TR-160, ICS-FORTH, Feb. 1996.

Real-time Detection of Maneuvering Objects by a Monocular Observer

Antonis A. Argyros, Manolis I.A. Lourakis,
Panos E. Trahanias and Stelios C. Orphanoudakis

Institute of Computer Science, FORTH
PO Box 1385, Heraklion, Crete 711-10
Greece

and

Computer Science Department, University of Crete
PO Box 1470, Heraklion, Crete 714-09
Greece

e-mail: {argyros, lourakis, trahania, orphanou}@ics.forth.gr

Abstract

A method is proposed for real-time detection of objects that maneuver in the visual field of a monocular observer. Such cases are common in natural environments where the 3D motion parameters of certain objects (e.g. animals) change considerably over time. The approach taken conforms with the theory of *purposive vision*, according to which vision algorithms should solve many, specific problems under loose assumptions. The method can effectively answer two important questions: (a) whether the observer has changed his 3D motion parameters, and (b) in case that the observer has constant 3D motion, whether there are any maneuvering objects (objects with non-constant 3D motion parameters) in his visual field. The approach is direct in the sense that the structure from motion problem - which can only be solved under restrictive assumptions - is avoided. Essentially, the method relies on a pointwise comparison of two normal flow fields which can be robustly computed from three successive frames. Thus, it by-passes the ill-posed problem of optical flow computation. Experimental results demonstrate the effectiveness and robustness of the proposed scheme. Moreover, the computational requirements of the method are extremely low, making it a likely candidate for real-time implementation.

1 Introduction

Most of the research efforts to date in computational vision are influenced by the so called *reconstructionist approach*. Their basic assumption is that the general goal of computer vision is to produce an accurate, quantitative 3D representation of a scene. During the last decade, a new theory of vision has emerged, that of *active* and *purposive vision* [1]. According to the purposive theory, a vision system should be implemented by a set of processes which cooperate for achieving specific goals. Each process is dedicated to understanding certain aspects of the environment, that are immediately related to the goal to be achieved and, therefore, uses a partial representation of the world. The purposiveness of visual processes enables the statement and the solution of simpler problems. Such problems have a relative small number of solutions and can be treated in a qualitative manner.

A successful vision system should support two general goals: *navigation* and *recognition* in complex, dynamic environments. In both cases, the notion of *attention* is of central importance. Attention can be understood as the selective sensing in space, time and resolution. The role of attention is crucial for a vision system, since it drastically reduces the computational effort that should be spent in order to accomplish its tasks.

A lot of research effort has been devoted in determining features that can drive attention. These include static features such as color, texture and depth or dynamic features such as motion, illumination change and generally every kind of change in the field of view [2]. One of the visual cues that play an extremely important role in driving attention is motion. Most of the primitive, survival tasks of biological organisms are based on the perception of motion. Thus, the latter is crucial for many behaviors that an autonomous biological or man-made system should exhibit in real world environments.

Recognizing the importance of the visual perception of motion, this paper studies one of its aspects, namely the detection of maneuvering objects. This is related to the problem of independent motion detection by a moving observer. Due to the *egomotion* of the observer in the 3D space, the whole visual field appears to be moving in a specific manner, which depends on the observer's 3D motion parameters and the structure of the scene in view. The problem of independent 3D motion detection can be defined as the problem of locating objects that move independently from the observer in his visual field. Recently, this has been approached mainly by assuming some knowledge about the observer's motion. Thomson uses knowledge of certain aspects of egomotion and scene structure [3]; Sharma and Aloimonos [4] assume known translational egomotion; Nelson [5] requires *a priori* knowledge of egomotion parameters and assumes upper bounds on the depth of the scene.

The problem of independent motion detection takes a special form in the case where the observer does not move relative to the static 3D environment. In this case, the problem of detecting moving objects can be treated as a problem of *change detection* [2]. The situation is much more complicated when the observer is moving relative to the environment. This case is also the most interesting because both biological and man made seeing systems move. Even if the body of an observer is still, the eyes are continuously moving. In such a case, the points of both the environment and the moving objects project in different 2D locations on the image plane and simple detection of intensity changes cannot anymore handle the problem of detecting moving objects.

Although the general problem of independent 3D motion detection is difficult, we argue that important aspects of it, such as the detection of maneuvering objects, can be solved robustly by simple algorithmic techniques. Based on the principles of purposive vision, this is approached in this paper by employing only an adequate representation of visual motion, rather than trying to fully recover motion information. More precisely, we extract from image sequences the minimum information needed for the detection of maneuvering objects. Towards this goal, the developed scheme does not rely on the computation of *optical flow*, rather on the spatiotemporal derivatives of the image intensity function, known as *normal flow*. Normal flow, although less informative compared to optical flow, can be robustly computed from a sequence of images. Based on the choice of normal flow to represent visual motion information, a method is proposed for the detection of maneuvering objects. The method performs a pointwise comparison of the two normal flow fields that result from three successive image frames; this comparison signals changes in the 3D motion parameters of the observer or of the objects in the field of view. The method is in fact a *data driven*, *feature based* attention mechanism [6], which can be exploited by a monocular observer pursuing unrestricted 3D motion. The problem of detection of maneuvering objects has also been approached in [5] by assuming smooth observer motion and inexact knowledge of the motion field. In our approach the observer's motion is not restricted and, moreover, changes in his 3D motion parameters are signaled. In addition, no assumptions about the objects' motion are imposed, making it useful in practical applications where the detection of motion changes is desired.

The rest of the paper is organized as follows. Section 2 presents the imaging geometry and the motion representation employed in this work. In section 3, the method for the detection of maneuvering objects is presented. The method relies on a direct, pointwise comparison of normal flow values and is capable of answering whether there are changes in the 3D motion parameters either of the observer or of independent objects (maneuvering objects). In section 4, experimental results from the application of the maneuvering objects detection algorithm to image sequences are presented. Finally, section 5 presents concluding remarks that summarize the results of this work.

2 Imaging system and motion representation

Let a coordinate system $OXYZ$ adjusted to the optical center of a camera, such that the OZ axis coincides with the optical axis, as shown in Fig. 1. Let the camera focal length be f, i.e. the image plane is at distance f from O. Under perspective projection, a point $P(X, Y, Z)$ in 3D space projects on the image plane at point $p(x, y)$. If P is moving relative to $OXYZ$ with translational motion $\vec{t} = (U, V, W)$ and rotational motion $\vec{\omega} = (\alpha, \beta, \gamma)$[1], the equations describing the 2D velocity (u, v) of the image point $p(x, y)$ are written as [7]:

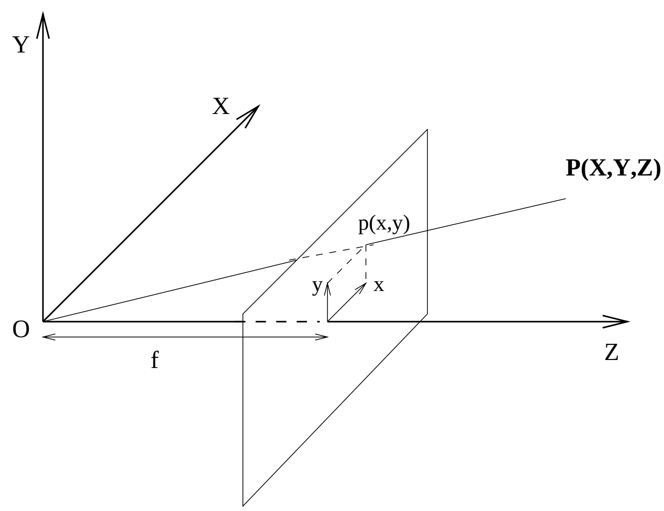


Figure 1: The camera coordinate system.

[1]This 3D motion may be due to a motion of the coordinate system (egomotion) and/or independent motion of point P.

$$u = \frac{(-Uf + xW)}{Z} + \alpha\frac{xy}{f} - \beta\left(\frac{x^2}{f} + f\right) + \gamma y \tag{1a}$$

$$v = \frac{(-Vf + yW)}{Z} + \alpha\left(\frac{y^2}{f} + f\right) - \beta\frac{xy}{f} - \gamma x \tag{1b}$$

2.1 Motion field - optical flow field

Equations (1) describe the 2D *motion field*, which relates the 3D motion of a point with its projected 2D motion on the image plane. The motion field is a purely geometrical concept and is not necessarily identical to the *optical flow* field [8], which describes the apparent motion of brightness patterns observed because of the relative motion between an imaging system and its environment. Verri and Poggio [9] have shown that the motion and optical flow fields are identical in specific cases only. Even in the cases that these two fields are identical, the problem of optical flow estimation is ill-posed [10]. This is often approached using regularization methods, which impose constraints on the solution. Such constraints are related to certain assumptions about the structure of the viewed scene. In practice - especially in the case of independent motion where motion discontinuities exist by definition - these assumptions are quite often violated, resulting in errors in optical flow estimation.

For the above reasons, the proposed scheme for the detection of maneuvering objects does not rely on the computation of optical flow, rather on normal flow, i.e. the projection of optical flow along the direction of intensity gradient. The normal flow field has been used in the past for both egomotion estimation [11, 12, 13] and independent motion detection [4, 5].

2.2 Normal flow field - normal motion field

Let the image sequence be modeled as a continuous function $I(x, y, t)$ of two spatial (x, y) and one temporal (t) variables. Assuming that irradiance is conserved between two consecutive frames, we get the well known *optical flow constraint equation*, originally developed by Horn and Schunk [14], in the form of a dot product:

$$(I_x, I_y) \cdot (u, v) = -I_t \tag{2}$$

where, I_x, I_y and I_t are the spatial and temporal partial derivatives of the image intensity function, respectively. Equation (2) gives a constraint for the components u and v of optical flow and enables the computation of the projection of the optical flow along the intensity gradient direction, namely the normal flow.

The normal flow field is not necessarily identical to the normal motion field (the projection of the motion field along the gradient), in the same way that the optical flow is not necessarily identical to the motion field [9]. However, normal flow is a good approximation to normal motion at points where the image gradient magnitude $||\nabla I||$ is large. Such points provide reliable information for motion perception.

Let (n_x, n_y) be the unit vector in the gradient direction. The magnitude u_n of the normal flow vector is given by $u_n = un_x + vn_y$ which, by substitution from eqs. (1), yields:

$$
\begin{aligned}
u_n &= (-n_x f)\frac{U}{Z} + (-n_y f)\frac{V}{Z} + (xn_x + yn_y)\frac{W}{Z} \\
&+ \left\{\frac{xy}{f}n_x + \left(\frac{y^2}{f} + f\right)n_y\right\}\alpha + \left\{-\left(\frac{x^2}{f} + f\right)n_x - \frac{xy}{f}n_y\right\}\beta + (yn_x - xn_y)\gamma
\end{aligned}
\tag{3}
$$

3 Detection of Maneuvering Objects

The method described in this section relies on motion information, acquired by a moving monocular observer and attempts to detect changes in the 3D motion parameters. To this end, two normal flow fields are computed from three successive image frames in time. The method decides whether the 3D motion of a certain point in the first pair of images (images acquired at instances $t-2$ and $t-1$), remains the same in the second pair of images (images acquired at instances $t-1$ and t).

3.1 Method description

Suppose that the 3D motion parameters of $P(X, Y, Z)$ remain constant over three frames that are acquired at time instances $t-2$, $t-1$ and t. Suppose also that we compute a normal flow field, from frame $t-1$ to frame t. According to eq. (3), the normal flow computed at point (x, y) with gradient direction (n_x, n_y), is equal to

$$
\begin{aligned}
{u_n}^{(t-1)\to t} &= (-n_x f)\frac{U}{Z} + (-n_y f)\frac{V}{Z} + (xn_x + yn_y)\frac{W}{Z} \\
&+ \left\{\frac{xy}{f}n_x + \left(\frac{y^2}{f} + f\right)n_y\right\}\alpha + \left\{-\left(\frac{x^2}{f} + f\right)n_x - \frac{xy}{f}n_y\right\}\beta + (yn_x - xn_y)\gamma
\end{aligned}
$$

where (U, V, W) are the translational motion parameters, (α, β, γ) are the rotational motion parameters and f is the focal length of the imaging system. Let us also compute the normal flow from frame $t-1$ to frame $t-2$. Because of the hypothesis of constant 3D motion, point P will again move from time $t-2$ to time $t-1$ with motion parameters (U, V, W) and (α, β, γ) or, equivalently, with parameters $(-U, -V, -W)$ and $(-\alpha, -\beta, -\gamma)$ from time $t-1$ to $t-2$. Therefore,

$$u_n{}^{(t-1)\to(t-2)} = -u_n{}^{(t-1)\to t} \tag{4}$$

Note that since both normal flow fields appearing in eq. (4) are computed with respect to time instant $t-1$, for a given point (x, y), the gradient direction (n_x, n_y) and the depth Z are the same for $u_n{}^{(t-1)\to(t-2)}$ and $u_n{}^{(t-1)\to t}$. Equation (4) provides a simple, yet effective criterion to check whether the 3D motion parameters of a point remain the same over three frames in time. Once the two normal flow fields are computed, then for each point the sum of the normal flow values should be equal to zero. A non-zero value signals a change in the 3D motion parameters of the corresponding point. In practical situations, the sum of normal flow values will not be zero due to errors in the computation of the time derivative. We may, however, require the absolute value of the sum to be small with respect to the sum of the absolute normal flow values, deriving the following criterion:

$$\frac{|u_n{}^{(t-1)\to(t-2)} + u_n{}^{(t-1)\to t}|}{|u_n{}^{(t-1)\to(t-2)}| + |u_n{}^{(t-1)\to t}|} < \delta_{un} \tag{5}$$

where δ_{un} is a threshold controlling the sensitivity to changes in motion in the three frames.

The satisfaction of criterion (5) over subsets of scene points, leads to four interesting cases; they are summarized below, where it is assumed that the majority of the scene points correspond to the static world. Let I_P be the set of image points for which reliable normal flow values have been computed; then:

1. **The criterion holds for all image points in I_P.** This is the case where neither the observer, nor any object(s) changed their motion parameters. Note that the change of motion parameters includes the case of previously static objects that have now started moving.

2. **The criterion holds for the majority of image points in I_P.** This is the case where the motion of the observer remained constant. Points where the criterion does not hold, are points of objects that changed their motion.

3. **The criterion holds for the minority of image points in I_P.** This is a special case where both the observer and the independently moving object(s) changed their motion in exactly the same way, so that no relative change can be detected.

4. **The criterion does not hold for any point in I_P.** The motion of the observer has been changed. It cannot be decided, however, whether some objects have also changed their motion.

Based on whether criterion (5) is satisfied or not at a certain point, a label may be assigned to that point, which describes whether its 3D motion parameters have been changed or not.

It is noted that by employing normal flows, only incomplete information about motion is used. A normal flow value is the projection of an optical flow vector at a certain direction. Infinite many other optical flow vectors may have the same projection in the same direction. Consequently, there are certain changes in the 3D motion parameters of a point that cannot be recovered through summations of normal flow values. However, in a region where 3D motion changed, it is expected that many different gradient directions exist and, therefore, the concentration of points that do not satisfy criterion (5) will be high. This observation, leads to the conclusion that some type of post processing is needed. Such a postprocessing is achieved through a simple majority voting scheme. The label of a point is changed to the label of the majority of the points in a small neighborhood. This allows isolated points to be removed.

There is a number of interesting analogies that can be drawn between change detection methods that are used to detect moving objects in the field of view of a static observer and the proposed method for the detection of changes in 3D motion. Table 1 summarizes these analogies.

3.2 Labeling of points through robust regression

Criterion (5) involves a threshold that controls the labeling of points as ones with constant motion parameters or not. Determining a threshold for such a labeling may result in inaccurate results in some cases, since the left hand side of criterion (5) may vary considerably. In our case, robust regression is a powerful alternative to thresholding, for deciding whether an image point belongs to a maneuvering object. Regression analysis (fitting a model to noisy data) is a

Table 1: Change detection in image intensities vs. change detection in normal flow fields

	Change detection in intensities	Change detection in normal flow
Goal:	Detection of position changes	Detection of changes in objects' 3D motion parameters
Assumption:	Constant observer position	Constant observer motion
Input:	2 image frames	2 normal flow fields (from 3 frames)
Approach:	Difference of image intensities	Difference of normal flows

very prominent statistical tool. In the general case of a linear model, given by the relation

$$y_i = x_{i1}\theta_1 + \ldots + x_{ip}\theta_p + e_i, \tag{6}$$

the problem is to estimate the parameters θ_k, $k = 1, \ldots, p$, from the observations y_i, $i = 1, \ldots, n$, and the explanatory variables x_{ik}. The term e_i represents the error present in each of the observations. Let $\hat{\theta}$ be the vector of estimated parameters $\hat{\theta}_1, \ldots, \hat{\theta}_p$. Given these estimations, predictions can be made for the observations:

$$\hat{y}_i = x_{i1}\hat{\theta}_1 + \ldots + x_{ip}\hat{\theta}_p \tag{7}$$

Thus, a residual between the observation and the value predicted by the model may be defined as:

$$r_i = y_i - \hat{y}_i \tag{8}$$

Traditionally, $\hat{\theta}$ is estimated by the popular least squares (LS) method. However, the LS estimator becomes highly unreliable in the presence of outliers, that is observations that deviate considerably from the model describing the rest of the observations. Robust regression methods [15] have been proposed in order to cope with such cases. The main characteristic of robust estimators is their high breakdown point, which may be defined as the smallest amount of outlier contamination that may force the value of the estimate outside an arbitrary range. A variety of robust estimation techniques have been used in computer vision. Some of them have been developed within the vision field (eg. [16, 17]); others have been borrowed from statistics (eg. [18, 19]).

The LMedS method (Least Median of Squares), which was proposed by Rousseeuw [20], involves the solution of a non-linear minimization problem, namely:

$$Minimize\{median_{i=1,\ldots,n} {r_i}^2\} \tag{9}$$

Intuitively, LMedS tries to find a set of model parameters such that the model best fits the *majority* of the observations. Once LMedS has been applied to a set of observations, a standard deviation estimate may be derived. Rousseeuw and Leroy [15] suggest a value of

$$\hat{\sigma} = 1.4826 \left(1 + \frac{5}{n-p}\right) \sqrt{med {r_i}^2} \tag{10}$$

Based on the standard deviation estimate, a weight may be assigned to each observation

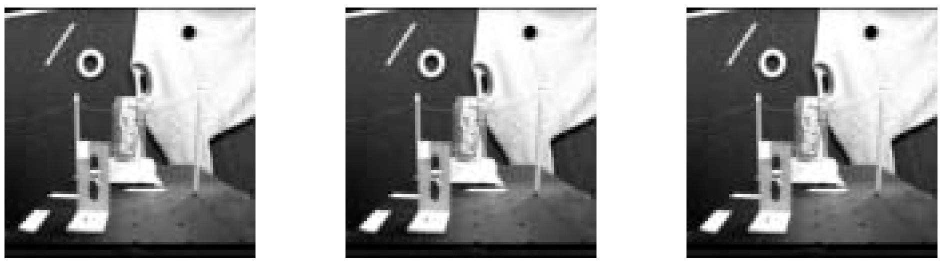
$$w_i = \begin{cases} 1, & \text{if } \frac{|r_i|}{\hat{\sigma}} \leq 2.5 \\ 0, & \text{if } \frac{|r_i|}{\hat{\sigma}} > 2.5 \end{cases} \tag{11}$$

All points with weight equal to 1 correspond to model inliers, while points with weight 0 correspond to outliers. The threshold 2.5 reflects the fact that in the case of a Gaussian distribution, very few residuals should be larger than $2.5\hat{\sigma}$.

LMedS in its simplest form (number of parameters $p = 1$) can be used to distinguish the populations of points corresponding to maneuvering objects, instead of applying thresholding to criterion (5). In this case, inliers will have very similar values and will correspond to the majority of points where reliable normal flow values were computed. These points provide information on the observer's motion. If the values are close to zero, the observer is moving with unchanged 3D motion parameters; in the opposite case, a change in his motion parameters has occurred. The outliers of the model will be points where normal flow changed considerably, with respect to the majority of points. These points signal the presence of an object (or objects) whose motion parameters changed in a different way from the observer's parameters. By employing the LMedS estimation technique, we are not interested in the value of the estimated (threshold) parameter, rather in the separation of the inliers from the outliers of the model. LMedS enables the automatic adaptation to the noise levels of the scene and, therefore, robust separation of the maneuvering objects. Note also that due to the fact that $p = 1$, the reported high computational complexity of LMedS is reduced to the complexity of finding the median within a set of numbers. This operation can be performed in linear time, without requiring sorting.

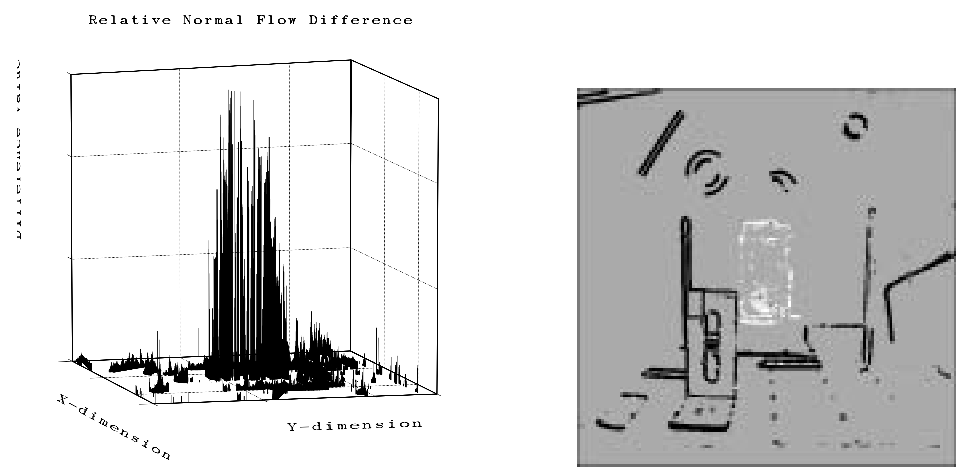
4 Experimental results

A set of experiments has been conducted in order to test the performance of the described method. Representative results from these experiments are given in this section. In the first experiment, the ‘‘coca-cola’’ image sequence was employed and modified with the addition of rotational motion. Fig. 2 shows three frames of this sequence, at time instances $t-2$, $t-1$ and t. The camera moves with translational motion approaching the scene. Since there is



(time t-2) (time t-1) (time t)

Figure 2: Three successive frames of the ‘‘coca-cola’’ sequence.

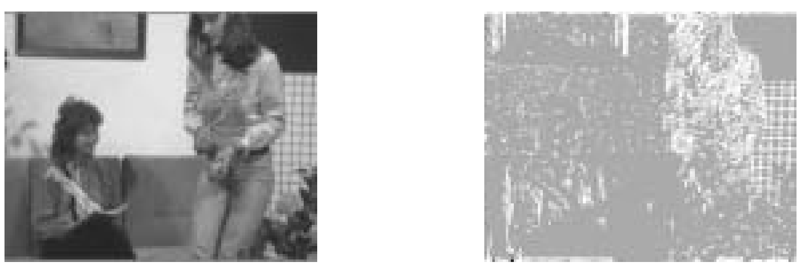


(a) (b)

Figure 3: (a) 3D plot of the image points with respect to criterion (5), and (b) characterization of points with respect to the constancy of their 3D motion parameters (see text for explanation).

no independent motion in the scene in view, a rotational motion has been synthetically added in the area of the coca-cola can. More specifically, in the third frame (frame at time t), the coca-cola can has been moved relative to the second frame by adding (synthetically) rotational motion (the observer's egomotion was left unchanged). After smoothing the images, the two normal flow fields were obtained and criterion (5) was computed for all image points with a reliable normal flow value. Figure 3(a) shows a three dimensional plot of the values of criterion (5). x and y dimensions of the plot correspond to the x and y dimensions of the image while the third dimension corresponds to the values of criterion (5). It is evident that in the points of the coca-cola can where a motion change occurs, criterion (5) gives distinguishably different values than in the other points of the image which move due to the constant egomotion. Figure 3(b) shows the final labeling of the pixels. White pixels correspond to points where the 3D motion has been changed, black pixels correspond to points which kept the same 3D motion parameters and gray pixels correspond to points where no reliable normal flow vectors could be computed.

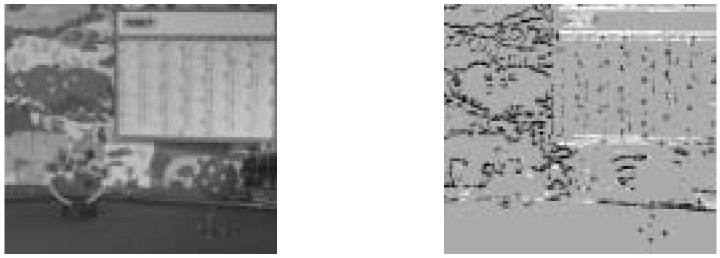
In another experiment, the ‘‘interview’’ sequence has been employed. In order to simulate a change in the motion parameters of the observer, four consecutive frames were selected and the third one was dropped. Thus, the frames used correspond to time instances $t-3$, $t-2$ and t. The omission of a frame is equivalent to a change in the observer's 3D motion parameters, since in the original sequence his motion is a constant one. Fig. 4(a) shows the frame at time instance $t-2$ (the middle of the three frames used) and the results after the application of criterion are illustrated in Fig. 4(b). Assuming the same coloring scheme as in Fig. 3(b), all points where normal flow has been reliably computed appear in white, signaling the change of the 3D motion parameters of the observer.



(a) (b)

Figure 4: (a) A frame of the ‘‘interview’’ sequence (b) characterization of points with respect to the constancy of their 3D motion parameters (see text for explanation).

A final experiment is presented regarding the ‘‘calendar’’ sequence. In this sequence, the calendar appearing on the top-right of the images is moving upwards and, subsequently, its motion is modified and oriented down and to the right with respect to the image frame. All other objects are moving with constant motion parameters. Fig. 5(a) shows the middle of the three frames used from this sequence and the results regarding motion changes are presented in Fig. 5(b). As can be verified, the points of the calendar that contribute to the normal flow field have been successfully detected as points where 3D motion parameters change.



(a) (b)

Figure 5: (a) A frame of the ‘‘calendar” sequence (b) characterization of points with respect to the constancy of their 3D motion parameters (see text for explanation).

The above experiments were carried out using the same algorithmic parameters (thresholds for normal flow rejection, outlier characterization threshold), which shows the robustness and wide applicability of the proposed method.

5 Summary

In this paper, a method for the visual detection of changes in the 3D motion parameters of objects has been described. The method is capable of answering two specific questions regarding a monocular observer and the scene being observed: (a) whether the observer moves with constant 3D motion parameters, and (b) whether some objects are maneuvering within the observer's visual field. Despite the high complexity of the general independent 3D motion detection problem, it has been shown that these two specific questions may be robustly answered by using a simple computational scheme. Interesting analogies have been drawn between classical change detection algorithms that operate on the image intensity function and the proposed method that operates on the normal flow field. In fact, the proposed method can be characterized as a *motion change detection* method. The method avoids the solution of the structure from motion problem and relies on the comparison of two normal flow fields that are computed from three successive image frames. Relying on normal flow and on its time-reversed computation, enables the method to also avoid the solution of the correspondence problem. The computational requirements of the proposed method are extremely low, facilitating real time implementation. This is due to the fact that the only operations involved are computations of normal flow values and LMedS estimation of a very simple, one-parameter regression model.

The experimental results presented serve as an indication of the effectiveness of the method, which answers important questions by employing minimum assumptions about the external world and the observer. Therefore, it provides solutions to specific problems under loose assumptions rather than trying to solve general problems which can be done under restrictive assumptions. Current research is targeted towards integrating the proposed method with other robust visual capabilities, in order to provide synergistic solutions to more complex vision problems.

REFERENCES

[1] Aloimonos Y and Badyopadhyay A. Active Vision. In *IEEE 1st Int. Conf. on Computer Vision*, pages 35--54, Jun. 1987.

[2] K. Skifstad and R Jain. Illumination independent change detection for real world image sequences. *Computer Vision Graphics and Image Processing*, 46:387--399, 1989.

[3] Thompson WB and Pong TC. Detecting Moving Objects. *International Journal of Computer Vision*, 4:39--57, 1990.

[4] Sharma R. Robust Detection of Independent Motion: An Active and Purposive Solution. Technical report, Center for Automation Research, University of Maryland, CAR TR-534, College Park, MD, 1991.

[5] Nelson RC. Qualitative Detection of Motion by a Moving Observer. *International Journal of Computer Vision*, 7(1):33--46, 1991.

[6] Swain J and Stricker MA. Promising Directions in Active Vision. *International Journal of Computer Vision*, 11(2):109--126, 1993. Written by the attendees of the NSF Active Vision Works., Univ. of Chicago, Aug. 5-7.

[7] Longuet-Higgins HC and Prazdny K. The Interpretation of a Moving Retinal Image. In *Proceedings of the Royal Society*, pages 385--397. London B, 1980.

[8] B. K. P. Horn. *Robot Vision.* MIT Press, Cambridge, MA, 1986.

[9] Verri A and Poggio T. Motion Field and Optical Flow: Qualitative Properties. *IEEE Trans. on Pat. Anal. and Mach. Int.*, PAMI-11(5):490--498, May 1989.

[10] Aloimonos Y, Weiss I, and Bandopadhay A. Active Vision. *International Journal of Computer Vision*, 2:333--356, 1988.

[11] Fermuller C. *Basic Visual Capabilities.* PhD dissertation, Center for Automation Research, University of Maryland, 1993.

[12] Aloimonos Y and Duric Z. Estimating the Heading Direction using Normal Flow. *International Journal of Computer Vision*, 13(1):33--56, 1994.

[13] Sinclair D, Blake A, and Murray D. Robust Estimation of Egomotion from Normal Flow. *International Journal of Computer Vision*, 13(1):57--69, 1994.

[14] Horn BKP and Schunck B. Determining Optical Flow. *Artificial Intelligence*, 17:185--203, 1981.

[15] Rousseeuw PJ and Leroy AM. *Robust Regression and Outlier Detection.* John Wiley and Sons Inc., New York, 1987.

[16] Boyer KL, Mirza MJ, and Ganguly G. The Robust Sequential Estimator: A General Approach and its Application to Surface Organization in Range Data. *IEEE Trans. on Pat. Anal. and Mach. Int.*, PAMI-16:987--1001, 1994.

[17] CV Stewart. MINPRAN: A New Robust Estimator for Computer Vision. *IEEE Transactions on PAMI*, 17(10):925--938, 1995.

[18] Kumat R and Hanson AR. Robust Methods for Estimating Pose and a Sensitivity Analysis. *CVGIP:Image Understanding*, 1994.

[19] Meer P, Mintz A, and Rosenfeld A. Robust Regression Methods for Computer Vision: A Review. *International Journal of Computer Vision*, 6(1):59--70, 1991.

[20] Rousseeuw PJ. Least Median of Squares Regression. *Journal of Amer. Stat. Assoc.*, 79:871--880, 1984.