Improving the Accuracy of Differential–Based Optical Flow Algorithms

R. Manduchi *
University of California at Berkeley
Technical Report UCB//CSD-93-776

November 12, 1993

Abstract

This work deals with the design of the low–level stages of any differential–based optical flow algorithm. The problem is the accurate estimaton of the spatio–temporal derivatives of the moving image, to be used in the solution of the gradient constraint equation. We claim that the "traditional" techniques for the partial derivatives estimation can be improved by i) exploiting all the information provided by both fields within a frame in the case of interlaced scanning systems, ii) adopting prolate spheroidal filters instead of gaussian filters in order to get rid of noise and aliasing, and iii) using larger size differentiators, designed by a weighted least squares technique.

1 Introduction

The measurement of the optical flow in an image sequence is a fundamental task in many Computer Vision and Television applications. Several methods for the optical flow estimation have been proposed. An updated and thorough overview of some of the best known techniques can be found in [1].

The optical flow algorithms may be roughly divided into two classes, namely those which are based on the estimation of the spatio–temporal derivatives of the image sequence, and those which are not (e.g., energy–based velocity techniques). In the present work, only the first class of algorithms are considered.

The effort of the researchers in the optical flow field has been aimed mostly toward the development of methods for solving the aperture problem [2]. Relatively less attention has been payed to the study of low–level techniques for the reliable estimation of the spatio–temporal derivative of the moving image. On

*This work was supported by a felloship offered by the University of Padova, Italy

the other hand, the accuracy of such measurement determines the reliability of the whole optical flow system.

This work presents an accurate analysis of the low–level processing stages of any differential–based optical flow algorithm (consisting of low–pass filters and of differentiators), and devises techniques capable of improving the optical flow estimation accuracy.

It is well known that the estimation of the derivative of a signal is an "ill posed" problem [3]. One cannot expect to obtain accurate results if the signal is noisy and undersampled. It is also true that the accuracy of an optical flow system depends also on other factors (e.g., the aperture problem). On the other side, the present work shows that it can be worth using schemes for the computation of the spatio–temporal derivatives, more sophisticated than the ones commonly adopted. The accuracy improvements attainable by such algorithms (especially for "high" image velocities, where differential optical flow algorithms typically fail to provide reliable results [1]) justifies the increased computational weight.

The computation of the spatio–temporal partial derivatives of an image sequence is subject to errors due to the noise, to the (spatio–temporal) discretization of the signal, and to the quantization. This work addresses mainly the problem of the discretization, and in particular considers the case (typically encountered in the practice) of images taken by an interlaced scanning system.

The issues considered in this work are the following:

- In the case of interlaced scanning systems, multidimensional interpolation techniques for the exploitation of the *whole* image information are described (note that Computer Vision algorithms typically discard one field over two in each interlaced frame, therefore loosing one half of the available information). The accuracy improvement attained in the case of images moving with "high" velocity, or characterized by a wide–band spatial description, is remarkable.

- Regarding the initial low–pass stage (necessary in order to get rid of noise and to reduce the discretization effects), it is shown that gaussian filtering (usually adopted in the Computer Vision practice) may not be the optimal solution. Instead, we propose the use of prolate spheroidal filters, which prove to be more efficient, due to their property of spectral concentration optimality.

- The use of optimal (in a weighted least–squares sense) differentiators, adapted to the initial low–pass stage, is proposed. The performances of such differentiators (especially in the image sequence temporal derivative estimation) is quite superior than those of the simple forward–backward differentiator.

The computational complexity of the proposed filters (interpolators, low–pass and differentiators) has been explicitely taken into account in the work.

In order to describe the improvement, in terms of optical flow estimation accuracy, offered by such techniques, an exhaustive analysis of the discretization effects on the computed optical flow, for both progressive and interlaced scanning systems, is presented in Section 2. Some previous results can be found in [4]. Our approach, though, is more general than the one in [4], and it considers the problem from a multidimensional sampling theory standpoint.

Also in Section 2, we recall some important notions concerning optical flow theory, in order to delineate the condition (namely, of "locally global translation") under which our analysis is consistent.

The techniques proposed for the accuracy improvement (vertico–temporal interpolation for interlaced systems, prolate spheroidal low–pass filtering and adapted differentiators), are described in Section 3.

Experimental tests showing the efficacy of the proposed schemes, in terms of optical flow measurement accuracy, are presented in Section 4.

2 Estimating the optical flow from discretized image sequences

In this section we deal with the problem of the optical flow computation starting from the samples of a discretized image sequence. Some basic notions of optical flow theory are first recalled in 2.1.

In Section 2.2 we address an important problem, namely the effect of the image brightness filtering on the computed optical flow. Such a non–linear problem arises in the analysis of the performances of any optical flow system. A simple case for which the solution is known (the filtering being uninfluential) is the "global translation" case. Any further analysis in this work assumes that such a condition is met, at least locally (as it is tacitally assumed in the analysis of most optical flow algorithms in the literature).

In Section 2.3 we consider the spectral aliasing produced by sampling a continuous moving image on a spatio–temporal lattice (some elementary multidimensional sampling theory notions are first recalled in 2.3.1). Useful bounds for the translational velocity in the cases of progressive and interlaced scanning systems (which allow for aliasing–free measurements) are found. They are related in a simple fashion to the spatial spectral description of the image. The effect of low–pass digital filtering on such bounds is considered as well. Such arguments generalize those found in [4], where the case of a moving one–dimensional sinusoid, using a progressive scanning system and 3–tap differentiators, was considered.

The results of this section enable the comparison between an optical flow system exploiting the information provided by both fields within each frame in an interlaced system, and one which discards one field per frame. It turns out that the first case is highly profitable in terms of robustness against motion–

induced aliasing.

2.1 Basic optical flow theory

The main hypothesis at the basis of the optical flow idea (which will be assumed to hold true throughout this paper) is that of *brightness constancy*. The brightness constancy concept can be stated in terms of an elementary vector field theory notion. Let S be the image space (unless otherwise specified, it will be assumed $S = R^2$, therefore neglecting the effects of the finite size of the "physical" focal plane), and let $l(x, y, t)$ be a differentiable function defined on $S \times R$. The physical meaning of $l(x, y, t)$ is that of brightness intensity signal as a function of the spatial coordinates x, y and of the temporal coordinate t. We will say that the brightness constancy hypothesis is satisfied if there exists a (smooth) vector field $\mathbf{v}(x, y, t) \in R^2$ such that $l(x, y, t)$ is a time–dependent first integral of the differential equation $\begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = \mathbf{v}(x, y, t)$ [5]. In such a case, the gradient constraint equation [2], referred to each point (x_0, y_0, t_0) of $S \times R$, is derived:

$$\nabla l|_{x_0,y_0,t_0} \cdot \mathbf{v}(x_0, y_0, t_0) + \frac{\partial l}{\partial t}|_{x_0,y_0,t_0} = 0 \tag{1}$$

where $\nabla l = \begin{pmatrix} \frac{\partial l}{\partial x} \\ \frac{\partial l}{\partial y} \end{pmatrix}$ is the spatial gradient of $l(x, y, t)$ and symbol "$\cdot$" represents the scalar product.

In the following, the reference to point (x_0, y_0, t_0) will be omitted whenever unambiguous. Furthermore, the notations l_x, l_y, l_t will be used denoting the partial derivatives of $l(x, y, t)$ with respect to x, y and t respectively.

The vector field $\mathbf{v}(x, y, t)$ described above will be called *optical flow associated to* $l(x, y, t)$. It is important to notice that in the considered 2-D case, the optical flow associated to $l(x, y, t)$ is not unique, as it can be easily verified from (1): a 1-D space of solutions for $\mathbf{v}$ may be found (unless $\nabla l = \mathbf{0}$, in which degenerate case any vector $\mathbf{v}$ is a solution of (1)). Such a fact is called *aperture problem* [2]: only the component of the optical flow in the direction of the brightness gradient at each point $\mathbf{x}_0$ of S (hereinafter called the *normal optical flow in* $\mathbf{x}_0$) can be estimated under the above hypothesis, and it is equal to

$$v^{\perp} = \frac{\mathbf{v} \cdot \nabla l}{\| \nabla l \|} = -\frac{l_t}{\| \nabla l \|} \tag{2}$$

(Note that in the 1–D case ($S = R$), the aperture problem does not arise, and the optical flow associated to a given function $l(x, t)$ is equal to $v = -l_t/l_x$ whenever $l_x \neq 0$). The normal optical flow associated to a given function $l(x, y, t)$ characterizes the whole family $\mathcal{V}_l(x, y, t)$ of optical flows associated to $l(x, y, t)$, as the optical flow component orthogonal to the image gradient at each point is arbitrary (given the smoothness condition for $\mathbf{v}(x, y, t)$ is satisfied). As pointed out in [1], in the literature there is a growing interest in the use of the normal optical flow.

In order to obtain a unique solution for the optical flow, it is necessary to regularize the problem. Such a task can be accomplished by imposing constraints on the optical flow, which may be "global" (such as the smoothness constraint of [2] or [6]) or "punctual" (such as the brightness gradient conservation constraint of [7]). Another interesting solution has been more recently proposed in [8]: the chosen optical flow is the intersection of the families $\mathcal{V}_{\tilde{l}1}$ and $\mathcal{V}_{\tilde{l}2}$ of optical flows associated to two differently filtered versions $\tilde{l}_1$ and $\tilde{l}_2$ of l (when it exists and it is unique). In [9] such a technique is extended considering a multi-scale approach for increasing the estimation robustness. Note incidentally that only the method of Verri et al. [7] chooses, for any input $l(x, y, t)$, an optical flow belonging to $\mathcal{V}_l$, i.e., such that it is a solution of (1).

The purpose of the present work is to present efficient solutions for the robust and accurate solution of the gradient constraint equation (1), starting from noisy observations of the brightness intensity signal. No regularization methods are therefore taken into exam, save in the experimental part (Section 4), where a scheme following the lines of [8] is adopted. Indeed, the techniques presented in this work regard only the low–level stages of the system, i.e., the processing of the discrete input signal in order to obtain reliable measurements of the spatio–temporal partial derivatives l_x, l_y and l_t.

The accuracy of the partial derivative measurements determines the accuracy of the whole system, as it is shown in the following. Let $\hat{l}_x, \hat{l}_y$ and $\hat{l}_t$ be the estimated partial derivatives at a given point (x, y, t), affected by errors $e_x = \hat{l}_x - l_x, e_y = \hat{l}_y - l_y$ and $e_t = \hat{l}_t - l_t$. In order to appreciate the effect of such errors, consider for example the estimated normal optical flow $\hat{v}^\perp = -\frac{\hat{l}_t}{\sqrt{\hat{l}_x^2+\hat{l}_y^2}}$. The relation between errors e_x, e_y and e_t and the normal optical flow estimation error $e_v = \hat{v}^\perp - v^\perp$ is not linear, as it is seen from (2). In the case of small errors e_x, e_y and e_t, the linearized form describing the error propagation from the partial derivatives to the estimated normal optical flow is given by terms

$$\frac{\partial v^\perp}{\partial l_x} = -\frac{2\ l_x}{\parallel \nabla i \parallel^2}\, v^\perp \quad , \quad \frac{\partial v^\perp}{\partial l_y} = -\frac{2\ l_y}{\parallel \nabla i \parallel^2}\, v^\perp \quad , \quad \frac{\partial v^\perp}{\partial l_t} = -\frac{1}{\parallel \nabla i \parallel} \tag{3}$$

A few observations may be drawn from (3). First, note that even if $v^\perp = 0$, the estimated normal optical flow may be not null, due to errors in the estimation of l_t (i.e., if $\hat{l}_t \neq 0$). It is perhaps worth recalling that the set of points of the image plane corresponding to null optical flow are of particular interest: consider for example the recovery of motion from the positions of singular points [10].
When $v^\perp \neq 0$, the errors occuring in the computation of the spatial derivatives contribute to e_v, too. In any case, the smallest the magnitude of the spatial gradient of $l(x, y, t)$, the higher the sensitivity to errors e_x, e_y and e_t.

Errors e_x, e_y and e_t are mainly due to i) the presence of noise added to the "true" brightness intensity signal, ii) the fact that only the samples of $l(x, y, t)$

on a discrete grid are given, and iii) the quantization of $l(x, y, t)$. The analysis of the discretization effects (in terms of spectral aliasing) is carried out in Section 2.3.

In order to reduce the effect of noise and discretization, a first low–pass stage is required, and it is present in any optical flow system. Concerning the errors due to the quantization, an analysis of the problem can be found in [11]. Although to the authors' knowledge no adequate technique for recovering such errors has been developed, it is felt that low–pass filtering the image (using floating–point arithmetic) may be effective also for this case.

2.2 Effects of image filtering on the optical flow

As pointed out in Section 2.1, low–pass filtering the input brightness intensity $l(x, y, t)$ is an effective way to improve the accuracy of the system, as it can reduce the effects of the noise, of the aliasing (consequent to the sampling of $l(x, y, t)$) and, hopefully, of the quantization. A linear filtering stage before the differentiation stage is adopted also in the schemes of [8] and [9], in order to solve the aperture problem. A multiscale approach using filter banks has been considered in [4],[9] and [12].

In any case, it is important to consider the effect of the brightness linear filtering on the resulting optical flow. The problem may be stated this way:

Problem 1 *Given a function $l(\mathbf{x}, t)$ and a filtered version $\tilde{l}(\mathbf{x}, t)$ of $l(\mathbf{x}, t)$, how are the optical flow families $\mathcal{V}_l(\mathbf{x}, t)$ and $\mathcal{V}_{\tilde{l}}(\mathbf{x}, t)$ (associated to $l(\mathbf{x}, t)$ and $\tilde{l}(\mathbf{x}, t)$ respectively) related to each other?*

As it can easily be seen from (1), the relationship between $\mathcal{V}_l(\mathbf{x}, t)$ and $\mathcal{V}_{\tilde{l}}(\mathbf{x}, t)$ is not linear, in spite of the fact that the application from $l(\mathbf{x}, t)$ to $\tilde{l}(\mathbf{x}, t)$ is linear. Such an observation follows from the inherent non–linearity of the relationship between $l(\mathbf{x}, t)$ and $\mathcal{V}_l(\mathbf{x}, t)$.

A special case is given by the *global translation* case, i.e., when it is possible to write

$$l(\mathbf{x}, t) = l_0(\mathbf{x} - \mathbf{v}t) \ \ , \ \ (\mathbf{x}, t) \in S \times R \tag{4}$$

for some function $l_0(\mathbf{x}) = l(\mathbf{x}, 0)$ and some vector $\mathbf{v}$. Condition (4) implies that, at each point $(\mathbf{x}, t)$, in the family $\mathcal{V}_l(\mathbf{x}, t)$ of optical flows associated to $l(\mathbf{x}, t)$, there exists a constant optical flow identically equal to $\mathbf{v}$.

It can be easily shown that, in the hypothesis of global translation by vector $\mathbf{v}$, the smoothness constraint method [2] and the brightness gradient conservation method [7] give the same constant optical flow, which is identically equal to $\mathbf{v}$.

The following result concerning the global translation case holds:

Fact 1 *Given a function $l(\mathbf{x}, t)$, satisfying the condition of global translation by vector $\mathbf{v}$, and any filter $h(\mathbf{x}, t)$, the filtered version $\tilde{l}(\mathbf{x}, t) = l * h(\mathbf{x}, t)$ of $l(\mathbf{x}, t)$ satisfies the condition of global translation by the same vector $\mathbf{v}$.*

(The writing $l * h(\mathbf{x}, t)$ stands for the convolutional integral, i.e. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} l(\bar{\mathbf{x}}, \bar{t})\, h(\mathbf{x}-\bar{\mathbf{x}}, t-\bar{t})\; d\bar{\mathbf{x}}\; d\bar{t}$)

Proof An easy proof can be found reasoning in the frequency domain. Let

$$L(\boldsymbol{\omega}_x, \omega_t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} l(\mathbf{x}, t) e^{-j(\boldsymbol{\omega}_x \cdot \mathbf{x} + \omega_t t)} d\mathbf{x}\; dt \tag{5}$$

be the Fourier transform of $l(\mathbf{x}, t)$. In the case of global translation by vector $\mathbf{v}$, it is [13]

$$L(\boldsymbol{\omega}_x, \omega_t) = L_0(\boldsymbol{\omega}_x)\; \delta(\omega_t + \mathbf{v} \cdot \boldsymbol{\omega}_x)/2\pi \tag{6}$$

where $L_0(\boldsymbol{\omega}_x)$ is the 2–D Fourier transform of $l(\mathbf{x}, 0)$ and $\delta(\cdot)$ is the Dirac impulse function. Therefore the support of the Fourier transform of $l(\mathbf{x}, t)$ lies on the plane of equation

$$\omega_t + \mathbf{v} \cdot \boldsymbol{\omega}_x = 0 \tag{7}$$

Let $H(\boldsymbol{\omega}_x, \omega_t)$ be the frequency response of filter $h(\mathbf{x}, t)$, and let $\tilde{L}(\boldsymbol{\omega}_x, \omega_t)$ be the Fourier transform of $\tilde{l}(\mathbf{x}, t)$. It is

$$\begin{aligned} \tilde{L}(\boldsymbol{\omega}_x, \omega_t) &= L(\boldsymbol{\omega}_x, \omega_t) H(\boldsymbol{\omega}_x, \omega_t) = \\ & L_0(\boldsymbol{\omega}_x) H(\boldsymbol{\omega}_x, -\mathbf{v} \cdot \boldsymbol{\omega}_x)\; \delta(\omega_t + \mathbf{v} \cdot \boldsymbol{\omega}_x)/2\pi \end{aligned} \tag{8}$$

Equation(8) shows that the support of $\tilde{L}(\boldsymbol{\omega}_x, \omega_t)$ and the support of $L(\boldsymbol{\omega}_x, \omega_t)$ lye on the same plane. Hence the condition of global translation by vector $\mathbf{v}$ is satisfied also by $\tilde{l}(\mathbf{x}, t)$ $\Box$

A corollary of Fact 1, is that, *in the case of global translation by vector* $\mathbf{v}$*, the normal optical flow associated to any filtered version* $\tilde{l}(\mathbf{x}, t)$ *of* $l(\mathbf{x}, t)$ *is, in general, different from the normal optical flow associated to* $l(\mathbf{x}, t)$. Indeed, at each point $(\mathbf{x}, t)$ of $S \times R$, the normal optical flow is the component of any optical flow vector associated to $l(\mathbf{x}, t)$ in the direction of the brightness gradient. Since the constant optical flow identically equal to $\mathbf{v}$ belongs to both $\mathcal{V}_l(\mathbf{x}, t)$ and $\mathcal{V}_{\tilde{l}}(\mathbf{x}, t)$, one has that the normal optical flows associated respectively to $l(\mathbf{x}, t)$ and $\tilde{l}(\mathbf{x}, t)$ coincide only if the brightness gradients of $l(\mathbf{x}, t)$ and $\tilde{l}(\mathbf{x}, t)$ at each point of $S \times R$ are parallel. But such a property does not hold true in general, due to the image filtering (unless, e.g., the spatial gradient of $l(\mathbf{x}, t)$ is constant).

Another consequence of Fact 1 is that, in the case of global translation by vector $\mathbf{v}$, the optical flows given by the methods of [8] and [9] are identically equal to $\mathbf{v}$.

If the global translation hypothesis does not apply, no general solution to Problem 1, to the authors' knowledge, has been studied (although it is intuitive that the narrower the pass–band of the filter, the larger the "difference" between $\mathcal{V}_l$ and $\mathcal{V}_{\tilde{l}}$). In our opinion, Problem 1 deserves special attention, as the knowledge of the effects of image filtering on the optical flow may give some insight into the behaviour of multiscale approaches [4], [9], [12], and it may provide useful information for the design of the low–pass filtering stage. Such a topic, however, goes beyond the scope of the present work, and we will not deal with it any further.

As to the purpose of the present work, the main indication of this subsection is that, in the hypothesis of global translation, assuming the estimation of l_x, l_y and l_t is not affected by errors, the measured optical flow is independent of the choice of the first stage filter. Such a notion is instrumental for approaching the problem of the low–pass and of the differentator filters design presented in the next sections, where the assumption of global translation will be assumed to be (at least "locally") verified. More specifically, our analysis assumes that, when measuring the optical flow in a point $(\mathbf{x}, t)$, the global translation condition is satisfied within the region around $(\mathbf{x}, t)$ corresponding to the support of the filter kernel. Clearly, in regions corresponding to high optical flow gradient, such a condition is not satisfied, and our analysis fails to be accurate.

The computation of the partial derivatives is performed by three independent differentiator filters along the three axes x, y and t respectively. Usually, the differentiators to be used in practical applications are designed so as to exhibit a low–pass behaviour, in order to remove noise in the high–pass region. In our case, it is seen that *such a low–pass behaviour should not be present independently in the three differentiators.* In other words, the low–pass action is to be performed entirely by the first–stage filter, while the three differentiators should be "as close as possible" to ideal differentiators. If it was not so, one would obtain something different from the partial derivatives of the low–pass filtered version of the input signal. The hypotheses of Fact 1 would not be satisfied, and the computed optical flow would be different from the global translation vector $\mathbf{v}$.

2.3 Sampling lattices and the motion–induced aliasing

In this subsection, the guidelines for the description of the spectral aliasing effects related to any sampling lattice are given, and two commonly adopted sampling lattices are taken as study cases. For a better understanding of the motion–induced aliasing phenomenon, the principal results concerning multidimensional sampling theory are recalled in Section 2.3.1 (for a deeper treatment, see [14] or [15]). In Section 2.3.2 a general model for the aliasing produced by the sampling of a moving image is presented (some preliminary results had already appeared in [16]). The results of this subsection are instrumental for the design of interpolator filters for signals sampled on non–orthogonal lattices (described in Section 3).

2.3.1 Sampling lattices

The brightness signal sampling structure (consequent to the scanning process) can be well modelled by a 3–D lattice. As it is well known, a N–D lattice Λ is the set of points of R^N formed by all the linear combinations by integer coefficients of a set of N linearly independent vectors of R^N. By lining up the components of such vectors with respect to the canonical basis of R^N into an $N \times N$ matrix $\mathbf{A}$ (called *basis* of the lattice), the lattice Λ can be expressed as

$$\Lambda = \left\{ \mathbf{An}, \mathbf{n} \in Z^N \right\}$$

The scanning process by a progressive scanning system can be modelled by the sampling of the brightness signal $l(x, y, t)$ on an *orthogonal* (ORT [14]) lattice (i.e., such that it admits a diagonal basis). The sampling structure is therefore

$$\Lambda_{\mathrm{ORT}} = \left\{ (k\Delta_x, m\Delta_y, nT_F)^T \ , \ k, m, n \in Z \right\} \tag{9}$$

where Δ_x is the distance between two adjacent pixels, Δ_y is the distance between two consecutive scanning lines and T_F is the temporal interval between two consecutive frames.

In the case of a 2:1 interlaced system (adopted by almost all commercial videocameras[1]), the sampling is performed on lattice (ALI [14])

$$\Lambda_{\mathrm{ALI}} = \left\{ (k\Delta_x, m\Delta_y, nT_f)^T \ , \ k, m, n \in Z \ , \text{ either both even or both odd } m, n \right\} \tag{10}$$

In this case, T_f is the temporal interval between the scanning of two subsequent fields, where by *field* it is meant the set of the odd– or of the even–ordered lines in a frame.

In the current practice, Computer Vision algorithms do not exploit the possibilities offered by the interlaced scanning systems; instead, the odd– (or the even–) ordered field of each frame is discarded. Hence, one half of the information provided by the scanning process is lost. It is apparent that algorithms which make use of both fields within a frame may considerably improve the dynamical parameters estimation accuracy.

The sampling of $l(x, y, t)$ on a given lattice Λ induces the periodical repetition of the Fourier transform of $l(x, y, t)$ on the points of lattice $2\pi\Lambda^*$ in the frequency domain, where Λ^* is the *dual lattice* of Λ. More specifically, if $L(\omega_x, \omega_y, \omega_t)$ is the Fourier transform of signal $l(x, y, t)$ (as defined in (5)), then the Fourier transform $\tilde{L}(\omega_x, \omega_y, \omega_t)$ of the sampled version of $l(x, y, t)$ on Λ is [14] (putting $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_t)^T$)

$$\tilde{L}(\boldsymbol{\omega}) = \sum_{\bar{\omega} \in 2\pi\Lambda^*} L(\boldsymbol{\omega} + \bar{\boldsymbol{\omega}}) \tag{11}$$

[1]Some commercial thermocameras adopt 4:1 interlace ratio.

If $\mathbf{A}$ is a basis of Λ, the dual lattice Λ^* admits the basis $\left(\mathbf{A}^{-1}\right)^T$. For the progressive scanning case (9), the frequency repetition lattice is

$$2\pi \cdot \Lambda^*_{\text{ORT}} = 2\pi \cdot \left\{ \left(\frac{k}{\Delta_x}, \frac{m}{\Delta_y}, \frac{n}{T_F} \right)^T \ , \ k, m, n \in Z \right\} \tag{12}$$

while for the 2:1 interlaced scanning case (10) it is

$$2\pi \cdot \Lambda^*_{\text{ALI}} = 2\pi \cdot \left\{ \left(\frac{k}{\Delta_x}, \frac{m}{2\Delta_y}, \frac{n}{2T_f} \right)^T \ , \ k, m, n \in Z \ , \ \text{either both even or both odd } m, n \right\} \tag{13}$$

If the spectral support of $l(x, y, t)$ is completely included in any elementary cell $\mathcal{P}$ of $2\pi\Lambda^*$ (that is, any region such that its versions translated on the points of $2\pi\Lambda^*$ form a non–overlapping tiling of R^N), the sampling of $l(x, y, t)$ on Λ produces no aliasing. In such a case, the information carried by $l(x, y, t)$ is (virtually) completely recoverable from its sampled version.

2.3.2 Motion–induced aliasing

As it will be clear in Section 3, in order to be able of estimating the three partial derivatives that appear in (1), starting from the signal samples on a lattice, it is necessary that an elementary cell of the frequency repetition lattice with a "simple" shape exists, such that the spectral support of the signal prior to sampling is entirely contained in it. In this subsection, we consider two simple polyhedral cells, and determine, given the spatial spectral description of the image, the range of global translation velocities for which such a condition is satisfied.

The case of the orthogonal sampling lattice Λ_{ORT} is taken into exam first. Consider the "rectangular" elementary cell of Λ^*_{ORT}

$$\mathcal{P}^{\text{R}}_{\text{ORT}} = \left\{ (\omega_x, \omega_y, \omega_t)^T , \omega_x \in \left[-\frac{\pi}{\Delta_x}, \frac{\pi}{\Delta_x} \right], \omega_y \in \left[-\frac{\pi}{\Delta_y}, \frac{\pi}{\Delta_y} \right], \omega_t \in \left[-\frac{\pi}{T_F}, \frac{\pi}{T_F} \right] \right\} \tag{14}$$

Assume the condition of global translation by vector $\mathbf{v} = (v_x, v_y)^T$ is satisfied. As shown by (6), the spectral support of $l(x, y, t)$ lies on the plane of equation $v_x\omega_x + v_y\omega_y + \omega_t = 0$. The tilt of the plane is univocally determined by vector $\mathbf{v}$. Let $\mathcal{I}_0$ be the spectral support (on the $\omega_x - \omega_y$ plane) of the 2–D signal $l(x, y, 0)$. The space spanned by the spectral supports of $l(x, y, t)$ for $\mathbf{v}$ ranging in R^2 is the generalized cylinder of equation $\left\{ (\omega_x, \omega_y, \omega_t)^T, (\omega_x, \omega_y)^T \in \mathcal{I}_0 \right\}$.

Suppose, for the time being, that $\mathcal{I}_0$ is contained in $\mathcal{P}^{\text{R}}_{\text{ORT}}$. Such a spatial bandlimitedness is a reasonable hypothesis in most cases, due to the low–pass behaviour of the Modulation Transfer Function [14] (MTF) of the videocamera (which can be regarded to as the transfer function of a spatial filter applied *before* the sampling). It is easy to see that the spectral support of $l(x, y, t)$ is contained in $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$ only for $\mathbf{v}$ belonging to set

$$V^{\mathrm{R}}_{\mathrm{ORT}} = \left\{ \mathbf{v} = (v_x, v_y)^T \ , \ |v_x \omega_x + v_y \omega_y| < \frac{\pi}{T_F} \ , \ (\omega_x, \omega_y)^T \in \mathcal{I}_0 \right\} \tag{15}$$

$V^{\mathrm{R}}_{\mathrm{ORT}}$ indeed represents the set of velocities (henceforth called *allowed velocities*) for which the supports of the spectral repetition centred on $(\omega_x = 0, \omega_y = 0, \omega_t = \pm\frac{2\pi}{T_F})$ do not intersect $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$, so that no aliasing occurs. Note that $V^{\mathrm{R}}_{\mathrm{ORT}}$ depends only on $\mathcal{I}_0$ and on the temporal sampling period T_F.

As an example, consider the case where region $\mathcal{I}_0$ is a disk in R^2 with radius equal to r_s. Such a situation is typically met in the practice, as the camera MTF can usually be considered circularly symmetric [14]. It is easy to see that in such a case it is

$$V^{\mathrm{R}}_{\mathrm{ORT}} = \left\{ \mathbf{v} = (v_x, v_y)^T \ , \ v_x^2 + v_y^2 < \left(\frac{\pi}{r_s T_F} \right)^2 \right\} \tag{16}$$

i.e., $V^{\mathrm{R}}_{\mathrm{ORT}}$ is a disk in R^2 with radius equal to $\frac{\pi}{r_s T_F}$.

The case of an image moving with $\mathbf{v} \notin V^{\mathrm{R}}_{\mathrm{ORT}}$ is exemplified by Fig. 1. The traces, on the $\omega_y - \omega_t$ plane, of the spectral supports of the still image and of the same image in global translation are depicted by bold and thin line respectively. The spectra are periodic on the points of lattice $2\pi\Lambda^*_{\mathrm{ORT}}$. The trace of $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$ on the $\omega_y - \omega_t$ plane is depicted by dotted line. The aliasing corresponding to the spectral repetitions centered in $(\omega_x = 0, \omega_y = 0, \omega_t = \pm\frac{2\pi}{T_F})$ is apparent.

A way to enlarge the set of allowed velocities is to suitably low–pass filter the input signal. Note that the filtering can be performed only *after* the sampling stage (since the "continuous" brightness signal is not available). Therefore, the frequency response of the adopted digital filter is periodic on the points of $2\pi\Lambda^*$, and it can be specified by its values within any elementar cell of $2\pi\Lambda^*$.

The purpose of the filter is to get rid of the aliased spectral repetitions in the rectangular cell $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$. In order to evaluate the effect of the filter on the set of allowed velocities, consider for example an ideal filter whose transfer function (within $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$) is vanishing for $(\omega_x, \omega_y)^T \notin \mathcal{B}_s$ and for $|\omega_t| \geq B_t$, where $\mathcal{B}_s \subset \left[-\frac{\pi}{\Delta_x}, \frac{\pi}{\Delta_x}\right] \times \left[-\frac{\pi}{\Delta_y}, \frac{\pi}{\Delta_y}\right]$ and $B_t < \frac{\pi}{T_F}$. The aliasing is avoided for velocities $\mathbf{v}$ belonging to set

$$V_H = \left\{ \mathbf{v} = (v_x, v_y)^T \ , \ |v_x \omega_x + v_y \omega_y| < \frac{2\pi}{T_F} - B_t \ , \ (\omega_x, \omega_y)^T \in \mathcal{B}_s \right\} \tag{17}$$

From (17) one can see that, for a fixed temporal bandwith B_t, the set V_H can be enlarged indefinitely by suitably narrowing $\mathcal{B}_s$. However, as pointed out previously, a too narrow pass–band region may cause undesired "distortion" of the estimated optical flow. Such an observation has suggested the multiscale approach of [9] for the robust estimation of the optical flow.

In the case $\mathcal{I}_0$ is not contained in $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$, i.e., if $l(x, y, 0)$ is not correctly bandlimited, then the spectral repetitions centred in $(\omega_x = \pm\frac{2\pi}{\Delta_x}, \omega_y = \pm\frac{2\pi}{\Delta_y}, \omega_t = \pm\frac{2\pi}{T_F})$ contributes to the aliasing too. Also in this case, a suitable digital low–pass filtering may be capable of reducing the aliasing effects on the velocity estimate.

Consider now the case of a 2:1 interlaced scanning system. The rectangular cell of $2\pi\Lambda^*_{\mathrm{ALI}}$ is[2]

$$\mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}} = \left\{ (\omega_x, \omega_y, \omega_t), \omega_x \in \left[-\frac{\pi}{\Delta_x}, \frac{\pi}{\Delta_x}\right], \omega_y \in \left[-\frac{\pi}{2\Delta_y}, \frac{\pi}{2\Delta_y}\right], \omega_t \in \left[-\frac{\pi}{T_f}, \frac{\pi}{T_f}\right] \right\} \tag{18}$$

Assume that $\mathcal{I}_0$, the spectral support of $l(x, y, 0)$, is contained in $\mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$. Then no aliasing occurs as long as the image moves by global translaton vector $\mathbf{v} = (v_x, v_y)^T$ belonging to set

$$V^{\mathrm{R}}_{\mathrm{ALI}} = \left\{ \mathbf{v} = (v_x, v_y)^T \ , \ |v_x\omega_x + v_y\omega_y| < \frac{\pi}{T_f} \ , \ (\omega_x, \omega_y)^T \in \mathcal{I}_0 \right\} \tag{19}$$

Fig. 2 shows the traces on the $\omega_y - \omega_t$ plane of a still image (bold line) and of the same image in global translation by vector $\mathbf{v} \in V^{\mathrm{R}}_{\mathrm{ALI}}$, sampled on lattice Λ_{ALI}. The trace of the rectangular cell $\mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$ is depicted by dotted line.

It is instructive to consider the case of an orthogonal sampling structure, obtained by discarding one field per frame in an interlaced scanning system (as usually performed by Computer Vision algorithms). Call $\bar{V}^{\mathrm{R}}_{\mathrm{ORT}}$ the set of allowed velocities in this case. In Fig. 2 the traces of the supports of the spectral repetitions consequent to such a subsampling operation are depicted by dashed line. Since $T_F = 2T_f$ in this case, from (15) and (19) it is seen that, for a given spectral support $\mathcal{I}_0 \subset \left[-\frac{\pi}{\Delta_x}, \frac{\pi}{\Delta_x}\right] \times \left[-\frac{\pi}{2\Delta_y}, \frac{\pi}{2\Delta_y}\right]$ (note that the distance between two scanning lines within a frame becomes $2\Delta_y$ after the subsampling), the sets of allowed velocities $\bar{V}^{\mathrm{R}}_{\mathrm{ORT}}$ and $V^{\mathrm{R}}_{\mathrm{ALI}}$ are related as

$$V^{\mathrm{R}}_{\mathrm{ALI}} = \left\{ 2 \cdot \mathbf{v} \ , \ \mathbf{v} \in \bar{V}^{\mathrm{R}}_{\mathrm{ORT}} \right\} \tag{20}$$

Relation (20) expresses quantitatively the intuitive notion that discarding one field per frames induces the loss of up to one half of the available information.

However, the hypothesis $\mathcal{I}_0 \subset \mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$ is not quite realistic. The spatial spectral support of any real image typically contains frequencies with vertical component beyond $\omega_y = \frac{\pi}{2\Delta_y}$. It is therefore convenient to consider the diamond–shaped elementary cell

$$\mathcal{P}^{\mathrm{D}}_{\mathrm{ALI}} \ = \ \left\{ (\omega_x, \omega_y, \omega_t)^T \ , \ \omega_x \in \left[-\frac{\pi}{\Delta_x}, \frac{\pi}{\Delta_x}\right], \omega_y \in \left[-\frac{\pi}{\Delta_y}, \frac{\pi}{\Delta_y}\right], \right. \tag{21}$$

[2] Note that one can consider also the rectangular cell $\mathcal{P}'^{\mathrm{R}}_{\mathrm{ALI}} = \left\{ (\omega_x, \omega_y, \omega_t)^T, \omega_x \in \left[-\frac{\pi}{\Delta_x}, \frac{\pi}{\Delta_x}\right], \omega_y \in \left[-\frac{\pi}{\Delta_y}, \frac{\pi}{\Delta_y}\right], \omega_t \in \left[-\frac{\pi}{2T_f}, \frac{\pi}{2T_f}\right] \right\}$. However, the choice of $\mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$ turns out to be more suitable for the present analysis.

$$\omega_t \in \left[-\frac{\pi}{T_f}\left(1 - |\omega_y|\,\frac{\Delta_y}{\pi}\right), \frac{\pi}{T_f}\left(1 - |\omega_y|\,\frac{\Delta_y}{\pi}\right)\right]\Bigg\} \tag{22}$$

The correspondent set of allowed velocities will be denoted by $V^{\mathrm{D}}_{\mathrm{ALI}}$. Fig. 3 shows the traces on the $\omega_y - \omega_t$ plane of the spatial support $\mathcal{I}_0$ of a still image with $\mathcal{I}_0 \subset \mathcal{P}^{\mathrm{D}}_{\mathrm{ALI}}$ but $\mathcal{I}_0 \not\subset \mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$ (the trace of $\mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$ is depicted by dotted line), and of the same image in global translation by vector $\mathbf{v} \in V^{\mathrm{D}}_{\mathrm{ALI}}$. The trace of the diamond–shaped cell $\mathcal{P}^{\mathrm{D}}_{\mathrm{ALI}}$ is depicted by dashed line. From Fig. 3 one can notice that if the spatial spectral support of the image contains frequencies with vertical component ω_y larger than $\frac{\pi}{2\Delta_y}$, the set of allowed velocities corresponding to $\mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$ is empty, but not the one corresponding to $\mathcal{P}^{\mathrm{D}}_{\mathrm{ALI}}$ (as long as no vertical component ω_y exceeds value $\frac{\pi}{\Delta_y}$). However, if $\mathcal{I}_0 \subset \mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$, it is always $V^{\mathrm{D}}_{\mathrm{ALI}} \subset V^{\mathrm{R}}_{\mathrm{ALI}}$.

In order to give a quantitative description of such arguments, recall the former example, where $\mathcal{I}_0$ was a disk with radius r_s. Assume $\mathcal{I}_0 \subset \mathcal{P}^{\mathrm{D}}_{\mathrm{ALI}}$. One can show that

$$V^{\mathrm{D}}_{\mathrm{ALI}} = \left\{ \mathbf{v} = (v_x\,, v_y)^T \;,\; v_x^2 + v_y^2 < \left(\frac{\pi}{r_s T_f}\right)^2 \cdot \left(1 - \frac{r_s \Delta_y}{\pi}\right)^2 \right\} \tag{23}$$

In this case, the set of allowed velocities is a disk with radius $\left(\frac{\pi}{r_s T_f}\right)\cdot\left(1 - \frac{r_s\Delta_y}{\pi}\right) = \frac{\pi}{r_s T_f} - \frac{\Delta_y}{T_f}$.

The results obtained in the examples are summarized in Fig. 4. The radius r_v of the disk of the allowed velocities is given as a function of the radius r_s of $\mathcal{I}_0$. It is assumed that $\Delta_x = \Delta_y$ (such a term is denoted by *pixel* in the figure), so that the disk on $\omega_y - \omega_t$ with radius $r_s = \frac{\pi}{\Delta_y}$ is included in $\mathcal{P}^{\mathrm{D}}_{\mathrm{ALI}}$. The variable r_s is given as *cycles/pixel* ("natural" frequencies, not angular! – we skip factor 2π), and r_v is given as *pixels*/T_F. The curve depicted by solid line represents the radius of $V^{\mathrm{D}}_{\mathrm{ALI}}$ (see (23)), while the one by dashed line represents the radius of $V^{\mathrm{R}}_{\mathrm{ALI}}$ (see (19)). The curve depicted by dotted line denotes the radius of the set of allowed velocity $\bar{V}^{\mathrm{R}}_{\mathrm{ORT}}$ for the case of the orthogonal lattice obtained by discarding one fields out of two (see Fig. 2). It should be clear that, for $r_s \geq \pi/2\Delta_y$(=0.25 *cycles/pixel*), sets $V^{\mathrm{R}}_{\mathrm{ALI}}$ and $\bar{V}^{\mathrm{R}}_{\mathrm{ORT}}$ are empty (because aliasing occurs).

The higher robustness with respect to motion–induced aliasing of a system making use of both the fields within a frame is apparent from Fig. 4. In particular, for images with "high" spatial frequencies (e.g., a spatial sinusoid with period less than 4 pixels), the technique of discarding one field per frame necessarily leads to measurements affected by aliasing errors.

3 Techniques for the accuracy improvement

In the previous section we developed an analysis of the effects of the discretization of moving images, and we found some necessary condition on the image velocity for the correct recovering of the optical flow. Here we analyze in detail the scheme for the computation of the partial derivatives that appear in (1), starting from the image sequence samples, and we devise efficient solutions that ensure a high degree of estimation accuracy.

In Section 3.1 the case of an interlaced system is dealt with. Non–separable design specifications for the differentiators, as well as an interpolation scheme, are derived.

The design of efficient low–pass filters for the signal regularization before the differentiation stage is considered in Section 3.2. It is shown that the use of prolate spheroidal filters (which are optimal in terms of spectral concentration) may be more convenient than adopting gaussian filters. Moreover, the first ones offer a higher degree of freedom to the designer.

Finally, in Section 3.3, we take into exam the "heart" of the system, that is, the differentiator filter. A weighted least squares criterion is described, which accounts for the spectral shape of the signal after the first low–pass filter stage. Such a differentiator design is highly feasible for wide–band signals. Hence, it is surely suitable for computing the derivative along the temporal axis, as the spectrum of a moving image spreads over the temporal frequency axis. It is also effective for the computation of the spatial partial derivatives, unless the image has a very poor spectral description, in which case it is shown that a simple forward–backward difference scheme is capable to yield good results.

Before introducing the single topics of this section, let us make one point clear. When we talk about low–pass (or band–pass) filters, here we mean digital FIR (i.e., non–recursive) *linear–phase* filters. The phase linearity is the most natural choice for a low–pass FIR filter (as long as no particular constraints are put [17]). The impulse response $h(\mathbf{x})$ of a multidimensional linear–phase filter (where $\mathbf{x}$ belongs to the definition lattice Λ) enjoies the symmetry property

$$h(\mathbf{x}) = h(2\mathbf{x}_0 - \mathbf{x}) \tag{24}$$

where $\mathbf{x}_0$ is the center of the smallest parallelepiped P_h with edges parallel to the cartesian axes, containing the support of $h(\mathbf{x})$ [18] (note that not necessarily $\mathbf{x}_0 \in \Lambda$). In particular, we consider zero–phase filters (i.e., such that their frequency response is purely real); in such a case, term $\mathbf{x}_0$ in (24) is null.

Assume the definition lattice Λ is orthogonal. It is straightforward that for any zero–phase FIR filter $h(\mathbf{x})$, P_h contains in each edge an odd number of samples of Λ (in the one–dimensional case, that simply means that the filter size of a zero–phase filter is odd). Note that this fact holds true also for non–separable filters.

In the applications of interest, Λ is the spatio–temporal lattice of definition of the signal. It is well–known that zero–phase filters cannot be implemented in practice: the support of the impulse response must belong to the half–space $\{(x,y,t)^T \in R^3\ ,\ t \geq 0\}$ in order for the temporal causality to be satisfied. Therefore, it is necessary to introduce a delay of $(n_0-1)/2$ temporal samples (frames) before the zero–phase filter, where n_0 is the number of samples in any edge of P_h along the temporal axis. In other words, the filter, designed as zero–phase, is implemented as a linear phase filter with $\mathbf{x}_0 = (0,0,T_F(n_0-1)/2)^T$ in (24).

Similar arguments hold for the case of differentiator filters, where (in the one–dimensional case), the antisymmetric property holds, and equation (24) becomes $h(x) = -h(-x)$.

3.1 Dealing with interlaced systems

In this subsection we deal with the estimation of the partial derivatives l_x, l_y and l_t of a signal $l(x,y,t)$, starting from its samples on the non–orthogonal lattice Λ_{ALI} (see Section 2.3). Let's examine first the trivial case of a signal sampled on the orthogonal lattice Λ_{ORT}. We are interested in computing the derivatives at the locations of the sampling points; without loss of generality, we can restrict our attention to the estimation, for instance, of the temporal derivative on the origin ($l_t(0,0,0)$). Recalling that the frequency response of the ideal derivation operator is $D_{\text{id}}(\omega) = -j\omega$, from the theorem of the initial value it is

$$l_t(0,0,0) = -\left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} j\omega_t\ L(\omega_x,\omega_y,\omega_t)\ d\omega_x\ d\omega_y\ d\omega_t \qquad (25)$$

where $L(\omega_x,\omega_y,\omega_t)$ is the Fourier transform of $l(x,y,t)$. The estimation of l_t from the samples of the signal on lattice Λ_{ORT} is performed by means of a digital filter, whose impulse response is defined on Λ_{ORT} as well. Let $D(\omega_x,\omega_y,\omega_t)$ be the frequency response of such a filter. $D(\omega_x,\omega_y,\omega_t)$ is periodic on the points of $2\pi\Lambda^*_{\text{ORT}}$, and it is completely determined by its values in any elementary cell of such lattice. The initial value theorem for sampled signals gives (putting $\boldsymbol{\omega} = (\omega_x,\omega_y,\omega_t)^T$)

$$\hat{l}_t(0,0,0) = \int_{\mathcal{P}} \tilde{L}(\boldsymbol{\omega})D(\boldsymbol{\omega})\ d\boldsymbol{\omega}$$

where $\hat{l}_t(0,0,0)$ is the estimated value for $l_t(0,0,0)$, $\tilde{L}(\boldsymbol{\omega})$ is the periodicized version of $L(\boldsymbol{\omega})$ on the points of $2\pi\Lambda^*_{\text{ORT}}$ (see (11)), and $\mathcal{P}$ is any elementary cell of $2\pi\Lambda^*_{\text{ORT}}$. If the support of $L(\boldsymbol{\omega})$ is entirely contained in $\mathcal{P}$, and if

$$D(\omega_x,\omega_y,\omega_t) = -j\omega_t\ ,\ (\omega_x,\omega_y,\omega_t)^T \in \mathcal{P} \qquad (26)$$

then $\hat{l}_t(0,0,0) = l_t(0,0,0)$, i.e., no estimation error occurs.

In the general case, the filter in (26) is not separable, in the sense that no triplet $(D_x(\omega_x), D_y(\omega_y), D_t(\omega_t))$ exists such that $D(\omega_x,\omega_y,\omega_t) = D_x(\omega_x)D_y(\omega_y)D_t(\omega_t)$. In the particular case $\mathcal{P}$ coincides with the rectangular cell $\mathcal{P}^{\text{R}}_{\text{ORT}}$ (see Section

2.3.2), $D(\omega_x, \omega_y, \omega_t)$ is separable, and its impulse response $d(k\Delta_x, m\Delta_y, nT_F)$ is null for $k \neq 0$ or $m \neq 0$. (This fact explains why we put so much attention on cell $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$ in Section 2.3.2).

Similar reasonings hold for the case of an interlaced system, where the signal is sampled on lattice Λ_{ALI}. In section 2.3.2, two elementary cells of $2\pi\Lambda^*_{\mathrm{ALI}}$ have been considered. In order to design the differentiator, one should first choose the elementary cell more suitable (in terms of robustness against aliasing) to the spectral characteristics of the image sequence (as outlined in Section 2.3.2). The frequency response of the differentiator filter $D(\omega_x, \omega_y, \omega_t)$ (whose impulse response is defined on Λ_{ALI}) should approximate, within the chosen elementary cell, the function $D_{\mathrm{id}}(\omega_x, \omega_y, \omega_t) = -j\omega_t$.

Two observations are necessary at this point. First, note that the filter $D(\omega_x, \omega_y, \omega_t)$ cannot be separable (a filter can be separable only if its frequency response is periodic on the points of an orthogonal lattice). Separability is advantageous in terms of computational complexity; however, several design techniques can be found in the literature (see [18],[17]) for non–separable filters with low overall complexity.

The second point regards the regularity of the ideal frequency response to be approximated (the "target" function). The frequency response $D(\omega_x, \omega_y, \omega_t)$ is periodic on the points of $2\pi\Lambda^*_{\mathrm{ALI}}$. One can easily see that, for any choice of the elementary cell $\mathcal{P}$ of $2\pi\Lambda^*_{\mathrm{ALI}}$, the target function is discontinuous on the boundary of $\mathcal{P}$, except at most in some isolated points. (Also in the case of lattice Λ_{ORT} with the rectangular cell $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$, the target function is discontinuous for $\omega_t = \pm n\frac{\pi}{T_F}$. Such a fact is unavoidable, as it will be discussed in Section 3.3.) The frequency response of any (stable) realizable filter is continuous (actually, it belongs to C^∞). Therefore, the filter $D(\omega_x, \omega_y, \omega_t)$ must exhibit a "low–pass" behaviour (more precisely, it will be $D(\omega_x, \omega_y, \omega_t) = 0$ on the boundary of cell $\mathcal{P}$).

In conclusion, we have seen that the three partial derivative estimates $\hat{l}_x, \hat{l}_y$ and $\hat{l}_t$ may be computed by three non–separable filters. The filters' sizes, in order to obtain a good aproximation of the ideal differentiators, are likely to be wide (as a consequence of the lack of regularity of the target function). Fig 5 (a) shows such a scheme.

Another possible solution is using an interpolation scheme. The idea is the following: interpolate the signal defined on lattice Λ_{ALI} to a signal defined on Λ_{ORT} (note that, putting $T_F = T_f$, it is $\Lambda_{\mathrm{ALI}} \subset \Lambda_{\mathrm{ORT}}$, see (9) and (10)). The derivative estimates are obtained from the interpolated signal by means of simple one–dimensional differentiators (like in the case of the orthogonal sampling lattice). The interpolation is performed by stuffing the original sequence by null samples on the points of $\Lambda_{\mathrm{ORT}}/\Lambda_{\mathrm{ALI}} = \{\mathbf{x} \in \Lambda_{\mathrm{ORT}}\ ,\ \mathbf{x} \notin \Lambda_{\mathrm{ALI}}\}$, and low–pass filtering the new sequence [14]. The impulse response of the interpolator filter is defined on lattice Λ_{ORT}. Its frequency response, within an elementary cell of $2\pi\Lambda^*_{\mathrm{ORT}}$, should approximate the indicator function of an elementary cell of

$2\pi\Lambda^*_{\text{ALI}}$, in order to get correct interpolation. Several design techniques have been studied in the television context [18],[19],[20]. The use of Nyquist filters is highly feasible, in terms of computational complexity, to the interpolation [20]: in such a case, only the new samples (the ones on $\Lambda_{\text{ORT}}/\Lambda_{\text{ALI}}$) are to be computed, the old ones being invariant. In the experimental tests of Section 4, we have adopted a Nyquist filter designed by means of the McClellan transform technique [18], which approximates, within $\mathcal{P}^{\text{R}}_{\text{ORT}}$, the indicator function of cell $\mathcal{P}^{\text{D}}_{\text{ALI}}$.

Choosing the interpolation scheme, the non–separable filtering is performed only once (in the interpolation stage), as it is represented by Fig. 5 (b). Hence, a benefit in terms of computational weight is gained. The price to pay is that now twice as many samples are to be stored (after the interpolation).

3.2 Prolate spheroidal versus gaussian filters

In this subsection we address the problem of the design of the first–stage low–pass filter $H(\omega_x, \omega_y, \omega_t)$. It will be assumed that the filter is separable:

$$H(\omega_x, \omega_y, \omega_t) = H_x(\omega_x)H_y(\omega_y)H_t(\omega_t) \tag{27}$$

The design problem is split into the independent design of the three 1–D filters $H_x(\omega_x)$, $H_y(\omega_y)$ and $H_t(\omega_t)$. Let $H(\omega)$ be the transfer function of any of such filters, and assume, without loss of generality, that the sampling period is unitary, so that $H(\omega)$ is periodic with period equal to 2π.

The low–pass filter $H(\omega)$ can be characterized by its stopband frequency ω_s. $H(\omega)$ should be "as close as possible" to zero for $\omega_s < |\omega| \leq \pi$. On the other side, in the region $0 \leq |\omega| \leq \omega_s$, it is not necessary to put "pass–band" constraints on $H(\omega)$. In fact, assuming global image translation, once the noise and the aliasing are removed, the presence of the filter is uninfluential with respect to the optical flow detection, as shown by Fact 1.

In light of such an observation, a suitable least–squares design criterion is the minimization of the *spectral concentration* [21] of $H(\omega)$, defined as

$$E_s = \frac{\int_{\omega_s}^{\pi} |H(\omega)|^2 \, d\omega}{\int_0^{\pi} |H(\omega)|^2 \, d\omega} \tag{28}$$

For FIR filters the solution to the minimization problem is known: the impulse response of the unit energy FIR filter of given size N whose transfer function $H(\omega)$ minimizes E_s is the discrete prolate spheroidal sequence[3] $\left\{ v_n^{(0)}(N, \frac{\omega_s}{2\pi}) \right\}$ [21],[22] (we use here the terminology adopted in [21]).

In order to find the terms of the prolate spheroidal sequence, one can either use the eigenvector approach described in [21] and [22] or, for reasonable large

[3] The prolate spheroidal functions have already been considered in the Computer Vision context [3]

N, use a closed form approximate solution based on the zero–th order Bessel function [23] (usually called "Kaiser window").

It is instructive to compare the transfer function behaviour of the prolate spheroidal filters with that of the commonly used gaussian filters. Digital gaussian FIR filters are typically designed by means of the window technique [24]. It is well known that, given a gaussian function $g(t) = \frac{2}{\sqrt{\pi}} e^{-t^2/2\sigma^2}$ of variance σ^2, its Fourier transform $G(\omega)$ is gaussian with variance $1/\sigma^2$: $G(\omega) = e^{-\omega^2\sigma^2/2}$. If $|G(\omega)|$ is negligible for $|\omega| > \pi$, then its periodicized version $\bar{G}(\omega) = \sum_{n=-\infty}^{\infty} G(\omega - 2n\pi)$ is such that $\bar{G}(\omega) \simeq G(\omega)$ for $-\pi < \omega < \pi$. Since $G(\omega) < 10^{-2}$ for $|\omega| > 3/\sigma$, we can assume that such an hypothesis holds true for $\sigma \geq 1$. In such a case, then, $G(\omega)$ is approximable within interval $-\pi < \omega < \pi$ by a trigonometric polynomial whose coefficients are obtained by sampling $g(t)$. It is a common practice using a simple rectangular window on the impulse response $g(n)$, and truncating it for $|n| > 3\sigma$. Note incidentally that the quadratic approximation error obtained by such a design technique [24] can be approximated by quantity $E_{\text{des}} = 2 \cdot \int_{3\sigma}^{\infty} |g(t)|^2 \, dt = 4\sigma \cdot \mathrm{erfc}(3)/\sqrt{\pi} \simeq 5 \cdot 10^{-5}\sigma$, where $\mathrm{erfc}(\cdot)$ is the complementary error function.

Given a gaussian filter with variance σ^2, its spectral concentration (28), related to a given stopband frequency ω_s, is given by $E_s = \mathrm{erf}(\omega_s \sigma)$, where $\mathrm{erf}(\cdot)$ is the error function. Assuming, as described above, to adopt the design constraint

$$N = 2 \cdot \lceil 3\sigma \rceil + 1 \tag{29}$$

where symbol $\lceil x \rceil$ represents the least integer greater than or equal to x, and N is the gaussian filter size, it is possible to relate the spectral concentration, referred to the stop–band frequency ω_s, to any given filter size N. Note incidentally the higher degree of freedom offered by the use of prolate spheroidal filters, which are parametrized by the filter order and by the stopband frequency independentely. In the case of gaussian filter design by the window technique, the filter order determines the minimum variance σ^2 as by (29).

Fig. 5 (solid line) shows the spectral concentrations of some gaussian filters versus their size N, for $\omega_s = \pi/6$. The same figure represents by dotted lines the spectral concentrations of the prolate spheroidal filters $\left\{ v_n^{(0)}(N, \frac{1}{12}) \right\}$ [21], for the same set of filter sizes N. It is apparent from Fig. 5 the better performance, in terms of spectral concentration, of the prolate spheroidal filters with respect to the gaussian filters, for the same filter size.

In Fig. 7(a) the transfer function magnitudes of the order 13 gaussian filter with $\sigma = 2$ (solid line) and of the order 13 prolate spheroidal filter $\left\{ v_n^{(0)}(13, \frac{1}{12}) \right\}$ (dotted line) are shown. Note that the filter energy is the same in both cases. The higher spectral concentration of the prolate spheroidal filter with respect to the gaussian filter results apparent from Fig. 7(a). The impulse responses of the two filters are shown in Fig. 7(b).

3.3 Adapted differentiators

The purpose of the differentiation stage is to give accurate estimates of the partial derivatives of the brightness signal, starting from the samples of its low–pass filtered version. In the case of images taken by an interlaced system, it will be assumed to adopt the interpolation scheme of Fig. 5(b).

The derivative of a signal is usually estimated from its samples by interpolation techniques. The estimate is taken as the derivative of an interpolating function (typically a piecewise polynomial function), which fits the given function in two or more adjacent samples (a least–squares fitting is sometimes considered). Smoothness constraints may be imposed on the resulting interpolation function. Examples of interpolation techniques are the Lagrangian, the Hermitian, and the splines interpolation [25].

It is possible to relate the resulting minimax estimation error with the Chebichev norm of the higher order derivatives of the signal and with the sampling interval [25]. Such relationships, however, are not much useful in practice. The *a priori* knowledge of the upper bounds of the higher derivatives is usually not available, neither can such upper bounds be set by linear techniques. The only hypothesis on the signal which can be assumed to hold true is that of bandlimitedness (under the conditions described in Section 2.3). On the other side, if the signal is correctly bandlimited, it is (virtually) possible to compute its "exact" derivative starting from its samples, by means of a suitable digital filter (note that we are interested in the signal derivatives only at the sampling points).

The frequency domain approach to the design of differentiators has been intensively studied (see for instance [23],[26],[27],[28]). The frequency response of the ideal digital full–band differentiator is

$$D_{\mathrm{id}}(\omega) = -j\omega \ , \ -\pi \leq \omega < \pi \tag{30}$$

Several design techniques for filters $D(\omega)$ which approximate (30) minimizing a given norm of $E(\omega) = D(\omega) - D_{\mathrm{id}}(\omega)$ can be found in the Signal Processing literature. Weighted least squares norms, as well as the minimax norm, have been considered. An interesting error criterion has been proposed in [27], which is particularly suitable for IIR differentiators However, in the present work only FIR differentiators are considered, for which the error criterion of [27] reduces to the least squares criterion.

It was early recognized [26] that, due to the discontinuity of the imaginary part of $D_{\mathrm{id}}(\omega)$ at $\pm\pi$, one cannot expect to obtain small approximation errors designing wide–band differentiators. Such an observation is in agreement with the results of [29]. A widely adopted solution to such a problem is to introduce the delay of half a pixel in the filter, in order to shift the phase discontinuity of π rad to the origin, where there is a zero of the transfer function. By such a "trick", the overall error lowers sensibly, as it can be seen in the examples 3.7 and 3.8 of [30]. Unfortunately, the half–sample delay is not admissible for the case of optical flow detectors. As a matter of fact, one has to relate the three partial derivatives of the brightness at the *same* point, in order to solve equation (1).

Another way for limiting the effects of the discontinuity of $D_{\mathrm{id}}(\omega)$, is to reduce the band of interest of the differentiator. Low–pass differentiators can be designed [23],[28],[31],[32], in order to get a simultaneous signal smoothing, which may reduce the effects of noise. However, as pointed out in Section 2.2, in the case of optical flow detectors, the signal must be low–pass filtered *before* the partial derivatives estimation. Indeed, if a low–pass behaviour is exhibited independentely by the single differentiators, the hypotheses of Fact 1 are not satisfied, and the global translation vector is not correctly recovered.

Nevertheless, the spectral shaping carried out by the low–pass filtering stage should be taken into account in the design of the differentiator. Let $H(\omega)$ be the transfer function of the low–pass filter of (27) corresponding to the direction along which the partial derivative is to be computed. The differentiator filter $D(\omega)$ is designed so as to approximate the ideal transfer function $D_{\mathrm{id}}(\omega)$ (30). Clearly, it is convenient that the error $E(\omega) = D(\omega) - D_{\mathrm{id}}(\omega)$ be smaller in regions where the signal is expected to conserve more energy after the low–pass filter. It seems therefore appropriate to use a weighted least squares criterion, i.e., to minimize the quadratic norm of $H(\omega)E(\omega)$:

$$\| H(\omega)E(\omega)) \|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H(\omega)\,(D(\omega)+j\omega))\,|^2\, d\omega \tag{31}$$

If $H(\omega)$ is the transfer function of an FIR filter, it is possible to find a closed form solution to the minimization problem, as shown in the Appendix.

By using such a design criterion in a multifrequency–multiscale approach [4],[9], it is possible to design the differentiators "adapted" to each considered (not necessarily low–pass) filter. The size of the differentiators should be chosen so as to get a small value of $\| H(\omega)E(\omega)) \|^2$. In Tab. 1 the "quadratic weighted errors" $\| H(\omega)E(\omega)) \|^2$ of differentiator filters of different sizes, adapted to a low–pass prolate spheroidal filter $H(\omega)$ of unitary energy with $\omega_s = \pi/3$ and size equal to 11, are shown. Choosing the 7–tap differentiator, the quadratic weighted error is equal to $4.28 \cdot 10^{-5}$.

It is instructive to compare such filters with the 3–tap (forward–backward difference) differentiator usually adopted in this context (see for instance [4]). The impulse response of such a filter is

$$d_3(n) = \begin{cases} \frac{1}{2} & , \quad n = -1 \\ 0 & , \quad n = 0 \\ -\frac{1}{2} & , \quad n = 1 \end{cases} \tag{32}$$

and its frequency response is

$$D_3(\omega) = -j\sin\omega \tag{33}$$

The quadratic weighted error of the 3–tap differentiator (related to the same filter $H(\omega)$) is equal to $3.59 \cdot 10^{-4}$, almost ten times higher than that of the 7–tap adapted differentiator (whose impulse response is shown in Fig. 8(b)). Fig. 8(a) shows the transfer function magnitude of the low–pass filter $H(\omega)$ (solid line), of the 7–tap adapted differentiator (dashed line) and of the 3–tap differentiator (dotted line). From Fig. 8(a) it can be seen that the transfer function of the chosen differentiator approximates pretty well the ideal differentiator transfer function (30) in the energy region of the low–pass filter.

In order to appreciate the versatility of the adapted differentiator design algorithm, consider also the case of a band–pass filter with frequency response $H(\omega)$ (depicted in Fig. 9(a) by solid line) centered on $\omega_0 = \pi/2$ (such a filter may be part of a filter bank for the multifrequency analysis of the input signal [9]). The quadratic weighted errors for adapted differentiators of different size are reported in Tab 2. The transfer function magnitude of the 9–tap adapted differentiator and of the 3–tap differentiator are depicted in Fig. 9(a) by dashed line and by dotted line respectively. In this case, the quadratic weighted error is equal to $4.19 \cdot 10^{-4}$ for the 9-tap adapted differentiator, and it is equal to 0.477 for the 3–tap differentiator. Note the dramatic performance improvement attainable by the adapted differentiator design.

However, the filters designed by minimizing form (31) may not be feasible in all cases. Experimental tests show that, for images characterized by very poor spectral content, the differentiation along the spatial axes by the simple 3–tap differentiator may sometimes give better results than using adapted differentiators of higher order. An explanation of such a phenomenon is outlined in the following.

Let $D(\omega)$ be the filter which minimizes form (31) for a given filter size N_D, and let $d(n)$ be its impulse response. Assume, for semplicity's sake, the input is one–dimensional. It is easy to see that, if the input $f(x)$ to the system is white noise, the filter $D(\omega)$ is optimal among the FIR filters of size N_D, in the sense that it minimizes the power of the error

$$e_d(n) = (f * h * d(n)) - \frac{d}{dx}(f * h(x))\,|_{x=n} \tag{34}$$

where $h(n)$ is the impulse response of filter $H(\omega)$. The writing $f * h(x)$ stands for $\sum_{n=-\infty}^{\infty} f(x-n)h(n)$ (note that $f(x)$ is defined for $x \in R$, while $d(n)$ and $h(n)$ are defined for $n \in Z$).

If the input signal is low–pass (i.e., its spectral density $F(\omega)$ is negligible for $\omega > \omega_l, \omega_l \ll \pi$) , then, unless the stop–band frequency ω_s of $H(\omega)$ is lower than ω_l, it is seen that the power of $e_d(n)$ may not be minimized by the adapted filter $D(\omega)$. If ω_l is small, the behaviour of the 3–tap differentiator can be better than that of adapted filters of higher size. Indeed, recalling (33), one has that, for $x \to 0$, $D_3(\omega) + j\omega = \mathcal{O}(\omega^3)$, i.e., the 3–tap differentiator approximates very well the ideal differentiator for signals with a narrow low-pass band.

This fact becomes apparent by looking at the frequency response of the "wide–band" least–squares optimal differentiator, which minimizes (31) for $H(\omega) = 1$ (i.e., no prefiltering). Assume the size of the differentiator is N_D (N_D must be odd, as stated in the Appendix). Its impulse response is

$$d_{\text{wb}}(n) = \begin{cases} -(-1)^n/n & , \quad -(N_D-1)/2 \leq n < 0 \\ 0 & , \quad n = 0 \\ (-1)^n/n & , \quad 0 < n \leq (N_D-1)/2 \end{cases} \tag{35}$$

Sequence $d_{\text{wb}}(n)$ is shown in Fig. 10(b) for $N_D = 7$, while Fig. 10(a) represents the frequency response of $d_{\text{wb}}(n)$ (dashed line), of the 3–tap differentiator (dotted line), and of the ideal differentiator (30) (solid line). It is apparent from Fig. 10(a) the bad behaviour of filter $d_{\text{wb}}(n)$ in the low–pass frequency region.

A more sophisticated design technique would consider the minimization of (31), with term $H(\omega)$ accounting also for the (supposed known) spectral distribution of the image (a complete treatment of this problem can be found in [28]).

As to our purposes, the differentiation is performed along the three axes x, y and t. If the spatial spectral behaviour of the input image is strictly low–pass, then the spatial partial derivatives can be computed with a good degree of accuracy by the 3–tap differentiators. On the other side, along the temporal axis, depending on the velocity of the objects in the scene, the spectral support of the signal is likely to be wide, as explained in Section 2.3. In this case, the adapted differentiator can provide much more accurate measurements.

Finally, consider the case the optical flow algorithm relies also on the estimation of the second–order derivative of the signal (as in the scheme of [7]). The design of second–order digital differentiators has been studied, for example, in [33] and in [31]. The frequency response of the ideal second–order digital differentiator is

$$D_{\text{id}}(\omega) = -\omega^2 \ , \ -\pi \leq \omega < \pi \tag{36}$$

Note that, in this case, the discontinuity of $D_{\text{id}}(\omega)$ for $\omega = \pm\pi$ is not present. For a given second–order differentiator transfer function $D_2(\omega)$, the error is defined as $E_2(\omega) = D_2(\omega) - \omega^2$. An "adapted" design technique would consider the minimization of

$$\| E_2(\omega)H(\omega) \|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| H(\omega)\left(D_2(\omega) + \omega^2\right)\right|^2 d\omega \tag{37}$$

In the Appendix it is shown how to design the FIR filter $D_2(\omega)$ of given size which minimizes form (37), if $H(\omega)$ is the transfer function of an FIR filter. Fig. 11(a) shows the transfer function magnitude of the low–pass prolate spheroidal filter $H(\omega)$ of unitary energy with $\omega_s = \pi/3$ and size equal to 11 (solid line), and of the 7–tap adapted second–order differentiator (dashed line). The impulse response of the latter is shown in Fig. 11(b).

4 Experimental tests

In this section, the results of some experimental tests concerning the optical flow measurement, using the techniques described in this work, are presented. They confirm the expectations offered by the theoretical arguments of the previous sections.

Before introducing the experimental set–up, let us stress a point, which derives from the observations of Section 2. In testing any system, one would like to have a set of "basis" test functions, that enabled him to draw conclusions on the general behaviour of the system from the knowledge of the system's output to such test signals. In the present case, as pointed out previously, it turns out that the optical flow associated to a brightness function $l(x, y, t)$ is *not* linearly related to $l(x, y, t)$. For example, a system that gives an accurate optical flow estimate for a certain input signal, may yield a large estimation error for a (linearly) filtered version of the same input signal. Therefore, the analysis of the experimental results for optical flow systems is a delicate matter, and devising an effective "basis" test brightness function set remains an open issue.

Optical flow systems can be tested using either artificial images or "natural" images, i.e., taken by a camera in a real–world environment. The second solution is appealing because it usually offers many different situations (occlusions, shading, noise, etc.). On the other side, the knowledge of the "real" optical flow is sometimes problematic (especially for subpixel displacements between two frames), therefore natural images are more suitable to a "qualitative" evaluation of the system. Since the present work addresses the problem of the estimation accuracy under the brightness constancy hypothesis (see Section 2.1), artificial test images have been used for the experiments, as they are more appropriate to such measures.

We have tested the systems with one–dimensional ($l(x, t)$) and two–dimensional ($l(x, y, t)$) moving signals. Note that, for one–dimensional signals, the optical flow errors depend only on the partial derivative estimates, as the aperture problem does not arise in this case. Hence, one–dimensional signals are particularly suitable to test the methods described in this work. In both cases, only global translation motion has been considered.

In order to evaluate the system performance, a bit of post–processing of the system output is necessary. The measured optical flow typically exhibits error "spikes", especially in region of small spatial gradient magnitude (see (3)). In order to correctly analyze the results, it is necessary to get rid of such error impulses. Some sort of thresholding may be set on the measures (unless more sophisticated confidence measures are available [1]). We sidestepped the problem of the threshold choice by adopting a median filter instead. Median filters are known to be effective for the reduction of "spike" noise. The output of the filter at a certain location $\mathbf{x}$ is the median of the values within a small window centered on $\mathbf{x}$. We chose a small window of 3 samples in the one–dimensional case, and a 3×3 samples square in the two–dimensional case. No thresholding was necessary after such an operation. The small window size allows for a good rendition of the system even in regions of high optical flow gradient.

We have not considered any optical flow smoothing after the median filtering, being interested in testing the "local" system accuracy.

4.1 One–dimensional measurements

The input sequence in the one–dimensional case was $l(k,n) = l_0(x - vt)|_{x=k,t=n}$, where

$$l_0(x) = \sin\left(\frac{x^2}{T}\right) \tag{38}$$

with constant T. Such a "chirp" signal is suitable for our tests, as its derivatives and its instantaneous frequency (defined as the derivative of the argument of (38)) can assume values in a large range.

The sequence we used is shown in Fig. 12. The ratio between the standard error deviation and the actual velocity has been measured for different image velocities (such a ratio is called *relative error* in the figures). The results of the tests are presented in the following.

1) Prolate spheroidal filters and 3–tap differentiators (Fig. 13)

The first low–pass stage is a prolate spheroidal filter with length 9 and stop–band frequency $\omega_s = \pi/3$ in the x direction, and with length 7 and stop–band frequency $\omega_s = \pi/2$ in the t direction (see Section 2.3). The partial derivatives are computed by a 3–tap differentiator.

Note that the relative error becomes small for values of the velocity v next to value $v_0 = 1$ *pixel/frame*. Actually, it is easy to see that, for $v = v_0$, no error occurs using *any* algorithm for the partial derivative estimation (as long as the same diffeentiator is used along the x and the t directions).

2) Prolate spheroidal filters and adapted differentiators (Fig. 14)

In this case, the low–pass stage of case 1 has been adopted, but the partial derivatives are computed by length–7 differentiators, adapted to the low–pass filters (according to the method proposed in Section 3.3). The relative errors are now 3 to 10 times smaller than in the case 1.

3) Prolate spheroidal filters and wide–band differentiators (Fig. 15)

In order to show the effectiveness of the adapted differentiator design, the scheme with the low–pass stage of case 2, but with lenght–7 wide–band least–squares optimal differentiators (as described in Section 2.3), instead of adapted differentiators, has been implemented. The relative errors in this case are approximately 4 times higher than in the case of case 1 (3–tap differentiator). This shows the delicacy of the differentiators' design, and confirms the expectations drawn by the arguments of Section 3.3.

4) Gaussian filters and adapted differentiators (Fig. 16)

Here the low–pass stage has been implemented using gaussian filters of the same size as the low–pass filters of the previous cases. According to the design formula (29), the variance of the gaussian filters result equal to 16/9 in the x direction, and to 1 in the t direction. Length–7 differentiators adapted to such filters have been used. The relative errors are 1.5–3 times higher than those of case 2. Hence, the higher spectral concentrations of the prolate spheroidal filters with respect to the gaussian filters (of the same size) result in a higher accuracy degree, as expected from the arguments of Section 3.2.

5) 6) 7) Noise added (Fig. 17, Fig. 18, Fig. 19)

In order to analyze the effect of noise on the system's performance, we have tested the systems of case 1, 2 and 3 with the input signal of Fig. 12, corrupted by additive white gaussian noise. The signal–to–noise ratio was set to 25 dB. The following conclusions can be drawn from the test results: i) the system's performances are better using adapted differentiators than using 3–tap differentiators also for noisy input (in spite of the fact that the 3–tap differentiator exhibits a low–pass behaviour, which might be helpful against wide–band noise – but see Section 2.2); ii) the system with low–pass prolate spheroidal filters behaves better than the one with gaussian filters also for noisy inputs; the difference, however, is small.

8) Interlaced scanning case (Fig. 22)

In the previous cases, the input signal was sampled as if taken by a progressive scanning system. Here we consider the case the input is sampled in a one–dimensional interlaced fashion. In other words, the input sequence may be seen as the output of a two–dimensional interlaced camera for a fixed value of the coordinate x, assuming the image translates along the y direction. In order to compare the results of both cases, assume the input in the previous cases (progressive scanning) was obtained by discarding one field per frame from the output of the interlaced scanning system. In the present case, all the fields are available. The velocity is still expressed as *pixels/frame* (note that here $T_F = 2T_f$).

We have adopted the interpolation technique described in Section 3.1 (Fig. 5(b)). The frequency response of the interpolator filter approximates, within the elementary cell $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$, the indicator function of cell $\mathcal{P}^{\mathrm{D}}_{\mathrm{ALI}}$. The filter has been designed by the McClellan transform technique [18] starting from a half–band (Nyquist) one–dimensional low–pass prototype. The resulting 2–D filter (whose impulse response is represented in Fig. 20) is still a Nyquist filter. The frequency response magnitude of such a filter is shown by contour plot and perspective plot in Fig. 21 (a) and (b) respectively.

The interpolated signal has been processed by the system of case 2. Comparing Fig. 22 and Fig. 14, it is seen that the relative error corresponding to a given velocity v using the interpolation technique is very close to the relative error corresponding to $v/2$ of case 2. This confirms the expectations of Section 2.3.

4.2 Two–dimensional measurements

The input signal used in this tests (shown in Fig. 23) is a two–dimensional version of the one used in the one–dimensional case: $l(k, m, n) = l_0(x - v_x t, y - v_y t)|_{x=k, y=m, t=n}$, where

$$l_0(x, y) = \sin\left(\frac{x^2 + y^2}{T}\right) \tag{39}$$

with constant T. Such a function is frequently adopted as a test signal in the television context, where it is called *zone plate*. Only the global translation case (i.e., constant v_x and v_y) is considered.

In order to solve the aperture problem, we have adopted a scheme similar to the one of Srinivasan [8]. The input signal is spatially filtered by two different filters, whose frequency response magnitudes are represented in Fig. 24 (a) and (b) respectively (the two filters are factorable). The measured optical flow at each point is the intersection of the two optical flow families associated to the two filtered version of the input at the same point (see Section 2.1).

In the experiments, the velocity was set to ($v_x = -0.5$ *pixels/frame*, $v_y = 2.6$ *pixels/frame*). The optical flow fields measured using 3–tap and adapted differentiators are represented in Fig. 25 and Fig. 26 respectively. The low–pass filters are prolate spheroidal as in Section 4.1 – case 1. The better behaviour of the system using adapted differentiators is apparent.

Fig. 27 represents the optical flow field measured from the sequence of the interlaced fields. The interpolator filter (along the y and t directions) is the same as in Section 4.1 – case 8. The accuracy of the measurement is comparable to that of the case of Fig. 26. Note, however, that the velocity in this case is equal to ($v_x = -1$ *pixels/frame*, $v_y = 5.2$ *pixels/frame*), twice as much as in the previous case!

5 Acknowledgment

This work was conducted while at the Robotics Lab in the EECS Department of the University of California at Berkeley. I wish to thank Jitendra Malik for his kind hospitality and for addressing me to several Computer Vision issues.

I am grateful to Guido Maria Cortelazzo (University of Padova, Italy), to Alessandro Verri (University of Genova, Italy), and to Joe Weber (Robotics Lab, UCB) for useful discussions. Gian Antonio Mian (University of Padova. Italy) kindly read the manuscript.

A Appendix

In this Appendix it is shown how to design the FIR filter of length N_D whose transfer function $D(\omega)$ minimizes the weighted least–squares error norm

$$\int_{-\pi}^{\pi} |H(\omega)\,(D(\omega) + j\omega)|^2 \, d\omega \tag{40}$$

where $H(\omega)$ is the transfer function of a given zero–phase FIR filter.

In the following, the word "filter" will mean either the filter impulse response (denoted by small letter) or its frequency response (denoted by the correspondent capital letter), as it will be clear from the context. By *quadratic distance* between two filters $h_1(n)$ and $h_2(n)$ it will be meant the value

$$\sum_{n=-\infty}^{\infty} |h_1(n) - h_2(n)|^2 =$$

$$\text{(by Parseval's theorem)}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_1(\omega) - H_2(\omega)|^2 \;\, d\omega$$

The frequency response $D(\omega)$ of the sought for filter $d(n)$ must be purely imaginary (so that $d(n) = -d(-n)$). This can be proved as follows. Let $D(\omega)$ be the frequency response of a length–N_D filter which minimizes norm (40). Since $H(\omega)$ is real, it is

$$\int_{-\pi}^{\pi} |H(\omega)\,(D(\omega) + j\omega)|^2 \, d\omega = \int_{-\pi}^{\pi} \left| H(\omega) \left(-\overline{D(\omega)} + j\omega \right) \right|^2 d\omega$$

where $\overline{D(\omega)}$ is the complex conjugate of $D(\omega)$. Therefore filter $-\overline{D(\omega)}$ (whose impulse response is $-d(-n)$) minimizes (40). But there cannot exist two different length–N_D filters minimizing (40), because (as it is shown in the following) the solution corresponds to the orthogonal projection of a vector in a Hilbert space onto a subspace. Hence, $D(\omega) = -\overline{D(\omega)}$, i.e., $D(\omega)$ is purely imaginary, as claimed before. This means that $d(n) = -d(-n)$. N_D can be taken odd, and we will write $N_D = 2\bar{N}_D + 1$.

On the converse, since $H(\omega)$ is zero–phase, it is $h(n) = h(-n)$. Therefore its order N_H is odd as well, and we can write $N_H = 2\bar{N}_H + 1$.

The integrand in expression (40) can be rewritten as

$$|H(\omega)D(\omega) + j\omega H(\omega)|^2$$

Let $\mathcal{H}$ be the space of the length–$(N_H + N_D - 1)$ FIR filters and let $\mathcal{H}_1$ be the subspace of filters of $\mathcal{H}$ whose transfer function can be written in the form

$H(\omega)B(\omega)$, where $B(\omega)$ is any length–N_D FIR filter (note that $H(\omega)D(\omega)$ belongs to $\mathcal{H}_1$). It is clear that the dimension of $\mathcal{H}_1$ is N_H. Writing $D_{\rm tar}(\omega) = \sum_{n=-\infty}^{\infty} D_{\rm id}(\omega - 2n\pi)$, where

$$D_{\rm id}(\omega) = \begin{cases} -j\omega & , \quad -\pi \leq \omega < \pi \\ 0 & , \quad \text{otherwise} \end{cases} \tag{41}$$

the problem of minimizing (40) can be restated this way:

> Find the filter of $\mathcal{H}_1$ which minimizes the quadratic distance to $D_{\rm tar}(\omega)H(\omega)$

Since $\mathcal{H}_1$ is a subspace of $\mathcal{H}$, a two–step solution can be devised:

1. Find the filter $A(\omega)$ of $\mathcal{H}$ which minimizes the quadratic distance to $D_{\rm tar}(\omega)H(\omega)$
2. Find the filter $H(\omega)D(\omega)$ of $\mathcal{H}_1$ which minimizes the quadratic distance to $A(\omega)$

Step 1. The length–$(N_H + N_D - 1)$ impulse response $a(n)$ of $A(\omega)$ is obtained by truncating the impulse response of $D_{\rm tar}(\omega)H(\omega)$ [24]. Since $H(\omega)$ is a zero–phase filter, one can write

$$H(\omega) = \sum_{n=0}^{\bar{N}_H} \bar{h}(n)\cos\omega n \ , \ \ \bar{h}(n) = \begin{cases} h(0) & , \quad n = 0 \\ 2h(n) & , \quad n \neq 0 \end{cases}$$

where h(n) is the impulse response of $H(\omega)$. The impulse response $\hat{a}(n)$ of filter $D_{\rm tar}(\omega)H(\omega)$ is

$$\hat{a}(n) = \tag{42}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} -j\omega H(\omega)e^{j\omega n}\ d\omega \tag{43}$$

$$= \frac{1}{2\pi}\sum_{m=0}^{\bar{N}_H} \bar{h}(m)\int_{-\pi}^{\pi} \omega\cos\omega m \sin\omega n\ d\omega \tag{44}$$

$$= \sum_{m=0}^{\bar{N}_H} \bar{h}(m)\cdot \begin{cases} (-1)^{m+n}\frac{n}{n^2-m^2} & , \quad m \neq \pm n \\ \pm\frac{1}{4m} & , \quad m = \pm n \neq 0 \\ 0 & , \quad m = n = 0 \end{cases} \tag{45}$$

and the sought for impulse response $a(n)$ is

$$a(n) = \begin{cases} \hat{a}(n) & , \quad -\bar{N}_H - \bar{N}_D < n < \bar{N}_H + \bar{N}_D \\ 0 & , \quad \text{otherwise} \end{cases}$$

Step 2. It is convenient here to use a vectorial notation for filters: a generic impulse response $g(n)$ of length N will be represented as a vector $\mathbf{g} = (g_1\ g_2 \ldots g_N)^T$

of R^N (written $\mathbf{g} \equiv g(n)$). If $g(n)$ is null for $n < -(N-1)/2, n > (N-1)/2$ (we are interested only in symmetric and antisymmetric impulse response), then $g_n = g(n-(N+1)/2)$, $n = 1, 2, \ldots, N$.

Consider the circulant $(N_H + N_D - 1) \times (N_H + N_D - 1)$ matrix

$$\mathbf{H} = \begin{pmatrix} h(0) & 0 & 0 & \ldots & h(2) & h(1) \\ h(1) & h(0) & 0 & \ldots & h(3) & h(2) \\ & & \vdots & & & \\ h(N_H) & h(N_H-1) & & \ldots & 0 & 0 \\ 0 & h(N_H) & & \ldots & 0 & 0 \\ & & \vdots & & & \\ 0 & 0 & & \ldots & h(1) & h(0) \end{pmatrix}$$

$\mathbf{H}$ is diagunalizable by orthonormal Fourier matrices [17]. The eigenvalues of $\mathbf{H}$ are the samples of the length–$(N_H + N_D - 1)$ DFT $H_{\text{DFT}}(k)$ of the sequence $h(n)$ padded with $(N_D - 1)$ null samples. Assume $\mathbf{H}$ is full rank (i.e., $H_{\text{DFT}}(k) \neq 0$, $\forall k$); in such a case, $\mathbf{H}$ is a basis of $\mathcal{H}$. If $H_{\text{DFT}}(k) = 0$ for some k, it is necessary to increase the dimension of the filter $h(n)$ (by padding it out with null samples) to the smallest value N_H such that $H_{\text{DFT}}(k) \neq 0 \, \forall k$.

Note that, given filter $g(n) \equiv \mathbf{g} \in R^{N_H+N_D-1}$, it is $\mathbf{Hg} \equiv h \circledast g(n)$, where symbol $\circledast$ represents the circular convolution operator [24]. It is easily seen that the subspace $\mathcal{H}_1$ can be represented as

$$\mathcal{H}_1 = \left\{ \mathbf{g} \in R^{N_H+N_D-1} \ , \ \mathbf{g} = \mathbf{Hb} \ , \ \mathbf{b} \in R^{N_H+N_D-1} \ , \ b_n = 0 \text{ for } i > N_H \right\} \tag{46}$$

i.e., the first N_D columns of $\mathbf{H}$ generate $\mathcal{H}_1$.

In order to find the filter of $\mathcal{H}_1$ which minimizes the quadratic distance to $a(n)$ (i.e., the orthogonal projection of $a(n)$ on $\mathcal{H}_1$), it is convenient to find a basis matrix of $\mathcal{H}$ such that the first N_D columns generate $\mathcal{H}_1$, while the other $(N_H - 1)$ columns generate $\mathcal{H}_1^\perp$ (the orthogonal subspace, $\mathcal{H}_1 \oplus \mathcal{H}_1^\perp = \mathcal{H}$). Recalling (46), it is seen that a columnwise Gram–Schmidt orthogonalization of $\mathbf{H}$ is suitable for this purpose (since the subspace spanned by the first N_H columns of $\mathbf{H}$ turns out to be invariant with respect to such an orthogonalization). Let $\bar{\mathbf{H}} = \mathbf{HC}$, with upper triangular $\mathbf{C}$, be the orthogonalized basis matrix, and put $\mathbf{b} = \bar{\mathbf{H}}^{-1}\mathbf{a}$, where $\mathbf{a} \equiv a(n)$ is the filter described in the first step of the algorithm. Then $\bar{\mathbf{H}}\bar{\mathbf{b}}$, with

$$\bar{b}_n = \begin{cases} b_n & , \quad 1 \leq n \leq N_D \\ 0 & , \quad \text{otherwise} \end{cases}$$

is the filter of $\mathcal{H}_1$ which minimizes the quadratic distance to $a(n)$. The sought for filter $\mathbf{d} \equiv d(n)$ whose transfer function $D(\omega)$ minimizes (40) is therefore

$$\mathbf{d} = \mathbf{H}^{-1}\bar{\mathbf{H}}\bar{\mathbf{b}} = \mathbf{C}\bar{\mathbf{b}}$$

The algorithm is easily adaptable to the case the norm to be minimized is

$$\int_{-\pi}^{\pi} \left|H(\omega)\left(D(\omega)+\omega^2\right)\right|^2 d\omega \tag{47}$$

where $H(\omega)$ is the transfer function of a given FIR filter. The only part of the algorithm to be modified concerns the determination of the impulse response $\hat{a}(n)$ in Step 1 (see (42)). In the present case, one needs to find the impulse response of the filter $D_{\text{tar}}(\omega)H(\omega)$, where $D_{\text{tar}}(\omega)$ is defined as before substituting in (41) term ω^2 to term $-j\omega$:

$$\hat{a}(n) = \tag{48}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} -\omega^2 H(\omega)e^{j\omega n}\ d\omega \tag{49}$$

$$= \frac{1}{2\pi}\sum_{m=0}^{\bar{N}_H} \bar{h}(m)\int_{-\pi}^{\pi} -\omega^2 \cos\omega m \cos\omega n\ d\omega \tag{50}$$

$$= \sum_{m=0}^{\bar{N}_H} -\bar{h}(m)\cdot \begin{cases} 2\cdot(-1)^{m+n}\frac{m^2+n^2}{(m^2-n^2)^2} & , \quad m \neq \pm n \\ \frac{1}{4m^2}+\frac{\pi^2}{6} & , \quad m = \pm n \neq 0 \\ \frac{\pi^2}{3} & , \quad m = n = 0 \end{cases} \tag{51}$$

The remaining of the algorithm is completely similar to the former case.

References

[1] J.L. Barron, D.J. Fleet, and S.S Beauchemin. Performances of optical flow techniques. Technical Report 299, Department of Computer Science, University of Western Ontario, July 1992.

[2] B.K.P Horn and B.G. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–204, 1981.

[3] V. Torre and T.A. Poggio. On edge detection. *IEEE Trans. Pattern Anal. Machine Intell.*, 8(2):147–163, March 1986.

[4] R. Battiti, E. Amaldi, and C. Koch. Computing optical flow across multiple scales: an adaptive coarse–to–fine strategy. *International Journal of Computer Vision*, 6(2):133–145, 1991.

[5] V.I. Arnold. *Ordinary Differential Equations*. The MIT Press, Cambridge, Mass., 1990.

[6] H.H. Nagel and W. Enkelmann. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Trans. Pattern Anal. Machine Intell.*, 8:565–593, 1986.

[7] S. Uras, F. Girosi, A. Verri, and V. Torre. A computational approach to motion perception. *Biological Cybernetics*, 60:79–97, 1988.

[8] M.V. Srinivasan. Generalized gradient schemes for the measurement of two–dimensional image motion. *Biological Cybernetics*, 63:421–431, 1990.

[9] J. Weber and J. Malik. Robust computation of optical flow in a multiscale differential framework. In *Proc. ICCV'93*, Berlin, May 1993.

[10] A. Verri, F. Girosi, and V. Torre. Mathematical properties of the two–dimensional motion fields: from singular points to motion parameters. *Journal of the Optical Society of America A*, 6(5):698–712, May 1989.

[11] B. Kamgar-Parsi and B. Kamgar-Parsi. Evaluation of quantization error in computer vision. *IEEE Trans. Pattern Anal. Machine Intell.*, 11(2):929–939, September 1990.

[12] K.M. Uz, M. Vetterli, and D. LeGall. A multiresolution approach to motion estimation and interpolation with application to coding of digital HDTV. In *Proc. IEEE ISCAS'90*, volume 2, pages 1298–1301, New Orleans, 1990.

[13] E.H. Adelson and J.R. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2:284–299, 1985.

[14] E. Dubois. The sampling and reconstruction of time-varying imagery with applications in video systems. *Proceedings of the IEEE*, 73:502–522, April 1985.

[15] F. Kretz and J. Sabatier. Echantillonnage des images de télévision: analyse dans le domaine spatio–temporal et dans le domaine de Fourier. *Ann. Télécommunications*, 36:137–144, February 1984.

[16] G.M. Cortelazzo, R. Manduchi, and C. Monti. On the relationships between motion and aliasing in typical video rasters. In *Proc. IEEE ISCAS'92*, volume 4, San Diego, May 1992.

[17] P.P. Vaidyanathan. *Multirate Systems and Filter Banks.* Prentice Halls, Englewood Cliffs, NJ, 1993.

[18] D.E. Dudgeon and R.M. Mersereau. *Multidimensional Digital Signal Processing.* Prentice–Hall, Englewood Cliffs, 1984.

[19] G. Tonge. Three-dimensional filters for television sampling. Technical Report 112, IBA, 1981.

[20] R. Manduchi, G.M. Cortelazzo, and G.A. Mian. Multistage sampling structure conversion of video signals. To appear on *IEEE Transactions on Circuits and Systems for Video Technology*, 1993.

[21] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty – V: the discrete case. *The Bell System Technical Journal*, 57(5):1371–1430, May–June 1978.

[22] A. Papoulis and M.S. Bertran. Digital filtering and prolate functions. *IEEE Trans. Circuit Theory*, 19(6):674–681, November 1972.

[23] F.F. Kuo and J.F. Kaiser. *System Analysis by Digital Computer.* John Wiley and Sons, New York, 1966.

[24] A.V. Oppenheim and R.W. Schafer. *Digital Signal Processing.* Prentice–Hall, Englewood Cliffs, NJ, 1975.

[25] P.M. Prenter. *Splines and Variational Methods.* John Wiley and Sons, New York, 1975.

[26] L.R. Rabiner and K. Steiglitz. The design of wide–band recursive and nonrecursive digital differentiators. *IEEE Trans. Audio Electroacoust.*, 18(2):204–209, June 1970.

[27] J. Spriet and J. Bens. Optimal design and comparison of wide–band digital on–line differentiators. *IEEE Trans. Acoust., Speech, Signal Processing*, 27(1):46–52, February 1979.

[28] B. Carlsson, A. Ahlèn, and M. Sternad. Optimal differentiation based on stochastic signal models. *IEEE Trans. Signal Processing*, 39(2):341–353, February 1991.

[29] S.D. Conte and C. de Boor. *Elementary Numerical Analysis*. McGraw–Hill, New York, 1980.

[30] T.W. Parks and C.S. Burrus. *Digital Filter Design*. John Wiley and Sons, New York, 1987.

[31] S. Usui and I. Amidror. Digital low–pass differentiation for biological signal processing. *IEEE Trans. Biomed. Eng.*, 29(10):686–693, October 1982.

[32] T. Vieville and O.D. Faugeras. Robust and fast computation of unbiased intensity derivatives in images. In *Proc. EECV'92*, pages 203–211, Santa Margherita Ligure, Italy, May 1992.

[33] Soo-Chang Pei and Jong-Jy Shyu. Eigenfilter design of higher–order digital differentiators. *IEEE Trans. Acoust., Speech, Signal Processing*, 37(4):505–511, April 1989.

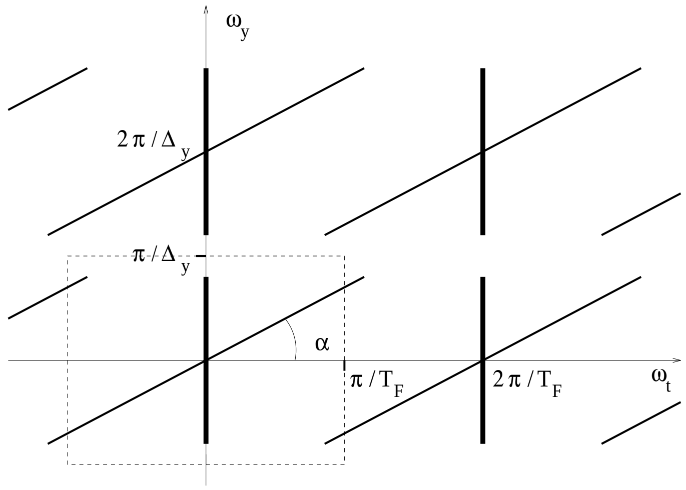


Fig. 1: Traces on the $\omega_y - \omega_t$ plane of the spectral supports of a still image (bold line) and of the same image in global translation by vector $\mathbf{v} = (v_x, v_y)^T$ with $\cot v_y = -\alpha$ (solid line), after sampling on lattice Λ_{ORT}. The trace of the rectangular cell $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$ is depicted by dotted line.

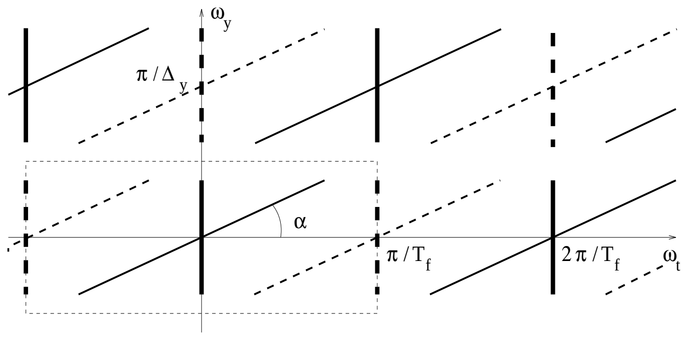


Fig. 2: Traces on the $\omega_y - \omega_t$ plane of the spectral supports of a still image (bold solid line) and of the same image in global translation by vector $\mathbf{v} = (v_x, v_y)^T$ with $\cot v_y = -\alpha$ (solid line), after sampling on lattice Λ_{ALI}. The trace of the rectangular cell $\mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$ is depicted by dotted line. The traces of the supports of the spectral repetitions introduced by discarding one field per frame are depicted by dashed line.

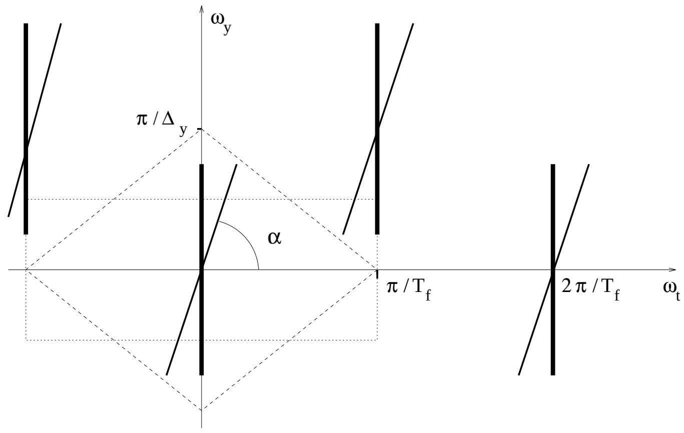


Fig. 3: Traces on the $\omega_y - \omega_t$ plane of the spectral supports of a still image (bold solid line) and of the same image in global translation by vector $\mathbf{v} = (v_x, v_y)^T$ with $\cot v_y = -\alpha$ (solid line), after sampling on lattice Λ_{ALI}. The traces of the rectangular and of the diamond–shaped elementary cells $\mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$ and $\mathcal{P}^{\mathrm{D}}_{\mathrm{ALI}}$ are represented by dotted line and by dashed line respectively.

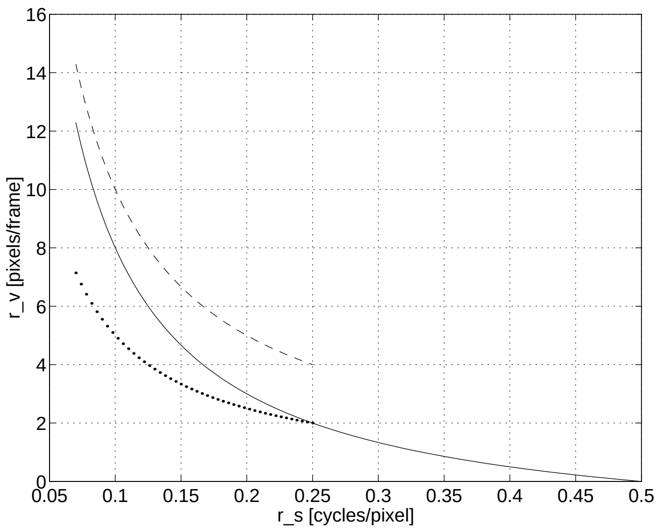


Fig. 4: Radius r_v of the disk of the allowed velocities as a function of the radius r_s of the disk of the image spatial spectral support (see Section 2.3.2). Solid line: case of cell $\mathcal{P}^{\mathrm{D}}_{\mathrm{ALI}}$. Dashed line: case of cell $\mathcal{P}^{\mathrm{R}}_{\mathrm{ALI}}$. Dotted line: case of cell $\mathcal{P}^{\mathrm{R}}_{\mathrm{ORT}}$ (after discarding one field per frame).

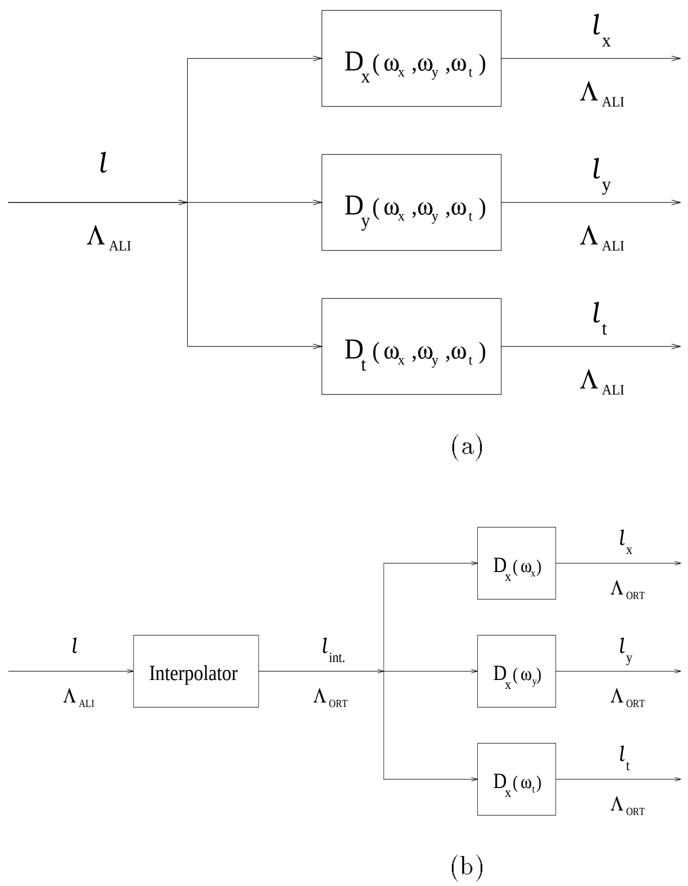


Fig. 5: Schemes for the computation of the partial derivatives from the samples of an image sequence on Λ_{ALI}
(a) Direct implementation
(b) Implementation via interpolation

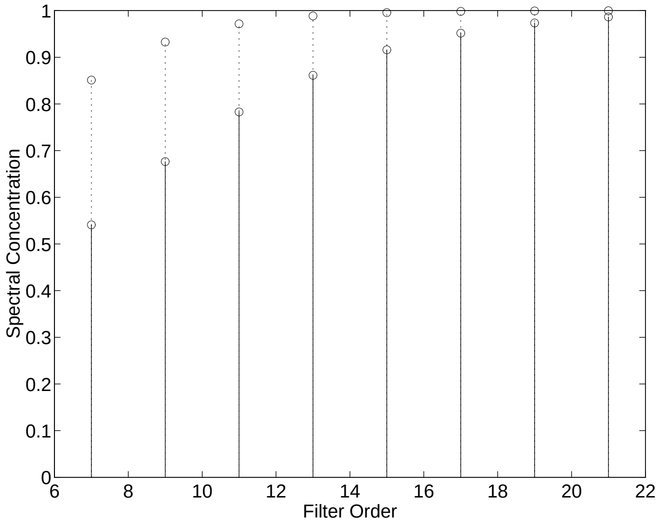
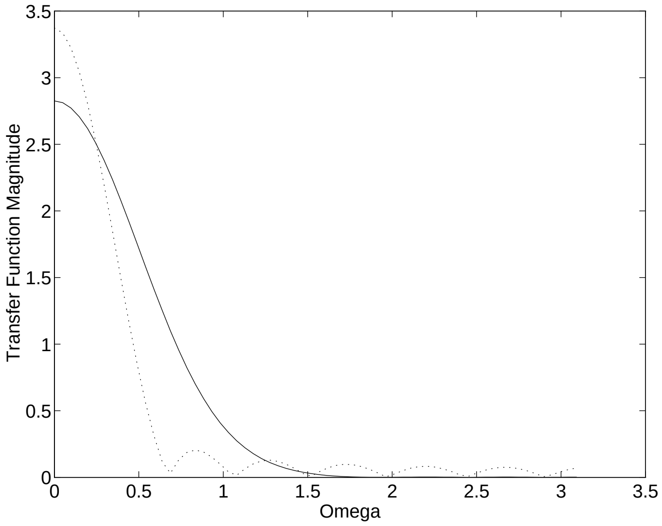
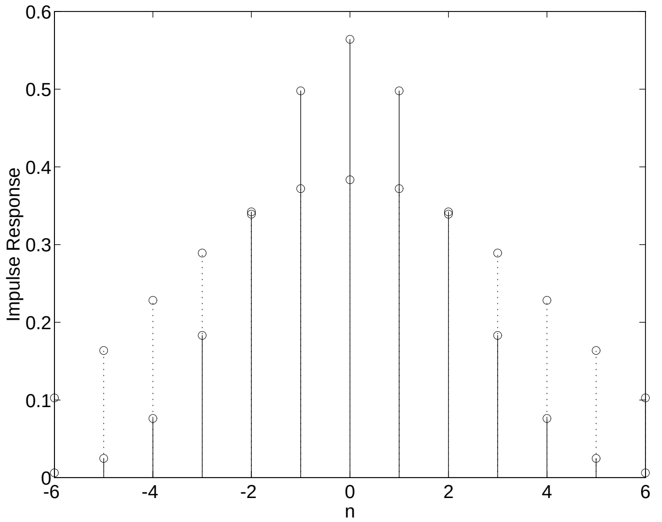


Fig. 6: Spectral concentration of gaussian filters (solid line) and of prolate spheroidal filters (dotted line) versus filter size for stop-band frequency $\omega_s = \pi/6$



(a)



(b)

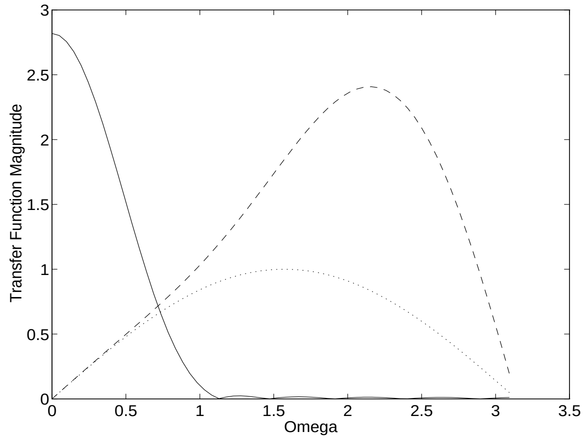
Fig. 7: Transfer function magnitude (a) and impulse response (b) of the 13–tap gaussian filter with $\sigma = 2$ (solid line) and of the prolate spheroidal filter $\left\{v_n^{(0)}(13, \frac{1}{12})\right\}$ (dotted line)

Tab. 1: Weighted error versus filter size for differentiators adapted to the 11–tap prolate spheroidal filter with $\omega_s = \pi/3$ (see Fig. 8)

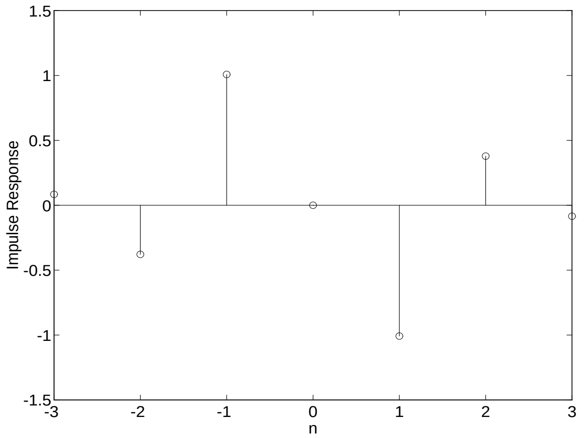
Derivator size	**Weighted error** $\parallel H(\omega)E(\omega)) \parallel^2$
3	$2.0186 \cdot 10^{-4}$
5	$7.5971 \cdot 10^{-5}$
7	$4.2807 \cdot 10^{-5}$
9	$3.2217 \cdot 10^{-5}$

Tab. 2: Weighted error versus filter size for differentiators adapted to the passband filter of Fig. 9

Derivator size	**Weighted error** $\parallel H(\omega)E(\omega)) \parallel^2$
3	0.1028
5	0.0109
7	0.0018
9	$4.19 \cdot 10^{-4}$

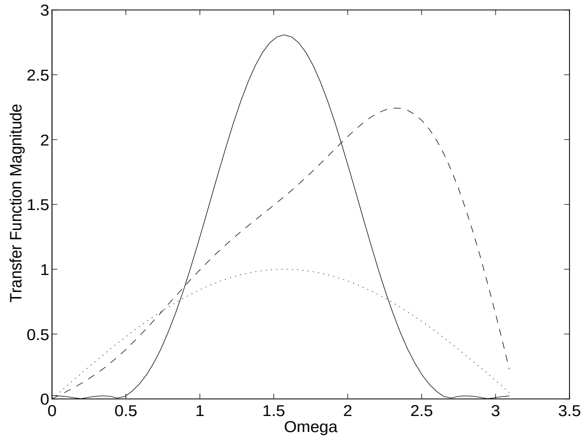


(a)

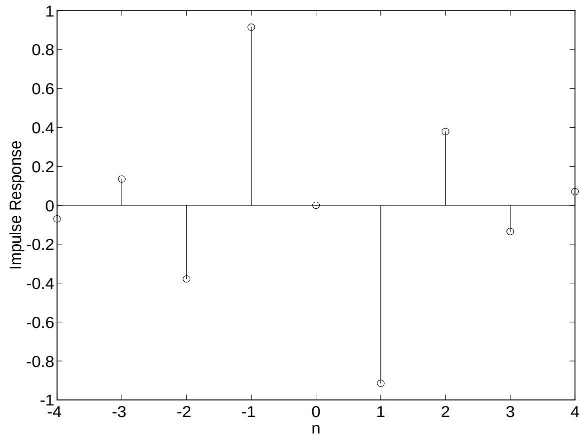


(b)

Fig. 8:(a) Transfer function magnitude of the 11–tap prolate spheroidal filter with $\omega_s = \pi/3$ (solid line), of the 7–tap adapted differentiator (dashed line) and of the 3–tap differentiator (dotted line); (b) Impulse response of the 7–tap adapted differentiator

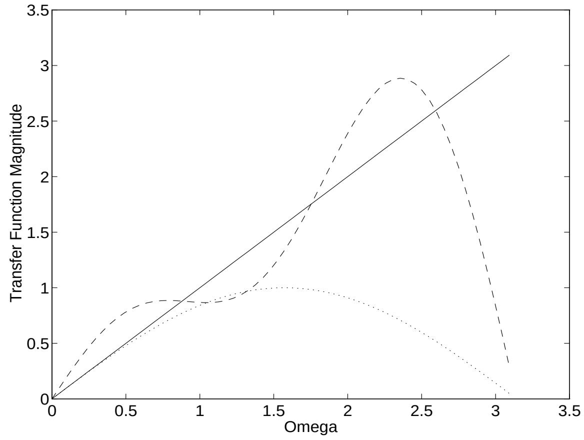


(a)

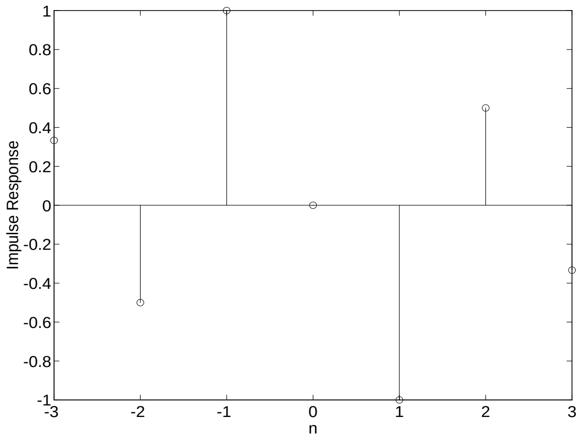


(b)

Fig. 9:(a) Transfer function magnitude of a passband filter (solid line), of the 9–tap adapted differentiator (dashed line) and of the 3–tap differentiator (dotted line); (b) Impulse response of the 9–tap adapted differentiator

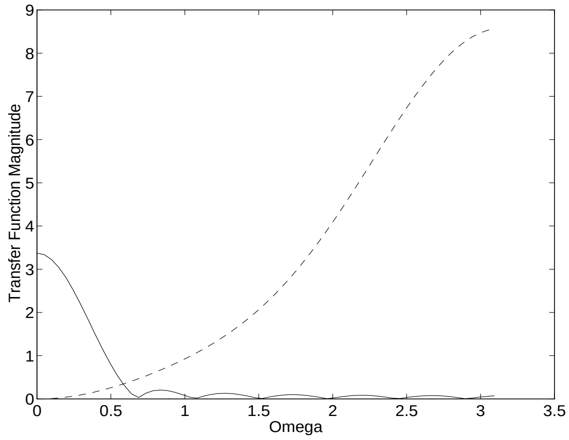


(a)

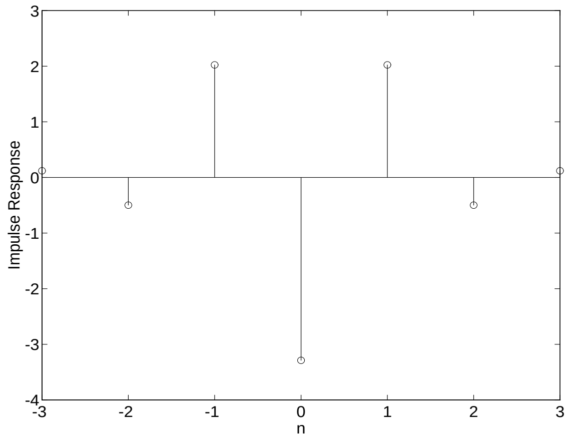


(b)

Fig. 10: (a) Transfer function magnitude of the ideal differentiator (solid line), of the least–squares optimal wide–band differentiator (dashed line) and of the 3–tap differentiator (dotted line)
(b) Impulse response of the least–squares optimal wide–band differentiator



(a)



(b)

Fig. 11:(a) Transfer function magnitude of the 11–tap prolate spheroidal filter with $\omega_s = \pi/3$ (solid line) and of the 7–tap adapted second–order differentiator (dashed line); (b) Impulse response of the 7–tap second–order adapted differentiator

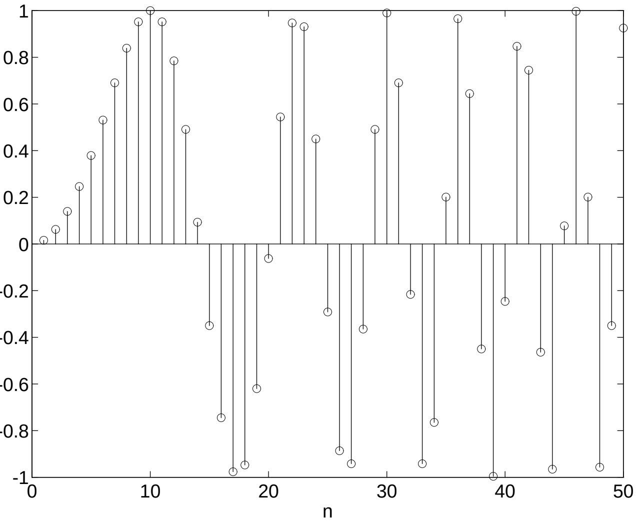


Fig. 12: One–dimensional test sequence

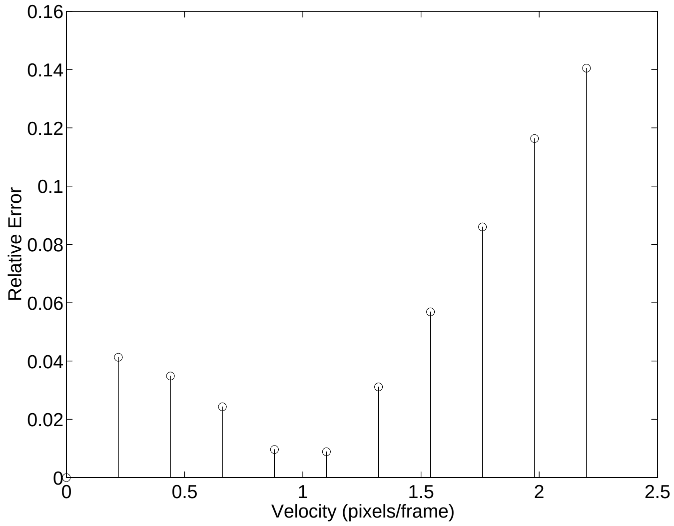


Fig. 13: Relative error versus velocity for the system of Section 4.1 – case 1

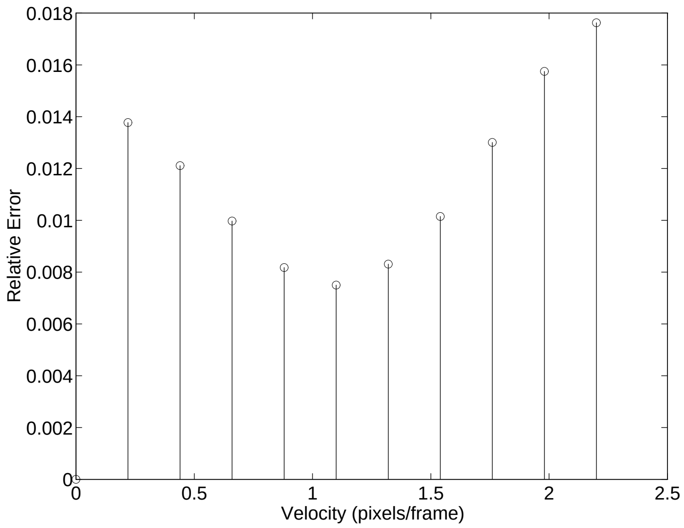


Fig. 14: Relative error versus velocity for the system of Section 4.1 – case 2

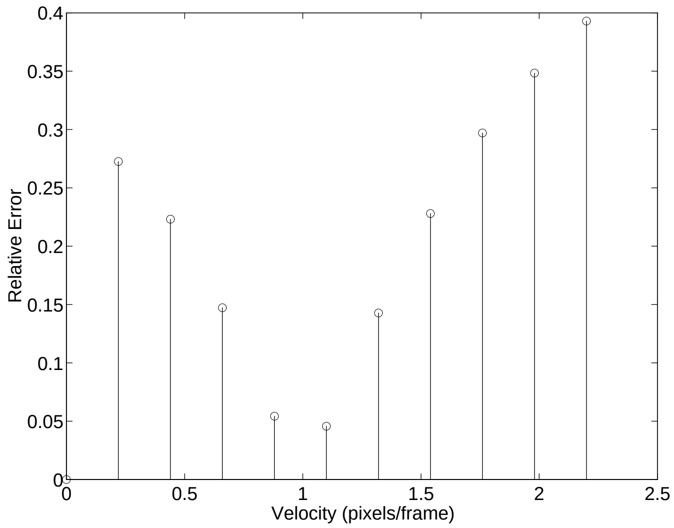


Fig. 15: Relative error versus velocity for the system of Section 4.1 – case 3

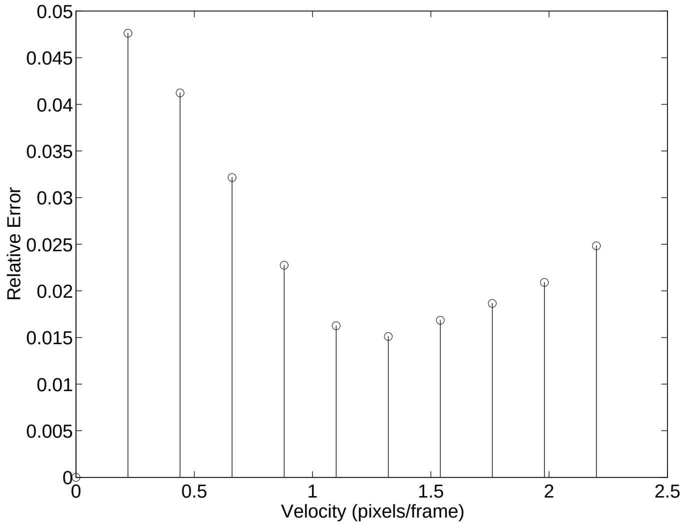


Fig. 16: Relative error versus velocity for the system of Section 4.1 – case 4

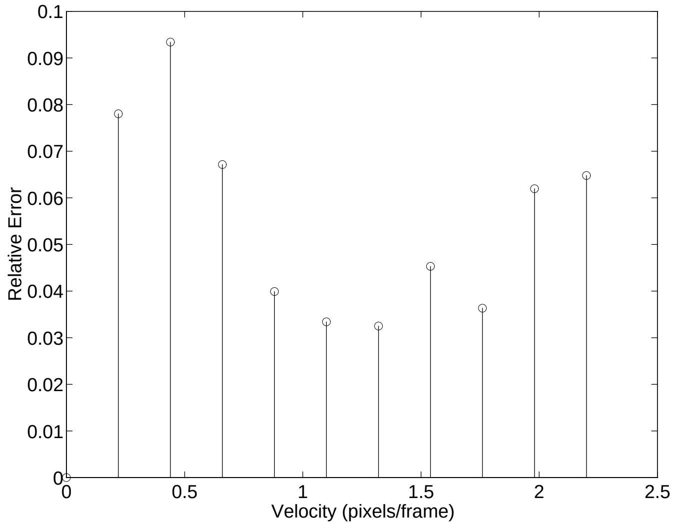


Fig. 17: Relative error versus velocity for the system of Section 4.1 – case 5

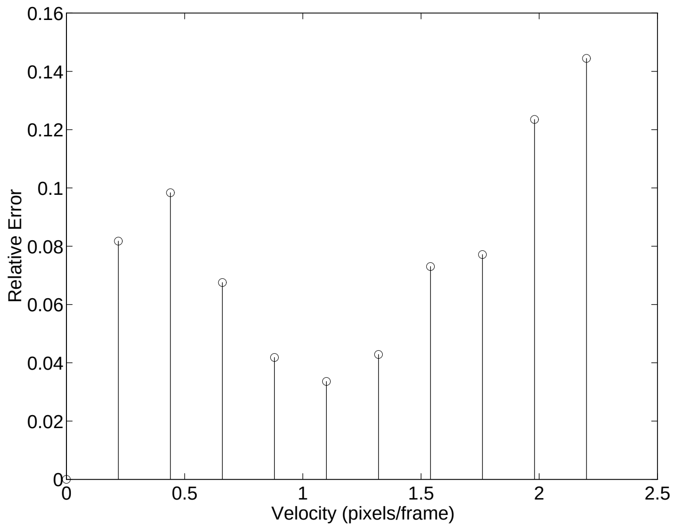


Fig. 18: Relative error versus velocity for the system of Section 4.1 – case 6

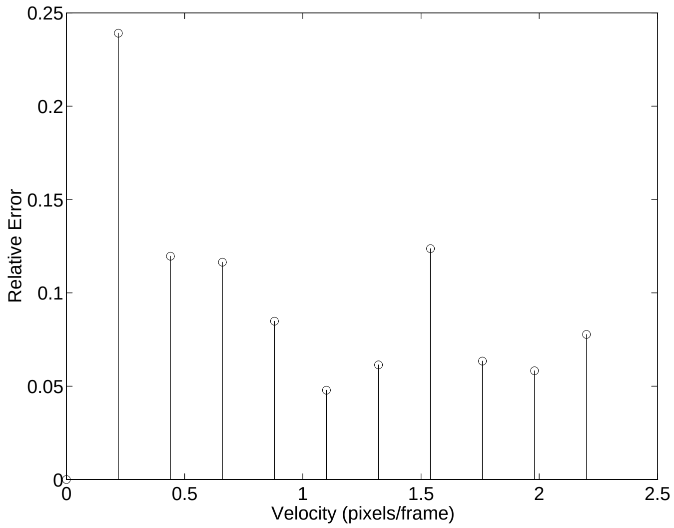
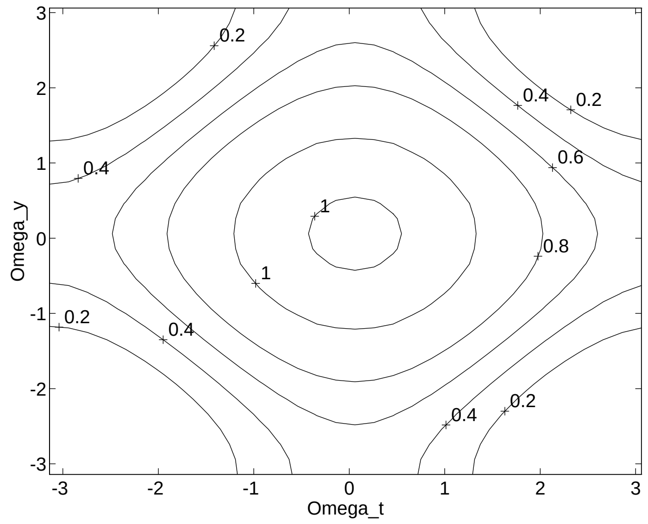


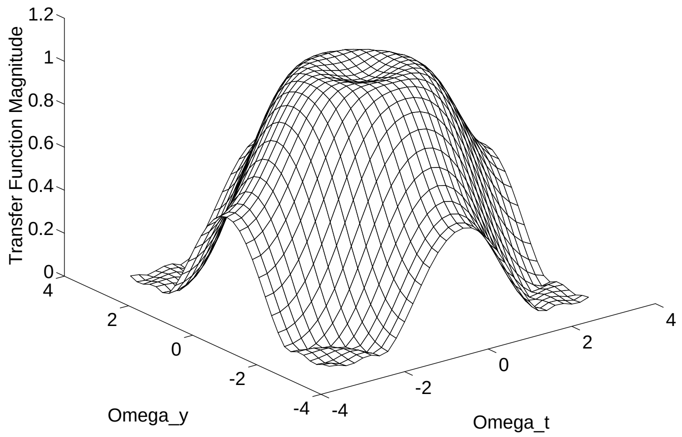
Fig. 19: Relative error versus velocity for the system of Section 4.1 – case 7

0	0	0	-0.079	0	0	0
0	0	-0.0238	0	-0.0238	0	0
0	-0.0238	0	0.1741	0	-0.0238	0
-0.0079	0	0.1741	0.5000	0.1741	0	-0.079
0	-0.0238	0	0.1741	0	-0.0238	0
0	0	-0.0238	0	-0.0238	0	0
0	0	0	-0.079	0	0	0

Fig. 20: Impulse response of the interpolator filter



(a)



(b)
Fig. 21: Transfer function magnitude of the interpolator filter:
(a) contour plot
(b) perspective plot

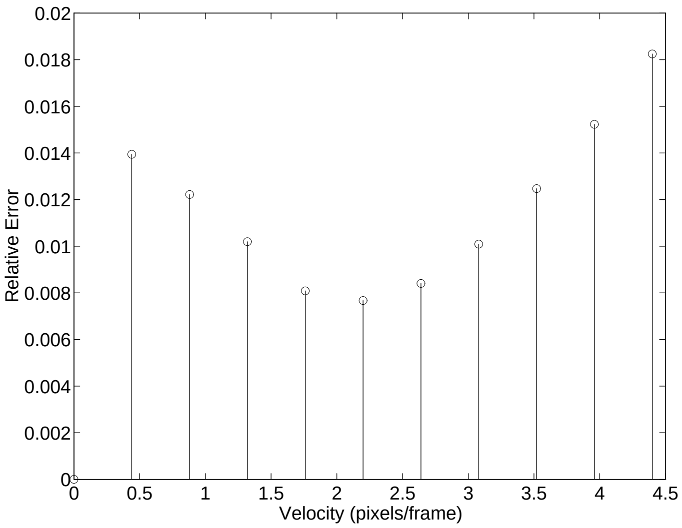


Fig. 22: Relative error versus velocity for the system of Section 4.1 – case 8

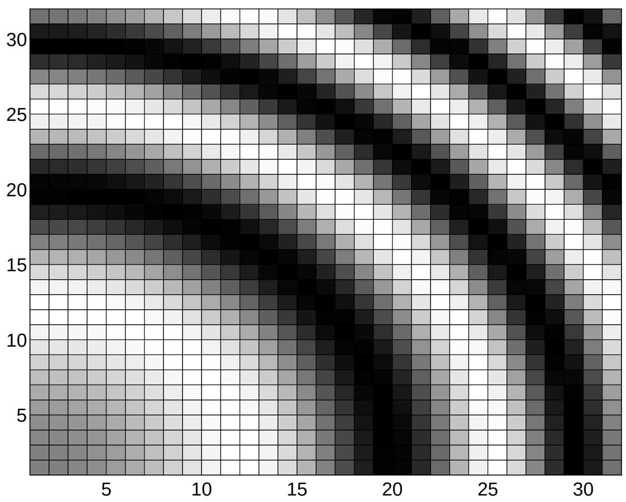
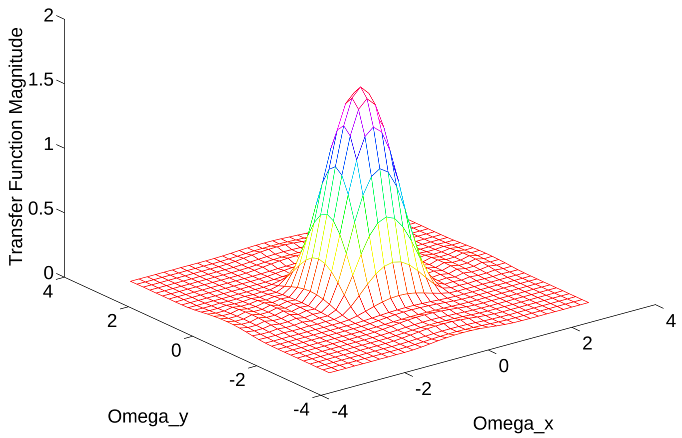
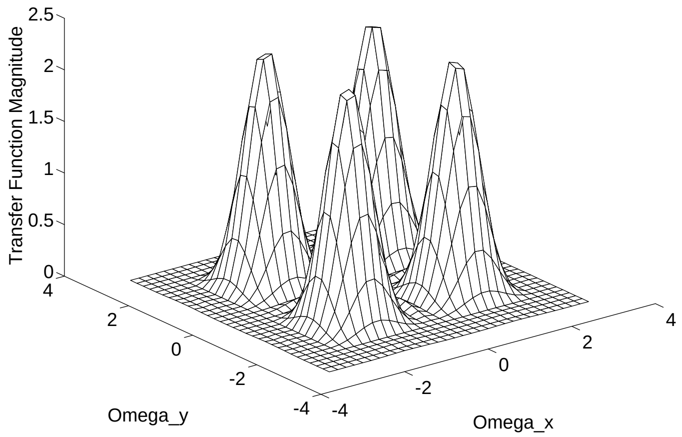


Fig. 23: Two–dimensional test sequence



(a)



(b)

Fig. 24 Transfer function magnitude of the two filters used in the two–dimensional test systems

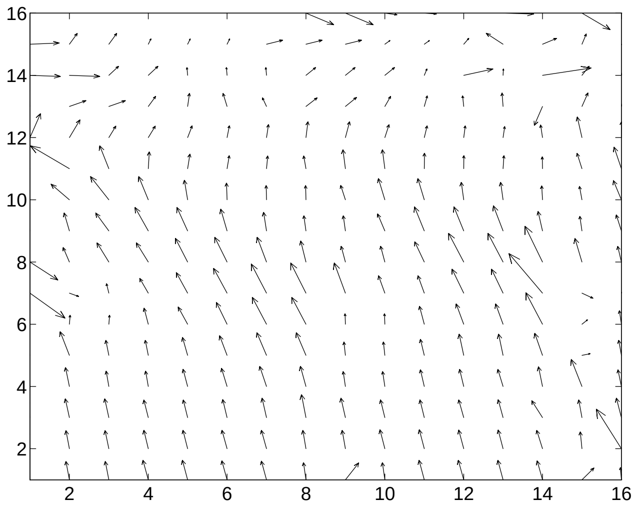


Fig. 25: Measured optical flow using 3–tap differentiators (v_x=-0.5 *pixel/frame*, v_y=2.6 *pixel/frame*)

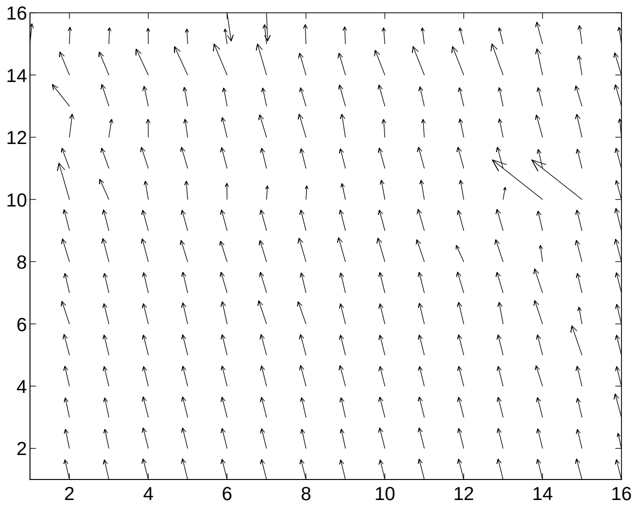


Fig. 26: Measured optical flow using adapted differentiators (v_x=-0.5 *pixel/frame*, v_y=2.6 *pixel/frame*)

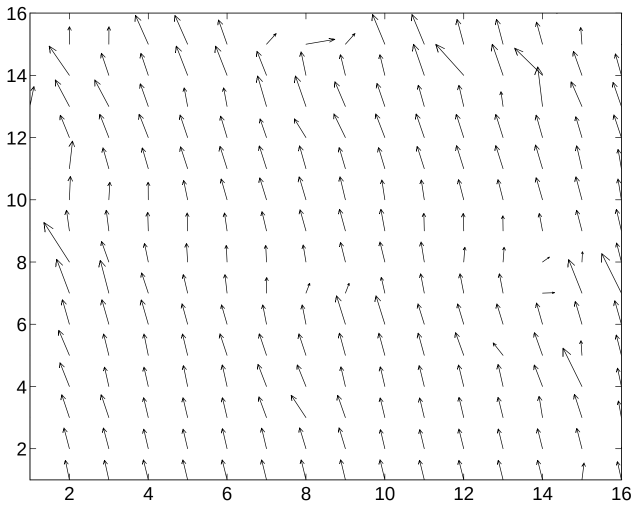


Fig. 27: Measured optical flow with interlaced input (v_x=-1 *pixel/frame*, v_y=5.2 *pixel/frame*)