Parallel Near-Real-Time Implementation of Multiple Light Source Optical Flow

Cristina E. Siegerist
Laboratory for Computational Intelligence
Department of Computer Science
The University of British Columbia
Vancouver B.C., Canada V6T 1Z4
e-mail: cristina@cs.ubc.ca

Abstract

Multiple light source optical flow is one method to compute a dense, local representation of optical flow. The basic idea is to overcome the aperture problem using multiple images of a moving object acquired simultaneously under different conditions of illumination. Each image provides one linear constraint equation. When the optical flow and the 2D motion field coincide these equations are in the same unknowns. Two equations are enough to calculate both velocity components, three or more images over-determine the system, allowing the calculation of confidence measures and making the computation more robust.

This paper presents a parallel near-real-time implementation (4 frames/sec) of multiple light source optical flow. The range of linear behavior of the system was studied. A comparison with other methods for a synthetic and a real image sequence is presented.

1 Introduction

Vision systems are designed to provide us with information about the three-dimensional structure of the world given the sequence of two-dimensional images continuously received by the retina. One important source of information is the change in time of the intensity pattern due to the three-dimensional motion of surfaces. The first stage in the processing of this information is to analyze the change in the intensity pattern to obtain the optical flow field, defined as the 2D velocity vector of the brightness values. The second stage is the use of the optical flow field, for example, to segment objects in the scene and to reconstruct the three-dimensional structure of the scene.

There are two important problems in the measurement of visual motion. One problem, known as the aperture problem[1], is that optical flow cannot be determined locally for images acquired under fixed conditions of illumination. Local measurements of motion in the brightness pattern provide only one component of the velocity vector. The second problem is that only under certain circumstances will the optical flow and the true motion field coincide[2].

This paper will focus on the implementation of a parallel near-real-time multiple light source optical flow method [3]. The basic idea is to use multiple images taken under different conditions of illumination to overcome the aperture problem. Each image is assumed to satisfy the standard optical flow constraint equation. When the optical flow and the 2D velocity field coincide, these multiple equations are in the same unknowns. Two equations are enough to determine the solution, however, three equations over-determine it, avoid local degeneracies and help to make the computation more robust[3].

The organization of this paper is as follows. Section 2 will describe in detail the optical flow problem and the approach followed in this implementation. Section 3 will focus on the implementation of the system using a network of C40 processors. Section 4 will present the results and a comparison with other methods and section 5 will present the conclusions.

2 Optical Flow

2.1 Introduction

The motion of an object in front of a camera produces intensity changes in the recorded 2D image. These intensity changes can be used to recover the relative motions and shapes of the objects in the scene and to segment the image into parts corresponding to different objects. A first step in solving these problems is to recover the motion field, which can be defined as the perspective projection onto the image plane of the 3D velocity field of moving objects in space. What is available to us is the motion of the brightness pattern in the image, known as optical flow.

There are some well known examples in which the optical flow does not give any information about the motion field. This is the case for a rotating sphere with no texture, under fixed illumination. The sphere rotates but the brightness pattern in the image does not change. The motion field coincides with optical flow when the moving surface undergoes pure translation, the illumination is uniform and the reflectance function is lambertian. In real images these conditions are never strictly satisfied but it is assumed that they hold locally.

2.2 The Optical Flow Constraint Equation

One of the most common assumptions in optical flow techniques is that the intensity of an image point does not change over time. This assumption can be stated formally as

$$E(x,y,t) = E(x+\delta x, y+\delta y, t+\delta t) \qquad (1)$$

Expanding the right hand side of equation(1) in a Taylor series [4] and keeping terms up to the first order we obtain the gradient constraint equation

$$\nabla E(\mathbf{x},t)\cdot\mathbf{v} + \partial E/\partial t = 0 \qquad (2)$$

where $\nabla E = (\frac{\partial E}{\partial x}, \frac{\partial E}{\partial y})$, $\mathbf{x} = (x,y)^T$ and $\mathbf{v} = (u,v)^T$. Equation(2) gives only the component of the image velocity in the direction of the intensity gradient. The component perpendicular to this direction can not be recovered unless more information is available. Equation(2) can be interpreted as a line in velocity space. Any velocity in this line will satisfy the equation. More constraint lines are needed to determine the full velocity vector.

2.3 Multiple Light Source Optical Flow

There are many ways to overcome the aperture problem. Horn and Schunck [4] use the standard optical flow equation together with a global smoothness constraint (regularization). Lucas and Kanade[5] combine optical flow constraint equations over a small window. Other methods combine second-order differential methods with global or local constraints. References [6] and [7] provide a survey of the different techniques used to estimate optical flow.

The approach followed in this work is to use multiple light sources [3] in order to get additional constraint equations. For three light sources, we obtain,

$$\begin{aligned} E_{1x}\,u + E_{1y}\,v + E_{1t} &= 0 \\ E_{2x}\,u + E_{2y}\,v + E_{2t} &= 0 \\ E_{3x}\,u + E_{3y}\,v + E_{3t} &= 0 \end{aligned} \qquad (3)$$

this equation can be written as

$$\mathbf{A}\mathbf{x} = \mathbf{b} \qquad (4)$$

where $\mathbf{x} = (u,v)^T$, $\mathbf{b} = -(E_{1t}, E_{2t}, E_{3t})^T$ and

$$\mathbf{A} = \begin{bmatrix} E_{1x} & E_{1y} \\ E_{2x} & E_{2y} \\ E_{3x} & E_{3y} \end{bmatrix} \qquad (5)$$

The standard least squares solution, $\mathbf{x}$, is given by,

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \qquad (6)$$

The solution is unique provided that the rank of $\mathbf{A}$ is 2. The assumption that equations(3) are defined in identical variables u and v is valid only when the optical flow and the motion field coincide[3].

2.4 Error measures

The fact that three light sources over-determine the system of equations(3) can be used to test the validity of the solution. A relative error term can be defined that takes into account the fit to the model (how good is the estimate of $\mathbf{x}$ in the equation $\mathbf{A}\mathbf{x} = \mathbf{b}$) and the redundancy of the measurements. Reference[3] defines this error as

$$\frac{\|\,\mathbf{r}\,\|}{\|\,\mathbf{b}\,\|} = \frac{\|\,\mathbf{b}-\mathbf{A}\mathbf{x}\,\|}{\|\,\mathbf{b}\,\|} \qquad (7)$$

where $\|\cdot\|$ denotes the l_2 norm. The error is zero if $A\mathbf{x} = \mathbf{b}$ in the case of three non-redundant measurements or when one of the intensity triplets

$(E_{ix}, E_{iy}, E_{it}, i = 1, 2, 3)$ is zero and the equation $\mathbf{Ax} = \mathbf{b}$ is reduced to a two light source case.

Another desirable feature of the system of equations (3) is to be well-conditioned so that the solution is stable. A measure of the stability of $\mathbf{x}$ is the condition number of $\mathbf{A}$, $\kappa(\mathbf{A})$, which can be calculated as[8]

$$\kappa(\mathbf{A}) = \sqrt{\kappa(\mathbf{A}^T\mathbf{A})} = \frac{\sqrt{\lambda_{max}}}{\sqrt{\lambda_{min}}} \qquad (8)$$

where λ_{max} and λ_{min} are the maximum and minimum eigenvalues of $\mathbf{A}^T\mathbf{A}$, respectively. We want $\kappa(\mathbf{A})$ to be small for the solution to be stable.

3 Implementation

3.1 Experimental Setup

The results presented in this paper were obtained using the Calibrated Imaging Facility (CIF) of the Laboratory for Computational Intelligence at UBC. The CIF consists of an optical bench (4 ft x 8 ft) with mounting hardware for positioning and motion of cameras, light sources and test objects. The equipment that was used includes a Sony DXC-755 3 CCD RGB camera with Fujicon 10-120mm (manual) zoom lens, three Newport MP-1000 Moire (white light) projectors with Nikon lenses and spectral filters, two linear Daedal rail tables plus one rotational, and associated controllers, motors, mounting hardware and power supplies.The CIF is integrated with a C40 network and image processing DigiColor and MaxVideo-200 subsystems.

This work requires multiple images of a scene acquired simultaneously under different conditions of illumination. This is achieved by multiplexing the spectral dimension. Three light sources with color filters, one red, one green and one blue, illuminate the work place from different directions. The filters are the Newport FS-225 set manufactured by Corion (Holiston, Mass.) and are Corion parts CA500(blue), CA550(green), and CA600(red). There is negligible overlap in the visible spectrum between the red-light source and either the green-light source or the blue-light source. There is a small overlap between the green and the blue sources for wavelengths between 500-520 nm (see [11] for more details). The Sony 3 CCD camera is used to acquire three separate black and white images simultaneously. These images correspond to the different conditions of illumination.

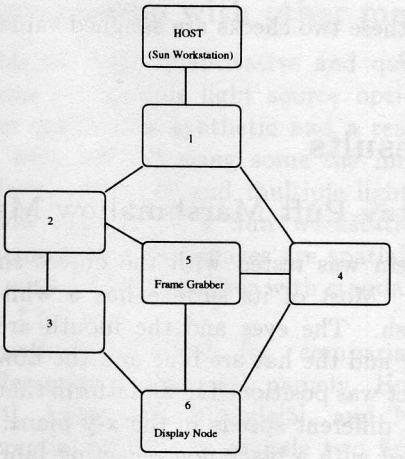


Figure 1: Topology of the C40 network.

3.2 Processing Scheme

Six C40 processors are connected in a network with a Sun workstation as a host. Figure(1) shows the topology of the network. Three black and white images corresponding to the different conditions of illumination are simultaneously acquired by the camera and sent to the frame buffer in node 5. All input images were quantized to 8 bits-per-pixel. The images are segmented into four smaller images and sent to nodes 1-4. The least squares solution for the velocity components is calculated in parallel in these four nodes and the optical flow components are sent to node 6. In this node the flow components are encoded using a color encoding and displayed in a color monitor. Hue encodes direction and brightness encodes magnitude.

The derivatives were estimated using three different schemes. (1) Central differences were calculated using a 3x3x3 cube of brightness values. The derivatives at any point were estimated using the mask $1/2(-1, 0, 1)$. (2) First differences were calculated using 2x2x2 cube with a mask $(-1, 1)$. In both cases spatial and temporal dimensions were treated symmetrically. (3) Four point (time) differences were calculated using a 3x3x5 cube, calculating central differences for the spatial dimensions and using a four point mask $1/12(-1, 8, 0, -8, 1)$ for the time derivatives. No pre-smoothing was performed in any of the images. Once the nine partial derivatives $(E_{ix}, E_{iy}, E_{it}, i = 1, 2, 3)$ are estimated, the value of the flow field is calculated following the least squares solution of equation(6). Two checks have to be done to guarantee that the computation is not degenerate. First, at least two of the spatial gradients must be non-zero. Second, the rank of $\mathbf{A}$, given by equation(5), must be two. Points that fail

either of these two checks are assigned values $u = 0$ and $v = 0$.

4 Results

4.1 Stay-Puft Marshmallow Man

The system was tested with the object shown in figure(2). Most of its surface has a white semi-gloss finish. The eyes and the mouth are black, the collar and the hat are blue and the bow is red. The object was positioned in a platform that can be moved at different speeds in the x-y plane. It was surrounded with a black non reflecting fabric. The

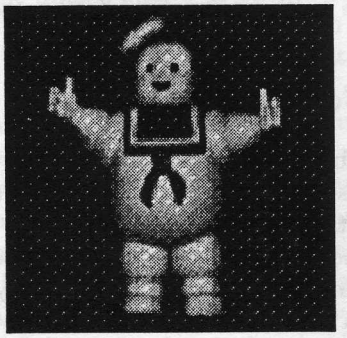


Figure 2: Object used in the experiments.

threshold for the brightness gradient was set to 2.5 ($\| \nabla E \| \geq 2.5$). Points that fail this test are discarded. The processing rate depends on the gradient threshold and on the number of moving points. In our experiments, for an image size of 240x240 pixels, the rate was between 3.5 and 4.0 frames per second. The estimates are good at the points where the surface is smoothly shaded. The estimates are inaccurate at discontinuities and where the brightness changes are too small.

The behavior of the system was studied for motion restricted to the horizontal direction. Since differentiation operators are not rotationally invariant we expect these results to vary slightly for motion in different directions. The object was moved at different speeds and the average optical flow (over the entire image) was measured. Figure(3,(a)) shows the magnitude of the measured average speed versus the real speed in pixels per frame for an image of 240x240 pixels and for this image averaged to 120x120 pixels. The statistics were done for 50−100 frames for each speed. The behavior is linear for velocities up to 1.7 pixels/frame for the 240x240 case and up to 2.5 pixels/frame (measured in the original image size) for the averaged image. The average relative error is displayed in figure(3,(b)). As expected the error is minimum at approximately 1 pixel/frame for the 240x240 case and at 2 pixels/frame for the 120x120 case. figure(4) shows the

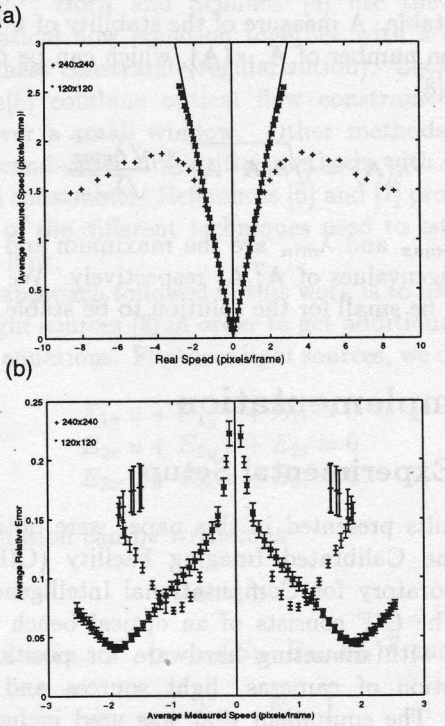


Figure 3: (a) Average measured speed vs real speed in pixels/frame for a 240x240 image and its average 120x120, the derivatives were estimated using central differences. (b) Relative error vs average measured speed

behavior of the system for the three different differentiation schemes described in subsection 3.2. The range of linear response is very similar in all cases. The relative error term is lower on average in the four point (time) differences, however, a difference in processing time of 200ms per frame suggests the use of central differences. All the experiments described in the following sections were done using central differences as the differentiation scheme.

4.2 Expanding and contracting balloon

Just as a test of the algorithm for a deforming surface, a balloon was inflated and deflated in front of the camera. The image size is 240x240 pixels and the processing rate was 4 frames/second. The velocity estimates are calculated on the well-shaded area between the edges and the center. The center is left mostly black because the intensity variation

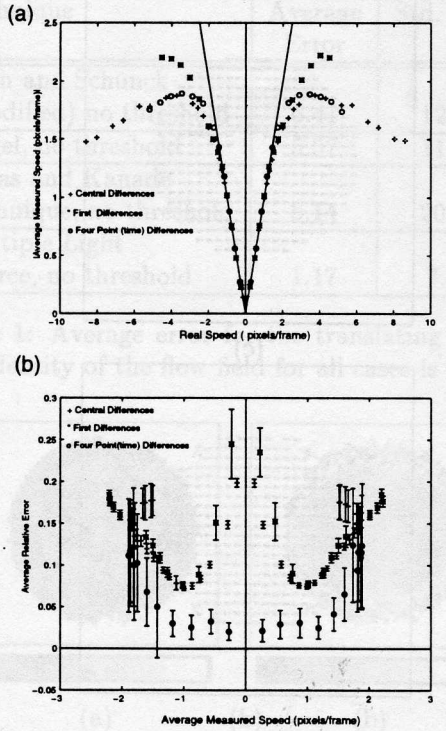


Figure 4: (a) Average measured speed vs real speed in pixels/frame for different derivative estimation schemes. (b) Relative error vs average measured speed.

is very small and these points are discarded in the calculation.

A short sequence of the expanding balloon was grabbed and the images were processed with a version of multiple light source optical flow running in Sun workstation. Figure(5) shows the optical flow, the flow field was subsampled every two pixels and it was scaled by two. The threshold for the magnitude of the brightness gradient was set to 2.5.

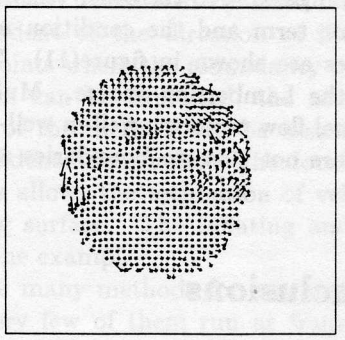


Figure 5: Multiple light source optical flow for a deforming balloon.

4.3 Comparison with other methods

To be able to make quantitative and qualitative comparisons of multiple light source optical flow with other methods a synthetic and a real image sequence were studied using some the implementations described in [6] and multiple light source optical flow running in a Sun workstation. In this implementation the images are spatially presmoothed with a Gaussian filter with standard deviation $\sigma = 1.5$.

The methods selected for comparison are three differential techniques, namely, Horn and Schunck[4], Lucas and Kanade[5], and Nagel[9]. The differential techniques include two first-order techniques and one second-order one, using global and local constraints to determine the full velocity vector. For details about these implementations see reference[6]. The error measure used to compare the results is the angular measure defined in references [6] and [10]. Let the estimated velocity $\mathbf{v_e} = [u, v]^T$ be written as $\mathbf{v_e} = (u^2 + v^2 + 1)^{-1/2}[u, v, 1]^T$. The error between the correct velocity $\mathbf{v}_c$ and the estimated velocity is defined as

$$error = arccos((\mathbf{v}_c)^T \cdot \mathbf{v}_e) \qquad (9)$$

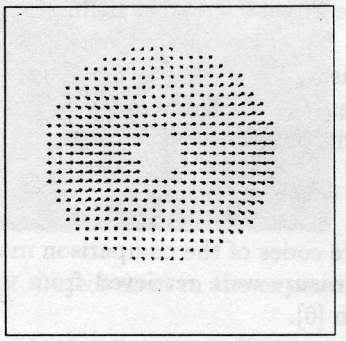
The source codes of the comparison methods and the error measure were retrieved from the ftp site referenced in [6].

4.3.1 Lambertian sphere

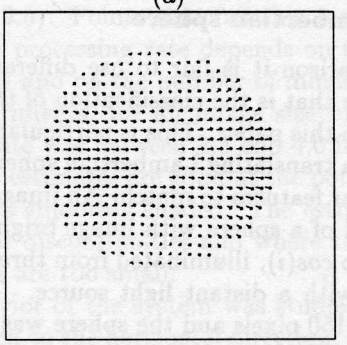
As a comparison it is fair to use differential techniques since that is the classification of the method described in this paper. This is particularly true for the case of a translating Lambertian sphere that has no particular features to match. An image sequence was created of a sphere with image brightness proportional to $cos(i)$, illuminated from three different directions with a distant light source. The image size is 150x150 pixels and the sphere was moved by 1.3 pixels/frame along the horizontal axis. Figures (6)-(7) show the optical flow for three differential techniques, Horn and Schunck, Lucas and Kanade, Nagel, and for multiple light source optical flow. The thresholds are indicated in the captions of the figure. The optical flow field is shown for every second point. Table(1) displays the average angular error for the unthresholded cases.

All the methods perform well on average. Horn and Schunck, and Nagel methods give pleasing, smooth fields but they give normal velocities at the edges. The local method of Lucas and Kanade gives good results for the body of the sphere. As expected, multiple light source optical flow performs

very well. This test case is exactly the type where this technique is known to give good results, i.e., smooth, well-shaded surfaces with no texture. Most of the errors are in the edges of the sphere, where the condition number of the system is larger. Areas of small variation of image brightness, like the center of the sphere, give larger relative errors. Figure(8) shows the relative error and the condition number for the translating sphere. Large relative errors and large condition numbers are encoded as white and low values as black, the relative error is shown in a scale from 0 to 1 and the condition number is shown in a scale from 1 to 25. Image points that are reached by two light sources have a relative error of zero, the two black areas on the edges of figure(8,(a)) correspond to this case. Points in the background are assigned white.



(a)

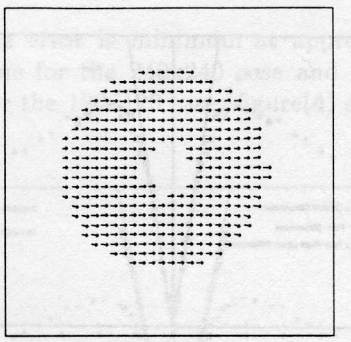


(b)

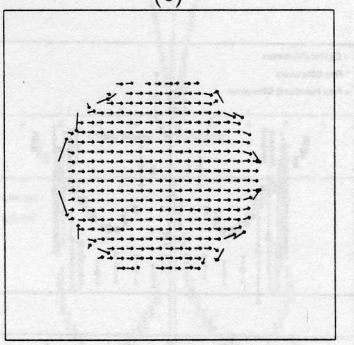
Figure 6: Translating Lambertian sphere. (a) Horn's method, $\| \nabla E \| \geq 2.5$, $\lambda = 0.5$, $\sigma = 1.5$, 100 iterations. (b) Nagel's method, $\| \nabla E \| \geq 2.5$, $\alpha = 0.5$, $\delta = 1.0$, $\sigma = 1.5$, 100 iterations.

4.3.2 Stay-puft Marshmallow Man

A real sequence of images was created by moving the object in the horizontal axis. Figures(9)-(10) show the results obtained with three differential



(c)



(d)

Figure 7: Translating Lambertian sphere. (c) Lucas' method, $\lambda_2 \geq 0.2$, $\sigma = 1.5$. (d) Multiple light source optical flow, $\| \nabla E \| \geq 2.5$

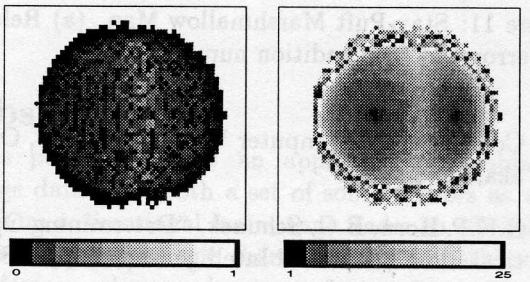
techniques, Horn and Schunck, Lucas and Kanade, Nagel, and multiple light source optical flow. The thresholds that were used are indicated in the captions of the figures. The choice of these values was such that a similar density was obtained for the flow field. The optical flow fields were subsampled to prevent cluttering and were scaled by two. In this case there is no correct optical flow to make quantitative comparisons but the results obtained with multiple light source optical flow are qualitatively good and comparable to the other techniques. The relative error term and the condition number for these images are shown in figure(11). The scales are as for the Lambertian sphere. Multiple light source optical flow performs well in well-shaded regions that are not close to boundaries and surface markings.

5 Conclusions

The measurement of optical flow is the first step in the analysis of sequences of time varying images. The second step is to use the optical flow field to recover the 3-D structure and the relative motion

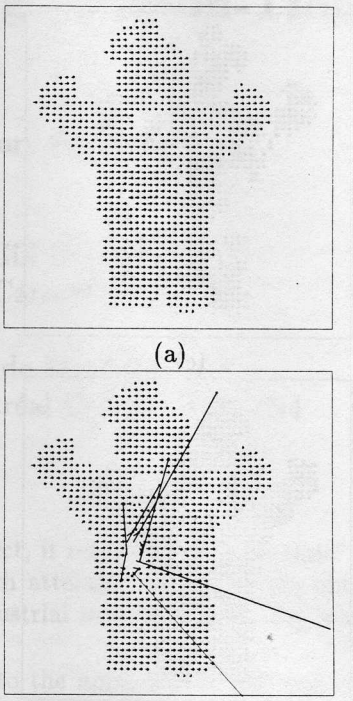
Technique	Average Error	Std. Dev.
Horn and Schunck (modified) no threshold	6.41	12.60
Nagel, no threshold	5.37	11.55
Lucas and Kanade Technique, no threshold	8.14	20.44
Multiple Light Source, no threshold	1.17	7.49

Table 1: Average error for the translating sphere. The density of the flow field for all cases is 100%.



(a) (b)

Figure 8: Translating Lambertian sphere, (a) relative error and (b) condition number.



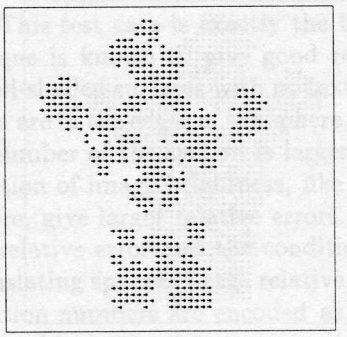
(a) (b)

Figure 9: Translating Stay-Puft Marshmallow Man, (a) Horn's technique, $\| \nabla E \| \geq 2.5$, $\lambda = 0.5$, $\sigma = 1.5$, 100 iterations. (b) Nagel's technique,$\| \nabla E \| \geq 2.5$, $\alpha = 0.5$, $\delta = 1.0$, $\sigma = 1.5$, 100 iterations.

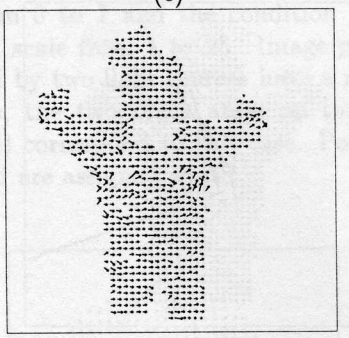
of the objects in the scene. Multiple light source optical flow is a method that allows the computation of a dense local representation of the flow field. A comparison with other methods shows that multiple light source optical flow performs very well when the surfaces are well-shaded, away from boundaries and surface markings. In that sense, multiple light source optical flow is complementary to matching techniques, it works well when matching methods fail and vice-versa. The method fails at surface markings because brightness changes at these points are dominated by local features and are independent of the direction of illumination. At image points where the redundancy of the three light sources can be exploited, the accuracy and robustness of the method can be assessed by the use of confidence measures. The locality of the calculations allows the estimation of velocities for a deforming surface. The inflating and deflating balloon is one example.

There are many methods for computing optical flow but very few of them run at frame rates on conventional hardware. This paper describes an implementation of multiple light source optical flow that allows the processing of 3.5 to 4.0 frames per second and gives a dense optical flow field. The method is particularly suited for real time processing for it does not involve any iterative technique, it is easy to parallelize and performs a uniform, local, point-by-point calculation. Even though not implemented in the real-time system, spatial smoothing poses no special problems. Data storage capability is the main limitation for performing temporal smoothing. A study of the response of the system for different speeds shows that the behaviour is linear for velocities up to 1.5 pixels/frame. This implementation has used three spectral channels in the visible spectrum. The use of other wavelengths is also possible. However, this would imply the use of a different camera.

A method related to multiple light source optical flow is photometric stereo[11]. Multiple images taken under different conditions of illumination are used in this case to solve the shape-from-shading problem. Multiple light source optical flow makes no assumptions about the reflectance properties of the surface undergoing motion. However, if the reflectance properties are known multiple light source can be applied simultaneously with photometric stereo to recover the 3-D structure



(c)



(d)

Figure 10: Translating Stay-Puft Marshmallow Man, (c) Lucas' technique, $\lambda_2 \geq 2.5$. (d) Multiple light source optical flow, $\| \nabla E \| \geq 2.5$, $\sigma = 1.5$.

and 3-D motion of the surface. Integrating these methods would be an important contribution to vision systems.

Acknowledgement

Stay-Puft Marshmallow Man is a registered trademark of Kenner Parker Toys International Inc. Support for this project was provided by the Institute of Robotics and Intelligent Systems (IRIS), a Canadian network of centres of excellence.

References

[1] E.C. Hildreth, "The Measurement of Visual Motion", MIT Press, Cambridge, Mass., 1984.

[2] A. Verri, T. Poggio, "Motion field and optical flow: qualitative properties", IEEE Trans. Pattern Analysis and Machine Intelligence (11), pp. 490-498, 1989.

[3] R.J. Woodham, "Multiple Light Optical Flow", Proceedings of the Third International Conference in Computer Vision, pp 42-45, Osaka, Japan,1990

[4] B.K.P. Horn, B.G. Schunck, "Determining Optical Flow", Artificial Intelligence(17), pp. 185-203, 1981.

[5] B.Lucas, T.Kanade, "An Iterative Image Registration Technique with Application to Stereo Vision", Proc. DARPA Image Understanding Workshop, pp. 121-130, 1981.

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

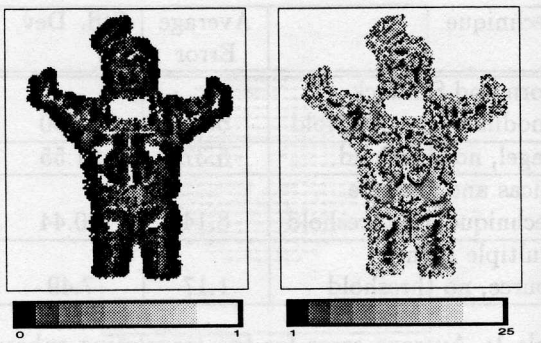
[7] S.S.Beauchemin andJ.L.Barron "The Computation of Optical Flow Techniques", ACM Computing Surveys, 27(3), pp. 433-467, September 1995.

[8] David S. Walkins, "Fundamentals of Matrix Computations", John Wiley & Sons, 1991.

[9] H.-H. Nagel, "On the Estimation of Optical Flow: Relations between Different Approaches and Some New Results", Artificial Intelligence", 33, pp. 299-324, 1987.

[10] D.J. Fleet, A.D. Jepson, "Computation of component image velocity from local phase information", IJCV, 5, pp. 77-104, 1990.

[11] R.J. Woodham, "Gradient and Curvature from Photometric Stereo Including Local Confidence Estimation", J.Opt.Soc.Am. A, vol 11, pp. 3050-3068, 1994.



(a) (b)

Figure 11: Stay-Puft Marshmallow Man. (a) Relative error and (b) condition number.