Qualitative Description of Camera Motion from Histograms of Normal Flow

Zoran Duric
Department of Computer Science
George Mason University
Fairfax, VA 22030
zduric@cs.gmu.edu

Ehud Rivlin
Department of Computer Science
Technion
Haifa, Israel 32000
ehudr@cs.technion.ac.il

Azriel Rosenfeld
Center for Automation Research
University of Maryland
College Park, MD 20742-3275
ar@cfar.umd.edu

Abstract

If we histogram the normal flow vectors in images of a scene viewed by a moving observer, we can use the time-varying histogram to derive qualitative information about the observer's motion—for example, whether it is (primarily) translational or rotational, and whether the direction of translation or axis of rotation is (roughly) parallel or perpendicular to the camera axis. This is illustrated using flow histograms obtained from a variety of real image sequences.

1 Introduction

A moving agent that obtains a sequence of images of a stationary scene can infer both its motion and the layout of the scene, up to a range/speed ambiguity, by analyzing the images. The classical "structure from motion" problem attempts to determine the motion and layout completely (except for the ambiguity). However, as is well known, it is quite difficult to obtain this complete information accurately. It may therefore be useful to attempt to derive only partial information.

A number of researchers have shown how to derive partial information about the agent's motion or the scene structure from images. Examples include relative distances of objects; time to collision with (or rate of approach toward) an object; and qualitative information about an object's shape. We show in this paper that a moving agent can obtain basic information about its motion by examining the sequence of two-dimensional histograms of normal flow vectors computed from successive pairs of images. (A related idea is the use of Hough transform methods in flow analysis (Kalvainen, et al. 1992; Bober and Kittler, 1994; Heikkonen, 1995).) Specifically, we show, using a variety of real image sequences, that these histogram sequences are quite different for different types of simple motions: translation along the camera axis, translation in a plane perpendicular to the camera axis, rotation around the camera axis, and rotation around an axis perpendicular to the camera axis. Obviously, combinations of these motions give rise to more complicated histograms; but it is well known that humans find such combined motions difficult to interpret. In any case, methods of selective stabilization (Duric and Rosenfeld, 1996; Yao and Chellappa, 1997) can be used to eliminate "unwanted" components—in particular, rotational components—from an image sequence.

[1]The support of the Office of Naval Research under Contract N00014-95-1-0521 is gratefully acknowledged, as is the help of Janice Perrone in preparing this paper. The authors also thank Brad Stuart for providing the panning sequence.

2 Flow Computation and Histogramming

The instantaneous velocity of the image point (x, y) under perspective projection is given by

$$\dot{x} = \frac{-Uf + xW}{Z} + \omega_x \frac{xy}{f} - \omega_y \left(\frac{x^2}{f} + f\right) + \omega_z y, \quad (1)$$

$$\dot{y} = \frac{-Vf + yW}{Z} + \omega_x \left(\frac{y^2}{f} + f\right) - \omega_y \frac{xy}{f} - \omega_z x. \quad (2)$$

where $\vec{T} = (U \ \ V \ \ W)^T$ is the translational velocity and $\vec{\omega} = (\omega_x \ \ \omega_y \ \ \omega_z)^T$ is the rotational velocity of the

0-7695-0750-6/00 $10.00

camera.

Let $\vec{\imath}$ and $\vec{\jmath}$ be the unit vectors in the x and y directions, respectively; $\dot{\vec{r}} = \dot{x}\vec{\imath} + \dot{y}\vec{\jmath}$ is the projected motion field at the point $\vec{r} = x\vec{\imath} + y\vec{\jmath}$. If we choose a unit direction vector $\vec{n}_r = n_x\vec{\imath} + n_y\vec{\jmath}$ at the image point $\vec{r}$ and call it the normal direction, then the *normal motion field* at $\vec{r}$ is $\dot{\vec{r}}_n = (\dot{\vec{r}} \cdot \vec{n}_r)\vec{n}_r = (n_x\dot{x} + n_y\dot{y})\vec{n}_r$. $\vec{n}_r$ can be chosen in various ways; the usual choice (and the one that we use) is the direction of the image intensity gradient $\vec{n}_r = \nabla I / \|\nabla I\|$.

Note that the normal motion field along an edge is orthogonal to the edge direction. Thus, if at the time t we observe an edge element at the position $\vec{r}$ the apparent position of the same edge element will at the time $t + \Delta t$ be $\vec{r} + \Delta t \dot{\vec{r}}_n$. This is a consequence of the well known *aperture problem.* We base our method of estimating the normal motion field on this observation. For an image frame (collected at the time t_0) we find the edges using an implementation of the Canny edge detector. For each edge element at $\vec{r}$ we resample the image locally to obtain a small window with its rows parallel to the image gradient direction $\vec{n}_r = \nabla I / \|\nabla I\|$. For the next image frame (collected at time $t_0 + \Delta t$) we create a larger window (typically twice as large as the maximum expected value of the magnitude of the normal motion field). We then slide the first (smaller) window along the second (larger) window and compute the difference between the image intensities. The zero of the resulting function is at the distance u_n from the origin of the second window; note that the image gradient in the second window at the positions close to u_n must be positive. Our estimate of the normal motion field is then $-u_n$ and we call it the *normal flow.*

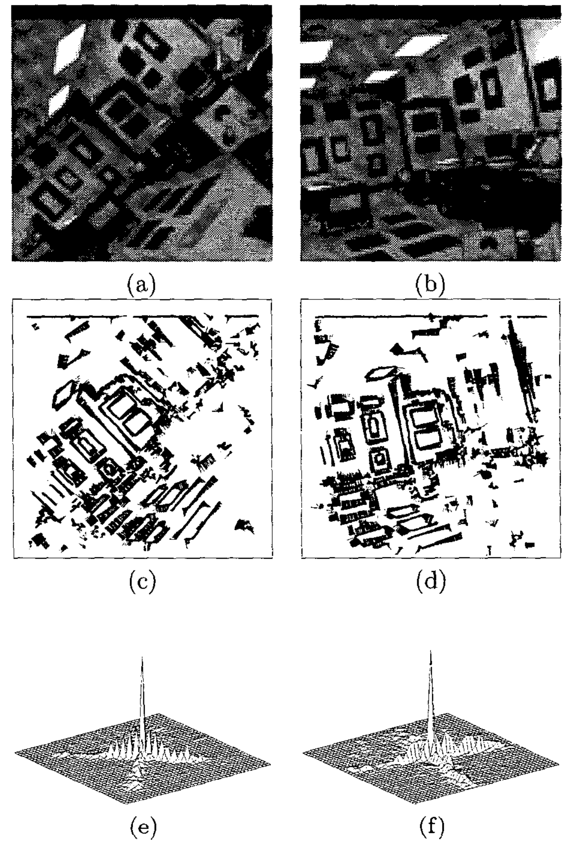

Figure 1. If the motion is z-axis rotation, the flow histogram rotates. (a-b) Frames 1 and 10 of the "Robot" sequence. (c-d) Flow fields for these frames. (e-f) Histograms of these flow fields.

3 Flow Histograms for Simple Motions

In this section we show examples of flow histogram sequences that result from four simple types of camera motion: z-axis rotation, z-axis translation, lateral translation, and pan. (We follow the usual convention that the z-axis is the optical axis of the camera; lateral translation is translation along an axis in the xy-plane, and pan is rotation around such an axis.) We also show flow histograms obtained from a forward-pointing camera carried by a ground vehicle moving on unpaved terrain; when the image sequence is not stabilized, the flow shows a mixture of effects due to translation, roll, and pitch, but stabilization can be used to remove the rotation effects.

Figures 1a-b show two frames from the "Robot" sequence, taken at the University of Massachusets; here the motion is essentially clockwise z-axis rotation. This motion results in counterclockwise image rotation around the center of the image plane; at each point of the image, the direction of motion is tangential, and its magnitude increases with distance from the center. If the scene contained edges oriented in all directions, the resulting normal flow would contain vectors of all magnitudes in all directions; but since most of the edges in the scene in Figure 1 are oriented in two perpendicular directions, most of the normal flow vectors are perpendicular to those directions (Figures 1c-d). This results in a +-shaped flow histogram (Figures 1e-f); as the camera rotates, this histogram rotates (since the edges are revolving around the image center).

Figures 2a-b show two frames of the "Coke can" sequence, taken at NASA Ames Research Center; here the motion is essentially z-axis translation toward the scene. This results in an expansion of the image; at each point of the image, the direction of motion is ra-

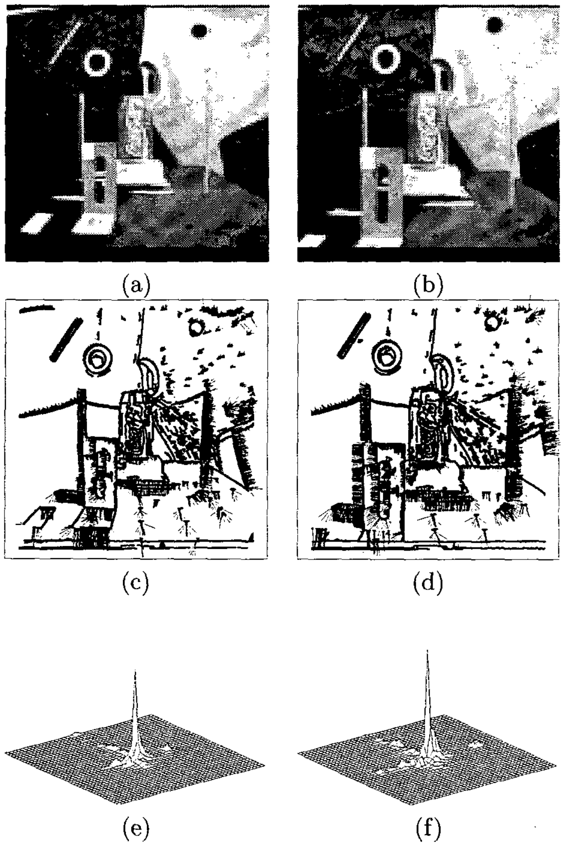

Figure 2. If the motion is z-axis translation, the flow histogram expands. (a-b) Frames 0 and 75 of the NASA Ames "Coke can" sequence. (c-d) Flow fields for these frames. (e-f) Histograms of these flow fields.

dial, and its magnitude increases with distance from the center (and with closeness of the scene point to the camera). If the scene contained edges oriented in all directions, the resulting normal flow would contain vectors of all magnitudes in all directions; but since most of the edges in the scene in Figure 2 are oriented horizontally or vertically, most of the normal flow vectors are horizontal or vertical (Figures 2d-f), once again resulting in a +-shaped flow histogram (Figures 2g-i). As the camera approaches the scene, the flow magnitude increases, so that the flow histogram expands.

Figure 3a shows one frame from the SRI "Tree" sequence, in which the camera is pointing forward and translating to the left. This results in rightward horizontal image flow whose magnitude increases with closeness of the scene point to the camera. Since the scene contains many near-vertical edges (the tree trunks), the normal flow is predominantly rightward

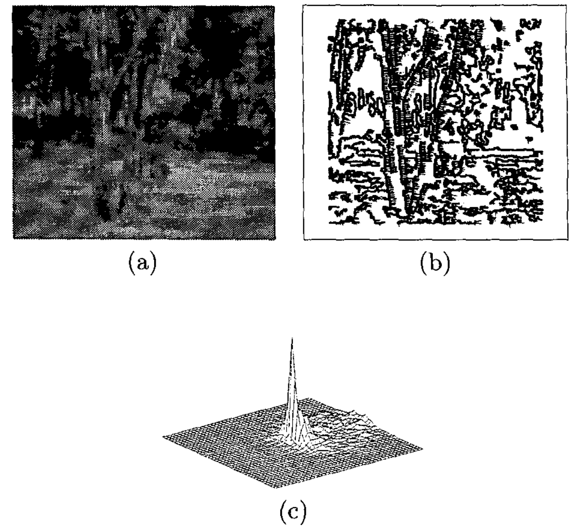

Figure 3. If the motion is lateral translation, the flow histogram is biased in the direction of motion. (a) Frame 0 of the SRI "Tree" sequence. (b) Flow field for this frame. (c) Histogram of the flow field.

(Figures 3b and c).

Figure 4a shows one frame from a panning sequence taken in our laboratory. Since the camera motion was leftward (counterclockwise around a vertical axis, as seen from above), the image flow is predominatly rightward (Figure 4b). The scene contains edges in many directions; for an image edge that makes an angle θ with the vertical, the magnitude of the normal flow is proportional to cos θ, and its direction is perpendicular to θ, so that the normal flow histogram displays a circular "crater" through the origin (Figure 4c).

Figures 5a-b show two frames from an image sequence taken by a forward-pointing camera mounted on a ground vehicle moving across rough terrain. Figures 5e,i,m show the normal flow at three intermediate frames of this sequence. Here the backward image flow due to the forward translation appears primarily in the lower part of the image, which shows nearby parts of the terrain. In the upper part of the image, showing terrain near the horizon, the translation produces negligible flow, but the rotational effects of the bumpy motion are quite apparent. In Figure 5e the vehicle is pitching downward, resulting in upward image motion along the horizon. In Figure 5i the pitching is upward, resulting in downward image motion at the horizon. Finally, in Figure 5m the vehicle is pitching downward and rolling counterclockwise; the roll causes the left side of the horizon image to move upward and the right side to more downward, while

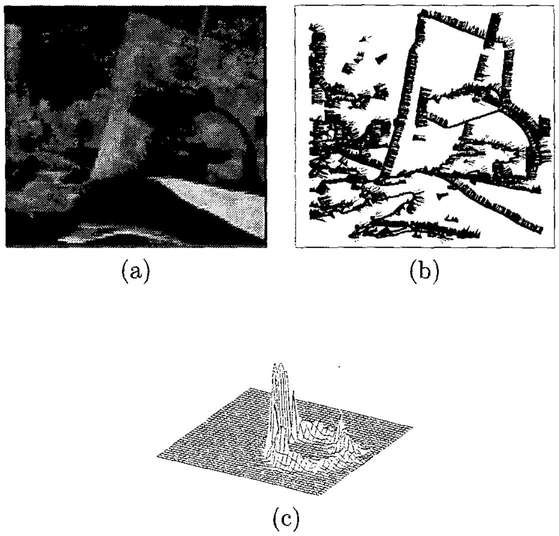

Figure 4. If the motion is panning, the image flow is approximately constant, so the normal flow at an edge of slope θ has magnitude proportional to cos θ; thus the flow histogram has a "crater". (a) A frame of a panning sequence obtained in our laboratory. (b) Flow field for this frame. (c) Histogram of the flow field.

the pitch causes the entire horizon image to move upward; at the left of the image these effects add, but at the right they cancel. All of these effects are apparent in the normal flow and the histograms (Figures 5f,j,n). The image sequence can be stabilized to eliminate the rotational effects, by detecting the horizon and warping the images so that the horizon remains stationary. Figures 5c,d show the same frames as in Figures 5a,b after stabilization; note the loss of portions of the image near the edges due to the stabilization. Figures 5g,h,k,l,o,p show the normal flow and the histograms corresponding to Figures 5e,f,i,j,m,n but for the stabilized image sequence; the rotational flow has been eliminated, leaving only the backward image motion due to the forward translation. Additional image sequences can be viewed at the following url: *http://www.cs.gmu.edu/~zduric/WebPages/Histograms.html*

4 Concluding Remarks

Normal flow histograms do not seem to have been used extensively in image sequence analysis. We have seen in this paper that they can provide useful qualitative information about the camera motion. Evidently they have many other uses; for example, when the motion is translational, they can provide useful information about scene depths (Jasinchi et al., 1991); when a stationary camera views a scene containing many moving objects, a normal flow histogram can provide useful information about the (apparent) velocities of the objects. The authors hope that this paper will serve to call attention to the value of flow histograms, and will stimulate further study of their uses.

References

[1] M. Bober and J. Kittler, Estimation of complex multimodal motion: An approach based on robust statistics and Hough transform, *Image Vision Computing* **12**, 1994, 661–668.

[2] Z. Duric and A. Rosenfeld, Image sequence stabilization in real time, *Real Time Imaging* **2**, 1996, 271–284.

[3] J. Heikkonen, Recovering 3-D motion parameters from optical flow field using randomized Hough transform, *Pattern Recognition Letters* **16**, 1995, 971–978.

[4] R. S. Jasinschi, A. Rosenfeld, P. Cucka, and K. Sumi, Discriminating 3D texture patterns: The velocity histogram method, CS-TR-2644, Center for Automation Research, University of Maryland, College Park, MD, 1991. (Published in part as R. Jasinschi, A. Rosenfeld, and K. Sumi, Perceptual motion transparency: The role of geometrical information, *J. Optical Society of America* **A9**, 1992, 1865–1879.)

[5] H. Kalvainen, E. Oja, and L. Xu, Randomized Hough transform applied to translational and rotational motion analysis, in Proc. Intl. Conf. on Pattern Recognition, Vol. 1, 1992, 672–675.

[6] Y.S. Yao and R. Chellappa, Selective stabilization of images acquired by unmanned ground vehicles, *IEEE Trans. Robotics Automation* **13**, 1997, 693–708.

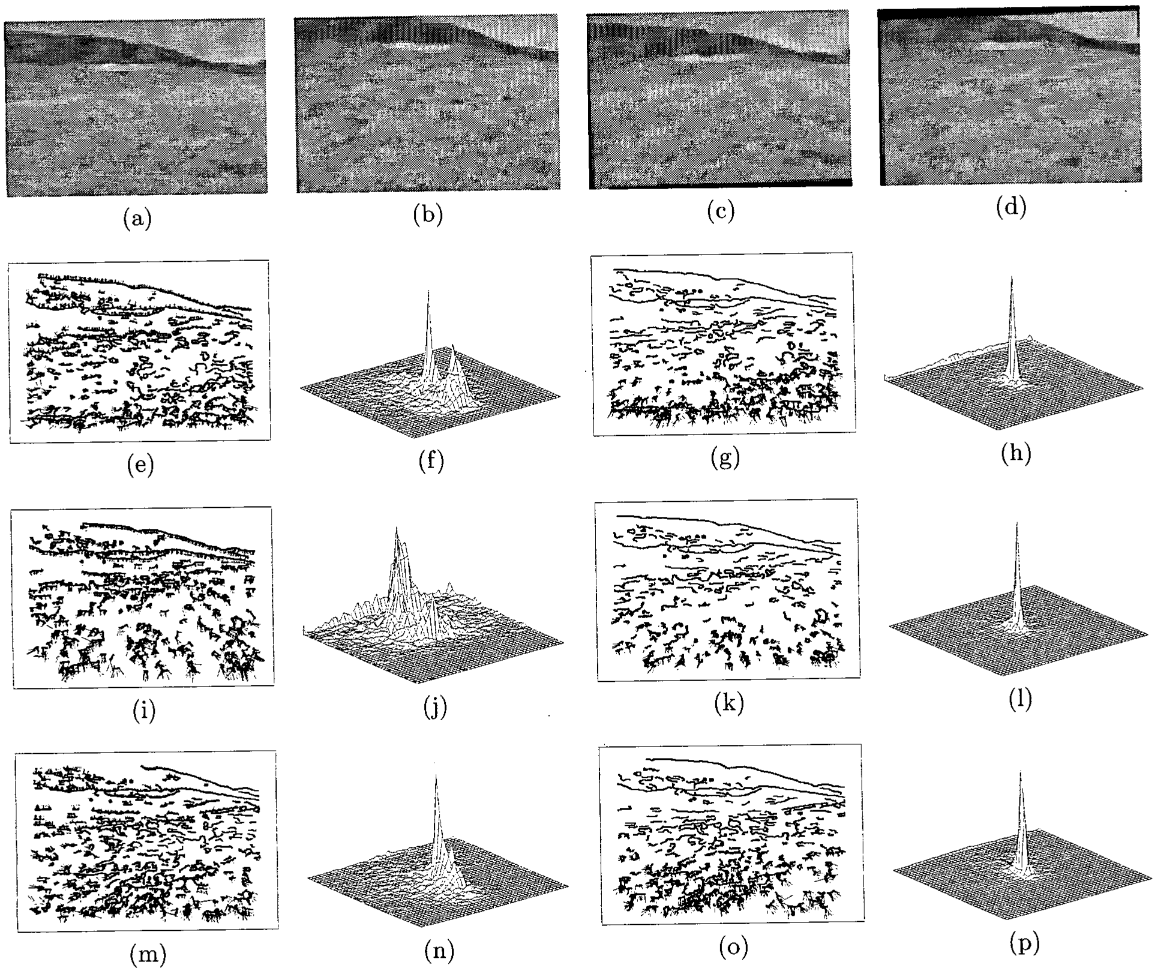

Figure 5. Mixed time-varying motions result in more complicated histograms, but stabilization can simplify them. (a-b) Two frames from an unstabilized image sequence obtained by a ground vehicle on rough terrain. (e,i,m) Flow fields for three frames intermediate between (a) and (b). The flow due to forward translation appears primarily in the lower (nearby) parts; the upper (distant) parts show the effects of (e) downward pitch, (i) upward pitch, and (m) counterclockwise roll combined with downward pitch. (f,j,n) Histograms of these flow fields. (c-d) Corresponding images, (g,k,o) flow fields, and (h,l,p) histograms after stabilization; the pitch and roll effects have been eliminated.