

A Novel Approach to Motion Segmentation

D. Sinclair, Andrew Blake, Paul Beardsley, David Murray

Robotics Research Group
Department of Engineering Science
Oxford University OX1 3PJ

Abstract

This paper offers a novel means of segmenting independently moving objects from rigid backgrounds. The camera motion can be arbitrary with the proviso that independently moving objects are not moving parallel to it. A new method of treating visual motion is presented. The method relies on describing the motion of points in the world in terms of their angular velocity relative to the camera. In the case of a static scene the method permits the recovery of the direction of camera motion and the camera's angular velocity from the optical flow and its first derivatives at only three points.

1 Introduction

There have been several approaches to the problem of determining the structure of the world from visual motion in a series of camera images. All methods depend on obtaining a reliable estimate of the camera's motion. The methods may be characterised by the varying number of correspondences required to solve for the camera motion. Longuet-Higgins' eight point algorithm functions by using optical flow at eight points to return the depth of the points together with the motion of the camera (up to the speed scale ambiguity). Waxman and Subbarao suggest the use of flow and its derivatives at one point to estimate both the motion of the camera and the slant and tilt of the surface under observation. Horn [1] demonstrated a method that functioned in the absence of rotation that used only the normal components of flow at numerous image positions. None of the above methods offer a particularly natural framework for segmenting independent motions from images and all are vulnerable to the two population problem. Nelson [4] demonstrated a motion segmentation procedure based on violation of the epipolar constraint by objects moving independently of the background. All of the above approaches are based on the equation for the motion of a projected point derived in [3]. This paper proposes using the angular velocity of a point relative to the camera (figure 1). This gives rise to a three dimensional angular velocity field rather than the conventional two dimensional flow field. Section 2 gives the derivation of the angular velocity field and its use in recovering observer motion, section 3 deals with the use of the angular velocity field for independent motion segmentation and section 4 shows how the angular velocity field may be estimated from the optical flow field.

2 The Angular Velocity Field Associated With a Moving Image

Conventionally in computer vision the optical flow field is the principle item of interest. In spherical projection [2] the motion V of projected points is given by;

$$V = \frac{((Q \times U) \times Q)}{|q|} - \Omega \times Q, \quad (1)$$

Where q is the vector from the optical center of the camera to a point on an object in the world, $Q = \frac{q}{|q|}$, U is the translational and Ω the angular velocity of the camera. An alternative description of the projected motion of a point may be given in terms of its angular velocity relative to a reference frame moving with the camera (figure 1). This gives

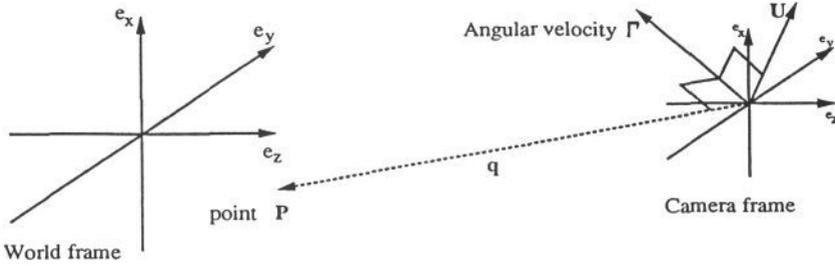


Figure 1: P is a fixed point in a static world frame at a position \mathbf{q} in a reference frame translating with velocity \mathbf{U} and not rotating. The angular velocity $\mathbf{\Gamma}$ of point P relative to the translating observer frame is, $\mathbf{\Gamma} = \frac{\mathbf{U} \times \mathbf{q}}{|\mathbf{q}|^2}$.

rise to a three dimensional vector field rather than the two dimensional one associated with optical flow. The angular velocity field is denoted here by $\mathbf{\Gamma}$ and is derived as follows. The angular velocity $\mathbf{\Gamma}$ of a point at a position \mathbf{q} with respect to a reference frame translating with velocity \mathbf{U} is given by;

$$\mathbf{\Gamma} = \frac{\mathbf{U} \times \mathbf{q}}{|\mathbf{q}|^2}. \quad (2)$$

Angular velocity is a vector quantity, hence if the angular velocity of a point is known with respect to one inertial reference frame (one undergoing straight translation) then the angular velocity of a point relative to a frame with an additional angular velocity $\mathbf{\Omega}$ is simply the initial angular velocity minus $\mathbf{\Omega}$.

$$\mathbf{\Gamma} = \frac{\mathbf{U} \times \mathbf{q}}{|\mathbf{q}|^2} - \mathbf{\Omega}. \quad (3)$$

$\mathbf{\Gamma}$ is then the angular velocity field of visible points in a rigid world, relative to a moving camera. The term in $\mathbf{\Gamma}$ arising from camera rotation is now a constant and not a function of position within the field of view (cf (1)). Unfortunately, in general, $\mathbf{\Gamma}$ is not directly measurable but, as we demonstrate in section 4, there exist generic conditions under which $\mathbf{\Gamma}$ is recoverable. These generic image positions correspond to where the shear in the parallax field is zero *ie* either the surface at the end of a viewing direction is 'fronto parallel' or a viewing direction is aligned with the motion direction. A minimum of three such positions (with non-coplanar viewing directions) are required to solve for an observer's rotational velocity.

Once the angular velocity field has been estimated, recovery of the observer's translation direction is straightforward. If the camera moves relative to a static rigid world then the points described by the vectors $\mathbf{\Gamma}$ all lie in a plane in angular velocity space (figure 2). If $\hat{\mathbf{u}} = \frac{\mathbf{U}}{|\mathbf{U}|}$, is defined to be a unit vector parallel to \mathbf{U} then taking the dot product of $\hat{\mathbf{u}}$ with (3) gives,

$$\mathbf{\Gamma} \cdot \hat{\mathbf{u}} = -\mathbf{\Omega} \cdot \hat{\mathbf{u}}, \quad (4)$$

which is a constant if $\hat{\mathbf{u}}$ and $\mathbf{\Omega}$ are constant, hence all $\mathbf{\Gamma}$ lie in a plane (see figure 2.)

3 Motion Segmentation With The Angular Velocity Field

If there is an object in the field of view moving independently of the background then provided that its velocity is not parallel to the camera's velocity, two planes will exist in angular velocity space (see figure 2).

If the distinct object is undergoing general motion then let $\mathbf{\Psi}$ be the angular velocity and \mathbf{U}_o the translational velocity of the object. Let \mathbf{r} be a vector from the centre of mass

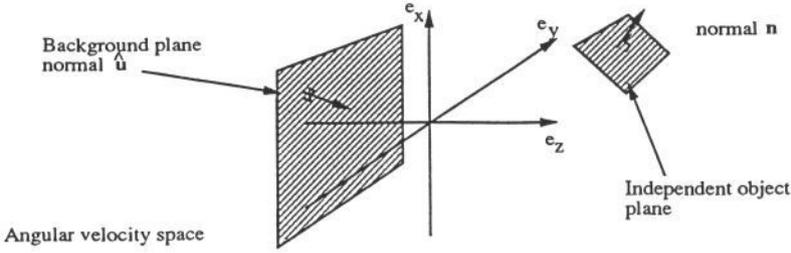


Figure 2: An object moving independently of the background will give rise to a distinct plane in Γ if \hat{u} is not parallel to $\mathbf{n} = (\mathbf{U} - \mathbf{U}_o + (\Psi \times \mathbf{q}_o))$ or $\Omega \cdot \hat{u} \neq \mathbf{n} \cdot (\Psi - \Omega)$

of the object to a point on its surface, each point on the object will have velocity \mathbf{v} where,

$$\mathbf{v} = \mathbf{U}_o + \Psi \times \mathbf{r}. \quad (5)$$

If the center of mass of the object is at \mathbf{q}_o relative to the camera then the motion of the object can be equivalently described by an angular velocity Ψ about the optical center of the camera and an additional translational velocity,

$$\mathbf{v} = \mathbf{U}_o + \Psi \times (\mathbf{r} + \mathbf{q}_o) - \Psi \times \mathbf{q}_o, \quad (6)$$

or substituting $\mathbf{q} = \mathbf{q}_o + \mathbf{r}$ gives,

$$\mathbf{v} = \mathbf{U}_o + \Psi \times \mathbf{q} - \Psi \times \mathbf{q}_o. \quad (7)$$

The motion of the object is now described by an angular velocity Ψ about the optical center of the camera and a translational velocity $(\mathbf{U}_o - \Psi \times \mathbf{q}_o)$. The angular velocity field associated with the rotating object then becomes,

$$\Gamma = \frac{(\mathbf{U} - \mathbf{U}_o + (\Psi \times \mathbf{q}_o)) \times \mathbf{q}}{|\mathbf{q}|^2} + \Psi - \Omega. \quad (8)$$

If an object in the field of view is rotating then the plane associated with it in angular velocity space will be separated from that associated with the background. The distance to the plane D will be given by,

$$D = (\mathbf{U} - \mathbf{U}_o + (\Psi \times \mathbf{q}_o)) \cdot (\Psi - \Omega), \quad (9)$$

and the normal to the plane \mathbf{n} will be parallel to, $(\mathbf{U} - \mathbf{U}_o + (\Psi \times \mathbf{q}_o))$. The normal to the second plane may thus be different from \hat{u} .

Segmenting a moving image then corresponds to finding planes in angular velocity space. The main advantage of this technique is that for an object in independent motion in a scene, points do not need to be grouped together as having come from a single rigid object before being used to solve for the object's motion relative to the camera.

4 Recovering the Angular Velocity Field

The angular velocity field is related to the spherically projected flow field by,

$$\Gamma \times \mathbf{Q} = \mathbf{V}. \quad (10)$$

Cross products are non-invertible and hence only the two components of Γ perpendicular to \mathbf{Q} may be directly recovered using (10); the third component, parallel to an image direction \mathbf{Q} , is not generally directly recoverable. However there exist generic image

positions where the component of Γ along a line of sight may be estimated. From (3) the component of Γ along \mathbf{Q} is,

$$(\Gamma \cdot \mathbf{Q})\mathbf{Q} = -(\Omega \cdot \mathbf{Q})\mathbf{Q}. \quad (11)$$

This component corresponds to local rotation of the image around a given image direction \mathbf{Q} . It is only possible to recover this component if the magnitude of the shear in the flow field at \mathbf{Q} is zero. It is proven in [5] that bounds can be placed on the magnitude of the rotation around the line of sight as follows,

$$\Omega_{min}^{max} = \frac{1}{2}(-\text{curl} \pm \text{shear magnitude}) \quad (12)$$

The magnitude of shear will be zero when either the surface in the field of view is perpendicular to an image direction or when the camera motion direction coincides with an image direction.

5 Results

The algorithm has been successfully tested as a motion direction estimator for the case of strict translation. A plane was fitted by least squares to a set of Γ estimated from a real scene using the motions of 'Droid' corners. It still remains to use the algorithm for independent motion segmentation and to implement the rotation recovery portion based on the first derivative of flow. It is envisaged that points in the flow field with zero shear would be used to solve for the camera's angular velocity relative to the 'background' of the scene and that subsequently this value of Ω be used in the recovery of Γ throughout the entire image. This assumed value will be incorrect for an independent object that is rotating but the difference in virtual velocity given in section 3 should be sufficient to segment it from the back ground.

6 Conclusion

Once estimated the angular velocity field is a much more tractable way of dealing with visual motion information than the conventional defined optical flow field. The former readily solves the motion segmentation problem and provides a natural framework for fixating visual systems (a fixating system is one in which an additional controlled angular velocity is added to the motion of the camera). There are difficulties associated with the estimation of the first derivatives of flow but these problem will be addressed in a future paper.

7 Acknowledgments

We acknowledge discussions with Professor J. M. Brady, Dr A. Zisserman, R. Cipolla and S. Smith of the Robotics Research Group. We are grateful for support from the SERC and the EEC Esprit program.

References

- [1] B.K.P.Horn, J.E.Weldon Jr 1987, Computationally Efficient Methods for Recovering Translational Motion, IEEE ICCV.
- [2] S.J.Maybank 1985, The Angular Velocity Associated the Optical Flow Field from Motion Through a Rigid Environment, Proc. Roy. Soc. pp 317 - 326.
- [3] H.G.Longuet-Higgins 1980, The Interpretation of a Moving Retinal Image Proc. Roy. Soc. pp 385 - 397.
- [4] R.C.Nelson 1991 Qualitative Detection of Motion by a Moving Observer Proc. Conf. Computer Vision and Pattern Recognition, 91.
- [5] M. Subbarao 1990 Bounds on Time to Collision and Rotational Component from First-Order Derivatives of Image Flow Computer Vision, Graphics and Image Processing 50, pp 329-341 (1990)