Proc. Computer Vision and Pattern Recognition, CVPR-91, Maui, Hawaii, June 1991, pp. 296-302.

Robust Dynamic Motion Estimation Over Time

Michael J. Black
Department of Computer Science
Yale University
New Haven, CT 06520–2158

P. Anandan
David Sarnoff Research Center
Subsidiary of SRI International
Princeton, NJ 08543–5300

Abstract

This paper presents a novel approach to incrementally estimating visual motion over a sequence of images. We start by formulating constraints on image motion to account for the possibility of multiple motions. This is achieved by exploiting the notions of weak continuity *and* robust statistics *in the formulation of a minimization problem. The resulting objective function is non–convex. Traditional stochastic relaxation techniques for minimizing such functions prove inappropriate for the task. We present a highly parallel* incremental stochastic minimization *algorithm which has a number of advantages over previous approaches. The incremental nature of the scheme makes it truly dynamic and permits the detection of occlusion and disocclusion boundaries.*

1 Introduction

This paper presents an approach for the incremental estimation of visual motion over time. The task of estimating visual motion involves specifying constraints which relate spatiotemporal intensity variations to image motion and express our assumptions about the spatiotemporal variation of the motion itself. We also require an effective procedure for computing the motion consistent with the assumptions.

In formulating constraints on image motion, the traditional Gaussian noise model assumes that within a small image region only a single motion is present. The assumption however, ignores the case of motion discontinuities [4] and results in either errors in the motion estimate or over smoothing across discontinuities. This paper formulates more realistic constraints which account for multiple motions occurring at surface boundaries by exploiting the notions of *weak continuity* [5, 7] and *robust statistics* [8].

The constraints are formulated as energy terms in an objective function which is minimized to estimate the motion. With the removal of the simplifying assumption of Gaussian noise the objective function becomes highly non-convex.

Stochastic methods, like simulated annealing [7, 12], are one approach for minimizing such complex functions with many local minima. While they are highly parallel, these approaches converge slowly making them ill suited to motion estimation which must be dynamic.

We propose a new *incremental stochastic minimization (ISM)* algorithm which has the benefits of simulated annealing without many of the shortcomings [2, 3]. The ISM approach, which is designed to minimize an objective function *changing slowly over time*, is parallel, incremental and robust. Additionally, the approach provides estimates of *occlusion* and *disocclusion* boundaries.

The next section formulates constraints which account for multiple motions. Section 3 presents the ISM approach and discontinuity detection. The algorithm is then extended to handle large motions in Section 4 and experimental results are presented in Section 5.

2 Multiple Motions, Robust Statistics and Weak Continuity

We specify our assumptions about the scene and the images in terms of constraints. The constraints are formalized as energy functions over local neighborhoods, or *cliques*, in a grid. For an image of size $n \times n$ pixels we define a grid of *sites:*

$$S = \{s_1, s_2, \ldots, s_{n^2} \mid \forall w \; 0 \leq i(s_w), j(s_w) \leq n-1\},$$

where $(i(s), j(s))$ denotes the pixel coordinates of site s. Horizontal and vertical image motion at a site s is denoted by the vector $\mathbf{u}(s) = (u(s), v(s))$.

For the remainder of the paper we focus on three constraints [2, 3]: *data conservation*, *spatial coherence*, and *temporal coherence*. Various approaches have been presented for formulating the the spatial coherence assumption to account for motion discontinuities; in particular, the notion of *weak continuity constraints* has been popular [5, 7] . Less attention has been paid, however, to relaxing the data conservation assumption. In fact, we observe that the two problems are both special cases of the more general statistical problem of *outlier rejection* encountered in *robust statistics* [8].

The general problem is one of finding the best fit of a model to data where we have some (possibly inaccurate) prior model of the statistics of the errors in the data. The least-squares fit of the sort typically employed with these constraints implies a Gaussian noise model. In the case of multiple motions, our prior Gaussian noise model is incorrect due to outliers. Our goal then is to find the best fit to the data while ignoring outlying data.

2.1 The Data Conservation Constraint

The data conservation constraint embodies the assumption that the intensity of a surface element remains constant over time, although its image location may change. We adopt a correlation based approach in which a correlation *surface* at a site s, $E_D(u, v, s)$, is defined over the space of possible displacements (u, v) with the height of the surface corresponding to an estimate of the data error of that displacement. The minimum of this surface corresponds to the best motion estimate with respect to the data conservation assumption.

Let s and t denote image locations, or sites, in S. We define a neighborhood of s, $\eta_D(s)$, for the data conservation constraint as a square "window" of sites about s. Data error is defined as the the difference between predicted and measured intensity values. Given image intensity functions I_n and I_{n+1} between two successive frames, the local contribution to the data conservation constraint $E_D(u, v, s)$ is defined as:

$$\sum_{t \in \eta(s)} \phi_D(I_n(i(t), j(t)) - I_{n+1}(i(t) + u, j(t) + v)). \quad (1)$$

The standard quadratic error measure ($\phi_D(x) = x^2$), which is a direct consequence of an additive Gaussian noise model, is not robust in the presence of outliers. As the magnitude of the data error increases, the contribution to the error term increases without bound. As a result, when multiple motions are present within the neighborhood of a site, the correlation computed for one of the motions is corrupted by the outliers corresponding to the other motion.

What is needed is a new error measure which takes into account the outliers which violate the Gaussian noise assumption. Heuristically, we would like such an error measure to behave like the quadratic measure when the data errors are small (and hence are more likely to have come from the consistent surface). We also want the influence of large errors (which correspond to the uncorrelated motion) to be reduced.

One way of characterizing the behavior of an error measure, $\phi(x)$, is by its *influence function*, $\psi(x) = \frac{d}{dx}\phi(x)$ [8]. For the standard quadratic error measure

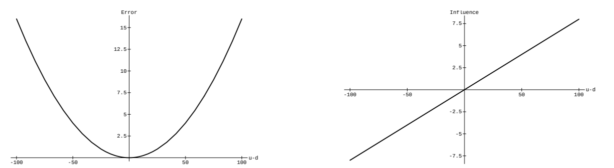

Figure 1: *a)* Standard quadratic error measure, *b)* Influence function for the quadratic error measure.

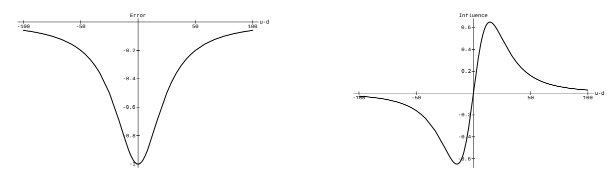

Figure 2: *a)* A robust error measure, ϕ_D, *b)* Influence function ψ_D for ϕ_D.

(figure 1*a*) the influence of errors increases linearly and without bound (figure 1*b*).

An error function with the desired saturating properties (figure 2*a*) is:

$$\phi_D(x) = \frac{-1}{1 + (x/\Delta_D)^2}, \quad (2)$$

where Δ_D is a constant scale factor. Examining the influence function of ϕ_D (figure 2*b*) we see it the influence of outliers tends to zero. This function ϕ_D is related to the *redescending* estimators used in robust statistics [8].

Sub-pixel Accuracy The data error term $E_D(u, v)$ as defined is discrete. Sub-pixel motion estimates can be obtained by interpolating the error surface [1, 11]. When the Gaussian noise assumption is violated the standard quadratic interpolation is incorrect. The new error surface can be interpolated by using *bi–cubic splines*[2].

2.2 The Spatial Coherence Constraint

The *spatial coherence* constraint is derived from the observation that surfaces have spatial extent and hence neighboring points on a surface will have similar motion. Once again, the spatial coherence assumption and its standard (quadratic) formulation [1] are invalid in areas containing multiple motions.

The neighborhood, $\eta_S(s)$, for the spatial coherence constraint is defined to be the nearest neighbors of a site s at location (i, j) in the grid. To cope with multiple motions, the spatial coherence constraint can be reformulated using *weak continuity constraints* [6]:

$$E_S(\mathbf{u}, s) = \sum_{t \in \eta_S(s)} \alpha(l)\|\mathbf{u}(s) - \mathbf{u}(t)\| + \beta(l), \qquad (3)$$

where $\mathbf{u}(s) = (u(s), v(s))$ is the motion vector at site s, l is a *continuous line process* variable, $0 \leq l \leq 1$, $\alpha(0) = 0$ and is increasing, and $\beta(0) = 0$ and is decreasing. The value of l can be thought of as indicating the likelihood of a discontinuity and $\beta(l)$ can be thought of as a penalty for introducing a discontinuity. This is a generalization of the Blake and Zisserman formulation [5].

The line process variables can be removed from the smoothness constraint by first minimizing over them [5, 6] resulting in an equivalent minimization problem:

$$E_S(\mathbf{u}, s) = \sum_{t \in \eta_S(s)} \phi_S(\mathbf{u}(s) - \mathbf{u}(t)), \qquad (4)$$

which is just a function ϕ_S of the difference in the neighbors' flow. For the appropriate choice of α and β (see [6]) we have

$$\phi_S(x) = \frac{-1}{1 + |x|/\Delta_S}. \qquad (5)$$

This error measure, like ϕ_D, saturates as errors increase thus performing outlier rejection.

2.3 The Temporal Coherence Constraint

Our current formulation of the temporal coherence constraint embodies the assumption that the image plane acceleration of a patch is constant over time. This can be regarded as a first approximation to a more accurate model, namely continuous 3-D motion.

Let $\mathbf{u}^-$ and $\Delta\mathbf{u}^-$ denote the predicted velocity and acceleration and $\mathbf{u}^+$ and $\Delta\mathbf{u}^+$ the estimated values. We predict the new velocity at time t of a given patch as the estimated motion at the previous time instant plus the predicted acceleration:

$$\mathbf{u}_t^- = \mathbf{u}_{t-1}^+ + \Delta\mathbf{u}_t^-. \qquad (6)$$

Since the estimated accelerations may be noisy, we predict the new acceleration to be a temporal average of previous estimates. This can be obtained by,

$$\begin{aligned} \Delta\mathbf{u}_t^- &= \alpha\Delta\mathbf{u}_{t-1}^+ + (1-\alpha)\Delta\mathbf{u}_{t-1}^- & (7) \\ \Delta\mathbf{u}_{t-1}^+ &= \Delta\mathbf{u}_{t-2}^+ + (\mathbf{u}_{t-1}^+ - \mathbf{u}_{t-2}^+), & (8) \end{aligned}$$

where $0 \leq \alpha \leq 1$ controls the rate at which new information replaces previous information.

Given a prediction of the new velocity of a patch $\mathbf{u}_t^- = (u_t^-, v_t^-)$, the temporal constraint is formulated as,

$$E_T(\mathbf{u}, t) = \phi_T(\mathbf{u} - \mathbf{u}_t^-), \qquad (9)$$

where ϕ_T is the same function used in the smoothness error term, with a possibly different Δ_T.

3 Recovering the Flow Field

The constraints of the previous section, which embody our assumptions about the world, can now be combined to form an objective function $H(u, v, t)$:

$$\beta_D E_D(u, v) + \beta_S E_S(u, v) + \beta_T E_T(u, v, t), \qquad (10)$$

where the $\beta_\star$ are constant weights which control the relative importance of the constraints. Based on our assumptions, the best interpretation of the motion, (u, v), is the minimum of this function.

The formulation of the constraints to account for multiple motions means that H has many local minima making the task of finding the (u, v) which minimize the function difficult. The definition of the constraints in terms of local neighborhoods on a grid allows the problem to be formalized as a *Markov Random Field (MRF)* [7, 10].

Each site in the MRF can be thought of as representing a small environmental surface patch. Associated with each site s is a continuous random vector $\mathbf{u}$ which represents the current image displacement of the corresponding surface patch. The discrete *state space* $\Lambda_s(t)$, at a site s, defines the possible values that the random vector can take on at a given time t.

For each site, we construct a probability density function Π defined over the range of possible displacements Λ using a *Gibbs distribution* [7] as follows:

$$\Pi(u, v, t) = Z^{-1} e^{-H(u,v,t)/T(t)}, \qquad (11)$$

where:

$$Z = \sum_{(u,v) \in \Lambda(t)} e^{-H(u,v,t)/T(t)}$$

and where t is the current time instance. The quantity $T(t)$ can be thought of as a *temperature* which serves to sharpen (or flatten) the distribution.

Standard *simulated annealing* techniques can be used to find the minimum (u, v) by sampling from Λ according to the distribution Π with logarithmicly decreasing temperatures[7].

While this simulated annealing approach is highly parallel, it suffers from two main problems. First, the Monte Carlo techniques used to sample Π assume a discrete state space while we need to solve a continuous minimization problem for arbitrary fractional displacements. Second, simulated annealing is computationally intensive, requiring hundreds of iterations to converge to reasonable results.

The first problem can be solved by using a *continuous* variant of simulated annealing [12]. The second problem requires a more radical solution. By tracking small patches of a scene over an image sequence, we will modify the basic annealing concept to work on changing data over time.

3.1 Continuous Annealing and Sub-Pixel Displacements

To solve continuous problems we allow the state space $\Lambda_s(t)$ to vary over time depending on the local properties of the function being minimized. At a given time t, we have an estimate of the motion $\mathbf{u}_t$, and consider making small changes $\Delta\mathbf{u}_t$ to the estimate in an attempt to minimize H. Vanderbilt and Louie[12] define a method which is *adaptive* in that the state space (defined by the step size, $\Delta\mathbf{u}_t$) adjusts to the local shape of the function being minimized.

We characterize the local shape of the function by its covariance matrix, S, computed at the current step size. We adjust the state space to best explore the function by choosing step sizes so that the covariance matrix, s, of the state space, Λ, is proportional to S. Intuitively, if the variance along a particular search direction is large, then we want to increase the step size in that direction to get a coarse view of the function. When the true minimum has been chosen at a coarse level, the variance will shrink. To explore the minimum more finely, the area covered by the state space should shrink resulting in smaller step sizes.

At a given site and at a given time, the state space Λ is always a 3×3 neighborhood of the current estimate, but the area covered by the neighborhood varies based on the current step size $\Delta\mathbf{u}_t = [\Delta u_t, \Delta v_t]$. Given a current estimate $\mathbf{u}_t = [u_t, v_t]$, at time t the state space Λ is defined as:

$$\Lambda = \{\mathbf{u} + \Delta\mathbf{u} \mid \Delta\mathbf{u} = \mathbf{Q} \cdot \mathbf{l},\ \mathbf{l} = [l_1, l_2]^T\} \qquad (12)$$

where, $l_1, l_2 \in \{-(3/2)^{\frac{1}{2}}, 0, (3/2)^{\frac{1}{2}}\}$, and where $\mathbf{Q}$ is a 2×2 matrix which controls the step size. Elements of the state space are all examined with equal probability, so the choice of trial steps is governed by a uniform probability distribution $\mathbf{g}(\mathbf{l})$ which over $\{-(3/2)^{\frac{1}{2}}, 0, (3/2)^{\frac{1}{2}}\}$ has zero mean and unit variance.

Since the mean of Λ is $\mathbf{u}$, the covariance matrix $\mathbf{s}$, of the state space is simply:

$$s_{ij} = \sum_{\Delta\mathbf{u} \in \Lambda} \Delta\mathbf{u}_i \Delta\mathbf{u}_j \mathbf{g}(\mathbf{l}). \qquad (13)$$

Vanderbilt and Louie [12] note that this can be expressed as:

$$\mathbf{s} = \mathbf{Q} \cdot \mathbf{Q}^T. \qquad (14)$$

Hence we can generate a state space with any desired covariance matrix $\mathbf{s}$ by solving for $\mathbf{Q}$ using Cholesky decomposition and then using $\mathbf{Q}$ to generate the state space in equation 12.

The actual step taken at a time t is determined by the probability distribution $\Pi(\mathbf{u}_t + \Delta\mathbf{u}_t)$ defined over the space of displacements. Using Π we can compute the mean μ at time t (note we drop t when it is constant across all terms):

$$\mu_i = \sum_{\mathbf{u} \in \Lambda} \Pi(\mathbf{u})\mathbf{u}_i. \qquad (15)$$

The covariance matrix $\mathbf{S}$ of Π given the current step size is:

$$S_{ij} = \sum_{\mathbf{u} \in \Lambda} (\mathbf{u}_i - \mu_i)(\mathbf{u}_j - \mu_j)\Pi(\mathbf{u}). \qquad (16)$$

We make the covariance matrix of the state space at time $t+1$ proportional to $\mathbf{S}^{(t)}$:

$$\mathbf{s}^{(t+1)} = \chi \mathbf{S}^{(t)}, \qquad (17)$$

where $\chi > 1$ is a scaling factor. Now solving $s^{(t+1)} = \mathbf{Q} \cdot \mathbf{Q}^T$ for $\mathbf{Q}$ gives the $\mathbf{Q}$ for determining the state space at the next time instant.

To prevent the state space from growing or shrinking too rapidly, we control the rate at which new information from $\mathbf{S}$ overwrites the previous information:

$$\mathbf{s}^{(t+1)} = \alpha\chi \mathbf{S}^{(t)} + (1 - \alpha)\mathbf{s}^{(t)},$$

where α can be viewed as a damping factor.

3.2 Incremental Minimization

The obvious disadvantage of simulated annealing is its computational expense. However, since we expect the changes in the images and in the scene to be gradual and predictable, the iterative minimization process can be extended over an image sequence. This will also allow the motion detection algorithm to exploit the wealth of information available over time to achieve greater sensitivity and robustness while minimizing the amount of computation between frames.

When a new image is acquired, the current motion estimate at a given site is used as the starting point for the continuous annealing algorithm and to compute the predicted motion used in the temporal coherence constraint. The current temperature at that site is used as the initial temperature, and is then lowered according to the annealing schedule.

After a fixed (usually small) number of iterations of the annealing process, each site has a new motion estimate and temperature. The various properties of the associated surface are then propagated to the new site where the patch has moved. These properties include the patch's motion, temperature, and state space. Additional properties like image intensity or higher level information about surface membership may also be propagated.

This propagation can be viewed as *warping* the sites according to the motion estimate[2, 9]. Since the motion is not discrete, the field is resampled using a weighted bi–linear interpolation, where the weighting reflects the confidence in the motion estimates.

Let $\eta(s)$ denote the neighbors of site s whose motion estimates place them within one pixel of s (we will extend this to large motions in the next section). Let ρ be a property of s. Then the new estimate of $\rho(s)$ is given by:

$$\rho(s) = \frac{1}{w(s)} \sum_{t \in \eta(s)} \Pi(\mathbf{u}(t))(1 - d(s,t))\rho(t) \quad (18)$$

$$w(s) = \sum_{t \in \eta(s)} \Pi(\mathbf{u}(t))(1 - d(s,t)), \quad (19)$$

where w is a normalizing term, and $d(s,t)$ is the distance between the projection of site t and the location of site s.

Occlusion and Disocclusion The propagation algorithm outlined above can be made sensitive to the presence of occlusion and disocclusion around each site. Observe that the normalizing factor w roughly measures the total flow into a site. In the absence of motion discontinuities this should be approximately unity. However, if occlusions are present within the neighborhood of a site, we may expect multiple sites to move towards it, thereby increasing the total in-flow. Similarly, if there is a disocclusion, we may expect the total flow to be less than unity.

Sites can be classified as locations of occlusion or disocclusion using two thresholds, one above and one below unity respectively. This simple scheme may prove insufficient for certain situations. For example, if there is significant divergence (or convergence) present within the neighborhood of a site, net flow will differ from unity, even if there are no motion discontinuities.

A disoccluded site indicates a new patch of the environment which was previously hidden from view. For this new patch, there is no prior motion estimate, hence the annealing process should be initially uncommitted about the motion. This is achieved by initializing the site to have a high temperature.

Unlike standard annealing, our algorithm uses different temperatures for the different sites and dynamically modifies the temperature according to the information available at a site. As a patch is tracked, its temperature will decrease over time. Hence, the temperatures of patches that have been tracked over many frames and whose motion is precisely known tend to be lower than those of more recently disoccluded patches.

Convergence Unlike simulated annealing, we have no theoretical convergence results for this new incremental minimization scheme. Empirical results indicate that the approach does in fact converge to the correct sub-pixel motion estimates. Obviously, the degree to which the constraints accurately reflect the physics of the world will affect both the convergence and the accuracy of the algorithm. The current model and the constraints used are first order approximations to the correct physical models. We expect, however, the framework presented here can be extended to incorporate more precise models of the scene and its geometry.

4 Spatio-Temporal Pyramid

The previous section described how small motions can be estimated over time. The most obvious way to estimate large motions is to expand the state space to be larger than 3×3 and increase the maximum allowed step size, but this results in a loss of efficiency and communication between distant sites. To achieve efficient and robust computation of large motions we adopt a multi-resolution strategy [1].

We construct a pyramid of spatially filtered and subsampled images so that at the highest level in the pyramid the largest motion is less than a pixel. Each level of the pyramid can be thought of as a MRF which is responsible for estimating motions of one pixel or less. The continuous annealing process described in the previous section is applied at each level in parallel so that each level estimates its motion simultaneously and independently.

To derive a global motion estimate, the motion estimates computed at each level are combined using a coarse-to-fine, flow-through, strategy without refinement in which large motions are determined solely at the lower spatial frequencies. The motion estimate for each site is taken from the highest level of the pyramid at which the motion is greater than one half pixel.

5 Experimental Results

The incremental algorithm has been tested on real and synthetic image sequences.

5.1 Synthetic Motion Experiments

To illustrate the convergence properties of the algorithm a synthetic image sequence was generated. The sequence consists of a 64×64 pixel uniform random signal over the range $[0, 255]$ which is undergoing a uniform translation of one half pixel to the right and down per frame.

Experiments with a noiseless signal illustrate the convergence of the algorithm over time. Figure 3 plots variance of the motion estimate as a function of the

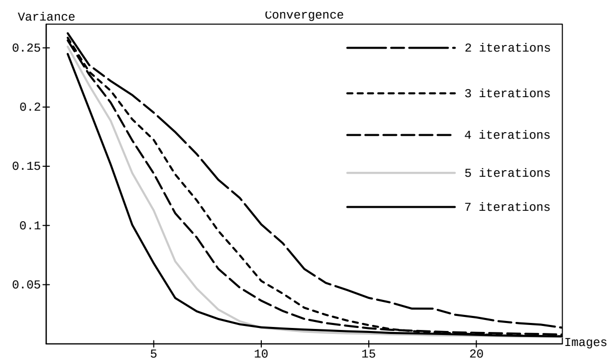

Figure 3: **Convergence Experiments.** Variance as a function of the number of frames in a 25 image sequence.

number of images examined in the sequence. The variance is plotted for trials using 2, 3, 4, 5 and 7 iterations per frame. Similar experiments with images corrupted by noise indicate that the algorithm can tolerate fairly large amounts of noise.

5.2 Motion Discontinuities

The following experiment involves an image sequence consisting of eight 64×64 square images; the last image in the sequence is shown in figure 4*a*. The flow field, computed to sub-pixel accuracy, is shown in figure 4*b*. Notice that over-smoothing does not take place and flow discontinuities are maintained.

Occlusion and disocclusion boundary estimates are shown in figure 4*c*. Bright areas correspond to occlusion, dark areas to disocclusion. It is important to remember, that while these results show only the final frames in the image sequence, both flow and discontinuity estimates are available at all times.

5.3 Nap-Of-the-Earth Experiment

The final experiment tests the full algorithm, including the multi-resolution strategy. The test sequence consists of 100 images of size 128×128 pixels. The images were acquired from a camera mounted on a helicopter in *Nap-Of-the-Earth* (NOE) flight. The sequence is challenging in many respects. The motion, ranging from 0 to approximately 4 pixels, is complex and changing; there is pitch, yaw and rotation in addition to translation. The actual motion is corrupted by jitter introduced by the camera mounting and turbulence.

Unfortunately, it is impossible to convey the dynamic behavior of the algorithm over the 100 image sequence in a static format for presentation here. Figure 5 shows snapshots of the processing after 45, 60, 75, and 90 frames.

6 Conclusion

This paper has presented a novel approach to incrementally computing motion estimates over a sequence of images. The starting point is the formulation of constraints on image motion which take into multiple motions. The resulting minimization problem is difficult to solve and traditional stochastic techniques are inappropriate for motion processing. To ameliorate these problems an incremental minimization algorithm was developed.

The approach has a number of advantages over previous approaches. The incremental and adaptive nature of the scheme makes it appropriate for dynamic motion processing. In particular, the local nature of the computations makes it possible to exploit the high degree of parallelism inherent in the problem. Additionally, the warping process allows the detection of occlusion and disocclusion boundaries.

Our current research is extending this scheme in a number of directions. First we are exploring new formulations of the temporal minimization problem and its relationship to Kalman filtering [9, 11]. We are also considering other possible constraints, for example rigid body motion. Additionally we are exploring the use of robust statistics in coping with multiple motions.

Finally, it should be noted that the usefulness of the model extends beyond motion estimation. The framework for tracking surface patches over time may permit the extension of traditional two frame algorithms to a sequence of frames.

Acknowledgements

This work was partially supported by the Defense Advanced Research Products Agency, contract number DAAA15–87–K–0001, administered by the Ballistic Research Laboratory and by a grant from the Whitaker Foundation. Portions of this work were performed while the first author was at the NASA Ames Research Center, Aerospace Human Factors Research Division, with the support of NASA RTOP 506-47, and while the second author was at Yale University.

References

[1] Anandan, P., "A computational framework and an algorithm for the measurement of visual motion," *Int. J. of Computer Vision*, 2, 1989, pp. 283–310.

[2] Black, M.J., and Anandan, P., "Robust dyanamic motion estimation over time," Yale University, Research Report YALEU/DCS/RR-835, November, 1990.

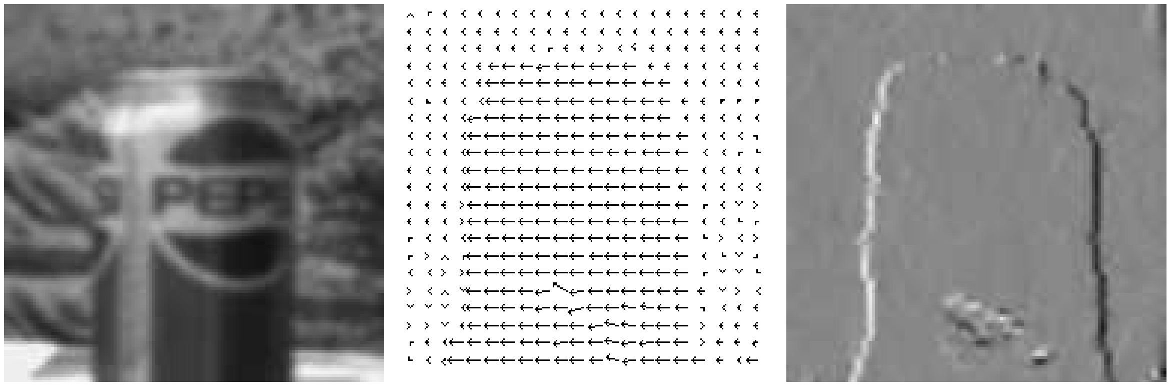

Figure 4: Pepsi can image sequence: *a)* Intensity image; *b)* Flow field; *c)* Occlusion/Disocclusion Boundaries.

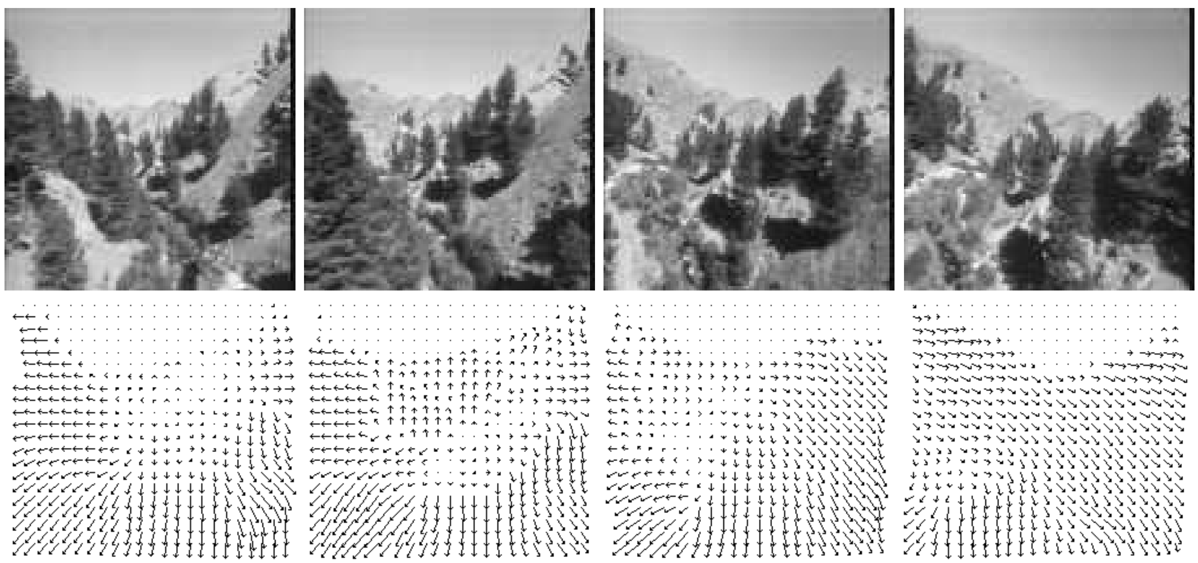

Figure 5: **Nap-Of-the-Earth Helicopter Sequence.** Snapshots of a 100 image sequence are shown with the current image above the motion estimate: *a)* after 45 images, *b)* after 60 images, *c)* after 75 images, *d)* after 90 images.

[3] Black, M.J., and Anandan, P., "A model for the detection of motion over time," *Proc. Int. Conf. on Comp. Vision*, ICCV-90, Osaka, Japan, Dec. 1990, pp. 33–37.

[4] Black, M.J., and Anandan, P., "Constraints for the early detection of discontinuity from motion," *Proc. AAAI-90*, Boston, MA, 1990, pp. 1060–1066.

[5] Blake, A. and Zisserman, A., *Visual Reconstruction,* The MIT Press, Cambridge, Massachusetts, 1987.

[6] Geman, D. and Reynolds, G., "Constrained restoration and the recovery of discontinuities," unpublished manuscript.

[7] Geman, S. and Geman, D., "Stochastic relaxation, Gibbs distributions, and Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. PAMI-6, No. 6, November 1984.

[8] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.,A, *Robust Statistics: The Approach Based on Influence Functions,* John Wiley and Sons, New York, 1986.

[9] Heel, J., "Temporally integrated surface reconstruction," *Proc. Int. Conf. on Comp. Vision*, ICCV-90, Osaka, Japan, Dec. 1990, pp. 292–295.

[10] Konrad, J., "Bayesian estimation of motion fields from image sequences," *Ph.D. Dissertation,* McGill University, Montreal, Canada, June 1989.

[11] Matthias, L., Szeliski, R., Kanade, T., "Kalman filter-based algorithms for estimating depth from image sequences," *Int. J. of Computer Vision*, 3(3), Sept. 1989, pp. 209–236.

[12] Vanderbilt D., and Louie S. G., "A monte carlo simulated annealing approach to optimization over continuous variables," *J. of Comp. Physics,* **56**, pp. 259–271, 1984.