Translational Decomposition of Flow Fields*

Daryl T. Lawton and **Warren F. Gardner**
College of Computing
Georgia Institute of Technology
Atlanta, Georgia 30332-0280

Abstract

We introduce a low-level description of image motion called the **local translational decomposition (LTD)**. This description associates with image features or small image areas, a three-dimensional unit vector describing the direction of motion of the corresponding environmental feature or small surface area. The local translational decomposition is derived by applying a procedure for processing purely translational motion to small overlapping image areas. This intermediate representation of motion considerably simplifies the inference of motion parameters for ego-motion and can support qualitative inferences for non-rigid motions. We first show how to compute the LTD from optic flow fields and then show how the LTD is used to recover the parameters of rigid body motions. We present three cases for which the recovery of motion parameters is particularly robust: motion constrained to a determined plane (the normal to the plane is known); motion constrained to an undetermined plane (the normal to the plane is not known); arbitrary motion relative to locally planar surfaces. We then discuss techniques for computing the local translational decomposition directly from real image sequences without the initial extraction of optic flow and other areas for future work.

1 Introduction

In previous work [Lawton, 1982], we developed a technique to process relative translational motion of a sensor with respect to a stationary environment or independently translating objects. This and related algorithms [Burger and Bhanu, 1989; Jain, 1983] are based on the strong geometric constraints on image motion in the case of translation – radial motion of image features from a focus of expansion (or contraction) determined by the intersection of the axis of translation with an imaging surface [Gibson, 1950; Lee, 1980]. The technique presented in [Lawton, 1982] was based on optimizing a measure which described the quality of feature matches restricted to lie along the radial flow paths associated with a potential axis of translation. The optimization process involved searching over the surface of a unit sphere where each point corresponded directly to a possible direction of translation. The optimization combined the determination of the direction of translation and the corresponding image displacements into a single, mutually constraining computation. It was possible to determine the direction of translation to within a few degrees in small image areas using a few distinctive features.

In this paper we extend the translational processing algorithm to work with general rigid body and other cases of motion by applying the translational procedure to local portions of a flow field. This processing associates a direction of relative environmental motion with a local portion of a flow field and also an error measure reflecting the validity of the translational approximation. We call this description of image motion the **local translational decomposition (LTD)**. Computing the LTD begins by decomposing a flow field into small overlapping neighborhoods and then approximating the motion for each neighborhood as being produced by translational motion of the corresponding portion of the environment. This approximation associates a unit vector describing the direction of environmental motion with local portions of a flow field. Each unit vector has an associated fit-value reflecting the validity of the translational approximation.

The LTD is a low level representation of environmental motion which considerably simplifies the recovery of the sensor motion parameters. The local directions of motion and corresponding error measures are used as constraints to determine the actual parameters of motion and to recover the structure and layout of environmental surfaces. This is broken into four cases. For motion constrained to a plane of a known orientation (See Section 2.1), the local translational approximation is recovered directly from the intersection of flow vectors with the horizon line determined by the plane of motion. For motion constrained to a plane of unknown orientation (See Section 2.2), all of the computed LTD vectors must be perpendicular to the normal of the unknown plane. This constraint leads to a direct fitting procedure to recover

*This research is supported by the Advanced Research Projects Agency of the Department of Defense and is monitored by the U. S. Army Topographic Engineering Center under contract No. DACA76-92-C-0016

Figure 1: Camera coordinate system

the plane of motion. For motion relative to locally planar surfaces (see Section 2.3), the combination of local planarity and rigidity is used. For arbitrary motion, rigidity between environmental points is used to recover motion parameters from a small number of image locations (See Section 2 and Section 3.1).

The reminder of this section introduces the notation used throughout this paper. Section 2 describes how the local direction of translation is estimated from a flow field and cases of motion for which this is particularly robust. Section 3 describes how the parameters of relative sensor motion can be recovered from the estimated local directions of translation. Section 4 discusses computing the local translational decomposition directly from real image sequences without the initial extraction of optic flow and other areas for future work.

1.1 Notation

The coordinate system used in this paper is shown in Figure 1. The origin of this right-handed coordinate system lies at the focal point of the camera. The image plane is parallel to the xy-plane and is centered on the point $(0, 0, f)$, where f is the focal length of the camera. A three-dimensional environmental point will be referred to as $p_{i,j} = (x_{i,j}, y_{i,j}, z_{i,j})$. The corresponding image point is $\tilde{p}_{i,j} = (\tilde{x}_{i,j}, \tilde{y}_{i,j})$. The first subscript i is used to differentiate between points. The second subscript denotes the time interval. Thus, $p_{i,j}$ refers to the ith point at time j. A three-dimensional displacement which transforms $p_{i,j}$ into $p_{i,j+1}$ forms a vector. This vector will be referred to as $v_{i,j}$. The corresponding optic flow vector on the image plane is $\tilde{v}_{i,j}$. In Section 2 a method for estimating $v_{i,j}$ is presented. This estimated vector will be referred to as $\hat{v}_{i,j}$. If $\hat{v}_{i,j}$ is correct, it will be parallel to $v_{i,j}$, but its depth will be unknown. $\hat{v}_{i,j}$ can be positioned anywhere along the rays of projection which pass through $\tilde{p}_{i,j}$ and $\tilde{p}_{i,j+1}$. Unless specified otherwise, $\hat{v}_{i,j}$ will be positioned at the image plane.

The motion of the camera can be described by six parameters. Let $r = (r_x, r_y, r_z)$ denote the axis of rotation, and $t = (t_x, t_y, t_z)$ the direction of translation. We assume the axis of rotation passes through the origin of the camera coordinate system. The magnitude of r is equal to the angle of rotation, and t is a unit vector.

2 Estimating Local Translation

In this section we show how to determine an axis of translation consistent with a local portion of a computed flow field. In section 4 we briefly discuss how to compute this directly from textured images without the initial extraction of a flow field.

Figure 1 shows that the plane formed by a flow vector and the focal point of the camera must include the estimated local translation vector (we refer to this as the *flow-vector plane* for a given flow vector). In the case of purely translational motion, the estimated local translation vector will be the same for all flow vectors in the neighborhood. Therefore, the estimated local translation vector is the vector which is parallel to all of the flow vector planes in the neighborhood. This observation leads directly to a method of solving for the estimated local translation.

The plane formed by $\tilde{v}_{i,j}$ and the focal point of the camera must include $\hat{v}_{i,j}$. Let this plane be designated by its normal $n_{i,j}$.

$$n_{i,j} = \tilde{p}_{i,j} \times \tilde{p}_{i,j+1} \tag{1}$$

Since $n_{i,j}$ is perpendicular to $\hat{v}_{i,j}$

$$n_{i,j} \cdot \hat{v}_{i,j} = 0 \tag{2}$$

In the case of purely translational motion, the direction of $\hat{v}_{i,j}$ is constant for all i. Therefore, Equation 2 can be rewritten as

$$n_{i,j} \cdot \hat{v}_j = 0 \tag{3}$$

where $\hat{v}_j = \hat{v}_{i,j}$ for all i. This equation is linear with three unknowns, and can be solved using a least squares technique.

An error measure is used to evaluate the validity of the local translation approximation. The error measure we use is the average, taken over the local neighborhood, of the angle between each flow vector plane and the local translation. Using the normals $n_{i,j}$ from Equation 1, the

Figure 2: Local translation associated with a rotating line

error measure is defined as

$$\frac{1}{m}\sum_{i=1}^{m} |\sin^{-1}\left(\frac{n_{i,j}\cdot\hat{v}_j}{\|n_{i,j}\|\|\hat{v}_j\|}\right)| \qquad (4)$$

where m is the number of flow vectors in the local neighborhood. Alternatively (and with greater expense), this measure could be optimized directly by a search procedure to determine an axis of translation.

In general, $\hat{v}_{i,j}$ is not constant for all i. However, in local areas $\hat{v}_{i,j}$ is approximately constant. For example, in Figure 2, points which are nearby on a line segment are shown to have approximately the same local translations when the line is rotated about its midpoint. Points near the axis of rotation would not have a good translational approximation as would be reflected in the corresponding error measure. Note that if the motion is composed of both a rotation and translation, the approximation will also be effected by environmental points at different depths, especially at occlusion boundaries. Since the flow vectors in the area of an occlusion boundary will not consistently emanate from a focus of expansion, the error measure given in Equation 4 returns a high value in these areas. Using the error measure, the unreliable occlusion areas can be avoided when computing the parameters of motion. Figure 3 shows the flow field for a scene containing multiple depths and undergoing an arbitrary motion. The error function derived from this flow field is shown in Figure 4. The scene contains two planes which occlude a planar background as well as each other. The planes, as well as the background, are skewed with respect to the image plane (i.e. the planes are receding in depth). The locations of the occlusion boundaries are obvious from the figure.

The method of LTD estimation discussed above was tested on several synthetic optic flow fields like the one shown in Figure 5. This flow field is the result of a rotation of 5.73^o about the axis $(5, 4, 1)$, followed by a translation of $(100, 25, -75)$. All units are given in pixels. The field of view of the camera is 90^o in both the X and Y directions. The image is 63x63, and the focal length is 31. The rectangle overlayed on the flow field represents the neighborhood over which the translational approximation is performed. The actual angles between the correct local translational vectors and the approximated local translational vectors at each position in the flow field is shown in Figure 6. The computed error measure-based upon Equation 4 is shown as a surface plot in Figure 7. Notice that the computed error measure

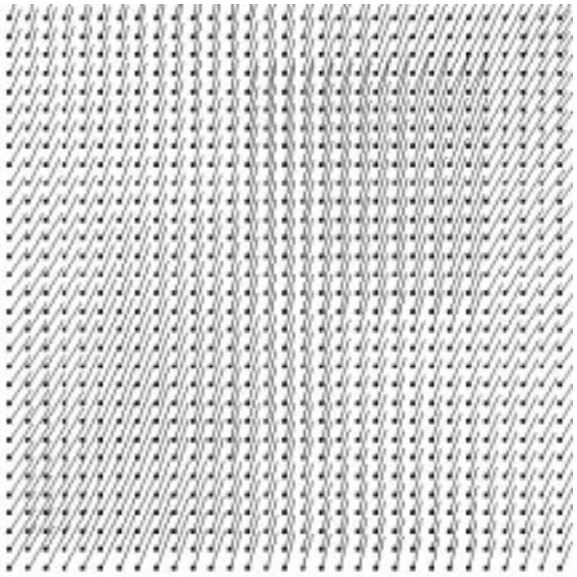

Figure 3: Flow field for an image containing occlusion

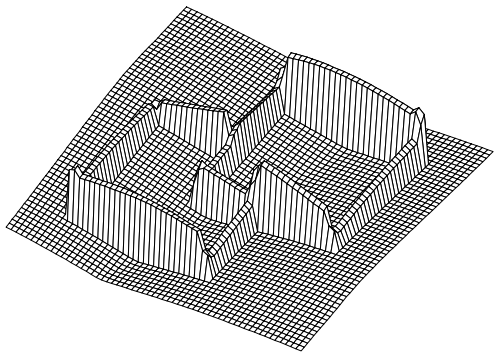

Figure 4: Error function for an image containing occlusion

in Figure 7 reflects a strong correspondence between the approximated translational vectors with the least error and the correct translational axes. This correspondence has been found to be typical. Figure 8 (a)-(c) shows the correct local directions of translation with the values of each component displayed as separate intensity plots. Since the translational vectors are represented as three-dimensional unit vectors with each component in the range of -1.0 to 1.0, Figure 8 displays the x, y, and z components of the local translation vectors with pure white corresponding to the value of 1.0 and pure black corresponding to -1.0. Figure 8 (d)-(f) shows the local translational values that were derived from the optic flow field using the approximation procedure. The

Figure 8: LTD vector components of an arbitrary rigid body motion (a) x-component (b) y-component (c) z-component (d) derived x-component (e) derived y-component (f) derived z-component

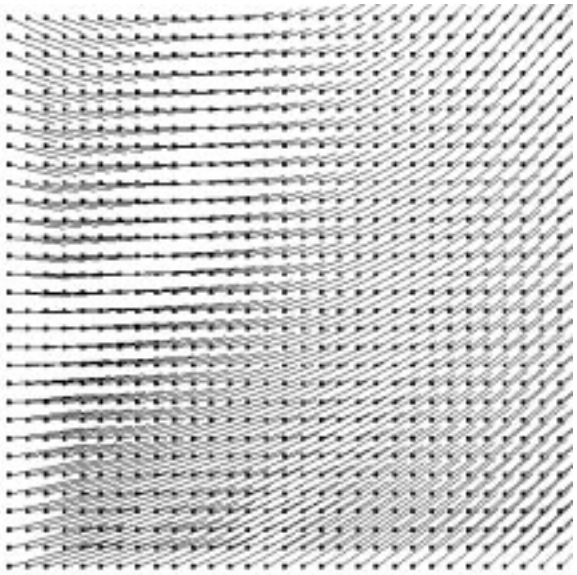

Figure 10: Optic flow field for a planar motion

Figure 9: Motion constrained to a plane

plane of motion. We know from Section 2 that $v_{i,j}$ also lies in the plane determined by its corresponding flow vector $\tilde{v}_{i,j}$ and the focal point of the camera. The estimated direction of motion lies along the intersection of these planes. The estimated direction of motion $\hat{v}_{i,j}$ can be determined by intersecting these planes. Figure 9 shows the geometry, where the plane of motion is positioned so that it intersects the image plane at the base of the flow vector $\tilde{v}_{i,j}$. In terms of image geometry, this corresponds to intersecting the horizon line, determined by the plane of motion through the focal point, with a flow vector. The point of intersection is a Focus of Expansion for the local axis of translation (or a Focus of Contraction: which depends on the direction of the flow vector relative to the point of intersection). Computing the LTD in this case has been found to give extremely low errors (small fractions of a degree) in the estimated local translations.

Motion constrained to a plane is typical in terrestrial circumstances. Several indoor robotic environments involve robot motion constrained to a plane. In vehicular environments, the translational approximation is usually valid due to limitations in vehicle turning radii, meaning that the overall motion of a vehicle can be locally approximated as a translation.

2.2 Motion Constrained to an Undetermined Plane

Processing in the case of motion constrained to an undetermined plane is similar to that of motion constrained to a determined plane. The only difference is that an estimate of the plane of motion must first be recovered. Using the technique described in Section 2 the local translation is computed at each flow vector. Since the motion that produced these local translations is constrained to a plane, each of the local translations must be parallel to this plane. This constraint can be written as

$$\tilde{v}_{i,j} \cdot n = 0 \tag{5}$$

where n is a vector normal to the plane of motion. Using this equation, n can be computed by a linear least squares technique.

An example of processing in this case is shown in Figure 10 to Figure 12. Figure 10 shows the flow field produced by a rotation of 4.58^o about the axis $(-1, 1, 2)$, followed by a translation of $(120, 20, 50)$. Units are given in pixels. This motion is constrained to lie in the plane perpendicular to the normal $(-1, 1, 2)$. However, the plane is unknown, so initially the local translation vectors must be computed by the method used for cases of arbitrary motion.

The angles between the correct local translational values and the derived local translational values shown are plotted in Figure 11. The error measure is shown in Figure 12. Since we have an error measure associated with each point describing the error of the translational approximation, we can select several positions of minimal error for use in Equation 5. Using the error measure from Equation 4 the best 15 local translations were selected for the least squares fit. The recovered plane normal is then $(-0.4107, 0.4129, 0.8129)$ which is off by an angle of 0.37^o from the correct value. We can then use this estimate to evaluate the directions of motion using the technique for motion constrained to a determined plane from the previous section. The computed directions of motion are then shown in Figure 13 (d)-(f). Like the case of motion constrained to a known plane, there is very little error in the derived LTD vectors. The mean angle between derived and actual LTD vectors was 0.176^o and the maximum angle was 1.274^o.

Figure 13: LTD vector components of an undetermined planar motion (LTD estimated using the determined planar motion technique) (a) x-component (b) y-component (c) z-component (d) derived x-component (e) derived y-component (f) derived z-component

gorithm in greater detail.

2.3.1 Local Planarity Assumption

Given a candidate LTD vector, we wish to solve for other nearby LTD vectors. In order to derive a relationship between LTD vectors within a neighborhood, we will assume that surfaces are locally planar. In this case directional derivatives of the LTD vectors along the image plane are constant. Let $\tilde{p}_{i-1,k}$, $\tilde{p}_{i,k}$, and $\tilde{p}_{i+1,k}$ be three collinear points on the image plane. Under the planar surface assumption, we have the following constraint

$$\frac{\hat{v}_{i+1,k} - \hat{v}_{i,k}}{\|\tilde{p}_{i+1,k} - \tilde{p}_{i,k}\|} = \frac{\hat{v}_{i,k} - \hat{v}_{i-1,k}}{\|\tilde{p}_{i,k} - \tilde{p}_{i-1,k}\|} \tag{6}$$

Letting $\hat{v}_{i,k}$ be the current candidate LTD vector, Equation 6 consists of two independent equations and six unknowns. The remaining equations needed to solve for these six unknowns can be provided by the LTD vectors' corresponding optic flow vectors. Figure 1 shows that the plane formed by a flow vector and the focal point

Figure 14: LTD vector components of an arbitrary rigid body motion (LTD vectors were derived using the local planar method) (a) x-component (b) y-component (c) z-component (d) derived x-component (e) derived y-component (f) derived z-component

determined, the solution for the parameters of motion becomes trivial.

Two LTD vectors $\hat{v}_{i,k}$ and $\hat{v}_{j,k}$ are assumed to have undergone identical rigid body motions. We wish to find the relative depth of these two vectors. Figure 15 shows the relationship between the two vectors. One of the vectors, $\hat{v}_{i,k}$, is fixed in depth so that it emanates from the image plane at the point $\tilde{p}_{i,k}$. The unknown depth of the other vector can be expressed as $\alpha\tilde{p}_{j,k}$ where α is some unknown scale factor. Since both of the LTD vectors are the result of the same rigid body motion, we have the following constraint

$$\|\alpha\tilde{p}_{j,k} - \tilde{p}_{i,k}\| = \|\alpha(\tilde{p}_{j,k} + \hat{v}_{j,k}) - (\tilde{p}_{i,k} + \hat{v}_{i,k})\| \quad (10)$$

Squaring both sides and solving for α, Equation 10 can be reduced to

$$\begin{aligned}(2\tilde{p}_{j,k} \cdot \hat{v}_{j,k} + \hat{v}_{j,k} \cdot \hat{v}_{j,k})\alpha^2 - \\ 2(\tilde{p}_{j,k} \cdot \hat{v}_{i,k} + \tilde{p}_{i,k} \cdot \hat{v}_{j,k} + \hat{v}_{i,k} \cdot \hat{v}_{j,k})\alpha + \\ (2\tilde{p}_{i,k} \cdot \hat{v}_{i,k} + \hat{v}_{i,k} \cdot \hat{v}_{i,k}) = 0 \end{aligned} \quad (11)$$

Figure 15: Relative depth of two LTD vectors

This equation is quadratic in α and results in two feasible solutions for the relative depth between two LTD vectors.

3.2 Inferring the Parameters of Motion

Once we have determined the relative depth between LTD vectors the estimation of the parameters of motion is trivial. The problem is equivalent to that of estimating the motion parameters from actual three-dimensional environmental surface positions. A rigid body motion can be expressed as

$$\alpha_{i,j}\hat{v}_{i,j} = r \times \alpha_{i,j}\tilde{p}_{i,j} + t \qquad (12)$$

where r is the axis of rotation and t is the direction of translation. This expression is linear and can be solved using a least squares technique. The expression consists of six parameters and two independent equations. Therefore, it can be solved using a minimum of three (non-collinear) LTD vectors.

3.3 Motion Parameter Inference Results

The rigidity constraints were used to compute the parameters of motion from the derived LTDs presented in Section 2. The results are shown for the case of arbitrary motion, motion constrained to a determined plane, motion constrained to an undetermined plane, and the rigidity-based method applied to arbitrary motion. In the previous section we noted that the parameters of motion can actually be estimated using only three LTD vectors. The feasibility of estimating the parameters of motion from a minimal set of data is demonstrated in the results presented below.

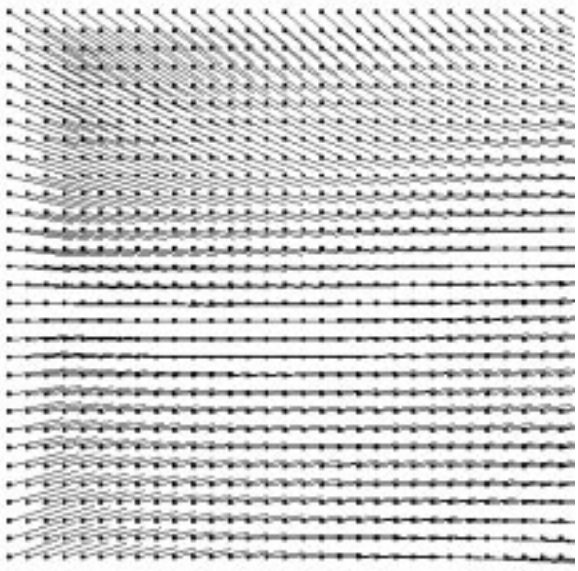

Figure 16: Optic flow field for motion relative to a curved surface

3.3.1 Motion Constrained to a Determined Plane

In the case of motion constrained to a determined plane, the LTD vector estimates tend to be highly accurate over an entire flow field. Typically, when using three LTD vectors selected at random from the derived local translations, the estimate of the axis of rotation and translation almost always are within a degree of the correct axes and the angle of rotation is determined to within a hundredth of a degree.

3.3.2 Motion Constrained to an Undetermined Plane

The case of motion constrained to an undetermined plane is similar to the case of motion constrained to a determined plane in that the LTD vector estimates are very good over the entire image. Three LTD vectors were selected at random from the derived local translations shown in Figure 13. The estimate of the axis of rotation was off by 0.99^o, the angle of rotation was off by 0.04^o, and the direction of translation was off by 0.83^o.

3.3.3 Local Planar Method

The rigidity-based method presented in Section 3.1 is also capable of accurate LTD estimates over the entire flow field. Three LTD vectors were selected at random from the derived local translations shown in Figure 14. The estimate of the axis of rotation was off by 2.26^o, the angle of rotation was off by 0.18^o, and the direction of translation was off by 2.84^o.

The camera was moved about a randomly curved surface. The optic flow field produced by this surface is shown in Figure 16. The three-dimensional environmental surface was reconstructed from this flow field. Figure 17 (a) shows a plot of the original surface. Figure 17 (b) shows the results of the surface reconstruction and Figure 17 (c) shows the resulting error in the

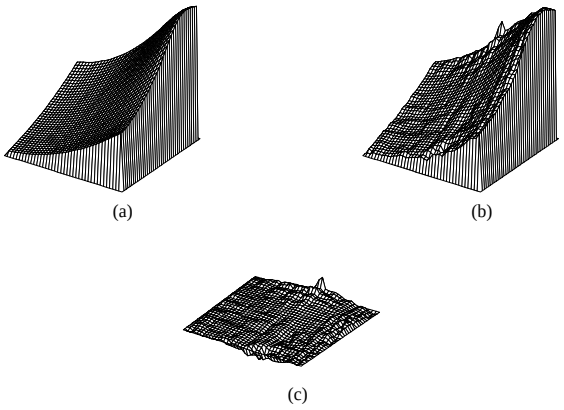

Figure 17: (a) Curved surface (b) Reconstructed surface (c) Error

reconstruction. The surface shown in this example is not planar. However, the reconstruction is fairly accurate, despite the violation of the planarity assumption. Experiments indicated that surfaces which are approximately planar in a local neighborhood can be successfully reconstructed. Therefore, any continuous surface can be reconstructed, given an appropriate density of optic flow vectors.

3.3.4 Arbitrary Motion

Using the error measure shown in Figure 7 and the derived LTD vectors shown in Figure 8, the three best LTD vectors were selected and used to compute the parameters of motion. The estimate of the axis of rotation was off by 8.13^o, the angle of rotation was off by 1.09^o, and the direction of translation was off by 12.02^o. In the previous section it was shown that the minimum number of LTD vectors which can be used to estimate the parameters of motion is three. However, we can use a larger set of LTD vectors in a least squares procedure to obtain more accurate results. For example, when the ten best LTD vectors were used, the axis of rotation was off by 3.65^o, the angle of rotation was off by 0.44^o, and the direction of translation was off by 9.32^o.

4 Summary and Future Work

We have introduced the local translational decomposition (LTD) as a low level representation of environmental motion which can simplify the inference of motion parameters from optic flow fields. We have found that this is particularly robust and simple for cases of motion constrained to a determined or undetermined plane, and motion relative to locally planar surfaces. In addition, It is possible to infer motion parameters from sparse LTDs.

Areas for further work include:

- Develop criteria to determine the the best set of estimated local translation vectors to estimate motion parameters in order to take advantage of the limited number of points for which the local translation needs to be determined to infer motion parameters.
- Investigate local translational analysis with the use of multiple cameras and longer image sequences.
- The local translation decomposition is similar to an array of localized looming detectors which determine whether things are coming towards or away from an observer at a particular image position. It may be possible to use such a distributed representation of motion relative to environmental surfaces to control navigation and other behaviors directly, without the inference of motion parameters from the LTD.
- The local translation approximation can be used as a criteria for computing flow to determine the LTD directly without the initial computation of a flow field. In the experiments presented above, we have assumed a uniformly dense flow field of high resolution. The translation procedure developed in [Lawton, 1982] was not applied to computed flow fields, but to successive images for which interesting points had been extracted from the initial image. Given distinctive features (at least two), it was possible to compute the direction of translation in a small image area. This use of the translational procedure can be seen as a local constraint on the determination of image displacements such that the corresponding environmental motion can be interpreted as being translational. For egomotion, this wouldn't require computation over the entire flow field since only three LTD vectors are needed. Where the translational approximation is poor there will be a large value in the error measure reflecting weaker confidence in the validity of the approximation.

References

[Burger and Bhanu, 1989] W. Burger and Bir Bhanu. On computing a 'fuzzy' focus of expansion for autonomous navigation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 563–568, 1989.

[Gibson, 1950] J.J. Gibson. *The perception of the visual world.* Houghton Mifflin, Boston, MA, 1950.

[Jain, 1983] R. Jain. Direct computation of the focus of expansion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5:58–64, 1983.

[Lawton, 1982] Daryl T. Lawton. Processing translational motion sequences. *Computer Vision, Graphics, and Image Processing*, 22:116–144, 1982.

[Lee, 1980] D.N. Lee. The optic flow field: The foundation of vision. *Philosophical Transactions of the Royal Society of London, Series B*, 290:169–179, 1980.