somehow "weak", in the sense that equation pairs do not intersect at right angles, or they are weakly well-conditioned.

In figure 3 a synthetic test scene is presented, containing two uniform squares (composing six corners and one gray value extremum) and three peaks: the one in the upper right corner is a Gaussian function, the other two in the lower left corner are a dot (formed by a 3x3 pixels window) and an inverse exponential function ($e^{-\frac{x}{x}}$). The dominance of the determinants for each image point is shown in figure 3 (c). Dark pixels correspond to a maximum for $\det \mathbf{M}_1$, the middle gray identifies pixels where $\det \mathbf{M}_3$ is maximum, and bright pixels correspond to a maximum for $\det \mathbf{M}_2$. The areas where $\det \mathbf{M}_2$ is prevalent are very small and limited to the peaks in the image intensity. It is now evident that not all the equations are "well tuned" to "sense" the temporal variation of every intensity pattern. But, rather, each equation is best suited to compute the motion of a particular intensity distribution.

How can we find the correct solution in this case? Let us consider the problem from another point of view. Suppose the constraint lines to be stochastic processes, where the intersection has a given probability to occur at a given point. Supposing two equation pairs to have a determinant above a given threshold Th (stating the intersection to be admissable), it is possible to consider the probability of the intersections of both line pairs and try to maximize the a posteriori probability. In the (u, v) space this corresponds to move one intersection point toward the other, according to their probability. If the two intersections have the same probability (or the three intersection points, obtained from the three equation pairs, are the vertexes of an ibsoscile triangle), then the most probable solution will be located at a point in the middle of the line connecting the two intersections. It is worth noting that this point does not correspond to the least squares solution.

Taking the value of the determinant for each equation pair, this corresponds to consider as the correct solution the "center of mass" of the two points, where the "mass" of each point is the value of the respective matrix determinant. In general the correct velocity will be given by:

$$\mathbf{V} = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} N_h^u + N_j^u \\ N_h^v + N_j^v \end{bmatrix} \cdot \frac{1}{D_h + D_j} \tag{7}$$

where $N_i^u = b_i^1 M_i^{22} - b_i^2 M_i^{12}$ and $N_i^v = b_i^2 M_i^{11} - b_i^1 M_i^{21}$ are the numerators of the expression for $\mathbf{V}_i = (u_i, v_i)$, as from equation (6), and $D_i = \det \mathbf{M}_i$. In

[4] By applying equations (3) also the corners in the image brightness contribute to the flow computation.

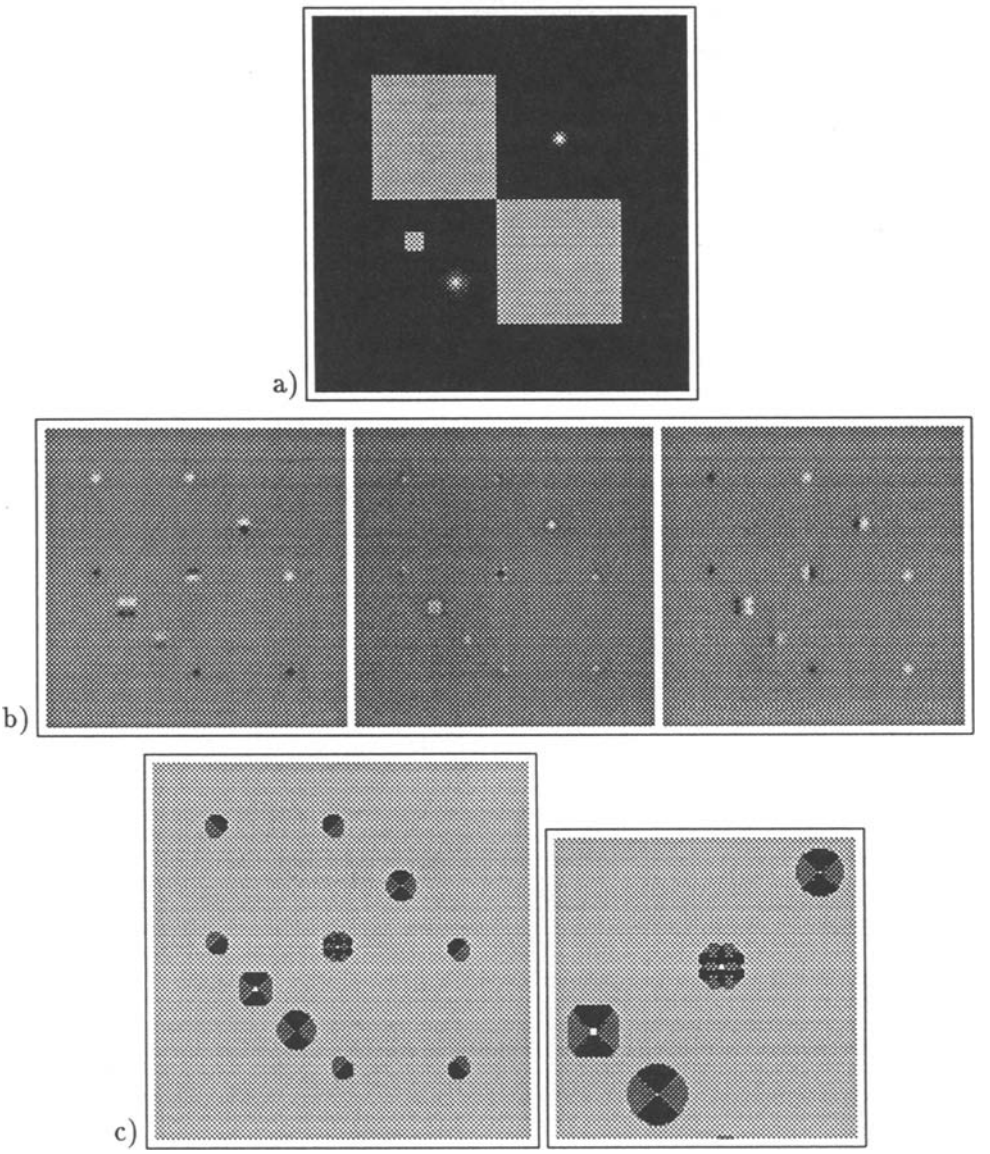

Fig. 3. (a) Synthetic image used to characterize the behaviour of the contraint equations in response to different intensity patterns. (b) Values of the determinants for the three equations (left $\mathbf{M}_1$, middle $\mathbf{M}_2$, right $\mathbf{M}_3$), the gray value codes the determinant values, light gray is 0, dark represent a negative value and bright a positive value. (c) Dominance of the determinants for the constraint equations applied to the test image. White values indicate that $\det \mathbf{M}_2$ is the greatest determinant, gray and black are points where $\det \mathbf{M}_1$ or $\det \mathbf{M}_3$ have the greatest values. The last image on the right shows the absolute value of the difference between the two greatest determinant values. The zero value is coded as bright gray.

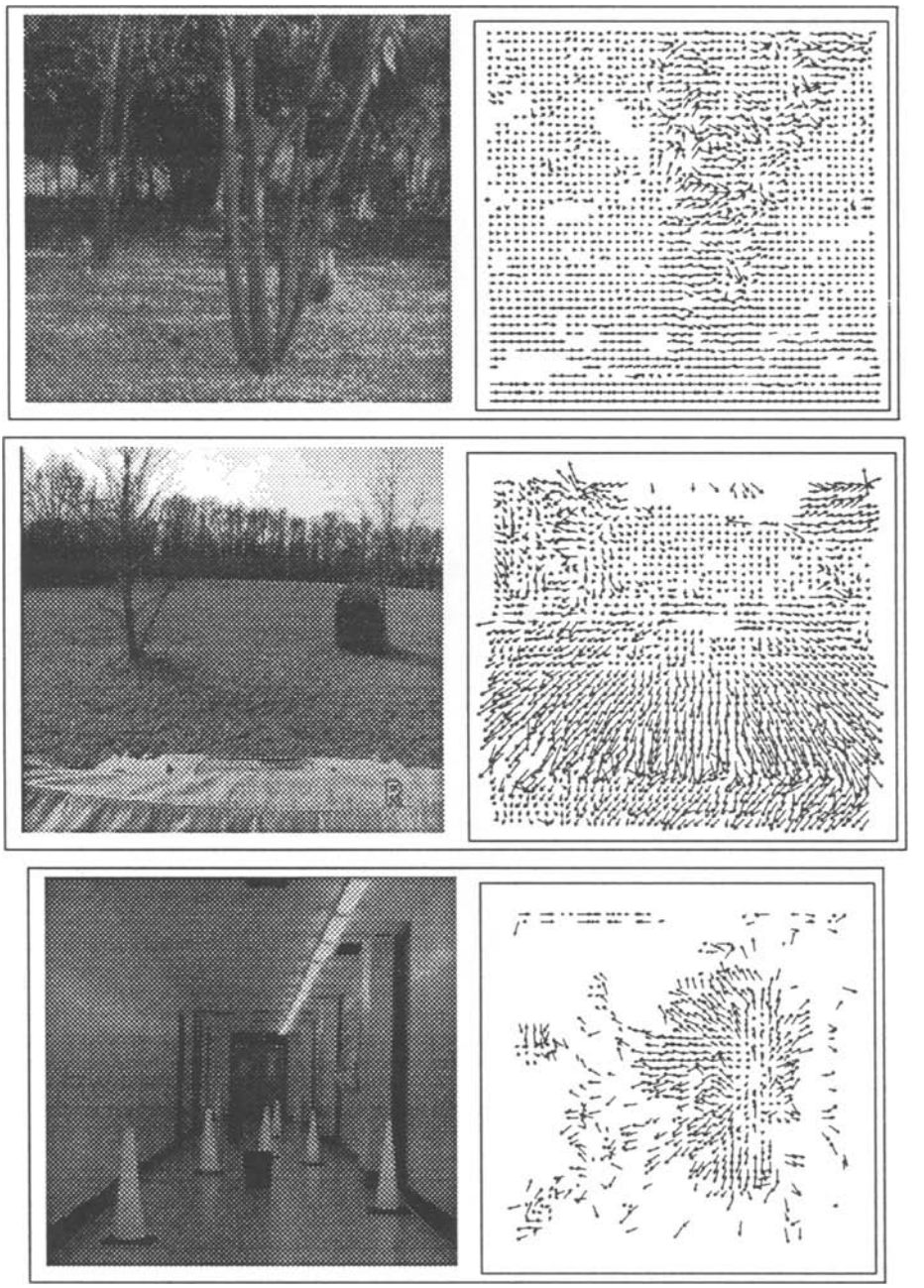

Fig. 4. Optical flows computed by using the weighted sum of the best intersections.

the general case, selected m constraints from the total n equations, we obtain:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_{i=l}^{m+l} N_i^u \\ \sum_{i=l}^{m+l} N_i^v \end{bmatrix} \cdot \frac{1}{\sum_{i=l}^{m+l} D_i} \tag{8}$$

where the selected m constraints are the best tuned to the local intensity profile.

The results of some experiments are reported in figure 4. They are the SRI *tree sequence* (left) a sequence acquired by an outdoor vehicle travelling along a direction almost parallel to the camera optical axis (courtesy of NIST) (middle) and the *cone sequence* acquired at UMASS (right). In the first sequence the camera was moving on the ground plane along a direction orthogonal to the optical axis. It is interesting to note that the pattern constituted by the branches of the foreground tree can not be tracked very well because of occlusions with the background. This is an intrinsic limitations of the constraints applied and could be overcomed by taking into account the occlusions explicitely or being able to detect them from the behaviour of the constraint equations.
The last sequence has been acquired from a camera moving along a corridor. There are two aspects in this sequence: the images have a poor resolution in intensity, about 6 bits (the image shown has been enhanced to improve visibility), and there are few features in the sequence, most of them are lines and corners in the cones, while the walls and ceiling are quite uniform in intensity. Nonetheless the optical flow is computed correctly and the position of the FOE agrees to the trajectory of the camera which can be estimated from the image sequence.

5 Conclusion

In this paper we have addressed the problem of combining multiple constraints to compute the optical flow field from an image sequence.

One of the main aspects which has been outlined in this paper is that the response of a given constraint, strictly depends on the local distribution of the image intensity. Therefore, the choice of the costraints to be applied should depend on the local structure of the image brightness and not only on the confidence associated to the measurement. In fact, there are examples where the local image structure does not allow to apply a given constraint at all, or the information obtained is completely wrong. These observations lead to the conclusion that, in order to compute the optical flow field from an image stream, the constraints to be applied to the image sequence should not be chosen only on the basis of the motion to be detected, but also on the local image structure. Not all the equations are equally suitable for computing the flow field at all image points. We have demonstrated, both analitically and with experiments, that the same equations applied to different brightness structures can give exact or wrong estimates.

At present we have set up a theoretical framework and performed several experiments, providing encouraging results to pursue the research along this direction. We are now making a quantitative analysis of the results, aimed to an evaluation of the real accuracy and reliability of the method.

Acknowledgements

This work has been partially funded by the Esprit project VAP and by the EUREKA project PROMETHEUS (sub-project Pro-Art). Thanks to E. Grosso and G. Sandini for the stimulating discussions and helpful insights in the developement of this work.

References

1. H. H. Nagel. Direct estimation of optical flow and of its derivatives. In G. A. Orban and H. H. Nagel, editors, *Artificial and Biological Vision Systems*, pages 193–224. Springer Verlag, 1992.
2. S. Uras, F. Girosi, A. Verri, and V. Torre. A computational approach to motion perception. *Biological Cybernetics*, 60:79–87, 1988.
3. A. Verri and T. Poggio. Motion field and optical flow: qualitative properties. *IEEE Trans. on PAMI*, PAMI-11:490–498, 1989.
4. H. H. Nagel and W. Enkelmann. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transaction on PAMI*, PAMI-8 1:565–593, 1986.
5. H. H. Nagel. On the estimation of optical flow: Relations between differenet approaches and some new results. *Artificial Intelligence*, 33:299–324, 1987.
6. J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *Proc. of second European Conference on Computer Vision*, pages 237–252, S. Margherita Ligure, Italy, May 19-22, 1992. Springer Verlag.
7. A. Verri, F. Girosi, and V. Torre. Differential techniques for optical flow. *Journal of the Optical Society of America A*, 7:912–922, 1990.
8. M. Tistarelli and G. Sandini. Estimation of depth from motion using an anthropomorphic visual sensor. *Image and Vision Computing*, 8, No. 4:271–278, 1990.
9. M. Tistarelli and G. Sandini. Dynamic aspects in active vision. *CVGIP: Image Understanding*, 56:108–129, July 1992.
10. B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17 No.1-3:185–204, 1981.
11. J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *Int. J. of Computer Vision*, also Tech. Rep. RPL-TR-9107, 1993.
12. D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *Int. J. of Computer Vision*, 5:77–104, 1990.