Estimating Multiple Independent Motions in Segmented Images using Parametric Models with Local Deformations

Michael J. Black
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304
black@parc.xerox.com

Allan Jepson
Department of Computer Science
University of Toronto
Toronto, Ontario Canada
jepson@vis.toronto.edu

Abstract

This paper presents a new model for optical flow based on the motion of planar regions plus local deformations. The approach exploits brightness information to organize and constrain the interpretation of the motion by using segmented regions of piecewise smooth brightness to hypothesize planar regions in the scene. Parametric flow models are fitted to these regions in a two step process which first computes a coarse fit and then refines it using a generalization of the standard area-based regression approaches. Since the assumption of planarity is likely to be violated, we allow local deformations from the planar assumption. This parametric+deformation model exploits the strong constraints of parametric approaches while retaining the adaptive nature of regularization approaches.

1 Introduction

Estimating the motion of independent or articulate objects necessitates the segmentation of the scene into regions of coherent motion. If the scene were segmented into roughly planar surfaces then the motion of each surface could be estimated using a parametric flow model. These models provide strong constraints on the motion within a region resulting in accurate flow estimates. In contrast to recent parametric approaches which assume that the entire image, or an arbitrary rectangular region, can be modeled by a single motion, we independently model the motion of segmented planar surface regions. But segmentation is a hard problem in its own right and, in particular, the recovery of segmented, or piecewise smooth, flow fields is notoriously difficult. Instead, this paper makes the simple hypothesis that *image regions of piecewise smooth brightness are likely to correspond to surfaces in the world.* These brightness regions are assumed to be planar surfaces in the scene and their motion is estimated using an eight-parameter flow model. In this way, information from image brightness is used to organize and constrain the interpretation of the optical flow. Since the assumption of planarity may be wrong, we allow local deformation from the planar assumption in the same spirit as physically-based approaches which model shape using coarse parametric models plus deformations [15]. The resulting model, in which optical flow is represented by the motion of planar image patches with local deformations, exploits the strong constraints of parametric approaches while retaining the adaptive nature of regularization approaches. Experiments with a variety of images indicate that the parametric+deformation model produces accurate flow estimates while the incorporation of brightness segmentation provides precise localization of motion boundaries.

The algorithm can be thought of as having low- and medium-level processing. At the low level there is a process which is always smoothing the image brightness while accounting for brightness discontinuities. There is another low-level process that is always providing coarse estimates of image motion. The medium level tries to organize and make sense of the low-level data by first finding connected regions of piecewise smooth brightness and then by estimating the motion of these regions. This medium-level motion-estimation process has two steps. The first fits a parametric model to the coarse motion estimates in each region. In the second step, the parametric fit from the initial estimate is used to warp the image regions into alignment. The optical flow constraints derived from the registered regions are used to refine the parametric fit by performing regression over each region. Finally, the planar patches are allowed to deform at the low-level subject to weak constraints from the optical flow constraints, the spatial coherence of the neighboring flow estimates, and the motion estimate for the planar patch.

2 Previous Work

Parametric Models of Image Motion. Parametric models of optical flow within an image region provide both a concise representation and enforce strong constraints on the interpretation of the motion. These techniques use regression or a Hough transform to estimate a few parameters (eg. two, six, or eight) given hundreds or thousands of constraints computed over the entire image or some pre-selected region [3, 13]; when the image motion conforms to the model assumptions this produces accurate flow estimates. The problem with this approach is that parametric motion models applied over arbitrary regions are rarely valid in real scenes due to surfaces at varying depths or the independent motion of objects.

Approaches have been devised which ameliorate some of the problems of global parametric models [5, 10, 11], but these approaches can cope with only a small number of motions and not with general flow fields. As global approaches, they do not address how to select appropriate image regions in which to apply the parametric models.

Another set of approaches apply parametric models to coarse flow fields by grouping the flow vectors into consistent regions [1, 17]. These approaches, like the regression approaches, are essentially global techniques in that they assume the image motion can be represented by a small number of layers. Additionally they fail to exploit information present in the image brightness about the nature of surfaces in the scene.

Exploiting Image Brightness. To improve motion segmentation a number of researchers have attempted to combine intensity and motion information [6, 9, 16]. In focusing on motion boundaries these approaches use weak models of optical flow (eg. regularization) and hence neglect one of the benefits of having a segmentation in the first place; that is, that the motion of a segmented region can often be described using a simple parametric model which allows many constraints to be integrated across the region.

There are numerous feature-based schemes which estimate motion by tracking points, edges, or region contours computed from the brightness image (eg. [18]). These approaches use information about image brightness to constrain the motion estimation, but brightness contours alone are an impoverished representation. The motion information available over an entire region, particularly if it is reasonably textured, provides additional constraints which can improve the accuracy of the recovered motion.

In the context of stereo reconstruction, Luo and Maître [14] use a segmented intensity image to correct and improve disparity estimates by fitting a plane to the disparities within a region of uniform brightness. This is similar to the first stage of our algorithm. The accuracy of this approach is affected by the accuracy of the initial disparity estimates. Koch [12] segments regions using disparity and brightness and then regularizes depth estimates within the regions. While this approach preserves depth boundaries it uses a weak model within regions instead of fitting a model with a small number of parameters.

3 Early Processing

At the low level there are two processes which examine the input images: segmentation and coarse motion estimation. The exact methods used for these early process are not crucial to the optical flow model described in this paper so the algorithms are described only briefly and the reader is referred to [8] for a complete description of the segmentation approach and to [7] for the coarse flow estimation.

3.1 Segmentation

For the experiments described here we have used a weak-membrane model of image brightness described in [8]. The goal is to reconstruct a piecewise smooth brightness image $\mathbf{i}$ given noisy data $\mathbf{d}$ by minimizing an objective function using a continuation method. Both spatial discontinuities and texture are treated as outlying measurements and rejected using analog outlier processes.

The approach is applied to the image in Figure 1*a*. Figure 1*b* shows the piecewise smooth reconstruction while Figure 1*c* shows the value of the spatial outlier processes (black indicates an outlier). The spatial outliers will be used for region segmentation at the medium level.

3.2 Coarse Optical Flow

Let $I(x, y, t)$ be the image brightness at a point (x, y) at time t and I_x, I_y, and I_t be the partial derivatives of I with respect to x, y, and t. To estimate the horizontal and vertical image velocity $\mathbf{u}(\mathbf{x}) = [u(\mathbf{x}), v(\mathbf{x})]^T$ at a point $\mathbf{x} = (x, y)$ we minimize an objective function $E_M(\mathbf{u})$ composed of data term and a spatial smoothness term [7]:

$$\sum_{\mathbf{x}} [\rho((\nabla I(\mathbf{x})\mathbf{u}(\mathbf{x}) + I_t(\mathbf{x})), \sigma_D) + \frac{1}{4} \sum_{\mathbf{z} \in \eta(\mathbf{x})} \rho(\|\mathbf{u}(\mathbf{x}) - \mathbf{u}(\mathbf{z})\|, \sigma_S)], \quad (1)$$

where $\nabla I = [I_x, I_y]$, $\eta(\mathbf{x})$ are the four nearest neighbors of $\mathbf{x}$ on the grid, and where ρ is a robust error norm. To obtain the coarse estimate, the values of the σ_* are chosen so that the objective function is convex and the function is minimized as described in [7]. A coarse to fine strategy, with warping between layers, is used to estimate large motions within the differential framework.

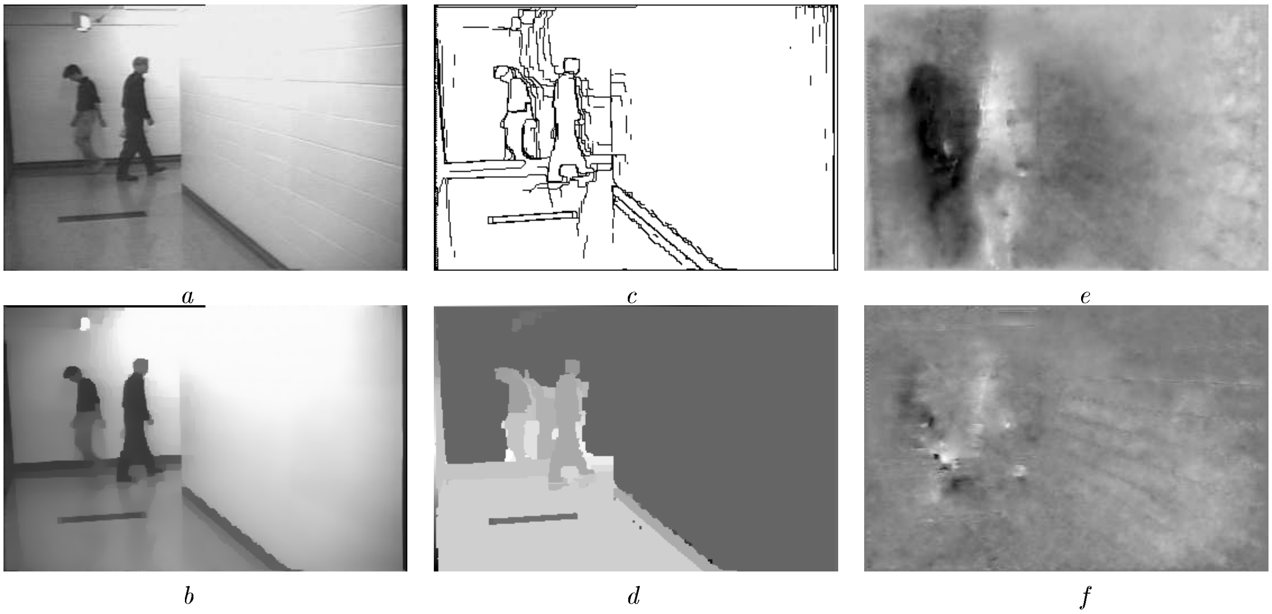

Figure 1: Hallway Sequence. *(a)* First image; *(b)* Piecewise smooth reconstruction; *(c)* Spatial outliers; *(d)* Connected component labels; *(e)* Horizontal component of flow (leftward motion = black; rightward motion = white); *(f)* Vertical component of flow (upward motion = black; downward motion = white).

Consider the hallway image sequence whose first image is shown in Figure 1a. The sequence contains two people walking in opposite directions while the camera is rotating and translating. The horizontal and vertical components of the coarse flow are shown in Figure 1e and 1f respectively. The results are very inaccurate and there is significant smoothing across motion boundaries.

4 Medium-Level Processing

The low-level processes described in the previous section are characterized by local processing and weak models of the scene. Medium-level processes can be seen as trying to find order and structure in the low-level data and, in doing so, impose more powerful models for interpreting the data.

We make a very simple hypothesis (which may be wrong) that regions of piecewise-smooth brightness in the image correspond to planar surfaces in the scene. The goal is to use information about image brightness to organize our interpretation of the motion in the scene. From the spatial outliers detected in the piecewise-smooth reconstruction of the image brightness (Figure 1c) we detect a set of connected regions R using a standard connected-components labeling algorithm. The connected components for the example image are shown in Figure 1d. These regions become our planar-surface hypotheses. Issues relating to under- and over-segmentation are addressed in Section 7.

4.1 Fitting Parametric Models to Flow Estimates

It is well known that the flow of a rigid planar region of the scene can be described by an eight-parameter model [1]. Using the notation from [3] let:

$$\begin{aligned} \mathbf{X}(\mathbf{x}) &= \begin{bmatrix} 1 & x & y & x^2 & xy & 0 & 0 & 0 \\ 0 & 0 & 0 & xy & y^2 & 1 & x & y \end{bmatrix} \\ \mathbf{a} &= \begin{bmatrix} a_0 & a_1 & a_2 & a_6 & a_7 & a_3 & a_4 & a_5 \end{bmatrix}^T. \end{aligned}$$

where the a_i are constants and where $\mathbf{u}(\mathbf{x}) = \mathbf{X}(\mathbf{x})\mathbf{a}$ is the flow at the image point $\mathbf{x} = (x, y)$.

To robustly estimate the motion $\mathbf{a}_r$ of a region $r \in R$ we minimize

$$\min_{\mathbf{a}_r} \sum_{\mathbf{x} \in r} \rho(\|\mathbf{X}(\mathbf{x})\mathbf{a}_r - \mathbf{u}_m(\mathbf{x})\|, \sigma), \qquad (2)$$

where $\mathbf{u}_m(\mathbf{x}) = [u_m(x, y), v_m(x, y)]^T$ is the coarse flow estimate and σ is a scale parameter. Since the coarse optical flow estimates are expected to have gross errors it is important that the estimation of the motion parameters be performed robustly. For this reason we take ρ to be a robust error norm with a redescending influence function. For the experiments in this paper the ρ-function was taken to be

$$\rho(x, \sigma) = \frac{x^2}{\sigma + x^2}. \qquad (3)$$

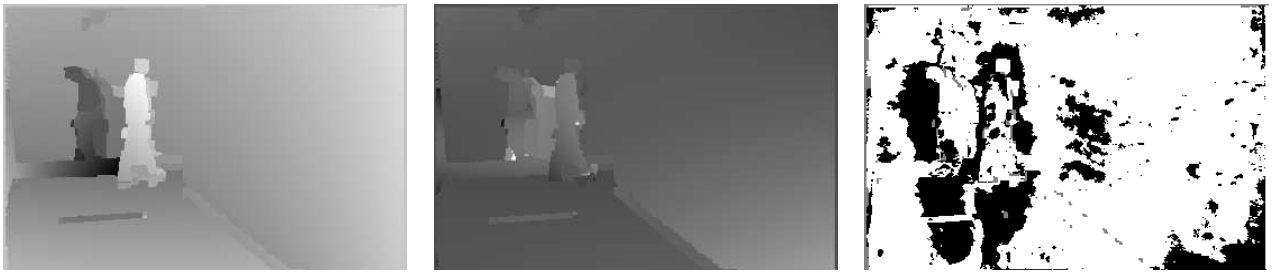

Figure 2: Hallway Sequence. *(a)* Horizontal component of flow; *(b)* Vertical component of flow; *(c)* Flow outliers.

Eqn. (2) is simply minimized using a continuation method in which σ starts at a high value and is gradually lowered. For each value of σ the objective function is minimized using one step of Newton's method. The effect of this is to track the solution while gradually rejecting outlying measurements.

The results of fitting the local parametric flow models to the coarse optical flow data for the Hallway image are shown in Figure 2. For regions smaller than a threshold, there is no attempt to fit a model to the flow and the coarse optical flow data is unchanged. For regions of size greater than or equal to 25, 500, or 1000 pixels we fit two-, six-, or eight-parameter flow models to the data respectively[1]. The parametric flow in these regions is then projected onto the image to produce the dense flow estimates in Figure 2. The results are a significant improvement over the coarse flow. Outlying coarse-flow vectors which are inconsistent with the parametric flow model are displayed in black in Figure 2*c*. The outliers predominantly correspond to regions where there was over-smoothing and to the reflections of the people in the floor.

4.2 Local Parametric Models of Image Motion

Fitting parametric models to the flow vectors in regions significantly improves the subjective quality of the flow field. Given the inaccuracy of the coarse flow estimates we would like to refine the motion estimates in each region by going back to the optical flow constraint equations at each pixel. The approach is a straightforward generalization of the approach described by Bergen *et al.* [3] for fitting a single global parametric motion to the entire image.

For each region $r \in R$ the brightness constancy assumption is

$$I(\mathbf{x}, t) = I(\mathbf{x} - \mathbf{u}(\mathbf{x}; \mathbf{a}_r), t+1) \quad \forall \mathbf{x} \in r$$

[1] Automatically choosing the right model for a given patch is addressed in Section 7.

where $\mathbf{u}(\mathbf{x}; \mathbf{a}_r) = \mathbf{X}(\mathbf{x})\mathbf{a}_r$ and $\mathbf{a}_r$ are the parameters for region r. Given the current fit $\mathbf{a}_r$ for a region we warp the image at time $t+1$ towards the image at time t. The original region at time t and this warped region are used to estimate the spatial and temporal derivaties I_x, I_y, and I_t. Let $\nabla I = [I_x, I_y]$, then to refine the current fit we minimize

$$\min_{\delta \mathbf{a}_r} \sum_{\mathbf{x} \in r} \rho((\nabla I(\mathbf{x}) X(\mathbf{x}) \delta \mathbf{a}_r + I_t(\mathbf{x})), \sigma), \tag{4}$$

and then the refined fit is $\mathbf{a}_r + \delta \mathbf{a}_r$.

To minimize Eqn. (4) we use exactly the same continuation method described above in which σ is gradually lowered and at each stage we apply one step in Newton's method. Since the initial flow estimates are fairly accurate we do not need to use a coarse-to-fine strategy as in [3].

The results of refining the flow for the hallway sequence are shown in Figures 3*a* and 3*b*. In this sequence it is difficult to see the improvement. Experiments in Section 6 with the synthetic Yosemite sequence illustrate the quantitative improvement that the region regression realizes.

One might ask "Why not start with this local-regression approach and ignore the coarse flow computation?" This approach will work for large, slow moving, regions. The problem with such an approach becomes apparent when trying to estimate the motion of a small region which is moving quickly. To deal with large motions using a differential technique, it is necessary to use a coarse-to-fine approach. But small regions may have little or no support at the coarse levels making it impossible to recover their motion.

5 Local Deformations

Local models of planarity are likely to be violated often in practice, particularly in natural scenes. For this reason we would like to use parametric models to provide a *coarse* description of the motion and allow *deformations* from the parametric model to account for errors

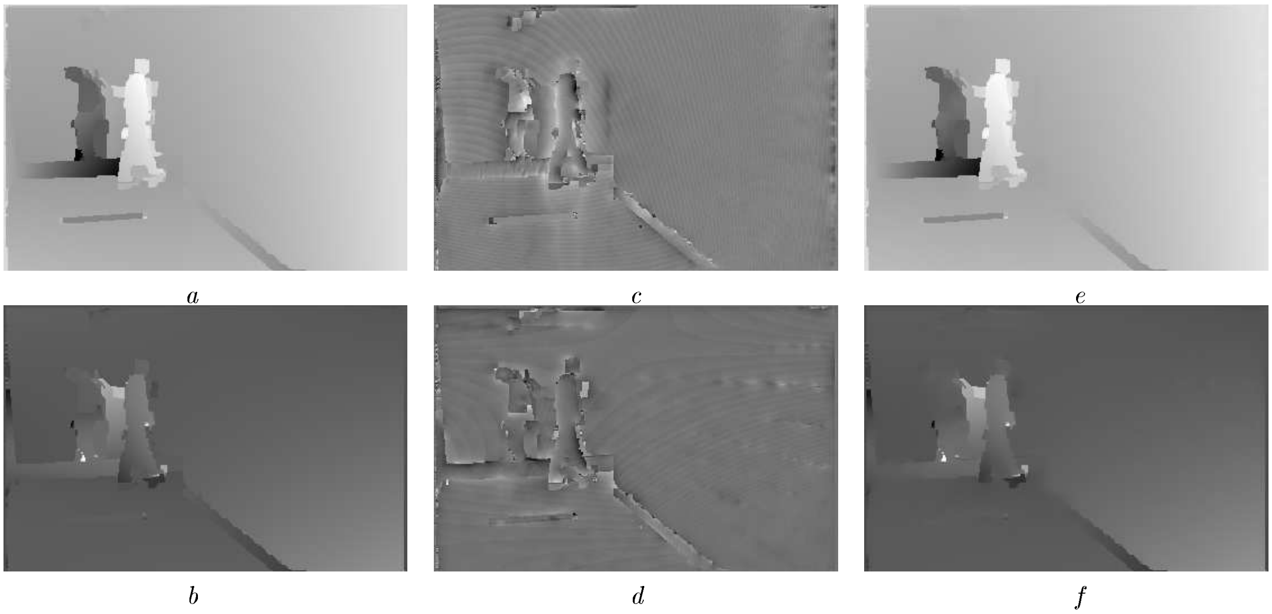

Figure 3: Hallway Sequence. *(a)* Refined flow: horizontal component; *(b)* Refined flow: vertical component of flow. *(c)* Horizontal flow deformation; *(d)* Vertical flow deformation. *(e)* Horizontal flow: Parametric plus deformations; *(f)* Vertical flow: Parametric plus deformations.

in the assumption. This is done by using the robust optical flow estimation technique described in [7] with the addition of a new term now coupling the flow estimate to the parametric-prediction of the flow. The flow estimate at each point can be thought of as being connected, via non-linear springs, to its neighbors, the data (optical flow constraint equation), and the motion of the planar-patch. The estimate is pulled by all these forces and the strength of the force is determined by the robust ρ-function. If the estimate gets pulled too far from its neighbors, the data, or the planar-patch estimate, the spring essentially goes "slack". This is equivalent to rejecting that measurement as an outlier.

Given the predicted flow in the planar patches, the image at time $t+1$ is warped back towards the image at time t to register them. The deformation $\delta\mathbf{u}$ is estimated to account for the discrepancy between the warped and original images. This physical model is implemented as the minimization of $E(\delta\mathbf{u}, \mathbf{u}, \mathbf{a})$ with respect to $\delta\mathbf{u}$:

$$
\begin{aligned}
\sum_{\mathbf{x}} [\rho((\nabla I(\mathbf{x})\delta\mathbf{u}(\mathbf{x}) + I_t(\mathbf{x})), \sigma_D) \\
+ \frac{1}{4} \sum_{\mathbf{z}\in\eta(\mathbf{x})} [\rho(\|(\mathbf{u}(\mathbf{x};\mathbf{a}) + \delta\mathbf{u}(\mathbf{x})) - \\
(\mathbf{u}(\mathbf{z};\mathbf{a}) + \delta\mathbf{u}(\mathbf{z}))\|, \sigma_S)] \\
+ \rho(\delta\mathbf{u}(\mathbf{x}), \sigma_M)], \qquad (5)
\end{aligned}
$$

where $\eta(\mathbf{x})$ are neighbors of $\mathbf{x}$, ρ is a robust error norm which rejects outlying measurements, $\mathbf{u}(\mathbf{x};\mathbf{a})$ is the refined flow from the medium-level processing, and where the spatial and temporal derivatives are computed with respect to the warped image.

The first term in E is a robust formulation of the standard optical flow constraint equation and enforces fidelity to the data. The second term pulls the deformation in a direction which minimizes the difference in the neighboring flow vectors. The final term forces the flow to be similar to the planar-patch estimate by penalizing for deformations.

Given the accurate initial estimate there is no need for a coarse-to-fine approach and in our experiments we simply minimize the objective function using Newton's method. A continuation method may be exploited using the scale parameters σ_* as was done in the previous section. We have not found this to be necessary since the estimates from the patches start the minimization near the global minimum.

The deformations from the refined parametric fit for the hallway sequence are shown in Figure 3c and 3d and the deformed flow is shown in Figure 3e and 3f. The most obvious improvement is in the body of the person on the left. Before deformation the body was split at the waist into two distinctly different planar motions. Local deformations bring the two motions into alignment.

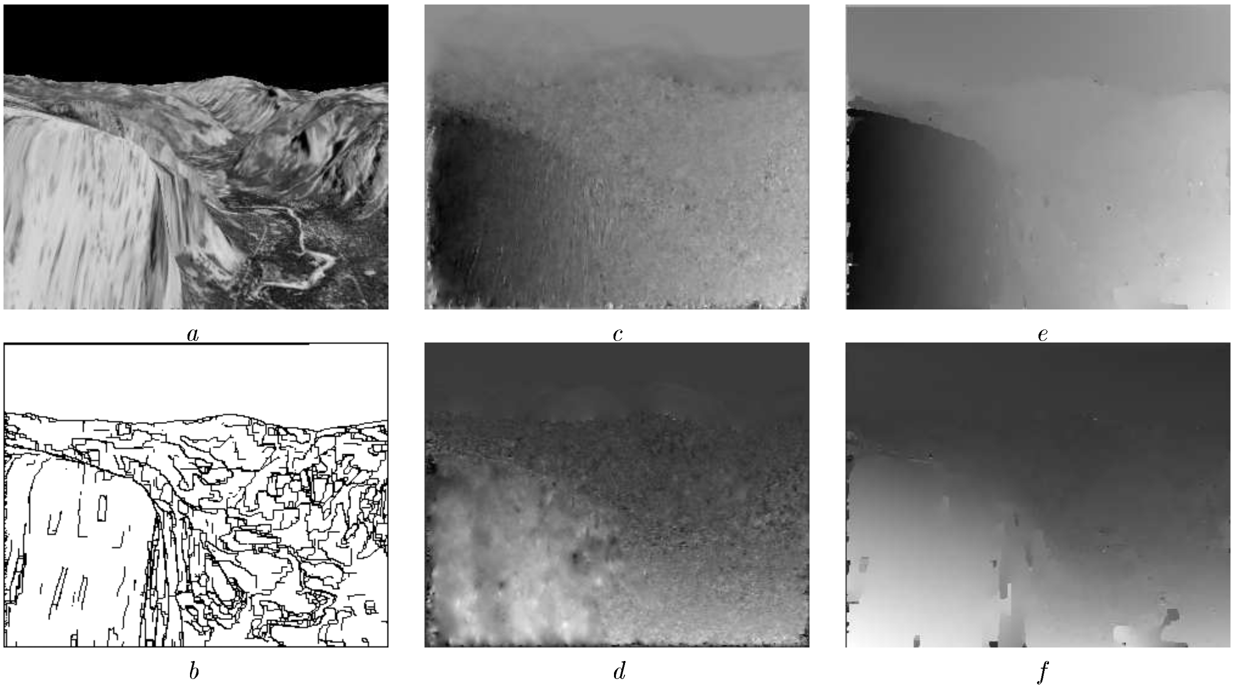

Figure 4: Yosemite sequence. *(a)* First image; *(b)* Spatial discontinuities; *(c)* Coarse horizontal displacement; *(d)* Coarse vertical displacement. *(c)* Final horizontal displacement; *(d)* Final vertical displacement.

	Average Error	Standard Deviation	Percent of flow vectors with error less than:				
			$< 1°$	$< 2°$	$< 3°$	$< 5°$	$< 10°$
Coarse:	8.0°	7.0°	3.6%	11.7%	21.4%	39.8%	72.6%
Parametric:	6.1°	4.6°	2.4%	10.8%	24.5%	53.5%	84.3%
Refined:	3.4°	3.2°	13.4%	36.5%	58.1%	81.5%	96.0%
Deformed:	2.5°	2.4°	21.5%	51.0%	71.6%	90.6%	98.5%

Table 1: Error results for the Yosemite fly-through sequence.

6 Experimental Results

The first experiment shows how the accuracy of the recovered flow is effected by each stage in the processing. The second experiment provides another illustration of the approach applied to scenes containing independently moving objects.

6.1 Yosemite Sequence

While the synthetic Yosemite image sequence[2] (Figure 4*a*) does not contain independently moving objects, it does allow the quantitative evaluation of the parametric+deformation model.

The edges detected in the piecewise smooth brightness image are shown in Figure 4*b*. The coarse flow estimates are shown in Figures 4*c* and 4*d*.

We quantify the accuracy of the results using the angular error measure of Barron *et al.* [2], with Table 1 showing the performance at each stage of the algorithm[3]. The accuracy of the initial coarse flow is quite poor. By fitting local parametric models to the coarse data some of the noisy estimates are removed and the mean accuracy of the flow improves but, given inaccurate estimates to start with, only a small percentage of the flow vectors achieve high accuracy. Refining the flow using local regression markedly improves the accuracy of the parametric fit, cutting the mean error nearly in half. Allowing local deformations to account for the coarse nature of the planar-patch approximation brings the mean angular error down to 2.5° with over half the flow vectors having angular errors less than 2.0°. These results are compared with other published results for the Yosemite sequence in Table 2.

[2] This sequence was generated by Lynn Quam.

[3] Error was not computed in the sky region since the Barron *et al.* images contained moving clouds and ours did not.

Technique	Average Error	Standard Deviation	Density
Horn & Schunck	32.43°	30.28°	100%
Anandan	15.84°	13.46°	100%
Singh	13.16°	12.07°	100%
Fleet & Jepson	4.17°	11.28°	34.1%
Weber & Malik [19]	3.42°	5.35°	45.2%
Black & Anandan [7]	4.47°	3.90°	100%
Black [4]	3.52°	3.25°	100%
Parametric+ Deformation	2.55°	2.40°	100%

Table 2: Comparison of various optical flow algorithms (adapted from [2]).

6.2 Walking Sequence

The second experiment shows another example of the approach applied to a sequence containing both camera motion and an independently moving object. Figure 5*b* shows the brightness discontinuities found in the first image of the sequence (Figure 5*a*). The coarse flow estimates are shown in Figures 5*c* and 5*d*. The final parametric+deformation results are shown in Figures 5*e* and 5*f*; the parameter settings for this sequence are the same as for the other examples.

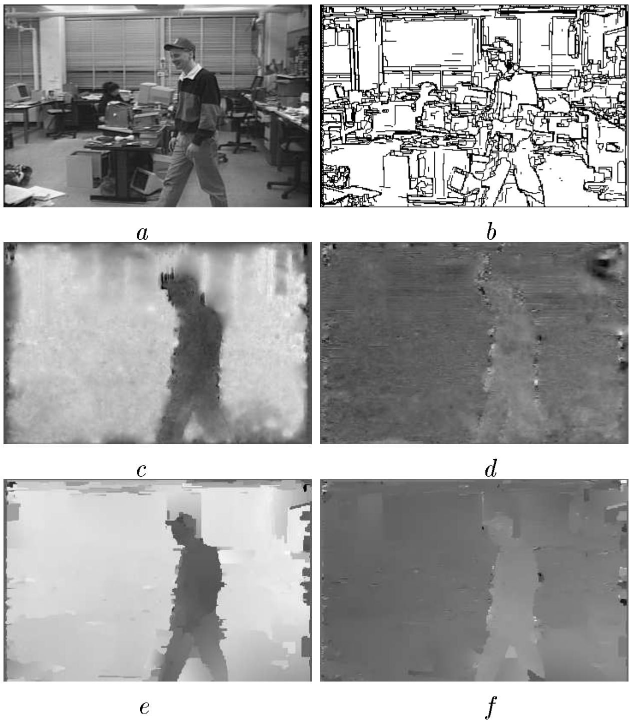

Figure 5: A cluttered scene in which the camera is panning and a person is walking. *(a)* First image; *(b)* Spatial discontinuities; *(c)* Coarse horizontal displacement; *(d)* Coarse vertical displacement; *(e)* Planar+Deformations: horizontal flow; *(f)* Planar+Deformations: vertical flow;

7 Open Questions

Due to over-segmentation based on brightness, the localization of objects, as opposed to surfaces patches, may require grouping patches together based on common motion. To some extent the local deformations provide this merging at a low level. It may be desirable to re-examine the segmented regions based on the deformed flow and group them into larger regions. Additionally, having the motion of segmented image regions means that the occlusion relationships between the regions can be analyzed. Moreover, it may be possible to incorporate a layered representation which can represent occluded portions of regions.

Under-segmentation is also an issue. For example, in our experiments with moving people, their legs are often segmented into a single region based on brightness. We would like to be able to detect that a single motion does not give a good fit to this region and break it into parts in the appropriate places. One possibility is to use the local deformation as a measure of strain and introduce breaks when the strain is too great. An alternative way to cope with undersegmentation is to allow multiple motions within a region and use either a robust estimation approach [5] or a mixture model approach [11] to recover the multiple motions.

Another issue which was not addressed in this paper is how to decide what is the right motion model to use for a particular region. The choice may be important to prevent over- or under-fitting the data. A common approach in the surface fitting literature is to start with simple surface models and, if their fit is poor, to fit increasingly complex models. Additionally, there is a tradeoff between model complexity and deformation; for example, complex motions can be accounted for either by using more expressive models or by allowing greater deformation. The detection of large deformations (or strains) over a patch might suggest the need to refit with a higher-order model; this issue deserves further study.

8 Conclusion

This paper has presented a new model for optical flow based on the motion of planar regions plus local deformations. The approach exploits brightness information to help organize and constrain the interpretation of the motion by using segmented regions of piecewise smooth brightness to hypothesize planar regions in the scene. Parametric flow models are fitted to these regions in a two step process which first computes a coarse fit and then refines it using a generalization of the standard area-based regression approaches. Since the planar-patch assumption is likely to be violated, we allow local deformations from the parametric flow using a physically-based model in which a regularized optical flow estimate is partially constrained by the parametric motion estimate.

Acknowledgements

We would like to thank David Fleet for discussions that helped to clarify this work.

References

[1] G. Adiv. Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(4):384–401, July 1985.

[2] J.L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. Technical Report Report No. 299, University of Western Ontario, July 1992.

[3] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In G. Sandini, editor, *Proc. of Second European Conference on Computer Vision, ECCV-92*, volume 588 of *LNCS-Series*, pages 237–252. Springer-Verlag, May 1992.

[4] M. Black. Recursive non-linear estimation of discontinuous flow fields. In J. Eklundh, editor, *European Conf. on Computer Vision, ECCV-94*, volume 800 of *LNCS-Series*, pages 138–145, Stockholm, Sweden, 1994. Springer-Verlag.

[5] M. Black and P. Anandan. The robust estimation of multiple motions: Affine and piecewise-smooth flow fields. Technical Report P93-00104, Xerox PARC, December 1993.

[6] M. J. Black. Combining intensity and motion for incremental segmentation and tracking over long image sequences. In G. Sandini, editor, *Proc. of Second European Conference on Computer Vision, ECCV-92*, volume 588 of *LNCS-Series*, pages 485–493. Springer-Verlag, May 1992.

[7] M. J. Black and P. Anandan. A framework for the robust estimation of optical flow. In *Proc. Int. Conf. on Computer Vision, ICCV-93*, pages 231–236, Berlin, Germany, May 1993.

[8] M. J. Black and A. Rangarajan. The outlier process: Unifying line processes and robust statistics. In *Proc. Computer Vision and Pattern Recognition, CVPR-94*, pages 15–22, Seattle, WA, June 1994.

[9] F. Heitz and P. Bouthemy. Multimodal motion estimation and segmentation using Markov random fields. In *Proc. IEEE Int. Conf. on Pattern Recognition*, pages 378–383, June 1990.

[10] M. Irani, B. Rousso, and S. Peleg. Detecting and tracking multiple moving objects using temporal integration. In G. Sandini, editor, *Proc. of Second European Conference on Computer Vision, ECCV-92*, volume 588 of *LNCS-Series*, pages 282–287. Springer-Verlag, May 1992.

[11] A. Jepson and M. J. Black. Mixture models for optical flow computation. In *Proc. Computer Vision and Pattern Recognition, CVPR-93*, pages 760–761, New York, June 1993.

[12] R. Koch. Automatic reconstruction of buildings from stereoscopic image sequences. *EUROGRPHICS'93*, 12(3):339–350, 1993.

[13] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. 7th IJCAI*, pages 674–679, Vancouver, B. C., Canada, 1981.

[14] W. Luo and H. Maître. Using surface model to correct and fit disparity data in stereo vision. In *Proc. IEEE Int. Conf. on Pattern Recognition*, pages 60–64, June 1990.

[15] D. Terzopoulos and D. Metaxas. Dynamic 3D models with local and global deformations: Deformable superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):703–714, July 1991.

[16] W. B. Thompson. Combining motion and contrast for segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2:543–549, 1980.

[17] J. Y. A. Wang and E. H. Adelson. Layered representation for motion analysis. In *Proc. Computer Vision and Pattern Recognition, CVPR-93*, pages 361–366, New York, June 1993.

[18] A. Waxman. An image flow paradigm. In *Proc. IEEE Workshop on Computer Vision: Representation and Control*, pages 49–55, Annapolis, MD, 1984.

[19] J. Weber and J. Malik. Robust computation of optical flow in a mulit-scale differential framework. In *Proc. Int. Conf. on Computer Vision, ICCV-93*, pages 12–20, Berlin, Germany, May 1993.