A New Framework for Optical Flow Estimation

Dr. W. A. Mahmoud* Dr. M. S. AbdulWahab H. N. AL-Taai****

Received On: 24/5/2004
Accepted On: 11/9/2004

Abstract

This paper presents a new method to estimate optical flow in a differential framework. The image sequence is first convolved with a spatiotemporal filter similar to those that have been used in other early vision problems such as texture and stereopsis. The brightness constancy constraint can then be applied to each of the resulting images, giving in general an over determined system of equations for the optical flow at each pixel. It is based on prediction of global flow field parameters, performs better than multi-resolution estimation methods and has been verified using standard test sequences as well as real-world data.

الخلاصة

يُقَدِّمُ هذا البحث طريقة جديدة لاستخلاص التدفّق البصريّ في هيكل تفاضليّ. يقوم أولا بعمل التفاف (convolution) لمتابعة الصور باستخدام مرشح وقتيّ-فضائي spatiotemporal مماثل إلى الذي اُستُعْمِلَ في مشاكل رؤية مُبكِّرة أخرى مثل النسيج texture and stereopsis ثمّ يُطبِّقَ قيد ثبات السطوع على الصور الناتجة من عملية الالتفاف، الذي يُسَلِّمُ بشكل عام إلى نظام معادلات للتدفّق البصريّ في كلّ pixel. إن هذه الطريقة تُستَنَدُ إلى التنبؤ إلى متغيرات مجال التدفق الشامل، وأثبتت أنها افضل من طرق الاستخلاص متعددة القرار، وأختُبِرَت على متتابعات صور قياسية إضافة إلى معلومات عالم حقيقي.

Keyword Optical flow estimation, gradient-based method, Smoothness constraints, brightness constancy assumption.

1. Introduction

Motion is one of the most important research topics in computer vision. It is the base for many other problems such as visual tracking, structure from motion, 3D reconstruction, video representation, and even video compression. Let us first define some basic concepts [1]:

Motion field is a 3D field of object velocities at point of space. The 3D motion of object in a time varying scene is defined completely by the motion field.Image flow is the visible portion of the 2D projection of the motion field onto the image plane. We would like to obtain the image flow as an intermediate result in the 3D motion estimation process, and then try to recover the 3D motion field from its 2D projection. However, this seems to be impossible without a prior knowledge on the motion field. Instead, we extract an optical flow, which is a 2D field of velocities associated with the variation of brightness patterns of the image. The evolution of the image in time is caused by two main factors [2]:

- Sudden changes between two successive sequences, that are usually relatively rare,
- The relative motion between objects in the scene and the camera.

The relative motion between

* Dept.of Electrical Eng., University of Baghdad, Baghdad-IRAQ.
** Dept.of Electrical Eng., University of Technology, Baghdad-IRAQ.

objects and the camera is a three-dimensional vector field in the reference frame of the camera which we call the real motion. The scene is then projected to the camera image plane, and we can thus define a second vector field: the projected motion field. A velocity vector is associated with each image point, and a collection of such velocity vectors is a 2D motion field.

In Fig. (1), p is a 3D point, i.e., $p = [X;\ Y;\ Z]^T$, and m is its image projection, i.e., $m = [x\ ;\ y]^T$. Then we have: $P = Z\,\hat{m}$

where $\hat{m}$ is the homogeneous coordinate of m. Then we have [3]:

$$\frac{dp}{dt} = \frac{dZ}{dt}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} + Z\begin{bmatrix} dx/dt \\ dy/dt \\ 0 \end{bmatrix} \qquad \ldots (1)$$

i.e., $\dot{p} = \dot{Z}\hat{m} + Z\dot{\hat{m}}$

$$V_P = \left(V_P^T k\right)\hat{m} + Z v_m \qquad \ldots (2)$$

where k is the unit vector of the depth direction, So,

$$v_m = \frac{1}{Z}\left(V_P - \left(V_P^T\ k\right)\hat{m}\right) \qquad \ldots (3)$$

which means that the 2D motion field v_m is a function of V_P/Z.

The purpose of optical flow measurement is only to estimate this motion in the image plane from the knowledge of the images sequence I(t;x). Optical flow is defined as the projection of velocities of 3D surface points onto the imaging plane of a visual sensor.

The following two examples help to understand the difference between an image flow and an optical flow. The first one is a uniformly painted ball rotating around its center in some way. In this case, the image flow is non-zero for every point of the ball projection on the image plane, while the optical flow is zero, since the image brightness does not change at all. The second example is a stationary scene with moving light source. Here the situation is exactly the opposite: The optical flow is non-zero due to intensity changes in the image, whereas absence of motion causes zero image flow.

Now the motion recovery problem can be introduced. It is formulated as follows: given a sequence of images of a dynamic scene, recognize moving objects and find their velocities (trajectories). The solution of this problem, like many others in computer vision, can be, somewhat artificially, divided into two main stages [1]:

- Low-Level Processing: During this stage a 2D field of velocities (the optical flow) associated with a velocity vector to each point of the image plane is determined.
- High-Level Processing: At this stage, the 3D velocity field (the true motion field) is estimated from the 2D field, determined at the previous stage, and analyzed in order to get the motion description of objects in the 3D scene.

While the high level stage of motion recovery assumes that it receives the image flow as its input, the low level stage, which is also called the optical flow estimation stage, can only produce the optical flow defined by the image sequence. It is immediately seen that there is a problem of equivalence of these two fields.

2. Optical Flow Constraint Equation

As defined above, the optical flow is a velocity field associated with brightness changes in the image. This suggests an assumption often made in methods for optical flow estimation, the brightness conservation assumption, which states that brightness of an image of any point on the object is invariant under motion.

We denote an image intensity function by I(x , y , t), and the velocity of an image pixel m=[x , y]T is:

$$v_m = \dot{m} = \begin{bmatrix} v_x & v_y \end{bmatrix}^T = \begin{bmatrix} dx/dt \\ dy/dt \end{bmatrix} \qquad \dots (4)$$

The initial hypothesis in measuring image is that the intensity structures of local time varying image regions are approximately constant under motion for at least a short duration (dt), i.e.:

$$\left.\begin{aligned} &I(x+dx,\ y+dy,\ t+dt) = I(x,y,t) \\ &I\left(x+\frac{dx}{dt}dt,\ y+\frac{dy}{dt}dt,\ t+dt\right) = I(x,y,t) \\ &I(x+v_x dt,\ y+v_y dt,\ t+dt) = I(x,y,t) \end{aligned}\right\} \qquad \dots (5)$$

If the brightness changes smoothly with x, y, and t, we expand the left-hand-side by Taylor series:

$$I(x,y,t)+\frac{\partial I}{\partial x}v_x dt+\frac{\partial I}{\partial y}v_y dt+\frac{\partial I}{\partial t}dt+O(dt^2)=I(x,y,t) \qquad \dots (6)$$

So, we have

$$\frac{\partial I}{\partial x}v_x + \frac{\partial I}{\partial y}v_y + \frac{\partial I}{\partial t} = 0 \qquad \dots (7)$$

i.e.

$$\nabla I \cdot v_m + \frac{\partial I}{\partial t} = 0 \qquad \dots (8)$$

Where $\nabla I = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\right]^T$ is image gradient at pixel m, which can be obtained from images. Also $\frac{\partial I}{\partial t}$ can also be obtained from images easily. We call this equation optical flow constrained equation.

Apparently, for each pixel, we have only one constraint equation, but we need to solve two unknowns, i.e., v_x and v_y, which means that we cannot determine optical flow uniquely only from such optical flow constraint equation. Fig. (2) gives a geometrical explanation of the constraint equation [3].

3. Smoothness Constraints

In case of rigid body, neighboring points of a body move similarly, their velocities differ only slightly. This results in a rather smooth optical flow. Horn and Schunk [4] were first to make this assumption and exploit it for determining an optical flow. As a measure of a field smoothness (or, more precisely, unsmoothness), they used the square of the magnitude of the velocity field gradient, i.e.:

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2, \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \quad \dots (9)$$

They transformed optical flow estimation into an optimization problem involving a combination of the two criteria:

- The error in the image brightness changes measurement:

$$E_b = I_x v_x + I_y v_y + I_t \qquad \dots (10)$$

- The quantity reflecting a "non-smoothness" of the velocity field [4]:

$$E_c^2=\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial v}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2 \quad \ldots (11)$$

A weighted sum of these two quantities summed over the image is to be minimized:

$$E^2 = \iint\left(E_b^2+\alpha^2 E_c^2\right)dx\,dy \quad \ldots (12)$$

or

$$E^2 = \int_D(\nabla I\cdot v+I_t)^2+\alpha^2\left(\|\nabla u\|^2+\|\nabla v\|^2\right)dx$$

Since the input image is corrupted by noise and quantization error, we cannot expect E_b to be identically zero. This quantity would have a magnitude proportional to the noise in the measurement, therefore the weighting factor α^2 in the sum should be chosen equal to the estimate of the noise variance in the image [4].

4. Optical Flow Techniques

Many methods for computing optical flow have been proposed and others continue to appear. Lacking, however, are quantitative evaluations of existing methods and direct comparisons on a single set of inputs. These can be roughly grouped into gradient-based, correlation-based, energy-based, phase-based and wavelet-based techniques [4-8].

A typical gradient-based approach was proposed by Horn and Schunck, which is mainly based on optimizing an energy function shown in equation (12) that is function of an image constraint and a smoothness constraint. In general, there are two possible ways to solve the optimization problem for the energy function shown in (12). The first is to convert the optimization problem into one of solving partial differential equations based on variational calculas. This kind of approach estimates flow vector iteratively. The other kind of approach directly uses the discrete version of (12) to calculate the flow vectors. The discretization process converts the original optimization problem into the problem of solving a linear system. However, such intuitive discretization might lose precision of the original energy function and information about the interaction between the image brightness function and the flow field.

Despite their difference, many of these techniques can be viewed conceptually in terms of three stages of processing [5,6]:

- Prefiltering or smoothing with low pass or band pass filters in order to extract signal structures of interest and to enhance the signal to noise ratio.
- Measurement extraction of the basic image structures, such as spatiotemporal derivatives (to measure normal components of velocity) or local correlation surfaces.
- The integration of these measurements either by regularization, correlation, or a least-squares computation aims to produce a 2D flow field, which often involves assumptions about the smoothness of the underlying flow field.

5. The Proposed Image Velocity Calculation

We have developed a new algorithm for computing optical flow in the differential framework which performs comparably to the Horn and Schunck approach, but with less computational cost and a higher density of estimates. The computation of image velocity can be viewed by these steps:

Step 1: presmoothing the images to reduce noise and aliasing effect using a spatiotemporal Gaussian filter with standard deviation of 1.5. By the term Gaussian filter, we mean a low-pass filter with a mask shaped similar to that of the Gaussian probability density function [9]. The term spatiotemporal means that the Gaussian filter is used for low pass filtering in both spatial and temporal domains.

Step 2: Computation of spatio-temporal intensity derivative: The numerical analysis contains many methods for approximating gradient filters. Most of the papers describing optical flow estimation through the brightness change constraint equation (12) apply simple gradient filters like $\frac{1}{2}[-1 \quad 0 \quad 1]$. In many papers, the choice of these filters is even not mentioned. In their original paper [4], Horn and Schunck proposed an approximation of the gradient filter with no pre-smothing. The gradient were obtained by averaging the first diffrences over a neighborhood of 2 x 2 in the image sequence. These gradient estimates refer to a center point of a 2 x 2 x 2 cube (which means that the estimated flow corresponds to points between pixels). No motivation or justification for this choice of gradient estimation is given. According to Barron et. Al. [6], these gradient filters are said to be "relatively crude form of numerical differentation and can be the source of considerable error". Barron et. Al. propose the application of a 5 x 5 x 5 spatio-temporal pre-smoother, construction using a sampled Gaussian filter with 1.5 variance at each axis. This variance was found empirically to give the best result. The gradient filters proposed by Barron is the 5-tap-1-D filter $\frac{1}{2}[-1 \quad 8 \quad 0 \quad -8 \quad 1]$, which is the result of the design procedure described in [10]. This 1-D gradient filter is used to produce 3 types of derivatives (x-derivative, y-derivative, and t-derivative) as described in Fig. (3). As can be seen, a 3-D [5 x 5 x 5] pre-smoothing kernel is first applied to the image sequence. Then, each axis is differentiated separately.

Step 3: Perform Iteration: Given the spatio-temporal derivatives, I_x, I_y and I_t computed as described in the previous step (and hence, the normal velocities), we integrate small neighbourhoods of these values into image velocities. First of all, we reorganize equation (12) into the form:

$$E^2 = \sum_x \sum_y \left(E_b^2 + \alpha^2 E_c^2\right) \qquad \ldots (13)$$

or

$$E^2 = \sum_x \sum_y \left(I_x \cdot u + I_y \cdot v + I_t\right)^2 + \alpha^2 \left(\nabla^2 u + \nabla^2 v\right)$$

The optical flow quantities u and v can be found by minimizing the total error. ∇^2 denotes the Laplacian operator. The Laplacian of u and v are approximated by

$$\begin{aligned} \nabla^2 u &= \bar{u}(x,y) - u(x,y) \\ \nabla^2 v &= \bar{v}(x,y) - v(x,y) \end{aligned} \qquad \ldots (14)$$

Equivalently, the Laplacian of u and v, $\nabla^2 u$ and $\nabla^2 v$, can be obtained by applying a 3x3 window operator, shown in Fig. (4), to each point in the u and v planes, respectively. The solution for velcity vector (u,v) is given as a set of Gauss Seidel equations, which are, solved iteratively:

$$u^{k+1} = \bar{u}^k - \frac{I_x\left[I_x\bar{u}^k + I_y\bar{v}^k + I_t\right]}{\alpha^2 + I_x^2 + I_y^2}$$
$$v^{k+1} = \bar{v}^k - \frac{I_y\left[I_x u^k + I_y v^k + I_t\right]}{\alpha^2 + I_x^2 + I_y^2} \quad \ldots (15)$$

where:

k : the iteration number.

u^0, v^0: Initial velocity estimates (set to zero).

$\bar{u}^k, \bar{v}^k$: Neighborhood averages of u^k, v^k.

There are two different ways to iterate; one way is to iterate at a pixel until a solution is steady. Another way is to iterate only once for each pixel. In the latter case, a good initial flow vector is required and is usually derived from the previous pixel.

6. Evaluation Test Of The Results

We applied here our method to estimate the optical flow on real sequences and synthetic sequences for which 2-D motion fields were known. All image sequences are downloaded from *ftp.csd.uwo.ca*.

6.1 Synthetic Image Sequences

The main advantages of synthetic input are that the 2-D motion field and scene properties can be controlled and tested in a methodical fashion. In particular, we have access to the true 2-D motion field and can therfore quantify performance. Our synthetic image sequences include:

Sinusoidal Inputs: This consists of the superposition of two sinsoidal plane waves:

$$\sin(k_1 \cdot x + w_1 t) + \sin(k_2 \cdot x + w_2 t) \quad \ldots (16)$$

The result reported is based on spacial wavelengths of 6 pixels, with orientations of 54° and –27°, and speeds of 1.63 and 1.02 pixel/frame respectively, and is called Sinusoid 1 as shown in Fig. (5). The resulting plaid pattern translates with velocity v=(1.5539,0.7837) pixel/frame.

Translating Squares: Our other simple test case involves a translating dark square (with a width of 40 pixels) over a bright background as shown in Fig. (6).

Yosemite Sequence: The Yosemite sequence is a more complex test case as shown in Fig. (7). The motion in the upper right is mainly divergent, the clouds translate to the right. This sequence is challenging because of the range of velocities and the occluding edges between the mountains and at the horizon. There is severe alliasing in the lower portion of the images however, causing most methods to poorer velocity measurements.

6.2 Real Image Sequences

Two real image sequences, shown in Fig. (8) and Fig. (9), were also used:

Rotating Rubik Cube: In this image sequence, a rubic's cube is rotating counterclockwise on a turntable. The motion field induced by the rotation of the cube includes velocities less than 2 pixel/frame.

Hamburg Taxi Sequences: In this street scene, there were four moving objects: 1) the taxi turning the corner; 2) a car in the lower left, driving from left to right; 3) a van in the lower right driving right to left; and 4) a predestrain in the upper left.

7. Conclusions

This paper has proposed a new algorithm for optical flow computation using a gradient-based methods. The original problem of minimizing the constraint function can be converted into one of solving partial differential equations based on variational calculas.

Observing that the first order difference is used to approximate the first order differentiation in Horn and Schunk's original algorithm, and regarding this as a relatively curde form and a source of error, our algorithm developed a modified version of Horn and Schunk method. It features a spatio-temporal presmoothing and a more advanced approximation of differentiation. Specifically, it uses a Gaussian filter as a spatiotemporal prefilter.

Our algorithm has achieved better performance than the original one. This success indicates that a reduction of noise in image (data) leads to a significant reduction of noise in optical flow (solution).

Expermental results show that in term of accuracy, our approach outperforms the existing methods which adopted the same objective function as ours. This method is quite stable under noise, though it has one drawback; it fails at sharp changes in image flow i.e. at edges of moving objects. This is explained by the fact that the smoothness assumption used by the method is clearly violated in such edge regions.

8. References

[1] P. Golland, "Use of Color for Optical Flow Estimation"; M.Sc. Thesis in Computer Science, Institute of Technology, Haifa, 1995.

[2] C. P. Bernard, "Wavelets and ill posed problems: optic flow and scattered data interpolation" Ph.D. thesis, Department of Applied Mathematics of Ecole Polytechnique, 1999.

[3] Y. Wu, "Optical Flow and Motion Analysis", ECE510- Computer Vision Notes Series 6, www.ece.northwestern.edu/~yingwu/teaching/ECE432/Notes/optical_flow.pdf.

[4] B. K. P. Horn and B. G. Schunck, "Determining Optical Flow" Artificial Intelligence, 17, pp.(185-204), 1981.

[5] J. L. Barron, D. J. Fleet and S. S. Beauchermin, "Performance of Optical Flow Techniques", IJCV, 12(1), pp. (43-77), 1994.

[6] S. S. Beauchermin and J. L. Barron, "The Computation of Optical Flow", ACM Computing Surveys, Vol. 27, No. 2, pp. (433-467), 1995.

[7] Y. T. Wu, T. Kanade, J. Cohn, and C.-C. Li, "Image registration using wavelet-based motion model," *Int. J. Computer Vision*, vol. 38, no. 2, pp. 129–152, 2000.

[8] Li-Fen Chen, Hong-Yuan Mark Liao, and Ja-Chen Lin, "Wavelet-Based Optical Flow Estimation", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No. 1, January 2002.

[9] Y.Q. Shi and H. Sun, "Image and Video Compression for Multimedia Engineering: Fundamentals, Algorithms, and Standards", CRC Press LLC, 2000.

[10]E. P. Simoncelli, "Design of Multi-Dimensional Derivative Filters", IEEE Int'l Conf. on Image Processing, Vol.1, Pages 790-793, November 1994.

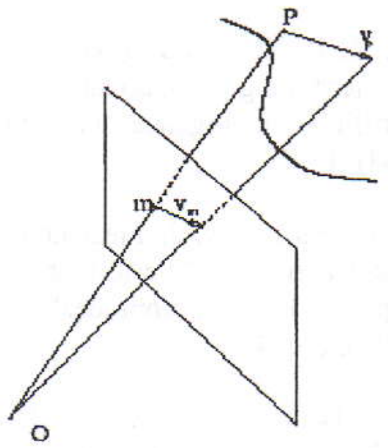


Fig. (1): Motion Field

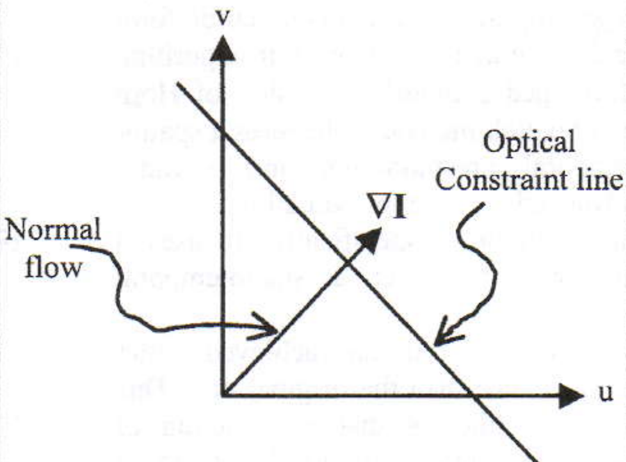


Fig.(2):Geometrical xplanation of the optical flow constraint

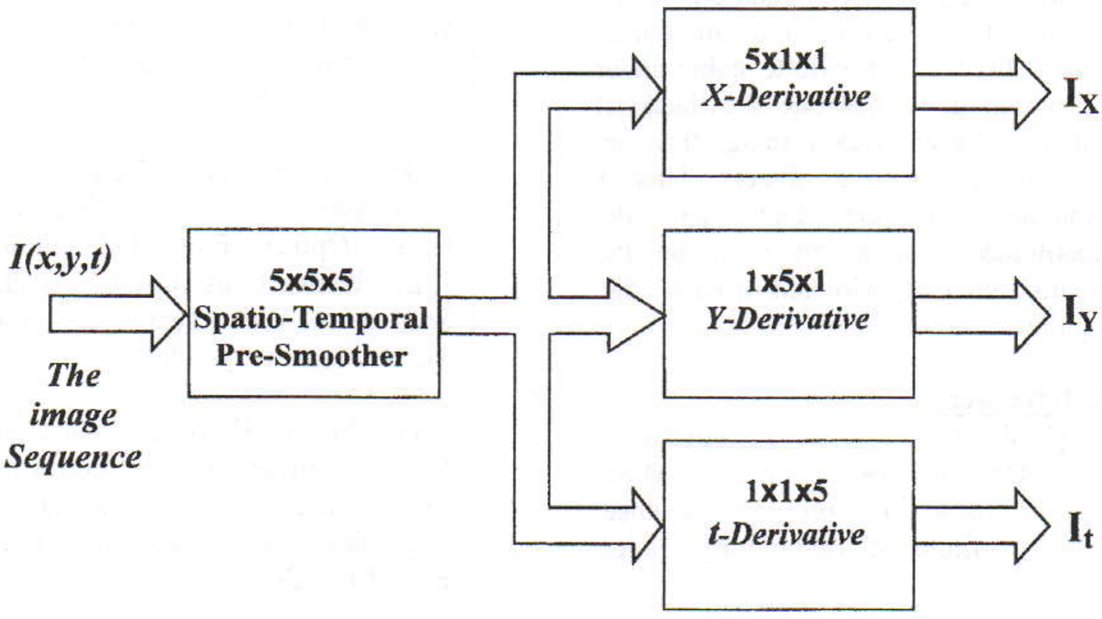
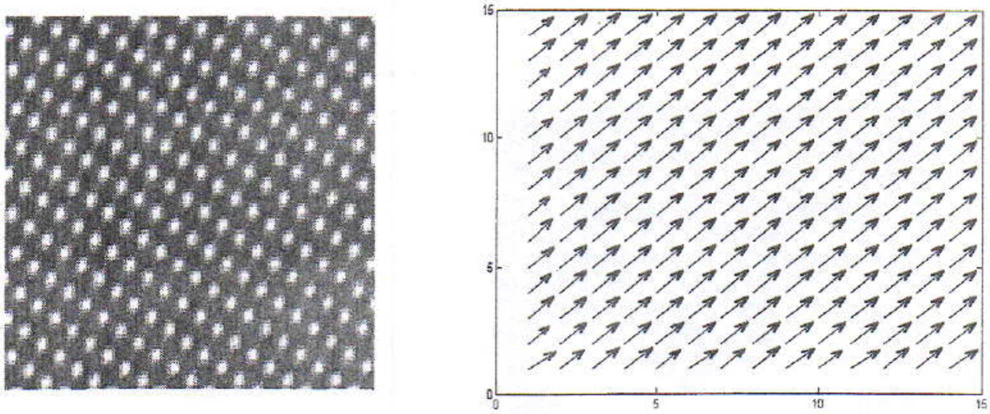


Fig. (3): Presmoothing and Spatio-temporal derivatives.

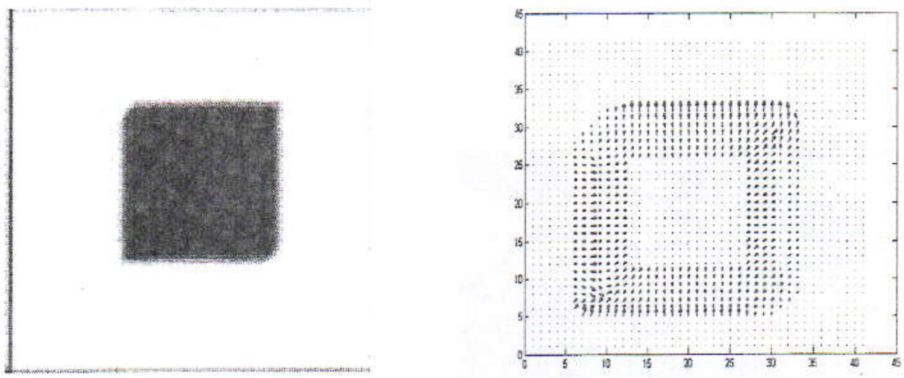
1/12	1/6	1/12
1/6	-1	1/6
1/12	1/6	1/12

Fig. (4): A 3x3 window operation for estimation of the Laplacian mask.



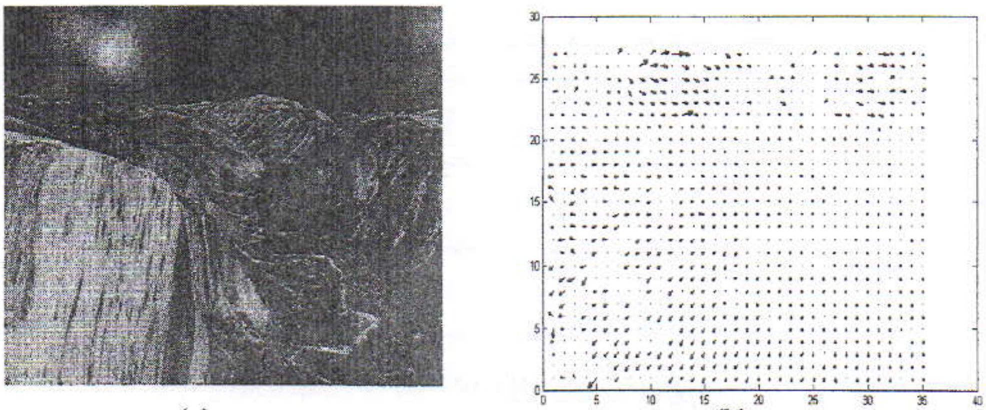
(a) (b)

Fig. (5): a) Sinusoid1 Image. b) Sinusoid1 Flow.



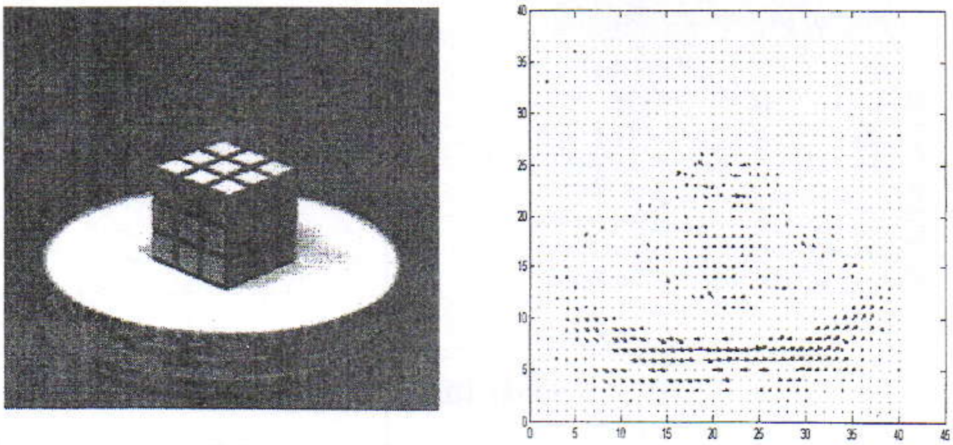
(a) (b)

Fig. (6): a) Square Image. b) Square Flow.



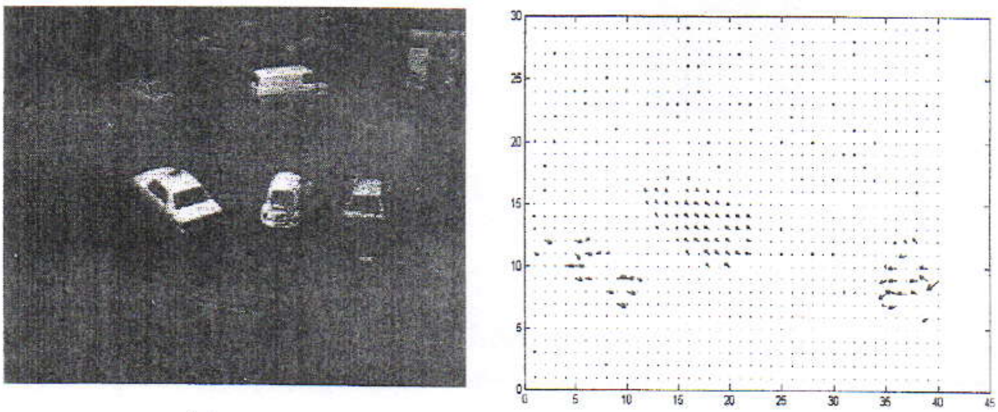
(a) (b)

Fig. (7): a) Yosemite Image. b) Yosemite Flow.



(a) (b)

Fig. (8): a) Rotating Rubik Image. b) Rotating Rubik Flow.



(a) (b)

Fig. (9): a) Hamburg Taxi Image. b) Hamburg Taxi Flow.