Universitat Autònoma de Barcelona

MASTER IN COMPUTER VISION AND ARTIFICIAL INTELLIGENCE
REPORT OF THE RESEARCH PROJECT
OPTION: COMPUTER VISION

Conditions Ensuring Accuracy of Local Optical Flow Schemes

Author: Patricia Marquez Valle
Date: 08/09/2010
Advisor: Debora Gil / Aura Hernandez

Contents

1. Introduction

Image sequence analysis involves, among others, recognizing specific objects, computing their position, tracking them or determining motion at each point of the image. Motion perception analysis of a sequence is an important issue for security aids (such as detection of anomalous and unpredicted agents in urban scenes), car driver assistance or pathology discrimination. For that, an appropriate computation of motion is a standout field in computer vision.

Motion can be rigid or elastic. On one hand, rigid motion is induced by the relative movement between the scene and the camera, so, objects have the same shape along the sequence. It is often present in natural scenes and its principal application is object tracking. Rigid motion is commonly analyzed by applying techniques such as particle filters. On the other hand, elastic motion keeps shape changes along the sequence. This implies that each pixel of an object has a different motion, thus, motion estimation has to be local. Elastic motion arises in biomedical images (tissue deformations). The milestone for computing point wise local estimation of motion is the Optical Flow (OF) introduced by Gibson (1950) and Gibson (1966).

OF is a tracking method for describing sequence motion. In particular, it is useful for elastic deformation and object segmentation. OF is defined as the velocity vector field that transforms one frame into the following one. It assumes that object appearance (given by image intensity) keeps constant along sequence frames (Brightness Constancy Constraint, BCC). Under such assumption, the vector field given by OF puts into correspondence pixels in consecutive frames that have the same appearance (intensity).

In mathematical terms, these requirements are formulated as follows. Let $I(x(t), y(t), t)$ (I for short) be a sequence and $(x(t_0), y(t_0), t_0)$ an image pixel at time t_0. If we assume the BCC, the equation to solve is the following:

$$I(x(t + t_0), y(t + t_0), t + t_0) = I(x(t_0), y(t_0), t_0) \tag{1}$$

Considering that movements are small and the image sequence varies smoothly along the spatial and temporal coordinates, we can use first-order Taylor expansion in time t_0, obtaining:

$$I_x x_t + I_y y_t + I_t = 0 \tag{2}$$

Since, OF is a vector field, we can define $W = (x_t, y_t) = (u, v)$ and re-write the equation (2) into the following compact form:

$$< \nabla I, W > + I_t = 0 \tag{3}$$

The above equation will be called OF equation.

The solution of (3) gives, for each pixel, the estimation of motion given by OF. The application of OF formulation to real life sequences presents three weighty limitations: BCC is not always fulfilled, sequences should have small temporal deformations and the aperture problem.

In real life sequences, BCC is not always fulfilled due to illumination changes, physical properties or image acquisition devices. Thus, the computed OF may not correspond to the real motion. One way of ensuring the BCC is to either change the feature to keep constant along sequence frames or the representation space of the sequence (Barron et al., 1994; Bruhn, 2006). There are two kinds of alternative features, region descriptors and edges and corners descriptors. Region-based techniques (Anandan, 1989; Singh, 1990) match shift region descriptors, while feature-based approaches (Buxton and Buxton, 1984; Waxman et al., 1988; Wills et al., 2006) seek correspondences of characteristic image features such as edges or corners. Besides, methods that change the image representation space replace brightness by a filter response. In this case, the velocity vector field is defined from the phase behavior of band-pass filter outputs in the Fourier domain and so they are called phase-based approaches (Fleet and Jepson, 1990; Felsberg, 2004; Garcia, 2009).

The second limitation of the OF formulation is that, independently of the feature to keep constant, OF equation is based on derivatives. Therefore, motion is not properly recovered for large deformations and a high temporal resolution sequence is needed.

Finally, the OF computation is an ill-posed problem. Equation (3) introduces one constraint with two unknowns so it can not be uniquely solved. Indeed equation (3) can only recover motion along the image gradient (normal to the image level sets). That is, if we express $W = \omega_1 \nabla I + \omega_2 \nabla I^\top$ and develop the scalar product, we have:

$$< \nabla I, W >= \cos\theta \cdot \|W\| \cdot \|\nabla I\| = \omega_1 \cdot \|\nabla I\| \tag{4}$$

for θ the angle between the motion vector and ∇I. By replacing (4) in (3) we obtain:

$$\omega_1 \cdot \|\nabla I\| + I_t = 0 \Leftrightarrow \omega_1 = \frac{-I_t}{\|\nabla I\|} \tag{5}$$

Figure 1 graphically shows the projection of W over ∇I.

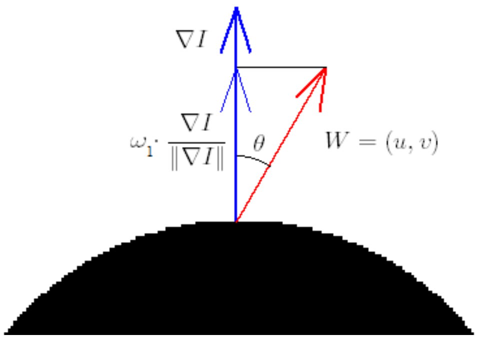


Figure 1: Geometric interpretation of the OF.

This phenomenon, called the aperture problem, arises as a consequence of trying to compute pixel wise motion with only one equation. We are converting a two dimensional problem into a one dimensional one. Therefore, depending on the geometry of the object and the kind of motion, OF equation can properly recover the motion or not. That is, in points where motion is perpendicular to the image level set, OF equation recovers the whole motion, while in points where motion is tangent to the image contours, the OF equation does not recover motion at all. The remaining possible motions in a point will be partially recovered. Figure 2 shows three different scenarios. Blue arrows correspond to the motion that OF equation recovers while red arrows correspond to the real motion of the object. In figure 2a motion is perpendicular to the edge; therefore, blue arrows coincide with the red ones, that is, we can recover the motion. In figure 2b motion is tangent to the edge, consequently no motion is recovered. Finally, in figure 2c motion is oblique to the edges, thus we will recover the normal component of the motion. Notice that, since in the corner there are two different equations of the OF, motion can be properly recovered.

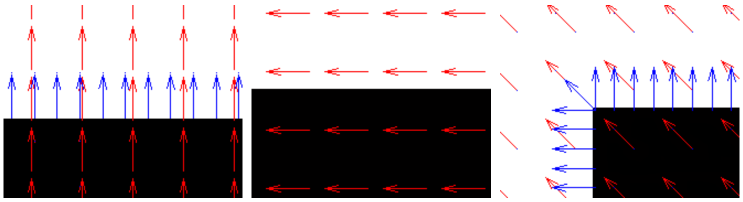

(a) (b) (c)

Figure 2: Consequences of the aperture problem. In red the motion of the object, and in blue the solution given by the OF.

In order to minimize the aperture problem some outstanding local techniques propose an equation system by assuming some properties of the vector field (Lucas and Kanade, 1981; Fleet and Jepson, 1990). Lucas and Kanade (1981) is a differential approach that assumes that motion is constant in a local neighborhood around each pixel and then, applying least squares, computes the solution. Fleet and Jepson (1990) is a phase-based technique that defines the velocity in terms of the gradient of the phase output of a Gabor filter. The vector is the least squares solution of an equation system. In Section 2 there is a detailed formulation of both methods.

Local techniques solve the aperture problem in some pixels but do not produce dense flow fields. In contrast, variational techniques, developed for the first time by Horn and Schunck (1981), produce dense flow fields by combining into a variational framework a data-term (which assumes constancy in the object appearance) and a smoothness-term (which models the behavior of the flow across the image). These approaches compute the OF (W) by finding the minimum of the following energy functional:

$$\varepsilon(W) = \iint \underbrace{\|f(\nabla I, W)\|^2_{L^2}}_{Data-Term} \,\mathrm{d}x\,\mathrm{d}y + \alpha \iint \underbrace{g(W, \nabla W, ...)}_{Smoothness-Term} \,\mathrm{d}x\,\mathrm{d}y \tag{6}$$

where f is the data-term, g the smoothness-term, α is a constant regularization parameter, $\|\cdot\|_{L^2}$ is the L^2 norm and ... denotes higher order terms. We remit the reader to Onkarappa and Sappa (2009) for a wide review of variational approaches.

The data-term is a function that puts into correspondence one frame with the following one. Existing approaches use either OF equation (Horn and Schunck, 1981; Nagel and Enkelmann, 1986) or the system equation provided by local techniques such as Lucas and Kanade (Bruhn et al., 2005) or Fleet and Jepson (Garcia, 2009). OF equation like data-term still presents the aperture problem, so it needs to be solved in the variational framework together with the smoothness-term. In the case of local techniques, since data-terms are own-solvable, the smoothness-term is introduced to regularize the velocity vector field. Thus, they are combining the robustness of local methods and the density of the variational ones.

The smoothness-term determines the global properties of the vector field. The first approaches defined the smoothness-term in the L^2 space to ensure differentiability (Horn and Schunck, 1981; Nagel and Enkelmann, 1986; Bruhn et al., 2005). The main limitation is that the solution can be over-regularized in cases that there are occlusions or discontinuities in the velocity vector field. In order to overcome that, total variational methods define g in the L^1 space (Bruhn et al., 2005; Wedel et al., 2009). The problem of those techniques is that there are functions from different spaces in the same variational framework, which is mathematically inconsistent and also there is no robust theory which assures the reliability of the solution.

Other techniques provide a trade-off between the data-term and the smoothness-term through the parameter α. The bigger the α is, the smoother the flow field becomes. In most cases, this parameter is constant and can be learned from a training set (Sun et al., 2010). Other works (Nagel and Enkelmann, 1986; Garcia, 2009) consider that the smoothness-term should play an important role in those points where the data-term does not provide motion information. Nagel and Enkelmann (1986) weight the variational by means of $\|\nabla I\|$. The main idea is that $\nabla I = 0$ indicates a flat region. Since in these points we can not recover any motion, the scheme gives more weight to the smoothness-term. However ∇I may not reflect all regions where the data-term can not properly recover motion. In this context we claim that α has to reflect the theoretical conditions that assure a good solution of the data-term. In this fashion, Garcia (2009) introduce the amplitude of the response of Gabor Filters in the weights with an evident improvement on the computation of the OF. This approach has been applied to medical images, in particular, to the assessment of the left ventricle motion.

1.1 Goal and Contributions

The goal of this project is to determine the theoretical requirements of the local system solutions of Lucas and Kanade (1981) and Fleet and Jepson (1990) to ensure their maximum precision in order to incorporate them as weights into their variational frameworks.

The main contribution of this master thesis is to define variable weights which force the above conditions and incorporate them into a variational formulation focusing on the approaches developed by Bruhn et al. (2005) and Garcia (2009). We will put special emphasis on the OF computation developed in the group of visualization, graphics and modeling (Garcia, 2009) in order to improve the weights and generalize them to urban scenes.

Moreover, in the same fashion of Barron et al. (1994), a database of synthetic images has been created in order to analyze the consistency of the theoretical assumptions needed to solve the local equation systems. Although there are well-known databases of sequences (based on real environments) with ground-truth (Mid), the difference between the computed OF and the ground-truth can come from the chosen technique or from errors inherent to the sequence (like discretization errors) (Baker et al., 2009). Thus, we propose a synthetic images set in order to avoid inherent errors.

2. Methodology

In this section we describe the local systems defined by Lucas and Kanade (1981) and Fleet and Jepson (1990). In addition, in order to define variable weights for the variational formulations, we determine the theoretical requirements that ensure maximum precision in the local system solutions.

2.1 Lucas-Kanade

Lucas-Kanade (LK) approach is based on the assumption that OF keeps constant in a neighborhood of a pixel of size ρ. Given an image sequence $I(x, y, t)$ where (x, y) is the pixel location and t denotes the time, we can compute the OF, $W = (u, v)$, by minimizing the function:

$$\varepsilon_{LK}(W) = K_\rho * \Big((I_x u + I_y v + I_t)^2 \Big) \tag{7}$$

where K_ρ is a Gaussian kernel of standard deviation ρ. It denotes a window function that gives more influence to the constraints at the center of the neighborhood than those at the boundary.

A minimum of equation (7) satisfies:

$$\left\{ \begin{array}{lcl} \frac{\partial \varepsilon_{LK}(W)}{\partial u} & = & 0 \\ \frac{\partial \varepsilon_{LK}(W)}{\partial v} & = & 0 \end{array} \right. \tag{8}$$

and then, we obtain the following linear system:

$$A_{LK} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} K_\rho * (I_x^2) & K_\rho * (I_x I_y) \\ K_\rho * (I_x I_y) & K_\rho * (I_y^2) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -K_\rho * (I_x I_t) \\ -K_\rho * (I_y I_t) \end{pmatrix} \tag{9}$$

If we define $\widetilde{W} = (u, v, 1)$, then, the LK approach can be re-written as follows:

$$\varepsilon_{LK}(W) = \widetilde{W} K_\rho * (\nabla I \nabla I^\top) \widetilde{W}^\top \tag{10}$$

where $\nabla I = (I_x, I_y, I_t)$.

Notice that A_{LK} is the Structure Tensor (ST). The ST (Jähne, 1993) describes the image local geometry and it is characterized by its eigenvectors and eigenvalues that are defined as follows:

$$A_{LK} = \begin{pmatrix} e_x^1 & e_x^2 \\ e_y^1 & e_y^2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} e_x^1 & e_x^2 \\ e_y^1 & e_y^2 \end{pmatrix}^{-1} \tag{11}$$

Figure 3 shows the relation between the eigenvectors and eigenvalues. The eigenvalues indicate the degree of anisotropy. In the case that eigenvalues are similar, i.e., $\lambda_1 \sim \lambda_2$, the eigenvectors of ST are isotropic, while different eigenvalues, i.e., $\lambda_1 \gg \lambda_2$ indicate that the eigenvectors of ST are anisotropic.

Observe that the ST matrix can be interpreted as an average between motion along direction of adjacent edges. In the corner of an image, since any motion has two well defined components, the eigenvectors of the ST are isotropic, thus the aperture problem is solved and motion can be recovered. On the contrary, in straight pure edges of an image, only the

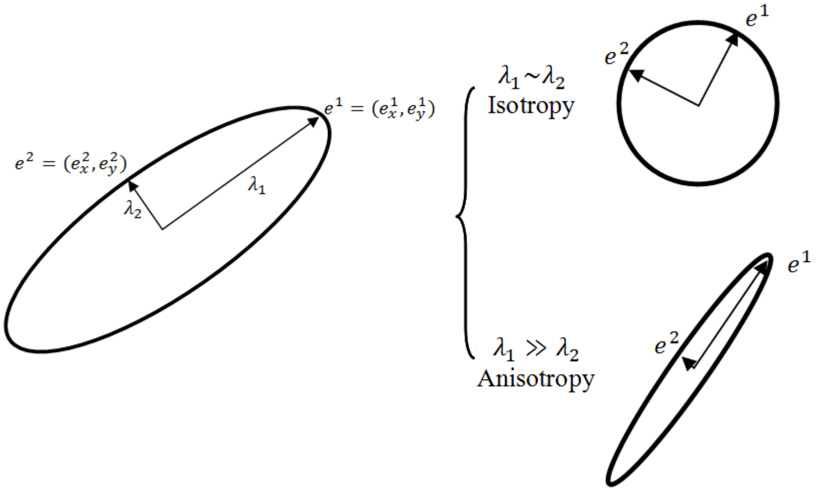


Figure 3: ST eigenvectors and eigenvalues. Isotropy and anisotropy.

component perpendicular to the edge direction is well defined, thus, the eigenvectors of the ST are anisotropic. In this case, only when $det(A_{LK})$ is bigger than a certain threshold the aperture problem is solved. So the ST not always solve the aperture problem, but at least is under control.

In addition, OF can not be computed in flat regions since there is no gradient.

From an algebraic point of view, LK system is solvable when the matrix A_{LK} is non-singular. For that, we propose as a Quality Measure (QM) to ensure the LK theoretical assumptions the normalized determinant of the system matrix:

$$QM_{LK} = \frac{|det(A_{LK})|}{\|det(A_{LK})\|_\infty} \tag{12}$$

The QM is normalized in order to take values between 0 and 1.

We propose the following incorporation of variable weights in the variational framework developed by Bruhn et al. (2005) that uses as a data-term the LK equation system:

$$\varepsilon(W) = \iint QM_{LK}\widetilde{W}K_\rho * (\nabla I \nabla I^\top)\widetilde{W}^\top + (1 - QM_{LK})\|\nabla\widetilde{W}\|^2 \,\mathrm{d}x\,\mathrm{d}y \tag{13}$$

In this manner, when QM_{LK} is close to zero, since the LK solution is not reliable, the data-term of the variational framework has a low weight.

2.2 Harmonic Phase Flow

Fleet-Jepson (FJ) approach defines the velocity vector field in terms of the gradient of the phase output of some Gabor Filters (GF) and then it solves the equation system by applying least squares. We present a simplified version that considers just two GF (Garcia, 2009).

$\theta_1 = 0 Rad$ $\theta_2 = \frac{\pi}{6} Rad$ $\theta_3 = \frac{\pi}{3} Rad$ $\theta_4 = \frac{\pi}{2} Rad$ $\theta_5 = \frac{2\pi}{3} Rad$ $\theta_6 = \frac{5\pi}{6} Rad$

Figure 4: GF varying the orientation.

Scale=1 Scale=2 Scale=3 Scale=4 Scale=5

Figure 5: GF varying the scale.

A GF (Garcia, 2009) is the product of a complex sinusoid (which determines the frequency), and a Gaussian kernel (which determines the scale):

$$G(x,y) = C \cdot \underbrace{e^{-i2\pi(\frac{\omega_x x}{d_1} + \frac{\omega_y y}{d_2})}}_{Frequency} \cdot \underbrace{e^{-\frac{(x')^2 + (\lambda y')^2}{2\sigma_{x'}^2}}}_{Scale} \tag{14}$$

where,

- Frequency Parameters:
  - $\omega = (\omega_x, \omega_y)$ is a vector that determines the frequency and orientation of the pattern.
  - d_1 and d_2 determine the width and the height of the window where the filter is defined.
- Scale Parameters:
  - $x' = x\cos\theta + y\sin\theta$ and $y' = -x\sin\theta + y\cos\theta$ define the orientation of the Gaussian envelope.
  - $\sigma_{x'}$ determines the size of the Gaussian envelope along the x' direction.
  - λ specifies the ellipticity of the Gaussian envelope.

A GF is determined by its orientation (θ) and scale (s). We denote G_i^s the GF with scale s and orientation $\theta_i = \frac{(i-1)\pi}{N} Rad$ for N the total number of orientations. In figure 4 we show G_i^4 for $i = 1, .., 6$ for visualizing different orientations, while in figure 5 we show the variation of scale by G_1^s for $s = 1, ..., 5$. We can observe how those parameters influence on a GF, the orientation indicates the rotation of the GF whereas the scale indicates the gaussian envelope size.

Convolving the image $I(x, y, t)$ with a GF G_n^s results in an harmonic image

$$I^n(x, y, t) = G_n^s * I = a^n(x, y, t)e^{i\Phi^n(x,y,t)} \tag{15}$$

where $a^n(x, y, t)$ is the amplitude, $\Phi^n(x, y, t)$ is the phase of the harmonic image and n denotes the orientation of the GF. In particular, we are projecting the GF over the contours of the image. Thus, the orientation and the scale of the GF affect the amplitude and the phase of the harmonic image. In the edges of an image, depending on its orientation and the GF orientation, there will be response or not. The response is higher when the GF orientation is perpendicular to the edge and lower when the GF orientation is parallel to it. Thus, the amplitude takes higher values in the edges perpendicular to the GF orientation, and lower values in the edges parallel to the GF orientation. If there is response in an edge, regardless of the GF orientation, the phase gradient is parallel to the edge gradient. Besides, in the corners of the image, regardless of the GF orientation, there is always response. The amplitude always take high values and the phase is always oriented as the GF orientation. Notice that the amplitude is an indicator of the response degree. Regarding the GF scale, the bigger the scale is, the bigger the region response is. Thus, if we use a big scale, the region of the image where the amplitude is non zero will be bigger. That is, using a big scale we obtain global information whereas using small scale we obtain local information.

We have computed the harmonic images of a synthetic triangle varying the orientation and the scale of the GF. In figure 6 there are the amplitudes of the harmonic images, whereas in figure 7 there are its phases. We have used GF with scales from 1 to 5 (columns) and 6 different orientations ($\theta_i = \frac{i\pi}{6}$ for $i = 0, ..., 5$) (rows). In red the contour of the triangle. We can see how the phase and the amplitude of an harmonic image vary depending on the orientation and the scale of the GF. For instance, we can observe that using a GF with orientation $\theta_i = \frac{\pi}{2}$, since the orientation is perpendicular to the horizontal edge, we have a high amplitude response, and, the bigger the scale is, the bigger the region of the amplitude response is. In addition, the regions where the phase is properly defined correspond to the ones where the amplitude has a high response.

Given two different harmonic images I^i and I^j, we replace in the OF equation the gradient of the phase of the harmonic images obtaining two equations. We call the following equation system Harmonic Phase Flow (HPF):

$$\left\{ \begin{array}{lcl} \varepsilon_1 = \Phi_x^i u + \Phi_y^i v + \Phi_t^i & = & 0 \\ \varepsilon_2 = \Phi_x^j u + \Phi_y^j v + \Phi_t^j & = & 0 \end{array} \right. \tag{16}$$

where Φ_x^k, Φ_y^k and Φ_t^k for $k \in \{i, j\}$ are the partial derivatives of the phase at time t.

This equation system can be formulated as:

$$A_{HPF} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \Phi_x^1 & \Phi_y^1 \\ \Phi_x^2 & \Phi_y^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -\Phi_t^1 \\ -\Phi_t^2 \end{pmatrix} \tag{17}$$

Notice that, in order to assure that A_{HPF} is invertible we need $\nabla\Phi^i$ and $\nabla\Phi^j$ to be linearly independent, for example perpendicular GF orientations. We consider two phases fulfilling that the orientations of the Gabor filters of the harmonic images I^i and I^j are perpendicular. We refer to those phases as complementary phases.

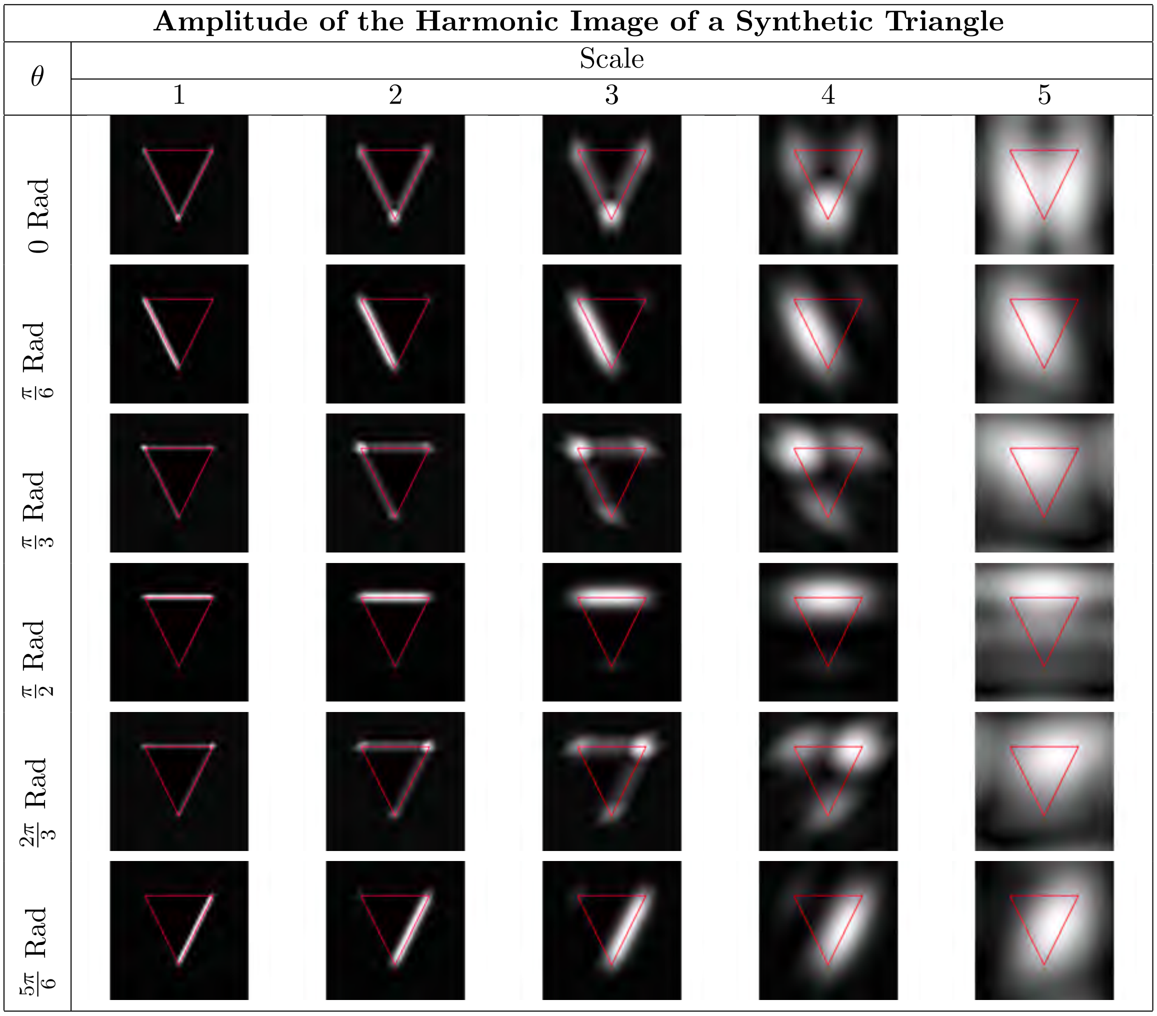


Figure 6: Amplitude of the harmonic image of a synthetic triangle. Varying the scale from 1 to 5 (columns) and with 6 different orientations (rows). In red the contour of the triangle.

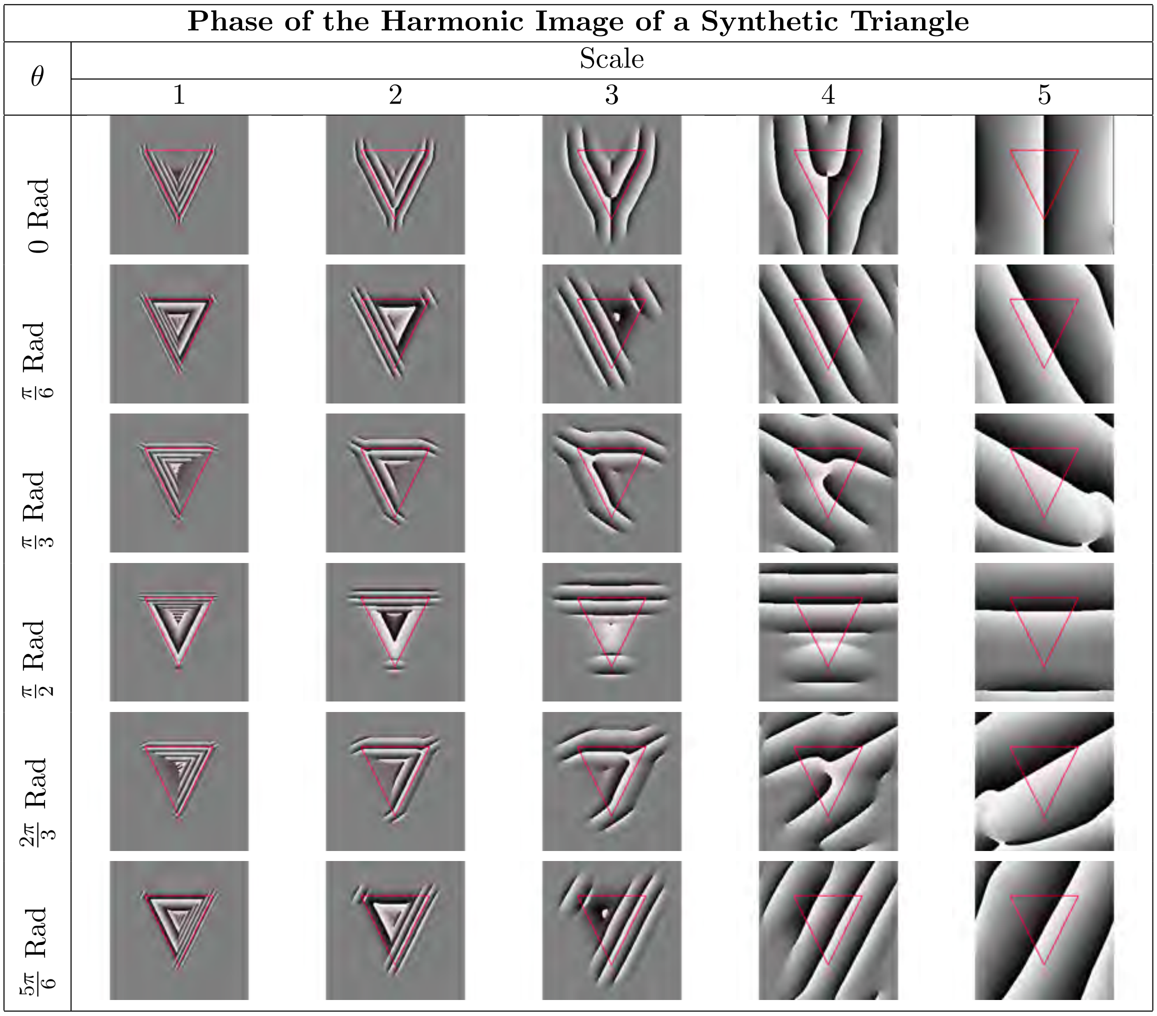


Figure 7: Phase of the harmonic image of a synthetic triangle. Varying the scale from 1 to 5 (columns) and with 6 different orientations (rows). In red the contour of the triangle.

In order to ensure the solution of this local equation system we need A_{HPF} to be invertible. As well, the information given by the phase of the GF also has to be reliable, so, the amplitude has to have a high response. We propose as a QM of the HPF local system

$$QM_{HPF} = min\left\{ \frac{|det(A_{HPF})|}{\|det(A_{HPF})\|_\infty}, min\left\{ \frac{a^i}{\|a^i\|_\infty}, \frac{a^j}{\|a^j\|_\infty} \right\}\right\} \tag{18}$$

where $\|\cdot\|_\infty$ is the maximum norm over the whole image. The determinant and the amplitudes are normalized in order to be comparable when considering the minimum. Taking the minimum between the determinant of the system matrix and the amplitudes we ensure that the determinant will be high enough and also the amplitudes.

We propose the following incorporation of variable weights for the Harmonic Phase Flow (HPF) developed by Garcia (2009):

$$\varepsilon(W) = \iint QM^2_{HPF}\Big(\varepsilon_1^2 + \varepsilon_2^2\Big) + \Big(1 - QM_{HPF}\Big)^2 \|\nabla W\|^2 \, \mathrm{d}x \, \mathrm{d}y \tag{19}$$

In the case that QM_{HPF} is close to zero the solution given by the local system HPF is not reliable, therefore, we weight more the smoothness-term.

In addition to the weights incorporation, we have changed the way of choosing the GF orientations. Initially only two orientations were chosen for the whole image, now, we compute for each GF orientation the corresponding harmonic image obtaining I^n for $n \in \{1, 2, ..., 12\}$, and its corresponding phase gradients. Then, for each pixel we choose, the harmonic image that in this pixel has the highest amplitude response and the harmonic image with complementary phase.

Notice that, unlike existing variational methods, our variable weights appear in both data and smoothness terms. We consider that they are complementary terms so they are linked in a variational scheme by this weight. The smoothness-term plays a strong role when the data-term does not properly recover the motion, so it has to vanish.

3. Experimental Settings

In order to analyze the QM consistency we have done two types of experiments:

- **Theoretical Requirements Consistency:** we study the relation between QM and the accuracy of the local systems solutions.
- **Optical Flow Accuracy:** we compare LK and HPF OF computation accuracy and we do a first qualitative experiment of the variational HPF.

Since we are not using any coarse-to-fine technique, we will use sequences which flow magnitudes are between 1 and 4 pixels per frame.

In order to measure the accuracy between the ground-truth and the computed OF we have used the End-Point Error (EE) , defined as follows:

$$EE = \|W_C - W_{GT}\|_2 \tag{20}$$

where W_C is the computed OF and W_{GT} is the ground-truth of the OF. We have chosen the EE instead of the Angular Error because of the latter might not be reliable in the case that the module of the OF is small (Baker et al., 2009).

Derivatives are computed using first four differences (Horn and Schunck, 1981) defined as follows:

$$\begin{aligned}
I_x(i,j,k) &= \tfrac{1}{4}\Big(I_{i,j+1,k} - I_{i,j,k} + I_{i+1,j+1,k} - I_{i+1,j,k} + I_{i,j+1,k+1} \\
&\quad - I_{i,j,k+1} + I_{i+1,j+1,k+1} - I_{i+1,j,k+1}\Big) \\
I_y(i,j,k) &= \tfrac{1}{1}\Big(I_{i+1,j,k} - I_{i,j,k} + I_{i+1,j+1,k} - I_{i,j+1,k} + I_{i+1,j,k+1} \\
&\quad - I_{i,j,k+1} + I_{i+1,j+1,k+1} - I_{i,j+1,k+1}\Big) \\
I_t(i,j,k) &= \tfrac{1}{4}\Big(I_{i,j,k+1} - I_{i,j,k} + I_{i+1,j,k+1} - I_{i+1,j,k} + I_{i,j+1,k+1} \\
&\quad - I_{i,j+1,k} + I_{i+1,j+1,k+1} - I_{i+1,j+1,k}\Big)
\end{aligned} \tag{21}$$

where $I_z(i,j,k)$ is the partial derivative over the coordinate $z \in \{x,y,t\}$ of the frame I in the position (i,j) at time k and $I_{i,j,k}$ is equivalent to $I(i,j,k)$.

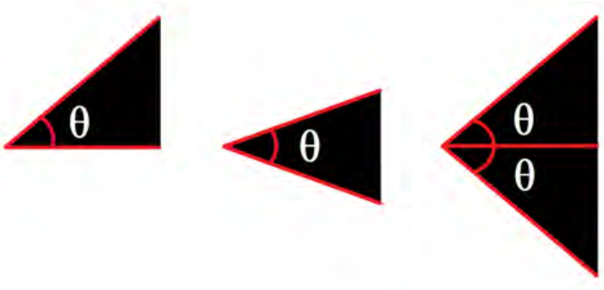
LK is computed with a Gaussian kernel of standard deviation $\rho = 4$. HPF GF are computed using scale 2. We chose the GF with maximum amplitude among 12 possible GF. The GF denoted by the number i for $i = 1, ..., 12$ corresponds to the GF with scale 2 and orientation $\theta_i = \frac{(i-1)\pi}{12}$.

In order to analyze the behavior of the solution of the local systems in a controlled framework we have created a database of synthetic sequences. In the literature there are also experiments with controlled synthetic sequences (Barron et al., 1994) but those sequences do not emphasize the geometry neither the movements of the figures, for that reason we have enlarged the range of movements and the corners profiles. We have created synthetic sequences of triangles with different angles $\theta_j = j \cdot 10^\circ$ for $j = 1, ..., 9$. Since not only the corner angle matters, also the aperture of the incident edges, we have considered three types of corners:

- **Unbalanced corner:** for a given angle θ_j we consider the rectangle triangle with an aperture of θ_j degrees (figure 8a). This kind of corners will be denoted by C^1_j.

- **Balanced corner:** for a given angle θ_j we consider the symmetric triangle with respect to the horizontal axis with an aperture of θ_j degrees (figure 8b). This kind of corners will be denoted by C_j^2.

- **Double Balanced corner:** for a given angle θ_{2j} we consider the symmetric triangle with respect to the horizontal axis and with an aperture θ_{2j} (figure 8c).



(a) (b) (c)

Figure 8: Synthetic data corner profiles. **(a)** Unbalanced Corner. **(b)** Balanced Corner. **(c)** Double Balanced Corner.

Notice that, since the Unbalanced and the Balanced corners have the same angle (θ_j) we can compare the behavior of the local system solutions in the corner, and also we can observe if the inclination of the edges affect the solution given by the local systems. Meanwhile, the Unbalanced and the Double Balanced corners are comparable in the inclination of the edge.

To create the sequences we have interpolated the images with a sample of movements covering all possible motions. The ground-truth for each pixel of the sequence is:

$$W_{GT} = (\cos(\alpha_i), \sin(\alpha_i))$$

where $\alpha_i = (i-1)5^\circ$ for $i = 1, 2, ..., 72$. We will denote a movement by α_i. Observe that, we are creating a continuous movement in a discrete space, so the ground-truth may not coincide with the real movement of the sequence (Baker et al., 2009). Since each pixel has only eight neighbors, if an object is moved in a sequence with a temporal resolution of one pixel, we have only eight possible movements that assure that there are no differences between the ground-truth and the real movement of the sequence (see figure 9).

We distinguish between two kind of sequences:

- **Discrete sequences:** sequences with movements $\alpha_i = (i-1)45^\circ$ for $i = 1, 2, ..., 8$, the ground-truth and the real movement of the sequence are the same.

- **Continuous sequences:** sequences with movements $\alpha_i = (i-1)5^\circ$ for $i = 1, 2, ..., 72$, the ground-truth and real movement of the sequence may not coincide.

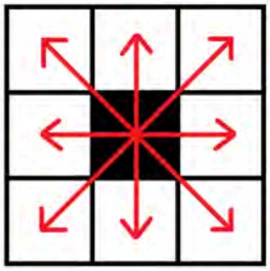


Figure 9: Exact movements of one pixel displacement.

3.1 Theoretical Requirements Consistency

In order to assess the consistency of the theoretical requirements, EE should be compared to QM for the synthetic discrete sequences. The stated requirements hold in the assumption that EE decreasingly depends on QM, such dependency is by means of the correlation between EE and QM (considered as random variables). Since the correlation might be non linear, a statistical way of analyzing it is to compute the Spearman correlation coefficient (ρ) (Pratt, 1991). ρ takes values from -1 to 1, indicating a positive correlation for value 1, and a negative one for value -1.

We compute the solution of the local systems and its QM in the profile of the figure and we study the correlation between EE and QM. In order to know if there is a decreasingly dependency we have computed the confidence intervals of the hypothesis test:

$$\begin{cases} H_0 : \rho \geq 0 \\ H_1 : \rho < 0 \end{cases}$$

rejecting the null hypothesis for $p - val < 0.05$, thus, we can assert that the dependency is negative.

3.2 Optical Flow Accuracy

In order to analyze the behavior of the local systems in a more realistic context, we compute for HPF and LK the ranges of Invalid Points (IP) and EE over the continuous sequences.

We denote IP that points of an image which have a $QM < tol$, notice that in these points we can not assure the reliability of the solution. We compute the percentage of IP in the contour of the corner obtaining $IP_{cont}^{C_j^k,\alpha_i}$. The range for all movements is the following:

$$IP_{C_j^k} = \mu\{IP_{cont}^{C_j^k,\alpha_i}, \forall i\} \pm \sigma\{IP_{cont}^{C_j^k,\alpha_i}, \forall i\}$$

And the range of IP for all movements and corners is computed as follows:

$$IP_{C^k} = \mu\{IP_{cont}^{C_j^k,\alpha_i}, \forall i, j\} \pm \sigma\{IP_{cont}^{C_j^k,\alpha_i}, \forall i, j\}$$

Besides, we compute the mean for the remaining non IP of the contour of a corner C_j^k with motion α_i, obtaining $EE_{cont}^{C_j^k,\alpha_i}$. The range for all movements is computed as follows:

$$EE_{C_j^k} = \mu\{EE_{cont}^{C_j^k,\alpha_i}, \forall i\} \pm \sigma\{EE_{cont}^{C_j^k,\alpha_i}, \forall i\}$$

A measure of the accuracy of the local system solutions for a sample of all possible movements and corners is the following:

$$EE_{C^k} = \mu\{EE_{cont}^{C_j^k,\alpha_i}, \forall i,j\} \pm \sigma\{EE_{cont}^{C_j^k,\alpha_i}, \forall i,j\}$$

Moreover, in order to illustrate the improvements of the varying weights in the variational framework, we have compared the original HPF and the improved HPF applying them to real image sequences. As a first step we present qualitative results by means of color flow coding (Baker et al., 2009). We have considered the direction of the OF solution, that is, all vectors of the computed OF are normalized.

We apply the original HPF and the improved HPF to the Army sequence (obtained from Mid) that contains several independently moving objects. This sequence flow magnitudes are not higher than 4 pixels per frame. Images are 584 × 388 pixels. In addition we apply both methods to the Tram sequence (obtained from Onv) that contains a tram's driver view. The temporal resolution is quite low although we do not know the maximum flow magnitude since there is no ground-truth for this sequence. Images are 704 × 576 pixels.

4. Results

4.1 Theoretical Requirements Consistency

Figures 10, 11 and 12 show the behavior of EE and QM for HPF and LK local systems in the contour of three different angles (30°, 60° and 90° respectively) and motion $\alpha_0 = 0°$ (horizontal motion). In the first column we show the two most representative harmonic images given by the selected GF at each edge of the figure. We also plot in red the depicted contour of the corner, and in the text we have the GF orientation. In the second column for HPF and in the third one for LK, we have over all points of the contour the QM in red and the EE in blue. The contour of the corner goes from the bottom right of the figure to the top right. For both plots the black vertical line indicates the corner and blue crosses indicate that the matrix of the local system is singular, that is, EE is NaN. In the second column, we have plot the number of the selected GF divided by ten in green. In order to compare across the different methods all measures have been normalized.

In these plots there are some features that deserve special attention. We expect to have singular equation systems in the edges of the images and be able to compute the OF only in the corners. But observe that we can recover the OF in the corners and also in some edges. In the corners, we have the best motion recovery and a high QM. Figure 12 for the unbalanced corner shows a clear example. Regarding to the edges, since the images are in a discrete domain, the oblique edges are stepped and the horizontal and vertical edges are straight pure lines. LK can recover motion in oblique edges but not in straight ones. Figure 10 for the balanced corner shows an example, since we have a oblique and a straight edge. Figure 12 for the double balanced corner is a straight edge, thus LK has blue crosses along the whole contour, the QM, since it is normalized, seems that takes high values, but its maximum value is 4.49×10^{-16} which is really low. Besides, in the case of HPF we have recovered motion along all contours, this is because we are computing the solution in the Fourier domain. So, we are computing the solution over a periodic image and we have the influence of the contiguous images, so we can recover motion over all edges.

In addition, since the geometric properties of the unbalanced and the balanced corner are the same, we expect the same behavior of the solution. But the images are in the discrete domain, thus, the unbalanced and the balanced corner differ in the edges geometry. The unbalanced corner has an oblique and a straight edge and the balanced corner has two oblique edges. So, the behavior of the local systems is different, for instance, LK is able to recover motion along the whole contour of the balanced corners but for the unbalanced corners recovers motion just for the oblique edge. Regarding to the balanced and double balanced figures, since they are symmetric figures, we expect a symmetric behavior of the equation systems solution, but observe that the behavior of the EE and the QM is not symmetric due to the derivatives are non symmetric.

Figure 13 shows EE against QM for the discrete sequences with corners $C^k_{30°}$ for $k = 1, 2, 3$. Both measures are normalized. We can appreciate a decreasing dependency between EE and QM (with the exception of some outliers), but this dependency is clearly non-linear.

In order to quantitatively analyze dependency between EE and QM, we have computed the Spearman correlation coefficient considering all corners and all discrete sequences. In tables 1 and 2 we show the Spearman correlation coefficient and its $p - value$. Since

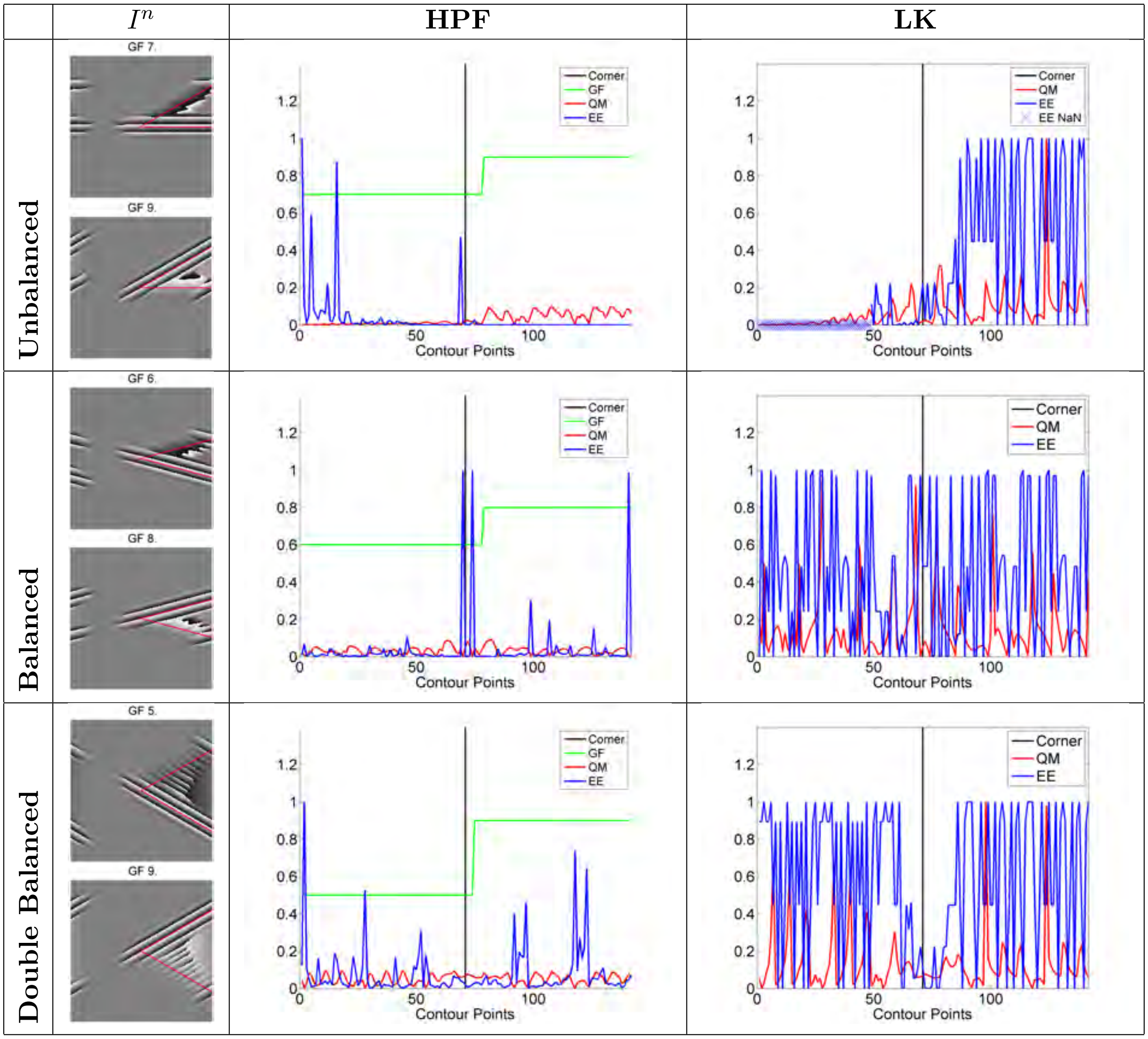


Figure 10: Behavior of the EE and QM for the HPF and LK for the corner $C_{30^\circ}^{k}$ with $\forall k$ with movement α_0. All shown measures are normalized.

$p - value < 0.05$ in all cases, we can reject the null hypothesis, and thus, there is a negative correlation between EE and QM for the HPF and LK local systems.

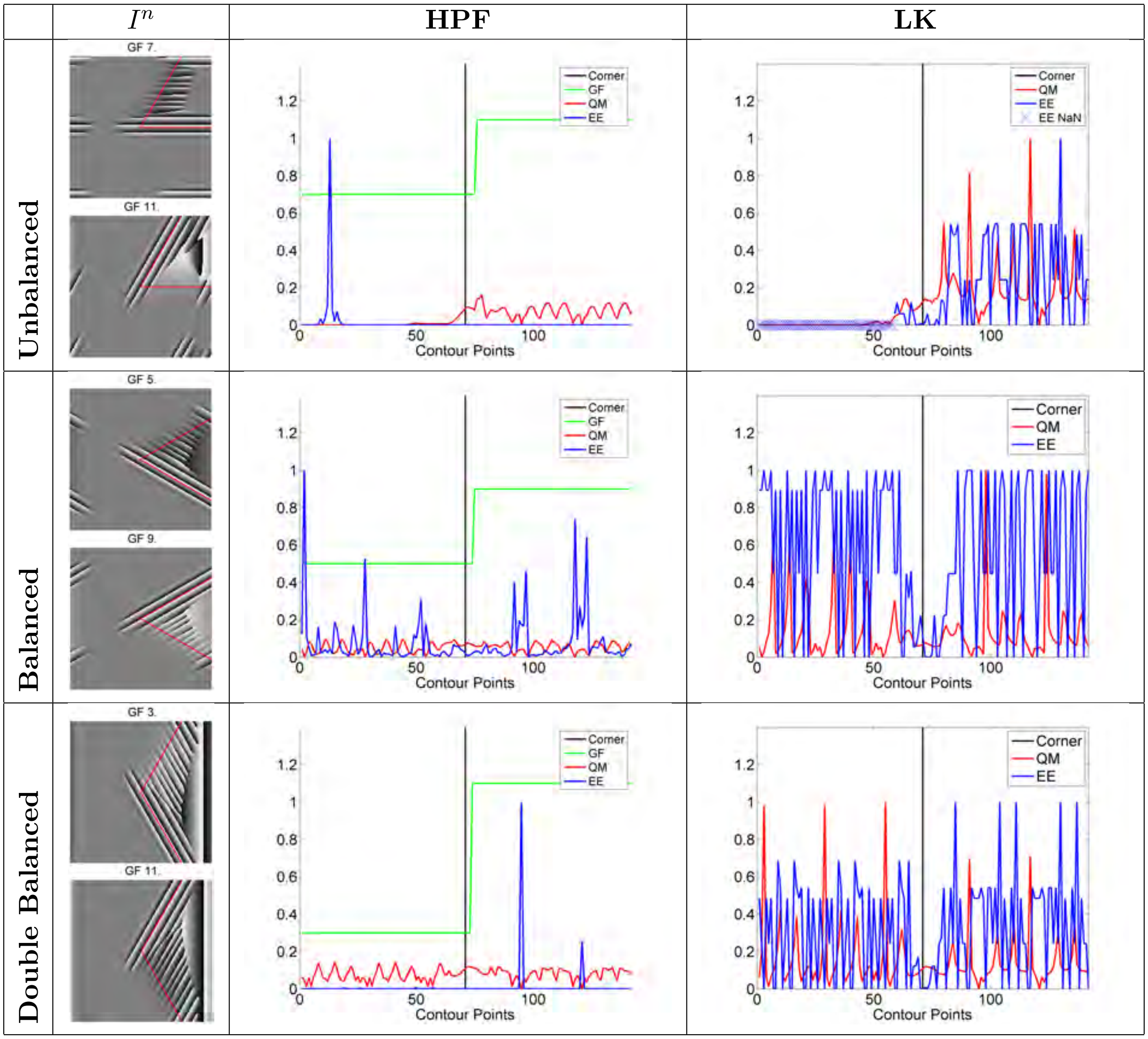


Figure 11: Behavior of the EE and QM for the HPF and LK for the corner $C^k_{60^\circ}$ with $\forall k$ with movement α_0. All shown measures are normalized.

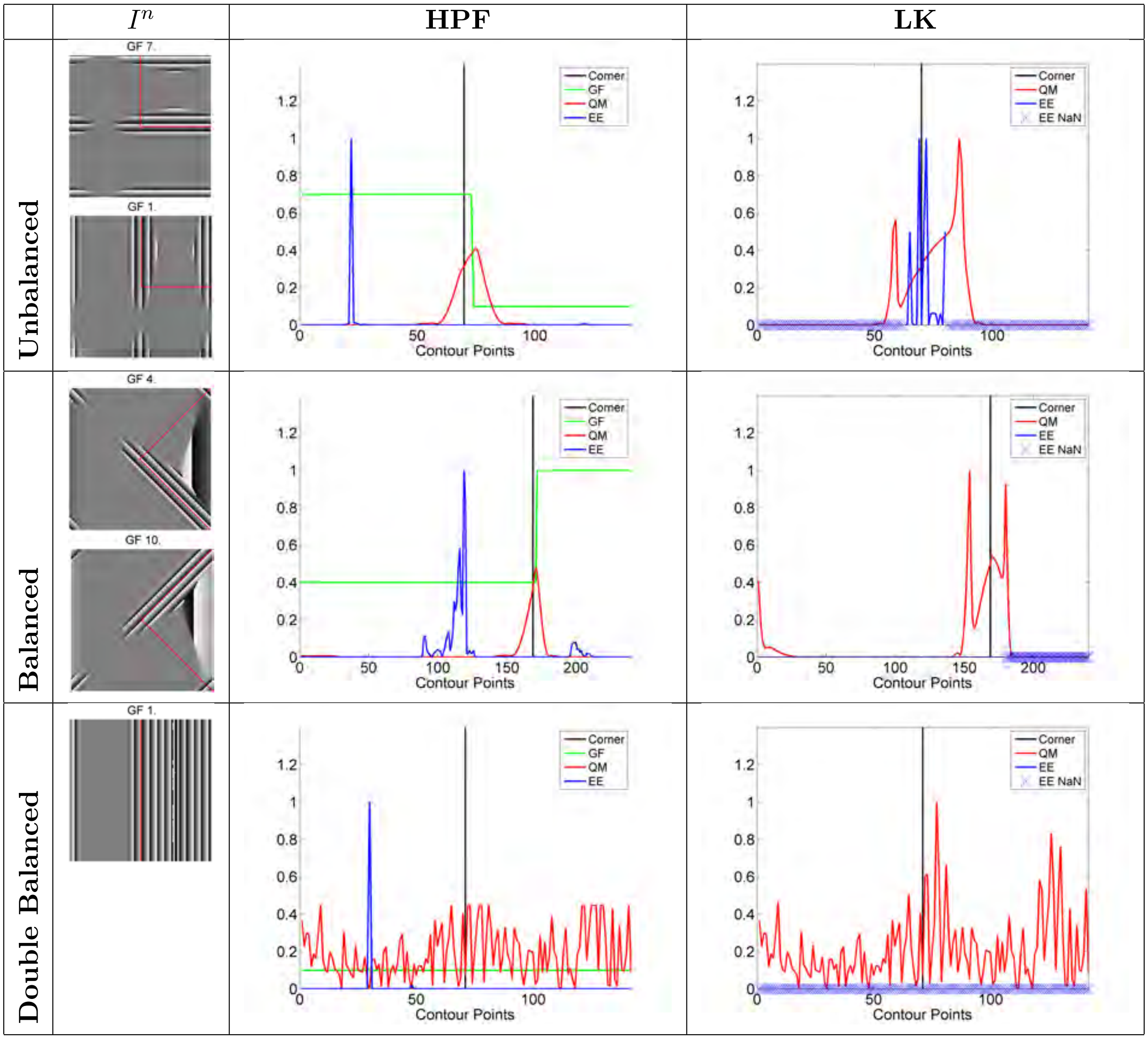


Figure 12: Behavior of the EE and QM for the HPF and LK for the corner $C^k_{90^\circ}$ with $\forall k$ with movement α_0. All shown measures are normalized.

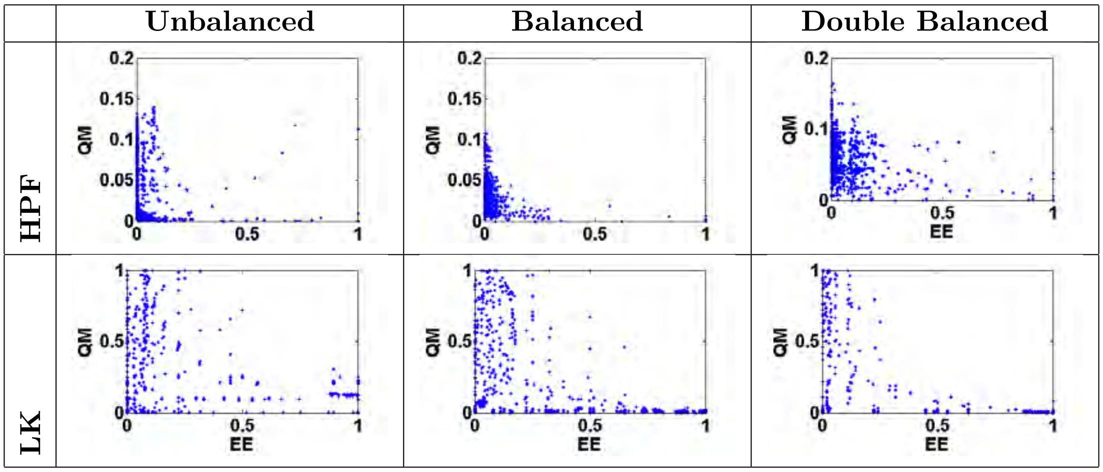


Figure 13: EE versus QM for HPF (first row) and LK (second row) equation systems applied to the corner $C_{30^\circ}^{k}$ with $\forall k$ for the discrete sequences.

HPF. Spearman Test					
Unbalanced		**Balanced**		**Double Balanced**	
ρ	$p - val$	ρ	$p - val$	ρ	$p - val$
-0.065	0.000	-0.038	0.000	-0.078	0.000

Table 1: Spearman test for the HPF equation system considering the discrete sequences and all corners.

LK. Spearman Test					
Unbalanced		**Balanced**		**Double Balanced**	
ρ	$p - val$	ρ	$p - val$	ρ	$p - val$
-0.304	0.000	-0.034	0.000	-0.196	0.000

Table 2: Spearman test for the LK equation system considering the discrete sequences and all corners.

4.2 Optical Flow Accuracy

In tables 3, 4, 5 and 6 we show the results for continuous sequences. In tables 3 and 4 we show, for the HPF and LK respectively, the ranges of the EE for all corners considering the continuous sequences. Besides, in tables 5 and 6 we show, for the HPF and LK, the percentage ranges of the IP with $tol = 10^{-2}$ for all corners considering the continuous sequences. For all tables we have also computed the ranges considering all corners from 10° to 80°. We have not considered the corner 90° because it is a particular case, the unbalanced corner has two pure edges and the double balanced corner is an edge, thus the errors and the percentage of IP for this corner are bigger than for the other ones.

Concerning the statistics shown in tables 3, 4, 5 and 6, when the angle of the corner is quite small or quite big, the edges inclination are close to be horizontal or vertical, so the equation system is singular in most of the points and also, when we can recover motion, the EE is higher. Notice that for the unbalanced and balanced corner 10° the percentage of IP is really low for the HPF method, this is because of the scale influence. Since the corner aperture is really small when we compute the OF in a concrete pixel, we have information from pixels where motion can be recovered properly.

Comparing the statistics shown in tables 3, 4, 5 and 6 for the unbalanced and balanced corners, we can observe that, since the unbalanced corner has a straight edge, the percentage of IP is higher. Besides, for the unbalanced and the double balanced corners of 70°, 80° and 90° statistics, we expect to have a similar EE range (since they have the same edge inclination) and a higher IP range for the unbalanced corner (since it has a straight edge). Notice that, as we expected, IP range is higher for the unbalanced corner. On the contrary, the EE range is quite different. The higher ranges of the EE for the double balanced corner are produced by the derivatives definition.

HPF. End Point Error Averages for all movements. $tol = 10^{-2}$.			
θ	Unbalanced	Balanced	Double Balanced
10°	0.20 ± 0.09	0.10 ± 0.05	0.18 ± 0.08
20°	0.20 ± 0.09	0.18 ± 0.08	0.16 ± 0.07
30°	0.19 ± 0.08	0.22 ± 0.10	0.14 ± 0.05
40°	0.19 ± 0.09	0.16 ± 0.07	0.18 ± 0.09
50°	0.19 ± 0.08	0.14 ± 0.05	0.08 ± 0.05
60°	0.17 ± 0.06	0.14 ± 0.05	0.28 ± 0.06
70°	0.20 ± 0.10	0.18 ± 0.08	0.66 ± 0.16
80°	0.19 ± 0.07	0.18 ± 0.09	0.75 ± 0.41
90°	0.29 ± 0.12	0.22 ± 0.07	3.55 ± 0.60
*All**	0.19 ± 0.08	0.16 ± 0.08	0.30 ± 0.29

Table 3: HPF EE ranges for all corners and considering the continuous sequences. *The range takes into account all corners from 10° to 80°.

Figures 14 and 15 show a qualitative comparison between the original HPF and the improved HPF applied to the Army (Mid) and to the Tram (Onv) sequences respectively. In the first row we show the two frames used to compute the OF. In the second row we show

LK. End Point Error Averages for all movements. $tol = 10^{-2}$**.**			
θ	Unbalanced	Balanced	Double Balanced
10°	0.14 ± 0.07	0.13 ± 0.07	0.19 ± 0.09
20°	0.18 ± 0.08	0.19 ± 0.09	0.18 ± 0.09
30°	0.22 ± 0.10	0.18 ± 0.10	0.26 ± 0.13
40°	0.17 ± 0.07	0.18 ± 0.09	0.22 ± 0.10
50°	0.14 ± 0.06	0.19 ± 0.06	0.21 ± 0.10
60°	0.19 ± 0.12	0.26 ± 0.13	0.21 ± 0.08
70°	0.16 ± 0.09	0.30 ± 0.15	0.16 ± 0.07
80°	0.12 ± 0.04	0.22 ± 0.10	0.15 ± 0.08
90°	0.28 ± 0.13	0.46 ± 0.15	0.55 ± 0.19
*All**	0.16 ± 0.09	0.21 ± 0.11	0.20 ± 0.10

Table 4: LK EE ranges for all corners and considering the continuous sequences. *The range takes into account all corners from 10° to 80°.

HPF. Invalid Points Percentage Averages for all movements. $tol = 10^{-2}$**.**			
θ	Unbalanced	Balanced	Double Balanced
10°	6.15 ± 2.24 %	0.84 ± 0.23 %	10.57 ± 1.34 %
20°	17.25 ± 0.76 %	10.57 ± 1.34 %	6.74 ± 0.74 %
30°	36.41 ± 1.24 %	9.98 ± 0.89 %	5.16 ± 1.56 %
40°	37.19 ± 0.67 %	6.74 ± 0.74 %	6.48 ± 0.39 %
50°	38.97 ± 0.26 %	4.19 ± 1.26 %	1.99 ± 0.06 %
60°	39.88 ± 0.81 %	5.16 ± 1.56 %	3.61 ± 1.87 %
70°	34.18 ± 0.33 %	6.69 ± 1.43 %	1.37 ± 0.65 %
80°	42.75 ± 0.90 %	6.48 ± 0.39 %	7.08 ± 3.80 %
90°	71.26 ± 0.25 %	78.62 ± 0.09 %	85.28 ± 1.77 %
*All**	31.60 ± 12.10 %	6.33 ± 3.11 %	5.37 ± 3.30 %

Table 5: HPF Invalid Points ranges for all corners and considering the continuous sequences. *The range takes into account all corners from 10° to 80°.

<table>
<tr><th colspan="4">LK. Invalid Points Percentage Averages for all movements. $tol = 10^{-2}$.</th></tr>
<tr><th>θ</th><th>Unbalanced</th><th>Balanced</th><th>Double Balanced</th></tr>
<tr><td>10°</td><td>27.02 ± 1.05 %</td><td>28.37 ± 13.06 %</td><td>61.65 ± 7.23 %</td></tr>
<tr><td>20°</td><td>35.24 ± 0.58 %</td><td>61.65 ± 7.23 %</td><td>53.49 ± 17.41 %</td></tr>
<tr><td>30°</td><td>37.74 ± 0.43 %</td><td>57.54 ± 15.21 %</td><td>68.85 ± 16.33 %</td></tr>
<tr><td>40°</td><td>38.39 ± 0.27 %</td><td>53.49 ± 17.41 %</td><td>45.12 ± 18.96 %</td></tr>
<tr><td>50°</td><td>39.38 ± 0.40 %</td><td>34.48 ± 12.68 %</td><td>37.84 ± 38.12 %</td></tr>
<tr><td>60°</td><td>40.43 ± 0.35 %</td><td>68.85 ± 16.33 %</td><td>35.22 ± 36.76 %</td></tr>
<tr><td>70°</td><td>40.64 ± 0.37 %</td><td>71.58 ± 16.68 %</td><td>28.51 ± 31.05 %</td></tr>
<tr><td>80°</td><td>40.92 ± 0.22 %</td><td>45.12 ± 18.96 %</td><td>34.83 ± 36.08 %</td></tr>
<tr><td>90°</td><td>84.71 ± 2.09 %</td><td>86.49 ± 1.46 %</td><td>93.10 ± 2.00 %</td></tr>
<tr><td><i>All*</i></td><td>37.47 ± 4.35 %</td><td>52.64 ± 20.99 %</td><td>45.69 ± 30.60 %</td></tr>
</table>

Table 6: LK Invalid Points ranges for all corners and considering the continuous sequences. *The range takes into account all corners from 10° to 80°.

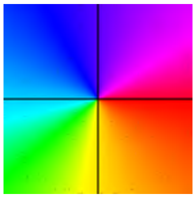
the color code (notice that it is normalized). In the third row we show the OF computed in color codification, the original HPF on the left image and the improved HPF on the right image. Finally, in the fourth row we show a detail of the OF in vectors codification, the original HPF on the left image and the improved HPF on the right image.

Since the Army sequence has several independently moving objects, we expect to recover different motions for different objects. In figure 14 the details of the color codification image are more defined in the improved HPF solution and we can distinguish the soldiers and other details of the sequence. The noise of the object contours is high due to the method does not over-regularize the velocity vector field. Notice that in the detail we can appreciate that OF field with the improved HPF distinguish the flow of different objects. Since the temporal resolution is quite high and the objects have independent movements, the scale that performs better when computing the OF is scale 1.

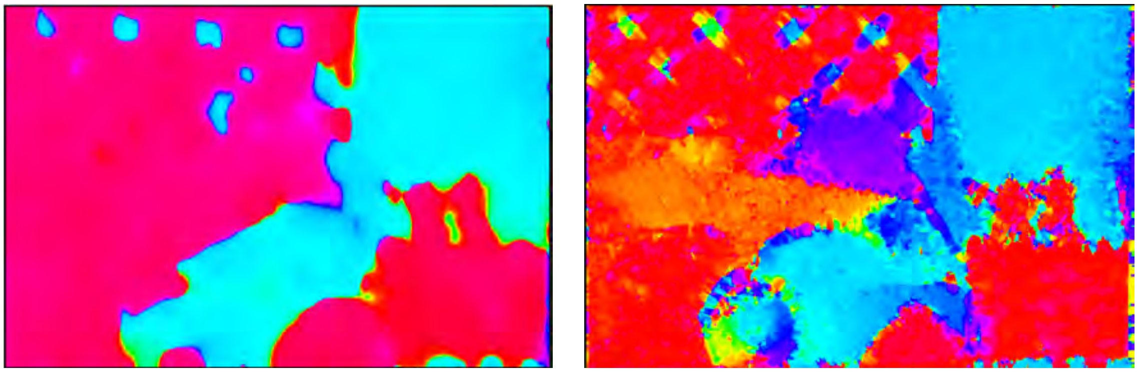
Since the Tram sequence contains a tram's driver, motion goes from the image frame to the vanishing point at the center. In figure 15 with the improved HPF we have obtained the vanishing point, which is an important point of the sequence. Also notice that the OF solution is too noisy, specially in the frame of the image, it might be produced because in this region the temporal resolution is the lowest of the whole image and also for the image resolution, which is really bad. Notice that in the detail we can see that even there is noise in the OF computation the velocity vector field is well defined. Since the temporal resolution is quite low and motion is continuous for the whole image, the scale that performs better when computing the OF is the scale 3.

Notice that in both sequences we can appreciate a clear improvement with the incorporation of the new weights and also, the improved HPF returns non over-regularized OF fields.

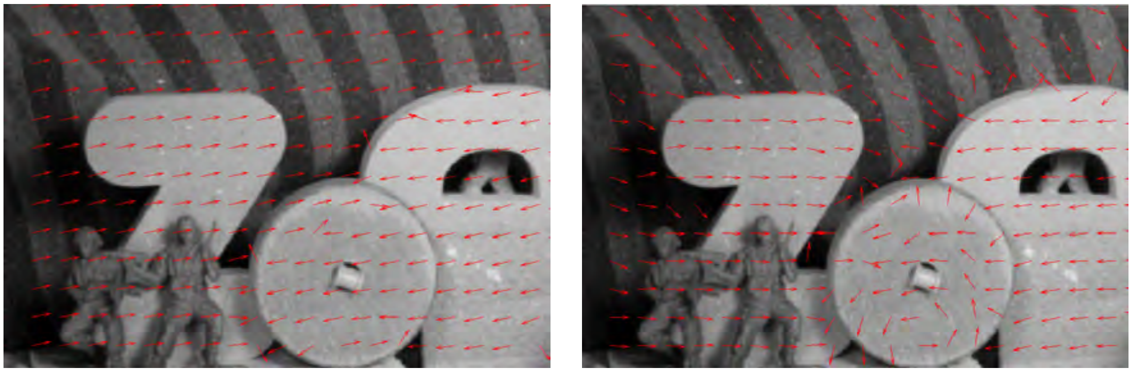
Frame 07 Frame 08



Flow Color Coding



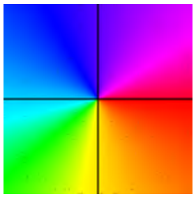
HPF OF Improved HPF OF



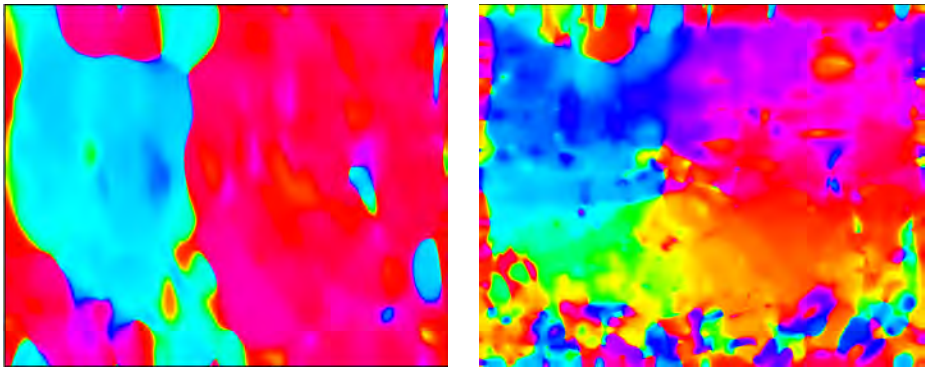
Detail HPF OF Detail improved HPF OF

Figure 14: Army sequence (Mid). With the original HPF and the improved HPF. Using scale 1 of the GF.

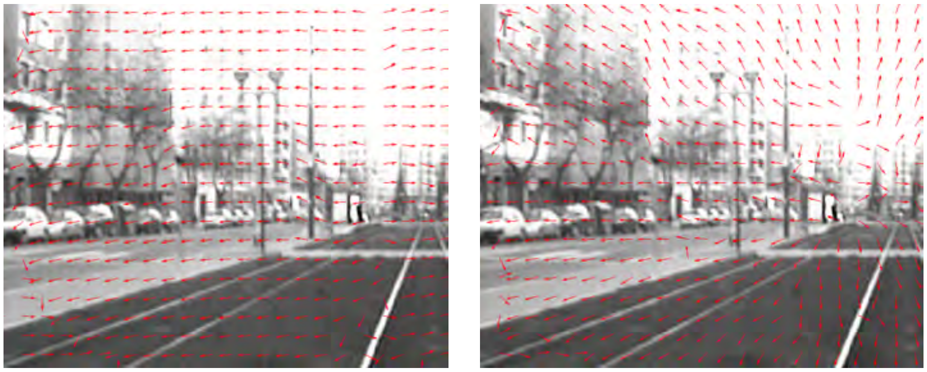
Frame 2 Frame 3



Flow Color Coding



HPF OF Improved HPF OF



Detail HPF OF Detail improved HPF OF

Figure 15: Tram sequence. With the original HPF and the improved HPF using scale 3 of the GF.

5. Conclusions and Future Work

In this project we have studied the conditions that ensure the accuracy of the local optical flow schemes of LK and HPF. In addition we have introduced those conditions as a weights in the variational framework of the HPF and we have done a first qualitative experiment to test the new weights. As far as we know, this is the first time that weights depend on the accuracy of the local system solutions and not on the image features.

Through the experiments with the discrete sequences, since QM and EE have a decreasing dependency, we have checked that QM that ensure the accuracy of the local systems solutions are well defined. In addition we have observed that even theoretically we can only recover motion in corners, due to images are in a discrete domain, we can also recover motion in some edges.

Regarding the experiment with the continuous database, we can observe that HPF is more stable to the corners geometry than LK. This is produced because HPF changes the space representation to a smoother one. In addition, since ranges of IP percentages are lower for HPF, HPF solves the aperture problem in more cases than LK. However, notice that HPF computes the solution over a periodic image and due to the inherent properties of our database we are favoring the OF computation. Therefore, we have planned to repeat the experiments in order to avoid giving more advantages to one system than to the other one. Nevertheless, we do not think that the results will be significantly different. Besides, since we are not computing EE ranges over the same points, LK EE and HPF EE ranges are not comparable, thus we have planned to repeat the experiment considering also the intersection of non IP of both methods.

In the case that applying the improved HPF to real sequences we can observe a clear improvement on the computation of the OF. In addition, we are not over-regularizing the solution. Therefore, implementing weights that take into account the theoretical requirements that ensure the solution of the local equation systems and that consider the data-term and the smoothness-term complementary terms, provide a real improvement in the OF computation.

We do not have any quantitative experiments of the variational HPF implementation because it suffers errors computing the magnitudes. This errors might be produced by the numerical scheme formulation. Thus we have planned to debug the magnitude computation error in the variational HPF.

Another point is to implement the weights in the variational framework developed by Bruhn et al. (2005) and compare qualitatively the results against the original implementation. In addition it would be interesting to analyze quantitatively the results obtained applying both improved variational frameworks to real life sequences.

We also have planned to implement coarse-to-fine techniques in order to avoid possible errors due to low temporal resolution of the sequence.

Finally, since the results are affected by the derivatives (which are non symmetric), we have planed to improve the derivatives definition.

References

Middlebury database. http://vision.middlebury.edu/flow/.

Onvisupra project. http://www.cvc.uab.es/ asappa/onvisupra.html.

P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *IJCV*, 2:283–310, 1989.

S. Baker, D. Scharstein, J.P. Lewis, S. Roth, M. J. Black, and R. Szeliski. A database and evaluation methodology for optical flow. Technical Report 179, Microsoft Research, 2009.

J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *IJCV*, 12(1):43–77, 1994.

A. Bruhn. *Variational Optic Flow Computation - Accurate Modelling and Efficient Numerics*. PhD thesis, Saarland University, Saarbrücken, Germany, 2006.

A. Bruhn, J. Weickert, and C. Schnörr. Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. *IJCV*, 61(2):221–231, 2005.

B. F. Buxton and H. Buxton. Computation of optic flow from the motion of edges in the image sequences. *Image and Vision Computing*, 2:59–75, 1984.

M. Felsberg. Optical flow estimation from monogenic phase. *IWCM*, LNCS 3417:1–13, 2004.

D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *IJCV*, 5(1):77–104, 1990.

J. Garcia. *Statistical Models of the Architecture and Function of the Left Ventricle*. PhD thesis, Autonomous University of Barcelona, Bellaterra, Spain, 2009.

J. J. Gibson. *The Perception of the Visual World*. Riverside Press, Cambridge, 1950.

J. J. Gibson. *The Senses Considered as Perceptual Systems*. Houghton-Mifflin, Boston, 1966.

B. Horn and B. Schunck. Determining optical flow. *Artificial Interlligence*, 17:185–203, 1981.

B. Jähne. *Spatio-Temporal Image Processing. Lecture Notes in Computer Science*, volume 751. Springer Berlin / Heidelberg, Berlin, 1993. ISSN:0302-9743.

B. Lucas and T. Kanade. An iterative image registration rechnique with an application to stereo vision. *Proc. DARPA IU Workshop*, pages 121–130, 1981.

H. H. Nagel and W. Enkelmann. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEE Transactions on Pattern Analysis and Machine Intelligence*, 8:565–593, 1986.

N. Onkarappa and A. D. Sappa. Advances in variational optical flow. *4th CVC Research & Development Workshop*, 2009.

W. Pratt. *Digital Image Processing 2nd Ed.* John Wiley & Sons, New York, 1991. ISBN:0-471-85766-1.

A. Singh. An estimation-theoretic framework for discontinuous flow fields. *Proc. IEEE ICCV*, pages 168–177, 1990.

D. Sun, S. Roth, and M. J. Black. Secrets of optical flow estimation and their principles. 2010.

A. M. Waxman, J. Wu, and F. Bergholm. Convected activation profiles and the measurement of visual motion. *Proc. IEEE CVPR*, pages 717–723, 1988.

A. Wedel, T. Pock, C. Zach, H. Bischof, and D. Cremers. An improved algorithm for tv-l1 optical flow. pages 23–45, 2009.

J. Wills, S. Agarwal, and S. Belongie. A feature-based approach for dense segmentation and estimation of large disparity motion. *IJCV*, 68(2):125–143, 2006.