Interpretation of Motion Trajectories Using Focus of Expansion

Krishnan Rangarajan and Mubarak Shah

***Abstract*—The focus of expansion (FOE) of a group of motion trajectories is defined to be a point in the image plane at which the trajectories intersect when they are extended. The FOE observed over a time sequenae defines the *locus of FOE*. We present an analytical approach for the study of dynamic events as they project on the image plane by analyzing the locus of FOE. We have found that the locus of FOE can be used to make qualitative assertions regarding the type of motion. An interesting behavior of the locus of FOE for various types of motion is observed. The cases include a single point, a horizontal, a vertical, and a slopped straight line. We can also determine whether the object has approaching and receding motion or when the object changes its direction of motion. This inference may be used in qualitative computer vision.**

***Index Terms*— Dynamic scene analysis, focus of expansion, motion, trajectories.**

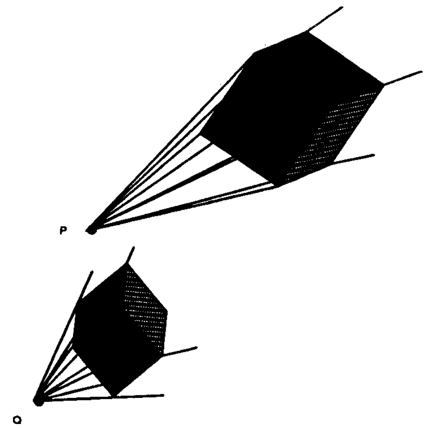

Fig. 1. Two moving cubes. The Cube at the top is moving with a velocity of $v_X = 5$, $v_Y = 5$, and $v_Z = 5$, and the trajectories of the corner points meet at vanishing point P with coordinates $(-100, -100)$. The other cube at the bottom is moving with a velocity $v_X = 6$, $v_Y = 3$, and $v_Z = 2$, and the trajectories of the corner points meet at vanishing point Q with coordinates $(-300, -150)$. The focal length f of the camera is assumed to be 100.

I. Introduction

Given n frames taken at different time instants and m points in each frame, the motion correspondence establishes a mapping of a point in one frame to another point in the next frame. This correspondence can be used to generate a path followed by a point lying on an object. A path can be generated by starting from a point in the first frame and ending at some point in the last frame, touching each frame at not more than one point, and by joining a point in a frame by a straight line with its corresponding point in the next frame. We call such a path a *trajectory*. Each trajectory is identified by a point in the first frame. A set of nonintersecting paths, which together involve all points in all frames, is a *trajectory set*.

Under perspective projection, lines that are parallel in space may not remain parallel in the image plane. With the camera pointing in the Z axis,[1] lines parallel in space and having components in the Z axis do not project to be parallel lines in the image plane. These lines meet at a common point in the x-y plane known as the *vanishing point*. Lines that are parallel in space and that do not have a component in the Z axis project to the parallel lines in the image plane, and hence, their vanishing point is at infinity.

The motion trajectories are representative of the motion of the underlying points. If we assume that the image plane x-y is parallel to the X-Y plane, the trajectories that have a nonzero motion component in the Z direction and are parallel in space will exhibit the same vanishing point. The trajectories of points lying on an object undergoing translation remain parallel in space. Hence, depending on the *vanishing point*, we can segment the trajectories belonging to the same object. Fig. 1 shows two cubes moving with different velocities in space. The *vanishing point* of trajectories belonging to one cube is point P $(-100, -100)$, and the vanishing point of trajectories of the other cube is point Q $(-300, -150)$. If the motions of objects are assumed to be independent, each vanishing point will represent one object.

The focus of expansion (FOE) is very similar to the vanishing point of the motion trajectories. The FOE is defined to be a point in the image from which, for a given forward direction of translating motion and direction of gaze, all image features seem to diverge radially. The FOE can be used to determine, for example, the direction of vehicle heading. Since the FOE has previously been used in motion research, we will also continue to use the FOE for the vanishing point of motion trajectories. Previous related research has been limited to computation of the FOE. In this correspondence, we will assume that there exists a reasonably good method for computing the FOE, and we will focus on the use of the FOE for interpretation of motion trajectories. Precisely, we will be dealing with an extended sequence of frames, which gives rise to a sequence of trajectories between two frames and, hence, a sequence of FOE's. We term this sequence of FOE's as the *locus of FOE*. The aim of this correspondence is to show that the locus of the FOE is closely related to the motion of the underlying object.

We present an analytical approach for the study of dynamic events as they project on the image plane by analyzing the locus of the FOE. We have found that the locus of the FOE can be used to make qualitative assertions regarding the type of motion. We can also determine whether the object has approaching and receding motion or when the object changes its direction of motion. This inference may be used in qualitative computer vision.

The organization of the rest of the correspondence is as follows. The next section deals with a survey of methods for computing the FOE. The projection model that will be used throughout this correspondence is described in Section III. In Section IV, we analyze the locus of the FOE for various types of motion. Finally, Section V deals with interpretation of motion trajectories using the locus of the FOE.

Manuscript received October 24, 1989; revised April 3, 1992. This work was supported by DOD/PMTRADE under contract no. N61339-88-G-0002/6. The content of the information herein does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred. Recommended for acceptance by Associate Editor N. Ahuja.

The authors are with the Computer Science Department, University of Central Florida, Orlando, FL 32816.

IEEE Log Number 9201899.

[1] We will be using uppercase X, Y, and Z to denote the world coordinates, whereas we will use lowercase x and y to denote the image coordinates.

II. Related Work

Earlier researchers have developed methods for computing the FOE. The FOE is defined only for pure translation. The work on *FOE* can be broadly categorized into two categories based on the density of the optical flow used: dense flow methods and sparse flow methods. The methods proposed by Ballard *et al.* [3], Negaharipour and Horn [9], Prazdny [10], Dutta *et al.* [5], and Burger *et al.* [4] are dense flow methods, whereas the methods proposed by Jain [6] and Lawton [7] are sparse flow methods based on a small number of point correspondences. These methods make use of the Hough transform or employ an error surface for a restricted region in the image plane and estimate the FOE by hill climbing. Each of these categories can again be divided into subcategories based on whether they address multiple independently moving objects or not.

Ballard *et al.* [3] use the fact that all the flow vectors converge at the FOE, and they estimate the FOE by a Hough formulation. Collinear flow vectors are detected by considering the (r,θ) space. A point (x,y) in the image that has a flow vector (u,v) votes for (r,θ), where $r = xcos\theta + y\sin\theta$, and $tan\theta = \frac{u}{v}$. The points in (r,θ) space of radial flow lines form a circle, each circle corresponds to one rigid object, and each circle is associated with a FOE. From (r,θ) space, the FOE is identified by the Hough method. Each r, θ votes for cells (a,b), obeying the constraint $\frac{r}{2} = acos\theta + b\sin\theta$. The FOE is given by $(2a, 2b)$. They concentrate on scenes having a single moving object, but they do suggest as to how it can be expanded to scenes with multiple moving objects.

Dutta *et al.* [5] address determination of the FOE under ego motion. Pure translation of the camera is idealistic, and in the sequences of real scenes, there is always a small rotation of the camera, which introduces error into the FOE determination. They consider the error due to rotation as a systematic error, develop expressions for the error, and undo the rotation. They show that error in the FOE estimation due to rotation is proportional to the depth of the points whose flow vectors are considered. The FOE estimated by using large flow vectors of the nearby points is closer to the actual FOE than the one found by using small flow vectors belonging to distant points. They verify their claims by comparing their method with the one that uses Anandan's method [2] for estimating the flow vectors and Adiv's method [1] for estimating motion parameters and, hence, the FOE. The method has been designed for sensor motion as in autonomous navigation and is not suitable for scenes with multiple independently moving objects.

The method proposed by Burger *et al.* [4] also takes a similar approach to estimating the rotation and correcting it to find the correct FOE. They define a fuzzy FOE that has an associated area around it and assume that the FOE falls inside it. This area is grown by a connected component algorithm around a point. This method has also been designed for sensor motion as in autonomous navigation and is not suitable for scenes with multiple independently moving objects.

Jain [6] has proposed a method for estimating the FOE without optical flow. He uses a simple function based on the geometry that, when maximized, leads to the FOE. The function that is maximized is the sum of the distance from an arbitrary point to the points in the two consecutive frames—the sum of the distances between points in frames. This is a continuous function, and a gradient method is used to find the minimum. There is no direct extension of this method to scenes having multiple independently moving objects.

Lawton has proposed a method [7] that can estimate the FOE given a sparse flow field. The main difference between this method and that of the method proposed by Jain is that it combines the token identification with the FOE estimation. The method, as it is, is applicable to images generated by moving sensor, and it cannot be extended to scenes with multiple independently moving objects in a straightforward manner.

There has been some work done in finding parallel lines in a perspective image using vanishing points. There is a similarity between vanishing points and the FOE. The vanishing point is the point of intersection of line segments in the scene that are the projection of parallel lines in 3-D space, and the FOE is the point of intersection of flow vectors in the image that are the projection of parallel 3-D flow vectors. Magee and Aggarwal [8] have proposed a method for finding the vanishing point. They consider end points of lines in the image and fit a plane with the focal point. The line of intersection between two such planes points toward the vanishing point of the two lines in the image. They map the unit vector of this line of intersection onto a point on the surface of the Gaussian sphere and cluster the points on the Gaussian surface based on distance. The lines corresponding to these clustered points are parallel in 3-D space and have a common vanishing point in the image plane.

Our correspondence deals with the analysis and use of the FOE, and we assume there exists some reasonably good algorithm for computing the FOE. The experiments reported in the correspondence use the FOE computed by finding the intersection of two straight line segments of trajectories analytically.

III. Projection Model

Let us assume that the origin of the world coordinate system coincides with the origin of the image coordinate system. In addition, let $(0,0,0)$ be the position of the camera in the world coordinates. Let (X_i^j, Y_i^j, Z_i^j) be the world coordinates of point i at time t_j, and the image plane coordinates (x_i^j, y_i^j) of point i at time t_j under perspective projection are given by

$$x_i^j = -\frac{f.X_i^j}{Z_i^j}$$
$$y_i^j = -\frac{f.Y_i^j}{Z_i^j}$$

where f is camera focal length. The FOE of trajectories belonging to an object is the point of intersection between the 2-D projection of the trajectories in the image plane. It is assumed that when the time interval between subsequent frames is small, these trajectories can be considered to be straight lines. Let $(\vartheta_x^j, \vartheta_y^j)$ denote the FOE of trajectories belonging to the same object between times t_j and t_{j+1}. Let v_X^j, v_Y^j, and v_Z^j denote the velocity of the object in the X, Y, and Z directions, respectively, in this time interval. The FOE of an object between two time frames is determined by finding the point of intersection of the projection of a pair of 2-D trajectories of points belonging to this object and is given by

$$\vartheta_x^j = -\frac{f.v_X^j}{v_Z^j} \tag{1}$$
$$\vartheta_y^j = -\frac{f.v_Y^j}{v_Z^j}. \tag{2}$$

These equations are derived in the Appendix. These equations imply that that the FOE is the perspective projection of the point in space whose coordinates are given by (v_X^j, v_Y^j, v_Z^j).

IV. Locus of FOE

For a sequence of frames 1 to n, the world coordinates of any point at time t_{j+1} belonging to an object is related to its coordinates at time t_j by its motion parameters at time t_j. The path traced by $(\vartheta_x^j, \vartheta_y^j)$, which is a FOE of the object, as we vary j from 1 to $n-1$ defines the locus of the FOE of the object. The type of motion and the nature of motion like constant velocity, constant acceleration,

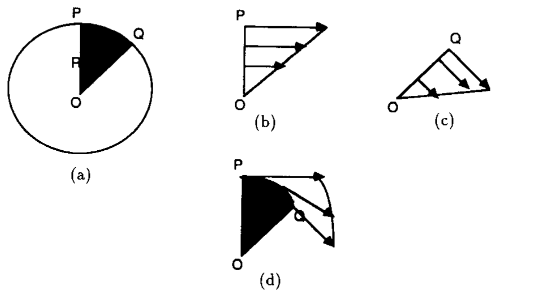

Fig. 2. (a) Sector POQ is rotating about O in the YZ plane; (b) velocity profile along OP. The points on OP are at different radii from the center of rotation O but are in phase with each other; (c) velocity profile along OQ; (d) velocity profile along arc PQ, where the radius remains constant but phase angle is changing.

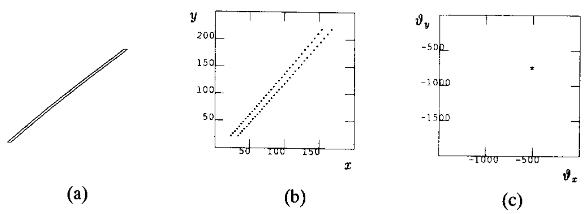

Fig. 3. (a) Isometric view of 3-D trajectories of a line with end points at (20,20,30), and (30,20,30) under translation with velocity components $v_X = 3$, $v_Y = 4$, and $v_Z = 5$; (b) perspective projection of trajectories in (a); (c) locus of FOE, which is a single point.

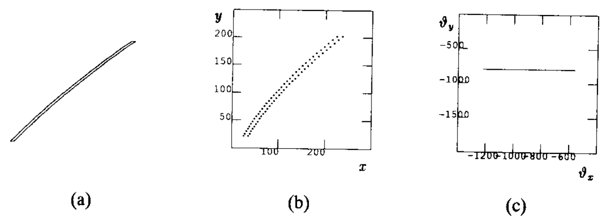

Fig. 4. (a) Isometric view of 3-D trajectories traced out by the end points at (20,20,30) and (30,20,30) of a line under translation with velocity components of $v_X = 3$, $v_Y = 4$, and $v_Z = 5$, and acceleration component $a_X = 0.1$, $a_Y = 0$, and $a_Z = 0$; (b) perspective projection of the trajectories in (a); (c) locus of FOE, which is a straight line parallel to the x axis.

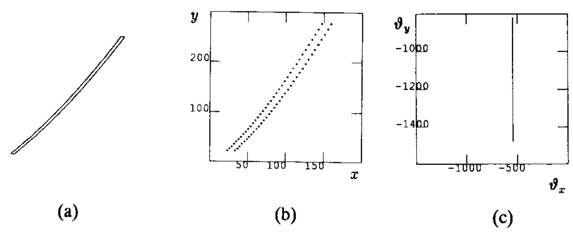

Fig. 5. (a) Isometric view of 3-D trajectories traced out by the end points at (20,20,30) and (30,20,30) of a line under translation with velocity components $v_X = 3$, $v_Y = 4$, and $v_Z = 5$, and acceleration component $a_X = 0$, $a_Y = 0.1$, and $a_Z = 0$; (b) perspective projection of the trajectories in (a); (c) locus of FOE, which is a straight line parallel to y axis.

etc. affect the locus of the FOE. In fact, the locus of FOE is the perspective projection of the space curve given by the equations

$$x = v_X$$
$$y = v_Y$$
$$z = v_Z.$$

In this section, we study the locus of the FOE under pure translation. For other types of motion, the FOE is not well defined. For instance, under rotation, all points belonging to the same object do not necessarily have the same FOE. Fig. 2(a) shows a sector OPQ rotating about O in the YZ plane. Points R and P have the same phase angle but different radii of rotation OR and OP, respectively, whereas points R and Q have a phase difference and different radii of rotation. The velocity profile along OP, where the radius alone varies but not the phase angle, is shown in Fig. 2(b). The velocity profile along OQ is shown in Fig. 2(c), and the velocity profile along arc PQ, where the radius remains constant but the phase angle varies, is shown in Fig. 2(d). The points on the same object will have the same FOE if and only if they are located at the same distance from the axis of rotation and they are in phase with each other.

The results in this section are reported with the aid of figures (Figs. 4–6) that have three parts. In (a), we plot the 3-D trajectories of the points in motion as in an *isometric view*; in (b), we plot the 2-D trajectories traced out in the image plane by the 3-D trajectories, and finally, in (c), we plot the locus of the FOE.

In this subsection, we will consider various cases related to translation.

1. When an object translates with a uniform velocity, the locus of the FOE is a point, that is, the FOE remains stationary. This follows from (1) and (2) as v_X^j, v_Y^j, and v_Z^j remain constant for all j. Fig. 3 shows an instance of this case.
2. If the object translates in X, Y, and Z but X alone has a constant acceleration a_X, the velocities at time instant j can be written as
$$v_X^j = v_X^1 + (j-1).a_X$$
$$v_Y^j = v_Y^1$$
$$v_Z^j = v_Z^1.$$
Since we are assuming acceleration to be constant, the locus of the FOE is given by the equation
$$\vartheta_x^j = -f.\frac{v_X^1 + (j-1).a_X}{v_Z^1}$$
$$\vartheta_y^j = -f.\frac{v_Y^1}{v_Z^1}.$$
The above equations represent a line parallel to the x axis. Fig. 4 shows an instance of this case.
3. If a rigid line translates in X, Y, and Z but Y alone has acceleration, which is constant a_y, then the velocities at time instant j can be written as
$$v_X^j = v_X^1$$
$$v_Y^j = v_Y^1 + (j-1).a_Y$$
$$v_Z^j = v_Z^1.$$
Since we are assuming acceleration to be constant, the locus of the FOE is given by the equations
$$\vartheta_x^j = -f.\frac{v_X^1}{v_Z^1}$$
$$\vartheta_y^j = -f.\frac{v_Y^1 + (j-1).a_Y}{v_Z^1}.$$
The above equations represent a line parallel to the y axis. Fig. 5 shows an instance of this case.

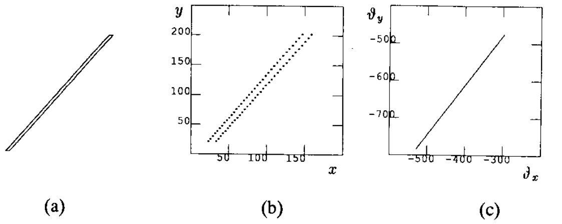

Fig. 6. (a) Isometric view of 3-D trajectories traced out by the end points at (20,20,30) and (30,20,30) of a line under translation with velocity components $v_X = 3$, $v_Y = 4$, and $v_Z = 5$, and acceleration component $a_X = 0$, $a_Y = 0$, and $a_Z = 0.1$; (b) perspective projection of the trajectories in (a); (c) locus of FOE, which is an inclined line.

4. If a rigid object translates in X, Y, and Z but Z alone has constant acceleration, then the velocities at time instant j can be written to be

$$v_X^j = v_X^1$$
$$v_Y^j = v_Y^1$$
$$v_Z^j = v_Z^1 + (j-1).a_Z.$$

Since we are assuming acceleration a_Z to be constant, the locus of the FOE is given by the equations

$$\vartheta_x^j = -f.\frac{v_X^1}{v_Z^1 + (j-1).a_Z}$$
$$\vartheta_y^j = -f.\frac{v_Y^1}{v_Z^1 + (j-1).a_Z}.$$

The above equations represent an inclined line with slope $\frac{v_Y^1}{v_X^1}$. Fig. 6 shows an instance of this case.

5. When an object translates in X, Y, and Z with constant acceleration of a_X, a_Y, a_Z in X, Y and Z, the velocities at time instant j can be written to be

$$v_X^j = v_X^1 + (j-1).a_X$$
$$v_Y^j = v_Y^1 + (j-1).a_Y$$
$$v_Z^j = v_Z^1 + (j-1).a_Z.$$

Since we are assuming acceleration to be constant, the locus of the FOE is given by the equations

$$\vartheta_x^j = -f.\frac{v_X^1 + (j-1).a_X}{v_Z^1 + (j-1).a_Z}$$
$$\vartheta_y^j = -f.\frac{v_Y^1 + (j-1).a_Y}{v_Z^1 + (j-1).a_Z}.$$

The above equations represent an inclined line with slope $\frac{(a_Y v_Z - v_Y a_Z)}{(a_X v_Z - v_X a_Z)}$. From these equations, we can see that the locus of FOE is an inclined line with slope $\frac{a_Y}{a_X}$ when $a_Z = 0$. When $a_X = 0$, $a_Y = 0$, and $a_Z \neq 0$, the slope of this line is $\frac{v_Y}{v_X}$, as in item 4 above.

V. Interpretation of Locus of FOE

In this section, we will summarize the observations regarding the interpretation of locus of the FOE under various types of motion. From the analysis, it will also be shown that the change of direction of motion and receding and approaching motion can be determined from the locus of the FOE.

1. If the locus of the FOE is a single point, then the object is undergoing translation without any acceleration.

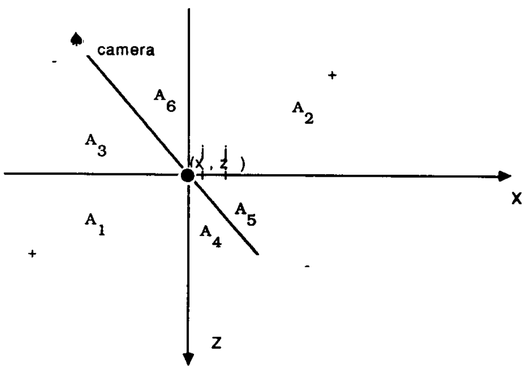

Fig. 7. Deciding whether a point i (shown by a filled circle) is approaching or receding. The figure shows the projection on the X-Z plane. The camera is shown as ♠. A reference axis is drawn at point i. Around the point i, six different areas A_1-A_6 are marked. The sign of the FOE is also marked in the quadrants formed by the reference axis.

2. If the FOE moves along a line, then the following holds:
   a. If the locus is a straight line parallel to the x axis, then the motion is translation with acceleration along the X direction only.
   b. If the locus is a straight line parallel to the y axis, then the motion is translation with acceleration along Y direction only.
   c. If the locus is an inclined straight line, then the motion is translation with acceleration in the$Z$ direction or along two or more of the X, Y, and Z directions. This includes cases 4 and 5 discussed in Section IV.
3. We can identify the instants at which the object changes its direction of motion along the Z axis. This is because the velocity along the Z direction appears in the denominator of the X and Y coordinates of the FOE. As the velocity along the Z direction changes sign, there is a discontinuity in the plot of ϑ_x and ϑ_y against time at the same time instant. This fact can be used in identifying instants at which the object changes its direction of motion along the Z direction.
4. We can also identify instants at which the object changes its direction of motion along the X and Y axes. The change in direction of motion along the X axis is seen as a smooth crossover from positive to negative values in the plot of the X coordinates of the FOE against time. This is because v_X, which is the velocity along the X direction, appears in the numerator of ϑ_x. Similarly, the change in direction of motion along the Y axis is seen as a smooth crossover from positive to negative values in the plot of the y coordinates of the FOE against time.
5. We can determine whether an object is approaching or receding at any instant of time using the sign of the FOE and the sign of image coordinates of a point. Fig. 7 shows the X-Z plane and a reference grid at the point of interest i at time instant j. In each quadrant, a '+' or '$-$' sign is marked, and it denotes the sign of the ϑ_x^j if the point moves into that quadrant in time instant $j+1$. These signs can easily be determined by using the fact that $\vartheta_x = -\frac{v_X}{v_Z}$. For instance, the point i shown in Fig. 7 will only move to the area A_1 if its v_X is negative and v_Z is positive; therefore, ϑ_x will be positive. We will consider the following cases:
   a. If $\vartheta_x^j > 0$, then the following hold:

No	X_1^1	Y_1^1	Z_1^1	X_1^2	Y_1^2	Z_1^2	x_1^1	y_1^1	x_1^2	y_1^2	ϑ_x^1	Approach/Recede	Area
1	50	50	50	40	70	80	-100	-100	-50	-87.5	33.33	Recede	A1
2	50	50	50	60	70	30	-100	-100	-200	-233.33	50	Approach	A2
3	50	50	50	40	70	45	-100	-100	-88.89	-155.56	-200	Approach	A3
4	50	50	50	70	70	80	-100	-100	-87.5	-87.5	-66.67	Recede	A4
5	50	50	50	70	70	60	-100	-100	-116.67	-116.67	-200	Recede	A5
6	50	50	50	40	70	30	-100	-100	-133.33	-233.33	-50	Approach	A6

Fig. 8. Rigid line with end points at 1, 2 was moved in space, and the ϑ_x^1 was computed. Our decision procedure was used in deciding whether point 1 was approaching or receding. In this table, we show six cases listed in column 1. The columns 2–4 and 5–7, respectively, show the world coordinates of point 1 in frames 1 and 2, whereas columns 8 and 9 and 10 and 11 show their corresponding image coordinates. The ϑ_x of each case is given in column 12.

- If $x_i^{j+1} - x_i^j > 0$, the object is receding. The point should move to one of the areas A_1, A_3, or A_4 for $x_i^{j+1} - x_i^j > 0$. As $\vartheta_x^j > 0$, the point would have moved into area A_1 in the Fig. 7. See entry one in the table shown in Fig. 8 for an example.
- If $x_i^{j+1} - x_i^j < 0$, then the object is approaching. The point should move to one of the areas A_6, A_2, or A_5 for $x_i^{j+1} - x_i^j < 0$. As $\vartheta_x^j > 0$, the point would have moved into area A_2 in Fig. 7. See entry two in the table shown in Fig. 8 for an example.

b. If $\vartheta_x^j < 0$, then the following hold:

- If $x_i^{j+1} - x_i^j > 0$, and $\vartheta_x^j \leq x_i^j$, then the object is approaching. The point should move to one of the areas A_1, A_3, or A_4 for $x_i^{j+1} - x_i^j > 0$. As $\vartheta_x^j < 0$, it could have moved to either A_3 or A_4. As $\vartheta_x^j \leq x_i^j$, the point would have moved to area A_3 in Fig. 7. See entry three in the table shown in Fig. 8 for an example.
- If $x_i^{j+1} - x_i^j > 0$, and $\vartheta_x^j \geq x_i^j$, then the object is receding. The point should move to one of the areas A_1, A_3, or A_4 for $x_i^{j+1} - x_i^j > 0$. As $\vartheta_x^j < 0$, it could have moved to either A_3 or A_4. As $\vartheta_x^j \geq x_i^j$, the point would have moved to area A_4 in Fig. 7. See entry four in the table shown in Fig. 8 for an example.
- If $x_i^{j+1} - x_i^j < 0$, and $\vartheta_x^j \leq x_i^j$, then the object is receding. The point should move to one of the areas A_6, A_2, or A_5 for $x_i^{j+1} - x_i^j < 0$. As $\vartheta_x^j < 0$, it could have moved to either A_6 or A_5. As $\vartheta_x^j \leq x_i^j$, the point would have moved to area A_5 in Fig. 7. See entry five in the table shown in Fig. 8 for an example.
- If $x_i^{j+1} - x_i^j < 0$, and $\vartheta_x^j \geq x_i^j$, the object is approaching. The point should move to one of the areas A_6, A_2, or A_5 for $x_i^{j+1} - x_i^j < 0$. As $\vartheta_x^j < 0$, it could have moved to either A_6 or A_5. As $\vartheta_x^j \geq x_i^j$, the point would have moved to area A_6 in Fig. 7. See entry six in the table shown in Fig. 8 for an example.

We assumed that the piercing point (the point in an image at which the ray along which the camera was aimed pierces the image plane) is to the right of the point under consideration. A similar decision procedure can be developed when the piercing point is to the left of the point under consideration. We tested these cases using synthetic data generated by moving a rigid line in space. The results are tabulated in Fig. 8 for one of the end points of the line.

VI. Conclusion

In this correspondence, we have shown that the FOE's of motion trajectories carry rich information related to the motion of the objects. The locus of the FOE also helps in identifying instances at which the object changes its direction of motion along any principle axes. Since under translation all points on the same object have the same velocity, their trajectories have the same FOE. This property of the FOE can be used in developing an algorithm for segmenting trajectories into groups belonging to the same object, which will make use of the information available in the whole span of the trajectories.

Appendix

In this Appendix, we will derive the expressions for the FOE (1) and (2). Let (x_1^j, y_1^j) and (x_1^{j+1}, y_1^{j+1}) be the trajectory segment of point 1 in frames j and $j+1$, and let (x_2^j, y_2^j), (x_2^{j+1}, y_2^{j+1}) be the trajectory segment of point 2 in frames j and $j+1$. Further, assume that points 1 and 2 lie on the same object, which is translating in space with a velocity of v_X^j, v_Y^j, and v_Z^j in the X, Y, and Z directions, respectively. Let the 3-D coordinates of points 1 and 2 in frame j be (X_1^j, Y_1^j, Z_1^j) and (X_2^j, Y_2^j, Z_2^j), let and those of points 1 and 2 in frame $j+1$ be $(X_1^{j+1}, Y_1^{j+1}, Z_1^{j+1})$ and $(X_2^{j+1}, Y_2^{j+1}, Z_2^{j+1})$. Clearly, the following equations hold:

$$\begin{aligned} X_1^{j+1} &= X_1^j + v_X^j & X_2^{j+1} &= X_2^j + v_X^j \\ Y_1^{j+1} &= Y_1^j + v_Y^j & Y_2^{j+1} &= Y_2^j + v_Y^j \\ Z_1^{j+1} &= Z_1^j + v_Z^j & Z_2^{j+1} &= Z_2^j + v_Z^j. \end{aligned}$$

In addition, applying the projection equations

$$\begin{aligned} x_1^j &= \frac{-fX_1^j}{Z_1^j} \\ y_1^j &= \frac{-fY_1^j}{Z_1^j} \\ x_1^{j+1} &= \frac{-fX_1^j + v_X^j}{Z_1^j + v_Z^j} \\ y_1^{j+1} &= \frac{-fY_1^j + v_Y^j}{Z_1^j + v_Z^j}. \end{aligned}$$

The equations of the trajectory segments of points 1 and 2 are given by the following equations:

$$\begin{aligned} \frac{x - x_1^j}{y - y_1^j} &= \frac{x_1^j - x_1^{j+1}}{y_1^j - y_1^{j+1}} \\ \frac{x - x_2^j}{y - y_2^j} &= \frac{x_2^j - x_2^{j+1}}{y_2^j - y_2^{j+1}}. \end{aligned}$$

FOE $(\vartheta_x^j,\ \vartheta_y^j)$ is the point of intersection of these lines and is given by

$$\begin{aligned} \vartheta_x^j &= -\frac{f.v_X^j}{v_Z^j} \\ \vartheta_y^j &= -\frac{f.v_Y^j}{v_Z^j}. \end{aligned}$$

REFERENCES

[1] G. Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects," *IEEE Trans. Patt. Anal. Machine Intelligence*, vol. PAMI-7, pp. 384–401, 1985.

[2] P. Anandan, "Measuring visible motion from image sequences," Ph.D. thesis, Univ. of Mass., Amherst, MA, 1987.

[3] D. H. Ballard and O. A. Kimball, "Rigid body motion from depth and optical flow," *Comput. Vision Graphics Image Processing*, vol. 22, pp. 95–105, 1983.

[4] W. Burger and B. Bhanu, "On computing a fuzzy focus of expansion for autonomous navigation," in *Proc. Conf. Comput. Vision Patt. Recogn.* (San Diego), 1989, pp. 563–568.

[5] R. Dutta, R. Manmatha, E. M. Riseman, and M. A. Snyder, "Issues in extracting motion parameters and depth from approximate translational motion," in *Proc. DARPA IU Workshop* (Cambridge, MA), Apr. 1988, pp. 945–960.

[6] R. Jain, "Direct computation of the focus of expansion," *IEEE Trans. Patt. Anal. Machine Intelligence*, vol. PAMI-5, pp. 58–64, 1983.

[7] D. T. Lawton, "Processing translational motion sequences," *Comput. Vision Graphics Image Processing*, vol. 22, pp. 116–144, 1983.

[8] M. Magee and J. K. Aggarwal, "Determining vanishing points from perspective images," *Comput. Vision Graphics Image Processing*, vol. 26, pp. 256–267, 1984.

[9] S. Negahdaripour and B. K. P. Horn, "A direct method for locating the focus of expansion," *Comput. Vision Graphics Image Processing*, vol. 46, pp. 303–326, 1989.

[10] K. Prazdny, "Determining the instataneous direction of motion from optical flow generated by a curvilinearly moving observer," *Comput. Vision Graphics Image Processing*, vol. 17, pp. 238–248, 1981.

Separability of Spatiotemporal Spectra of Image Sequences

Michael P. Eckert, Gershon Buchsbaum, and Andrew B. Watson

***Abstract*— We calculated the spatiotemporal power spectrum of 14 image sequences in order to determine the degree to which the spectra are separable in space and time and to assess the validity of the commonly used exponential correlation model found in the literature. We expand the spectrum by a singular value decomposition into a sum of separable terms and define an index of spatiotemporal separability as the fraction of the signal energy that can be represented by the first (largest) separable term. All spectra were found to be highly separable with an index of separability above 0.98. The power spectra of the sequences were well fit by a separable model of the form**

$$P(k,f) = \frac{ab/(4\pi^3)}{((a/2\pi)^2 + k^2)^{3/2}((b/2\pi)^2 + f^2)}$$

where k is radial spatial frequency, f is temporal frequency, and a,b are spatial and temporal model parameters that determine the effective spatiotemporal bandwidth of the signal. This power spectrum model corresponds to a product of exponential autocorrelation functions separable in space and time.

Manuscript received November 7, 1990; revised March 6, 1992. This work was supported by the NASA Graduate Fellowship Program and by grants NSF 8351637 and AFOSR 91-0082. Recommended for acceptance by Associate Editor N. Ahuja.

M. P. Eckert and G. Buchsbaum are with the Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA 19104-6392.

A. B. Watson is with the NASA Ames Research Center, Moffet Field, CA 94035-1000.

IEEE Log Number 9204244.

I. INTRODUCTION

The statistics of images and image sequences have been extensively studied for image coding and compression applications [1], [2] as well as for the development of models of biological image processing [3], [4]. An exponential autocorrelation function has been shown to be a good model for temporal frame-to-frame correlations of image sequences, e.g., [5]–[8], and for spatial correlations within each frame, e.g., [2], [3], [9].

This paper focuses on the separability of the spatiotemporal statistics of image sequences and on the validity of using a separable exponential autocorrelation model for the spatiotemporal statistics. The autocorrelation function is uniquely related to the power spectrum via a Fourier transform, and either is valid as a description of the statistics.

The spectra of 14 image sequences were calculated. The sequences represented a small ensemble of possible motion activity. The sequences were selected for a range of motion activity. For example, a fast camera pan represents the maximum image motion activity, and a small moving object with a static background represents the least activity. Sequences with motion activity between these extremes had slight camera motion and some object motion.

II. CALCULATION OF IMAGE STATISTICS

We collected 14 image sequences (256 × 256× 64 @ 8 b/pixel, 30 frames/s with no scene cuts) from a video disc that contained scenes from a broadcast TV source. Each frame was originally sampled at 512 × 512 pixels/screen, but adjacent pixels were averaged, and the image was subsampled to 256 × 256 pixels/screen. The sample mean of each sequence was removed to reduce low-frequency bias in the calculations.

The sample power spectrum $P(k_1, k_2, f)$ of each sequence $x(n_1, n_2, t)$ is the squared magnitude of the discrete Fourier transform calculated as

$$P(k_1, k_2, f) = \frac{1}{256 \cdot 256 \cdot 64} \cdot \left| \sum_{n_1=0}^{255} \sum_{n_2=0}^{255} \sum_{t=0}^{63} x(n_1, n_2, t) e^{-j2\pi(k_1 n_1 + k_2 n_2 + ft)} \right|^2 \quad (1)$$

where k_1,k_2 are spatial frequencies, f is temporal frequency, n_1,n_2 are spatial locations, and t is time measured in frame number.

We converted the two spatial frequency dimensions k_1 and k_2 into one radial frequency dimension k by averaging in 32 annuli around the spatial frequency origin as illustrated in Fig. 1. In this manner, the spatial frequency range of 0–127 cycles/screen of k_1 and k_2 is represented by 32 annuli in bands of 4 cycles/screen. Averaging the spatial spectra in annuli is equivalent to assuming a circularly symmetric spatial autocorrelation function. This autocorrelation function is not separable in the two spatial dimensions but is considered a better fit than the corresponding separable autocorrelation function for most images [9].

The average magnitude of the power spectrum in each annulus can be obtained by summing over the power spectrum $P(k_1, k_2, f)$ in the annulus indexed by k and normalizing by the number of sample