

A Fast Object Flow Estimation Method based on Spacetime Image Analysis

Kenji Mase, Atsushi Sato, Yasuhito Suenaga and Ken-ichiro Ishii

NTT Human Interface Laboratories

Nippon Telegraph and Telephone Corporation

1-2356, Take, Yokosuka-Shi, Kanagawa 238-03, Japan

E-mail: mase@nttcvg.ntt.jp

Abstract

This paper presents a fast object flow estimation method based on spacetime image analysis. We introduce the *Ortho-sectioning* method which uses the object's geometry on two slices of spacetime volume to estimate the object's flow. The one slice image method which can estimate binary flow direction is also described. An experimental realtime system using the method is built on a workstation and experimental results are shown.

1 Introduction

The volume of automobile and pedestrian traffic provides important information for surveillance, market planning and traffic control. Periodical counts of visitors and customers at convention halls or department stores can assist in the planning of events and business decision making. If automatic pedestrian counting is possible with a TV-camera and image processing, the sensor can be very small.

In this paper, we newly point out the fact that the projection of a moving object on the *non-epipolar* spacetime plane contains the temporal and shape information of the object. We derive the *ortho-sectioning method* which estimates object flow from two orthogonal sectioning plane images of the spacetime volume. The above problems are transformed to two dimensional region analysis, the only constraint is that counting is done only at the sampling slit line which is used to form one of spacetime image. This constraint reduces the computational complexity and difficulty of counting moving objects in a sense.

Zheng and Tsuji introduced the "Dynamic Projection(DP)" of a scene, which is an example of non-epipolar plane image[1]. The projection is based on the concepts of "Anorthoscopic Perception[2]." They did not describe the relation between DP and object's velocity in the paper. Baker and Bolles introduced generalized EPIs[3]. The intermediate image data produced by arbitrary camera motion could be equated to non-EPI with the generalization. However, they did not utilize the shape information of generalized EPIs, but transformed the images to construct linear EPIs with the help of a known camera path. Several researchers have constructed spacetime surfaces of moving objects. They computed the flow on the surface from the partial derivatives of the surface and by examining features on the surface[3, 4, 5]. Peng and Medioni[4] introduced the slice analysis of the spacetime volume. Our method is akin to theirs. However, they computed only the normal component of motion. Moreover, the point-wise flow field was recovered by post-processing.

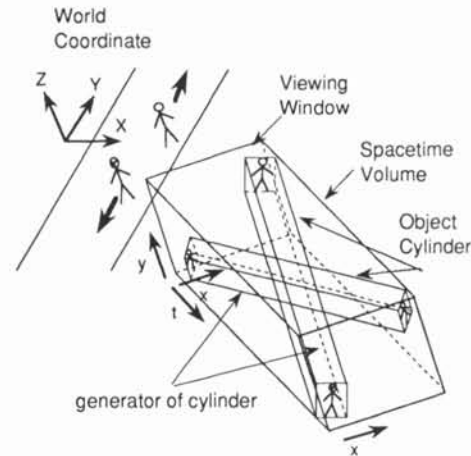


Figure1: Spacetime Representation of Image Sequence

If objects can be extracted as regions from the non-EPI, we only need to count the number of regions to find the number of objects passing through the sampling slit line. This frees us from the identification and tracking problems. Hwang and Takaba[6] used this idea to build a real time image processing system that counted pedestrians. They counted the number of regions on the image as the number of passing pedestrians. However, the fact that the skewed region contains flow information was overlooked. Thus, they counted the number of object without detecting the moving direction.

In this paper, we define ortho-sectioning of the spacetime image volume. Equations of object flow in terms of geometric features of two object projections on two ortho-sectioning planes are proposed. The detection method of binary direction is then presented as the one-plane method. Object flow estimation results are shown using a computer generated moving object and a real image sequence of walking pedestrians.

2 Flow from Ortho-Sectioning

2.1 Ortho-Sectioning of Spacetime Image

Spacetime image representations can be constructed in various ways. We consider the case of viewing moving objects with a stationary camera without loss of generality. Therefore, our spacetime image representation is a 3D volume, in which each object constructs a generalized cylindrical volumetric region within the background. The 3D volume has two spatial axes and one temporal axis as illustrated in Figure 1. A sectioned

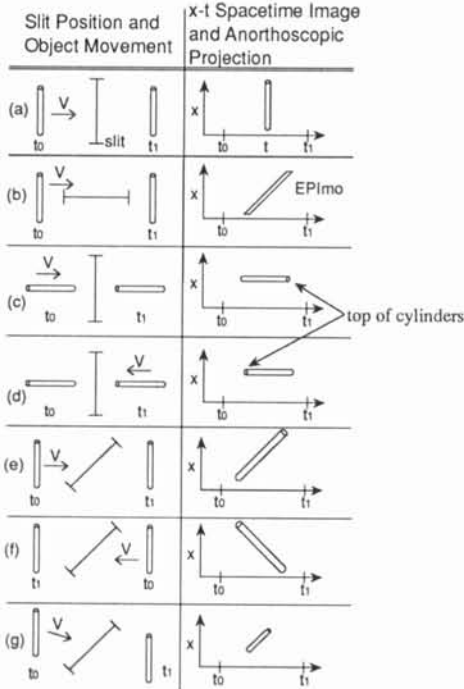


Figure 2: Slit Alignment and Anortho-Projection

plane image which is a slice of the 3D volume contains various kinds of information.

First, when the sectioning plane includes (i.e. is parallel to) both temporal axis and the flow of the moving object, the slice is called an Epipolar Plane Image (EPI_{mo}) in an analogy to the epipolar plane image (EPI) defined in motion stereo analysis. On the contrary, if the sectioning plane is skewed against the generator of the cylinder, the resulting image can contain both shape and time information. We call such an image the non-Epipolar Plane Image of moving object (non-EPI_{mo}, or non-EPI in this paper). If the object is rigid and its orientation against camera is fixed, its projection on non-EPI takes the form of a skewed duplicate of the original object projection against the camera view. Even in the case of non-rigid objects, the non-EPI projection reflects much of the original shape if the camera angle is adequately set. We call this projection the *Anorthoscopic Projection (anortho-projection or skewed projection)* using the analogy of anorthoscopic perception[2].

Since the anortho-projection of an object has both shape and flow information, we can use two anortho-projections to eliminate the ambiguity between spatial and temporal information. We use two orthogonal non-epipolar sectioning planes (*ortho-sections*) to get two anortho-projections of each object for flow estimation; this is called the *ortho-sectioning method*. The non-epipolar ortho-sections can be arbitrarily placed. We use a special and efficient configuration of ortho-sections; the usual x-y plane (camera view plane) as one section, and a spacetime section perpendicular to the x-y plane. With no loss of generality, the camera can be rotated around the optical axis, we define the latter plane to be the *x-t spacetime image* created by a sampling slit parallel to the x-axis.

The alignment of the sampling slit line against the

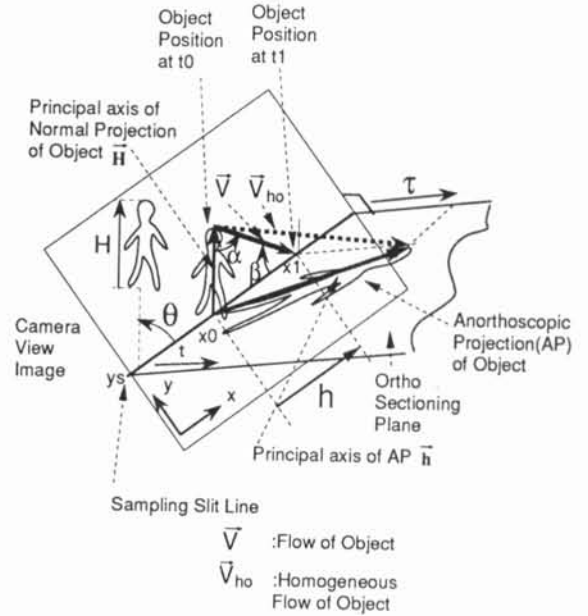


Figure 3: Object Flow by Ortho-Sectioning Method

moving object is the key to extract object flow. Figure 2 illustrates the anortho-projections of a simple object on the x-t spacetime image with various alignments. The example given in figure 2-(b) is equivalent to the EPI_{mo} case; the object flow is parallel to the sampling slit. Other examples show that anortho-projection can retain shape information to some extent. Although the directional information of flow merely consists of the appearance of the cylinder top in examples 2-(c) and 2-(d), examples 2-(e)-(g) shows that simple directional information (i.e. left or right) can be recovered very easily from the angle of anortho-projection when the sampling slit line is skewed against the object flow. The orientation of object flow, the sampling slit line and the object are important and will be discussed in the following section.

2.2 Object Flow from Ortho-Sections

The special configuration of ortho-sections, the camera view and x-t spacetime plane, is used to estimate object flow in this section. The object's geometries in the normal projection and the anortho-projection are also used in the estimation. Figure 3 shows the ortho-sectioning of an object O moving with flow V . We assume that the normal projection of the object has a primal axis which is skewed by θ ($0 \leq \theta < \pi$) against the sampling slit line and that the object's length is H . We regard the primal axis as a vector of finite length in this paper. Suppose the object O moves across the slit taking time t_0 to t_1 . The duration $\tau = t_1 - t_0$. The movement of the object generates an anortho-projection O' which has the dimensions of width h and height τ on the x-t spacetime image. The width h is the length between the point (x_0, y_s) where the object begins traversing the slit and (x_1, y_s) where the traverse finishes ($h = x_1 - x_0$). $y = y_s$ is the position of the sampling slit line. We define the *object flow*, V , by the following equations.

$$V \equiv (V_x, V_y) = \frac{1}{\tau}(h - H \cos \theta, -H \sin \theta), \quad (1)$$

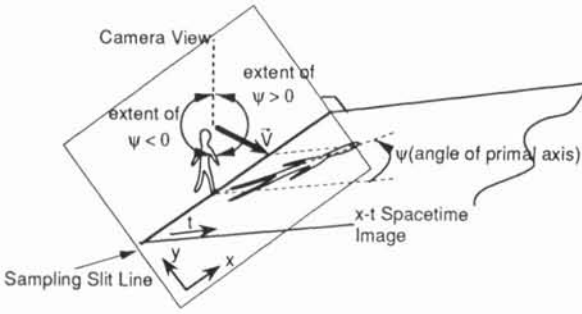


Figure4: Binary Direction from Anortho-Projection

or by a homogeneous representation;

$$\mathbf{V}_{ho} \equiv (u, v, w) \equiv \mathbf{h} - \mathbf{H} = (h - H \cos \theta, -H \sin \theta, \tau), \quad (2)$$

where $\mathbf{H} \equiv (H \cos \theta, H \sin \theta, 0)$ and $\mathbf{h} \equiv (h, 0, \tau)$ are the primal axes of normal projection and anortho-projection of the object, respectively. Also, $(V_x, V_y) = (u/w, v/w)$.

As a matter of fact, the object flow is the transposition vector from the primal axis of normal projection \mathbf{H} to the primal axis of anortho-projection \mathbf{h} . This is naturally understood if we regard the vector \mathbf{V}_{ho} as the trace of a object tip from t_0 to t_1 in the spacetime volume. Therefore, the estimation accuracy depends on how well the primal axis of projection reflects the primal axis of the object. The dimensions of anortho-projection on x-t spacetime image can be expressed by the following equations which constrain computation possibility and the sampling slit alignment;

$$h = \frac{|\mathbf{H}| \sin \alpha}{\sin \beta}, \quad (3)$$

$$\tau = \frac{|\mathbf{H}| \sin \theta}{|\mathbf{V}| \sin \beta}, \quad (4)$$

where, $\alpha (-\pi < \alpha < \pi)$ is the angle between flow vector \mathbf{V} and primal axis vector \mathbf{H} of normal projection and $\beta (0 \leq \beta < \pi)$ is the angle between flow vector and the sampling slit line vector. Here, $\alpha + \beta + \theta = \pi$ (if $0 < \alpha + \theta < \pi$) or 2π (otherwise).

If $\sin \beta = 0$, h and τ become infinite and the EPImo appears as an x-t spacetime image. The tangent, h/τ , of projected object then gives the x-axis component of flow. If $\sin \theta = 0$, τ becomes zero and the anortho-projection does not contain any temporal information for flow computation.

2.3 One Sectioning Plane Method for Detecting Binary Direction

We can easily perceive the binary direction of flow from the orientation of the anortho-projection as illustrated in Figures 2(e) and (f). Figure 4 illustrates the relation of the sign and direction. Using α , which is the angle of object flow against the primal axis of normal projection;

$$\text{sign}(\sin \alpha) = \text{sign}(\psi). \quad (5)$$

This can also be confirmed with the following equation derived from equation (3).

$$\sin \alpha = \frac{h}{|\mathbf{H}|} \sin \beta, \quad (6)$$

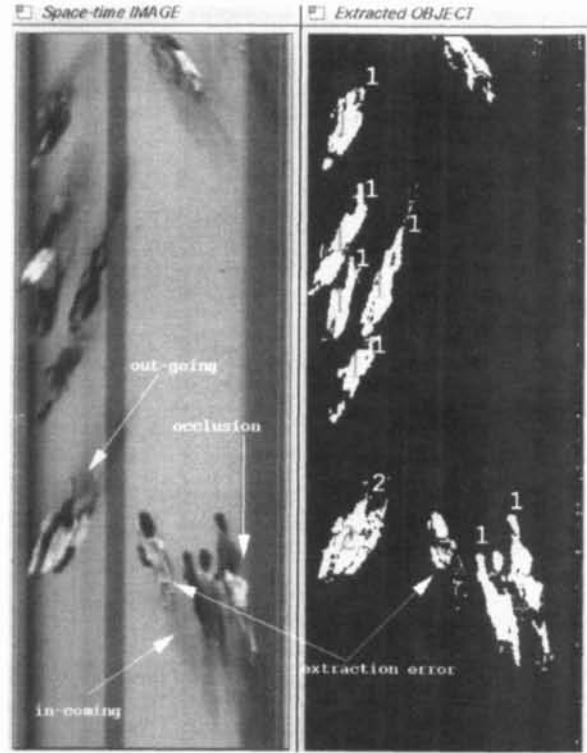


Figure5: Experimental result of counting

where $h (= x_1 - x_0) > 0$ if $\psi > 0$, $h < 0$ if $\psi < 0$, and $\sin \beta \geq 0$.

It is interesting and useful that this calculation does not care which side of the object crosses the sampling slit line first. Thus, we only need to compute the signed orientation of object region against the positive t-axis. This is possible, for instance, by computing the second degree moments of each region.

2.4 Pedestrian Counting System

The proposed flow estimation method can construct a system that automatically counts moving objects such as pedestrians and automobiles. We have built a workstation based pedestrian counting system. Figure 5 is a part of the x-t spacetime image of the image sequence. The counting result is shown as the numbers in the image.

3 Experimental Results of Object Flow Estimation

We evaluated the ortho-sectioning method (two plane method) with two experiments to demonstrate its ability and possible applications. First, rectangular graphic objects were used in simulation tests. The estimation error is evaluated. Then, a real image sequence of walking pedestrians was used in a flow estimation test.

3.1 Simulation with Graphic Object

Single objects were moved at constant speed along straight lines that passed through the center of the window. Speed, direction and object's orientation were ran-

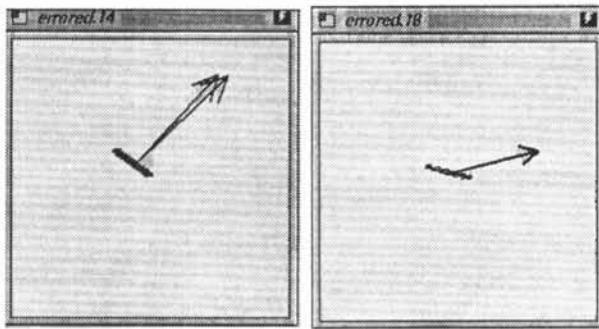


Figure6: Simulation of Object Flow Estimation

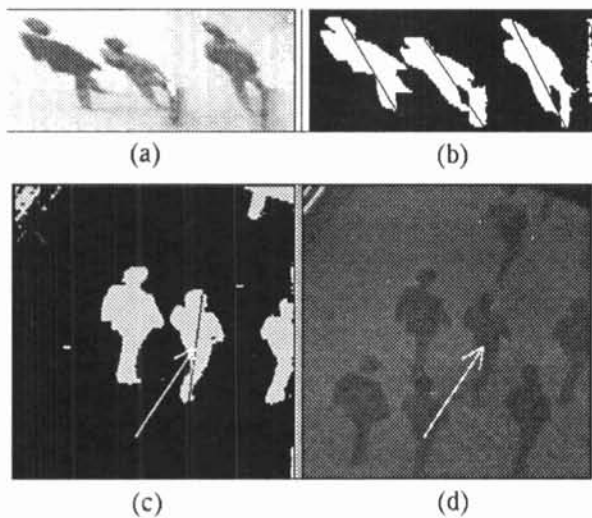


Figure7: Examples of Flow Estimation Result

domly selected. Figure 6 are examples of simulation results in which real flow and estimated flow are illustrated by arrows. The figures in which the two arrows overlap indicate that our ortho-sectioning method works well for simple objects. The average estimation error by the orientation of object, θ , against sampling line with two sizes of object in the 500 simulation tests was investigated. The method works well with less than 0.5 % of error if the orientation is between 20 and 160 degrees.

3.2 Pedestrian Flow

The second experiment analyzed a sequence of walking pedestrians. The image sequence was taken with a fixed TV camera mounted high above a walkway. We used a long focus lens to take images that roughly satisfied the parallel projection assumption. The alignment of the horizontal slit and almost vertical object orientation satisfies the conditions of equations (3) and (4). In order to extract object regions in projected images, we use the color space to eliminate the background. Local smoothing was employed to reduce the effect of noise and to construct object regions. The objects, pedestrians in this case, are not rigid. Figures 7-(a) and (b) show the x-t spacetime image and extracted region of

an ortho-projection of the objects. The primal axes are also shown. Figure 7-(c) shows an extracted region of normal projection of the objects and one primal axis. Figure 7-(d) is a composite image constructed by overlaying two frames one second apart. The estimated object flow for the object in the figure is shown as an arrow. The flow accurately matches the actual movement. The horizontal line is the sampling slit line. The flow vector can be projected onto the walkway plane in order to recover the actual walking velocity.

4 Conclusion

We have presented a method for object flow estimation, the spacetime image ortho-sectioning method, which uses a spacetime volumetric representation of an image sequence. The simplified version using one sectioning plane is also presented for the purpose of fast pedestrian counting with moving directions. The object flow equation utilizes geometric information of object projections onto non-epipolar sectioning planes of volume.

Experimental results of flow estimation have shown the efficiency of the two plane method with both synthesized images and real images. Estimation accuracy is improved by using skewed spacetime sectioning plane images derived from several non-parallel sampling slits.

The advantages of this method are; (1) the object can be non-rigid provided the primal axes are stable, (2) object tracking is unnecessary because the two projections on the ortho-sections connect at the sampling slit which makes correspondence checking very easy, (3) the computational cost is very low because only two dimensional space is processed.

Acknowledgements The authors wish to thank Drs. Takaya Endo and Yukio Kobayashi for their support of this work. Special thanks are due to Yasuko Takahashi, Masaki Nitta and Masaaki Fukumoto for their assistance in image acquisitions. Thanks also to the members of the Human Interface Laboratories for their valuable discussions.

References

- [1] J. Y. Zheng and S. Tsuji: "From Anorthoscopic Perception to Dynamic Vision", 1990 IEEE Int. Conf. on Robotics and Automation, pp. 1154-1160(1990).
- [2] I. Rock: "Anorthoscopic perception", Scientific American, pp. 103-111(1981).
- [3] H. Baker and R. Bolles: "Generalizing Epipolar-Plane Image Analysis on the Spatiotemporal Surface", IJCV, 3, pp. 33-49(1989).
- [4] S. L. Peng and G. Medioni: "Interpretation of image sequences by spatio-temporal analysis", Proc. IEEE workshop on Visual Motion, pp. 344-351(1989).
- [5] M. Allmen and C. R. Dyer: "Computing spatiotemporal surface flow", Proc. ICCV'90, pp. 47-50(1990).
- [6] B. W. Hwang and S. Takaba: "Real-Time Measurement of Pedestrian Flow Using Processing of ITV images", trans. IEICE of Japan, J66-D, 8, pp. 917-924(1983), (in Japanese).