

Object and Motion Recognition using Plane Plus Parallax Displacement of Conics

Douglas R. Heisterkamp
University of South Alabama
Mobile, AL 36688-0002, USA
dheister@jaguar1.usouthal.edu

Prabir Bhattacharya
University of Nebraska-Lincoln
Lincoln, NE 68588-0115, USA
prabir@cherokee.unl.edu

Keywords: Invariants, Conics, Parallax, Motion Segmentation, Object Recognition

Abstract

Parallax displacement is used to determine a relative 3D conic projective structure. This value is invariant between any number of views in time if the conic is not moving with respect to the plane of the homography. It can be used to determine conic correspondence between three simultaneous views. The corresponding conics may then be used to determine the epipolar geometry. This method of determining conic correspondence works with unknown and even changing epipolar geometry. The relative 3D conic projective structure may be used to segment groups of conics which have consistent motion.

1 Introduction

Traditional computer vision has had success using structured environments such as manufacture lines and using expensive high quality calibrated cameras. To take advantage the availability of low cost uncalibrated camera, the existing techniques need to be adapted and new techniques created. Projective invariants and image warping are two such techniques that remove the need for expensive calibrated cameras. This work is focused on the geometric image feature of conics in the plane plus parallax framework. The plane plus parallax framework is a subset of image warping. Conics are a fundamental image feature. Many natural and man-made object have circular shapes. In addition, many curves may be approximated with conics. Conics are always mapped to conics under a general projective transformation. Conics have been very useful in graphics and modeling [2]. Conics have numerous uses in vision [11, 14, 15].

Image warping has been used in a number of vision areas [4, 17, 16, 1, 8]. A reference view is used. The other views are warped (transformed) back to this view. In the plane plus parallax approach, the transformation is based on one plane in space. The points on this plane are mapped to their corresponding points on the reference image. This plane in space is often called the *dominant plane* as it is selected because it has the most points (i.e. a background plane such as a wall).

Optical flow is often used to calculate this transformation [10, 9, 7, 12, 3]. The points not on this plane in space are not mapped to their corresponding points in the reference image. The displacement between the reference image point and the warped point is called the *planar parallax displacement* (it is also called the *parallax error*, or just parallax).

Previous work has been concerned with points and in some cases lines [1, 18, 17, 20]. This work is concerned with conics. A visible conic in space is projected to a conic in each view. If the conic is not on the dominant plane, then there is a parallax displacement of the warped conic on the reference image. This work is concerned with representing and using that displacement.

2 Plane Plus Parallax Framework

The relative 3D projective structure of points was present in [8]. Similar approach may be found in [16]. The plane plus parallax framework is based on the planar homography that exists between a plane in space and an image plane. Between two image planes, there exist a homography, \mathbf{H} , that maps the projected points of the space plane in one image onto the corresponding points in the second image. For points not on the space plane, the homography does not map the points in one image to their corresponding points in the second image. The error between the homography mapped point and the corresponding point is called the *planar parallax displacement*.

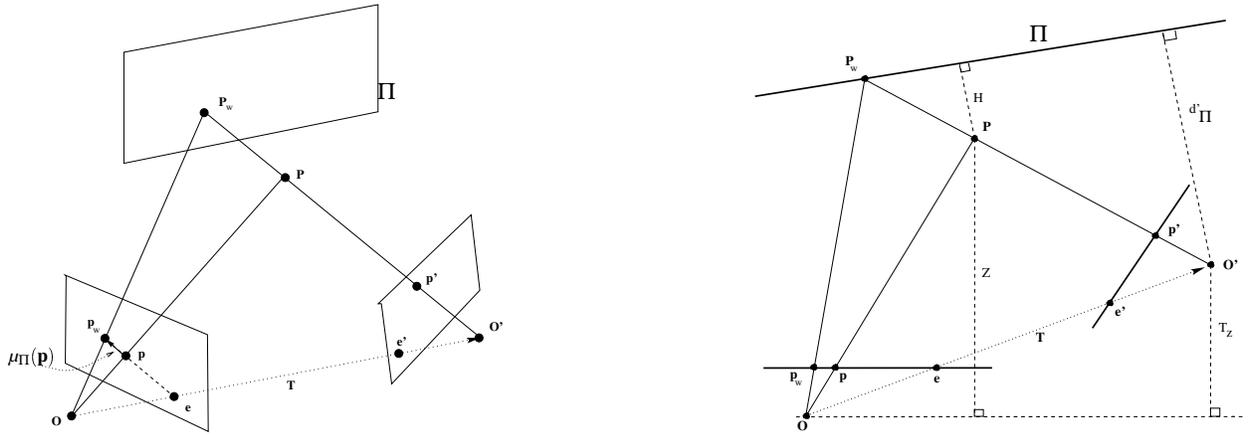


Figure 1: Geometry of Plane plus Parallax

The geometry of the planar parallax is presented in Figure 1 using the notation from [8]. The planar parallax displacement $\hat{\mu}$ can be expressed as $\hat{\mu} = \gamma \frac{T_z}{d_{\Pi}(\hat{\mathbf{e}} - \hat{\mathbf{p}}_w)}$ where T_z is the z component of $\hat{\mathbf{T}}$ and γ is a measure of the 3D shape of the point \mathbf{P} , in particular, $\gamma = \frac{H}{Z}$. When $T_z = 0$ then $\hat{\mu} = \frac{\gamma}{d_{\Pi}} \hat{\mathbf{T}}$.

If $\hat{\mu}_1$ and $\hat{\mu}_2$ are the parallax displacement vectors of two points that belong to the static back-

ground, then their *relative 3D projective structure* $\frac{\gamma_2}{\gamma_1}$ is given by

$$\frac{\gamma_2}{\gamma_1} = \frac{\hat{\mu}_2^T(\Delta\hat{\mathbf{p}}_w)_\perp}{\hat{\mu}_1^T(\Delta\hat{\mathbf{p}}_w)_\perp} \quad (1)$$

where $(\Delta\hat{\mathbf{p}}_w)_\perp$ is a vector in the direction that is perpendicular to $\hat{\mathbf{p}}_{w2} - \hat{\mathbf{p}}_{w1}$. If the two points belong to the static background over three frames then the *parallax based rigidity constraint* is the statement that the relative 3D projective structure does not change. Based on the parallax motion, the rigidity constraint can be used to detect inconsistencies in the 3D motion of two points without knowledge of camera geometry, camera motion or structure parameters. Irani and Anandan (see [8]) used the rigidity constraint to detect consistently moving 3D object using an uncalibrated camera by clustering.

3 Planar Parallax Displacement of Conics

3.1 Point Correspondence

The point correspondence of C and C_w is determined by the epipolar lines. That is, corresponding points will lie on the same epipolar line. The first theorem will allow us to make the point correspondence without explicitly using the epipole. Corresponding points will lie on the same line of the pencil determined by a pair of tangents common to both C and C_w . Of course, Theorem 1, shows that the line is an epipolar line. The proofs of all Lemmas and Theorems of this section may be found in [5].

Theorem 1 *Let C be the original conic on the reference image, let C_w be the conic C' warped back to the reference image. If the conic C' corresponds to the conic C , (C and C' are images of the same space conic at the same time instance), then a pair of tangents, common to both C and C_w , will intersect at the epipole.*

Two conics will have four common tangents (real or imaginary). That means at most six possible points of intersection. We want the constraint that when a line of the pencil formed by two common tangents has a real intersection with one conic, it has a real intersection with the other conic. Sometimes there may be two or more pairs of tangents that satisfy the constraint. In that case we either pick the pair whose point of intersection is consistent with the epipolar flow fields of the rest of the image data or just repeat the calculations of the 3D structure for each pair of common tangents.

For example see Figure 2, the four common tangents are \mathbf{v}_0 , \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . They have six intersections with the first three, \mathbf{e}_0 , \mathbf{e}_1 , and \mathbf{e}_2 labeled. Looking at the possible epipole \mathbf{e}_2 , the

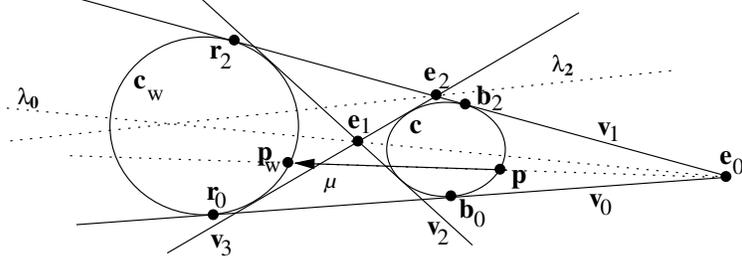


Figure 2: Reference Image Geometry of Conics under Plane plus Parallax

epipolar line λ_2 intersects conic \mathbf{c}_w at real points but intersects conic \mathbf{c} at imaginary points. Thus it is not a possible epipole and we can discard it (this corresponds to the constraint that will be presented in Equation (2)). Similarly for the other three nonlabeled intersection points. The leaves only \mathbf{e}_0 and \mathbf{e}_1 as possible epipoles. The geometry for \mathbf{e}_1 to be an epipole is such that the dominant plane must be behind one of the camera's. In most cases this is not a possibility. When a plane not based on image data (such as using the reference image plane as the dominant plane) this case could be valid. Thus \mathbf{e}_1 can be discarded (this corresponds to the constraints that will be presented in equation (3)). Of the six intersections, only \mathbf{e}_0 is left and it is the epipole.

When a pair of common tangents are imaginary, but intersect at a real point, then that point is in the interior of both conics. Then any real lines of the pencil will intersect both conics in real points. When both common tangents are real and intersect at a real point, then the next couple of constraints are needed. The notation used will follow from that used in Figure 2, \mathbf{b}_0 , \mathbf{b}_2 , \mathbf{r}_0 , and \mathbf{r}_2 are the four contact points of the common tangents, \mathbf{v}_0 and \mathbf{v}_1 . A $\hat{\cdot}$ over a vector means that it is normalized to a fixed representation. For points, this is $z = 1$ and for lines it is a unit vector.

The first constraint enforces the restriction that an epipolar line which intersects one conic in a real point must intersect the other conic in a real point. The second constraint is that the points $\hat{\mathbf{b}}_0$ and $\hat{\mathbf{r}}_0$ lie on the same side of line $\hat{\mathbf{v}}_1$ and similarly, that the points $\hat{\mathbf{b}}_2$ and $\hat{\mathbf{r}}_2$ line on the same side of the line $\hat{\mathbf{v}}_0$.

$$\text{Constraint 1: } (\hat{\mathbf{v}}_1^T \hat{\mathbf{b}}_0) (\hat{\mathbf{v}}_0^T \hat{\mathbf{b}}_2) \leq 0 \quad \text{and} \quad (\hat{\mathbf{v}}_1^T \hat{\mathbf{r}}_0) (\hat{\mathbf{v}}_0^T \hat{\mathbf{r}}_2) \leq 0 \quad (2)$$

$$\text{Constraint 2: } (\hat{\mathbf{v}}_1^T \hat{\mathbf{b}}_0) (\hat{\mathbf{v}}_1^T \hat{\mathbf{r}}_0) \geq 0 \quad \text{and} \quad (\hat{\mathbf{v}}_0^T \hat{\mathbf{b}}_2) (\hat{\mathbf{v}}_0^T \hat{\mathbf{r}}_2) \geq 0 \quad (3)$$

A conic may be parameterized by using the Bernstein form of a conic (see, [2] or [5], for background). Using that approach, we find $\mathbf{c}(t)$, the parametric form of \mathcal{C} , and $\mathbf{c}_w(t)$, the parametric form of \mathcal{C}_w , and the complementary segments, $\bar{\mathbf{c}}(t)$, $\bar{\mathbf{c}}_w(s)$ as

$$\mathbf{c}(t) = (1-t)^2 \hat{\mathbf{b}}_0 + 2t(1-t) \alpha \hat{\mathbf{e}} + t^2 \hat{\mathbf{b}}_2 \quad \mathbf{c}_w(s) = (1-s)^2 \hat{\mathbf{r}}_0 + 2s(1-s) \alpha_w \hat{\mathbf{e}} + s^2 \hat{\mathbf{r}}_2 \quad (4)$$

$$\bar{\mathbf{c}}(t) = (1-t)^2 \hat{\mathbf{b}}_0 - 2t(1-t) \alpha \hat{\mathbf{e}} + t^2 \hat{\mathbf{b}}_2 \quad \bar{\mathbf{c}}_w(s) = (1-s)^2 \hat{\mathbf{r}}_0 - 2s(1-s) \alpha_w \hat{\mathbf{e}} + s^2 \hat{\mathbf{r}}_2. \quad (5)$$

where $\mathbf{e} = \mathbf{v}_0 \times \mathbf{v}_1$ and the scale factors α and α_w are

$$\alpha = \pm \sqrt{\frac{(\hat{\mathbf{b}}_0 \times \hat{\mathbf{b}}_2) \mathbf{A}^{-1} (\hat{\mathbf{b}}_0 \times \hat{\mathbf{b}}_2)}{(\hat{\mathbf{b}}_0 \times \hat{\mathbf{b}}_1) \mathbf{A}^{-1} (\hat{\mathbf{b}}_1 \times \hat{\mathbf{b}}_2)}}. \quad (6)$$

As t ranges from 0 to 1, $\mathbf{c}(t)$ traces the segment of the conic from point $\hat{\mathbf{b}}_0$ to point $\hat{\mathbf{b}}_2$ that is closest to $\hat{\mathbf{e}}$ and $\bar{\mathbf{c}}(t)$ traces the segment of the conic from point $\hat{\mathbf{b}}_0$ to point $\hat{\mathbf{b}}_2$ that is farther from $\hat{\mathbf{e}}$. Now we can set up the correspondence between the points of C and C_w .

Theorem 2 *With the parametric forms, $\mathbf{c}(t)$ and $\bar{\mathbf{c}}(t)$, of the conic C , and the parametric forms, $\mathbf{c}_w(s)$ and $\bar{\mathbf{c}}_w(s)$ of the conic C_w (see equations (4), and (5)), then the point $\mathbf{c}_w(s_\lambda)$ corresponds to the point $\mathbf{c}(t_\lambda)$ and the point $\bar{\mathbf{c}}_w(s_\lambda)$ corresponds to the point $\bar{\mathbf{c}}(t_\lambda)$ with λ ranging from zero to one. The parameter values, t_λ and s_λ , are the solutions to the quadratic equations $L(\lambda)^T \mathbf{c}(t) = 0$ and $L(\lambda)^T \mathbf{c}_w(s) = 0$ that lie in the range zero to one.*

The correspondences of points between the conics C and C_w is defined completely by the two common tangents and their points of contact.

3.2 Relative 3D Conic Projective Structure

Over the points of a conic, the 3D shape measure γ has a maximum value γ_{\max} and a minimum value γ_{\min} . A point of the conic with γ_{\max} will have the maximum value of the γ -ratio of the points of the conic with respect to any other point. A similar result holds for a point with γ_{\min} . The γ -ratio of γ_{\min} to γ_{\max} can be recovered from the γ -ratios of the points of the conic with respect to any point \mathbf{p} since $\frac{\gamma_{\min}}{\gamma_{\mathbf{p}}} / \frac{\gamma_{\max}}{\gamma_{\mathbf{p}}} = \frac{\gamma_{\min}}{\gamma_{\max}}$.

We propose using $\frac{\gamma_{\min}}{\gamma_{\max}}$ as the *relative 3D conic projective structure*. Since the relative 3D conic projective structure is the same with respect to any point \mathbf{p} , the calculations are simplified by using the points of contact of the common tangents as \mathbf{p} . The calculation of the relative 3D conic projective structure consists of finding the maximum and minimum γ -ratio of the points on the conic with respect to a point. For numeric stability, the point $\hat{\mathbf{b}}_0$ was used when $0.5 \leq \lambda \leq 1$ and the point $\hat{\mathbf{b}}_2$ was used when $0 \leq \lambda \leq 0.5$. Though in this section an iterative method was used to solve for the minimum and maximum, with a little more analysis a closed form solution for the minimization and maximization problems may be determined.

Since the relative 3D conic projective structure is based on the relative 3D projective structure of points, the parallax based rigidity constraint holds. Thus it may be used to detect inconsistent 3D motion of a conic over three or more frames without the need for camera geometry, camera motion, or shape parameters in the same way that the point based rigidity constraint is used. In addition, the specialization of a single point is removed. In the point case, all of the relative 3D

projective structure is calculated with respect to a single reference point. Thus it depends on the accuracy of the reference point and of its correspondences throughout the image sequence. The relative 3D conic projective structure does not depend on a reference point. It is the same with respect to any reference point and is, in particular, being calculated with respect to points that are known to correspond – the points of contact of the common tangents.

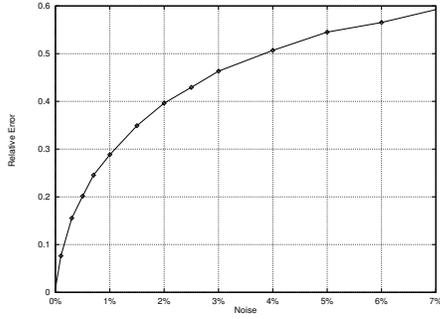
4 Experimental Results

4.1 Results under Noise using Synthetic Data

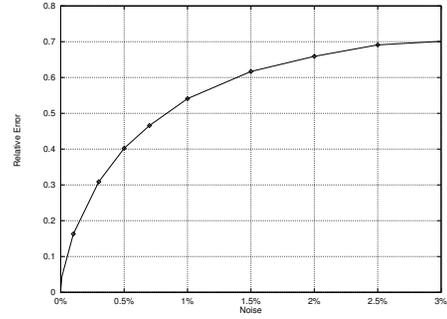
By Theorem 1, the epipole of the reference image with respect to a second image may be determined by the intersection of a pair of tangents common to both the reference conic and the warped conic. The stability of calculating epipoles using this method is investigated following a method similar to [13]. Noise was added to the geometric properties of the image conics. For a noise percent a , the lengths of the major and minor axis were perturbed by a uniform random value in the range $\pm a$ percent of axis length. The location of the center of the ellipse was varied by the $\pm a$ percent of the average of the axis lengths. The orientation of the axis was varied by $\pm a$ percent of 45 degrees. Noise was added to the planar homography by adding noise to the points used to calculate the homography. The noise level of 1% corresponds to applying a uniform random noise in the range of ± 1 pixel to the points in a 512×512 image. The relative error for each coordinate of the epipole, as defined by [13], is $\min \left\{ \frac{|x-x_0|}{\min(|x_0|, |x|)}, 1 \right\}$. The relative error of the epipoles was calculated for five thousand space conics. The graph of relative error when noise is applied to image conic only is presented Figure 3(a) and when noise is applied to both image conics and the planar homography is presented in Figure 3(b).

The stability of the epipoles calculated by the conic planar parallax is comparable to stability the epipoles calculated by the linear method but worse than the non-linear methods (see Figure 6 in [13]). The stability is better than the single plane method. The results presented by [13, Table 3] show a 0.486 relative error for as small as 0.4 pixel Gaussian image noise. From Figure 3(b), a similar relative error occurs at 0.75% noise level.

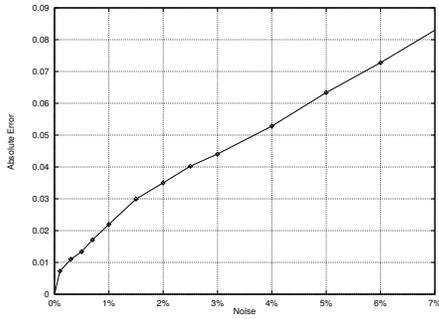
The stability of the $\gamma_{\min}/\gamma_{\max}$ ratio was investigated. The experiment setup is the same as the epipolar calculation, but instead of calculating the epipole, the relative 3D conic projective structure was calculated. The stability was tested with five hundred space conics. The graphs of the absolute error is presented in Figures 3(c) and 3(c).



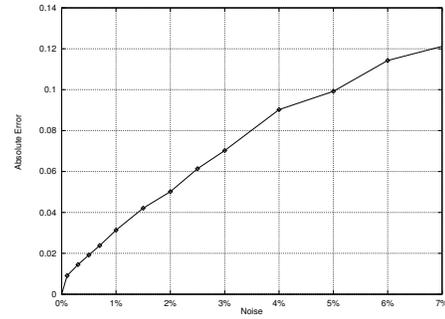
(a) Relative error of epipoles calculated by intersecting common tangents



(b) Relative error of epipoles calculated by intersecting common tangents with noise applied to both image conics and homography



(c) Absolute error of $\gamma_{\min}/\gamma_{\max}$ with noise added to image conics



(d) Absolute error of $\gamma_{\min}/\gamma_{\max}$ with noise to both conic and homography

Figure 3: Stability of epipoles and relative 3D conic projective structure

4.2 Results using Real Images

The experimental setup similar to that used in [6]. Images were taken from three different camera orientations and four different times. Between each time, the box with the attached conics was moved a small distance. The three different camera views are labeled A, B, and C. View A at time 1 is picked as the reference image. A subset of the images are presented in Figures 4 and 5. The end of the large box was used as the dominant plane.

Conic 9 is the outer edge of the CD in the images. It remains stationary with respect to the dominant background plane. The value of the relative 3D conic projective structure of conic 9 remains basically the same over the view and times, see Table 1. Conic 4 is the conic on the right side of the box. The box is moving between different times. From Table 1, it is seen that at time 1, the value is basically the same for the different views, but at later times the value changes significantly. Thus with three simultaneous views, we can determine the correspondence of moving conics and over a sequence of time segment the different types of motion.

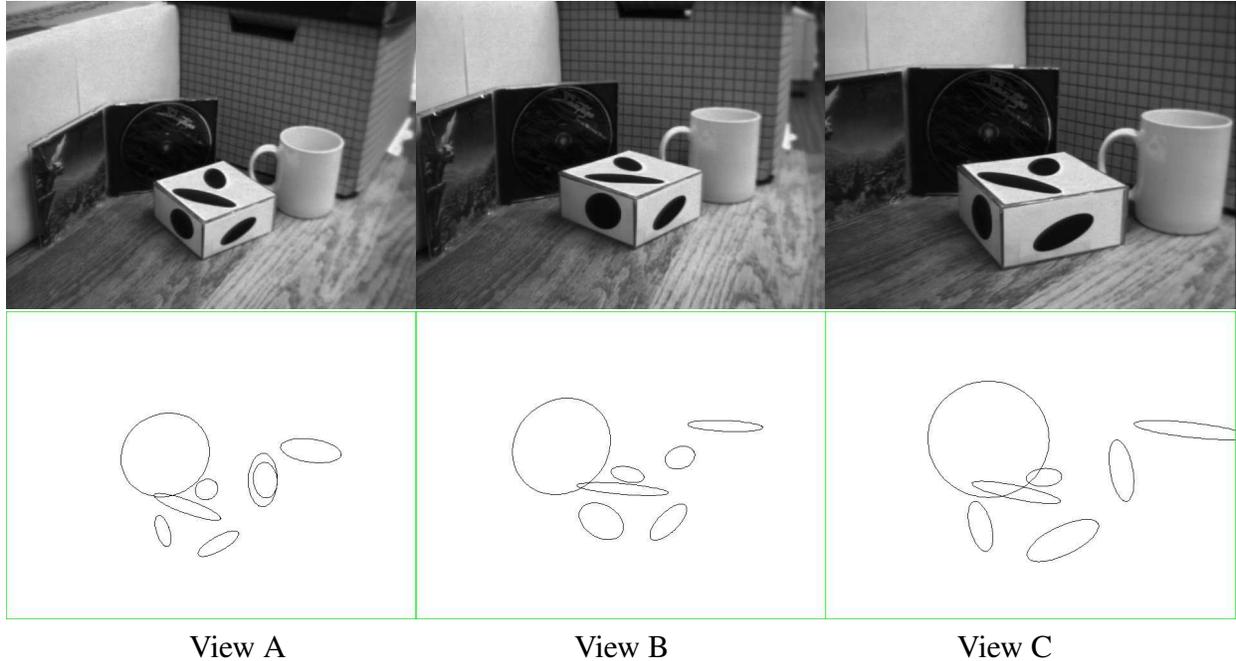


Figure 4: Three views with detected conics at time 1

An example of the relative 3D conic projective structure calculated for conic 9 as it was matched against each conic from view C is presented in Table 3(a). From the table, it is seen that the relative 3D conic projective structure can discriminate between conics. The cluster of epipoles presented in Table 3(b) is also tight.

Table 1: Relative 3D conic projective structure ($\gamma_{\min}/\gamma_{\max}$)

Reference View	View B	View C	View C	View C	View C
Time 1	Time 1	Time 1	Time 2	Time 3	Time 4
(Fig. 5a)	(Fig. 5b)	(Fig. 5c)	(Fig. 6a)	(Fig. 6b)	(Fig. 6c)
Conic 9	0.382	0.378	0.379	0.378	0.378
Conic 4	0.870	0.868	0.835	0.796	0.752

5 Summary

It is useful to exploit the geometry of conics under the plane plus parallax framework. Conic correspondence and epipolar geometry may be determined from three views. Motion analysis may then be conducted as in [19]. Motion segmentation may be conducted as in [21]. The relative 3D conic projective structure is defined based on the relative 3D projective structure of points. The relative 3D conic projective structure has an advantage over the relative 3D projective structure of points in that it is not dependent on the correspondence of a single point. The relative 3D

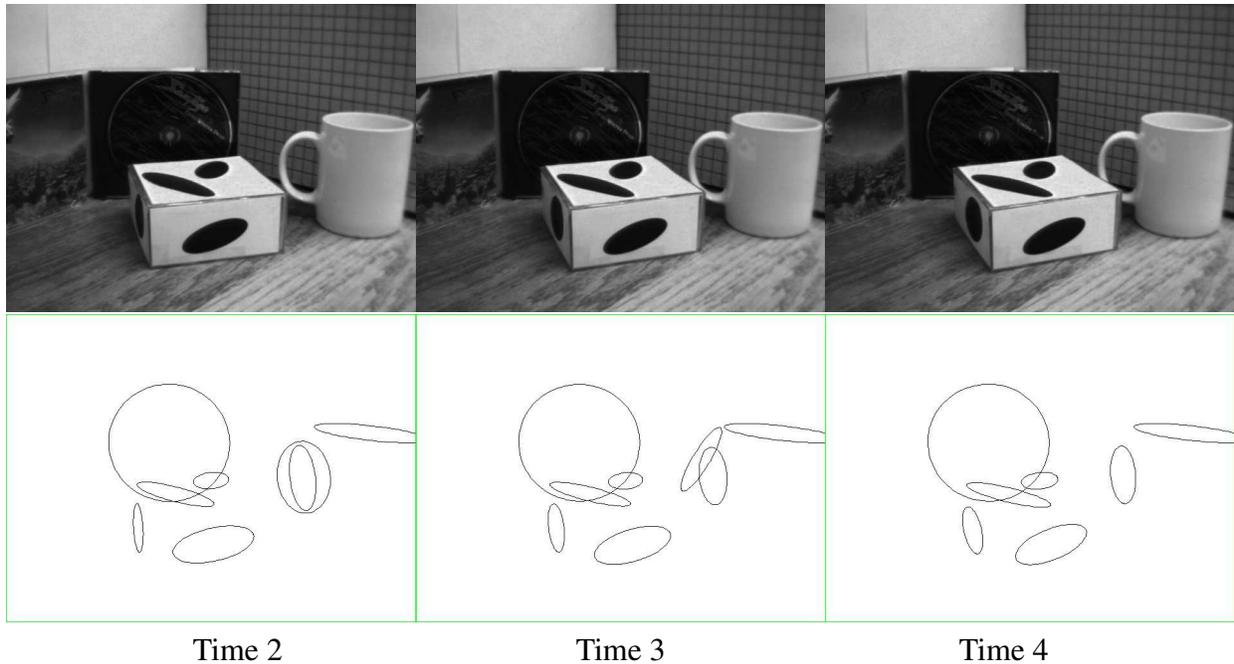


Figure 5: View C at three additional times

conic projective structure is invariant to camera motion and may be used to segment multiple moving objects as done for the point structure in [8]. The relative 3D conic projective structure can discriminant be conics in real images and has a stable response to noise.

References

- [1] P. Anandan, K. Hanna, and R. Kumar. Shape recovery from multiple views: A parallax based approach. In *ICPR*, pages A:685–688, Jerusalem, Israel, October9–13, 1994.
- [2] G. Farin. *NURBS Curves and Surfaces from Projective Geometry to Practical Use*. A K Peters, Wellesley, MA, 1995.
- [3] P. Fornland and C. Schnorr. Determining the dominant plane from uncalibrated stereo vision by a robust and convergent iterative approach without correspondence. In *Symposium on Image Analysis*, pages 11–15, 1997.
- [4] G. Hager and K. Toyama. X vision: Combining image warping and geometric constraints for fast visual tracking. In *ECCV*, pages II:507–517, Cambridge, England, April 1996.
- [5] D. R. Heisterkamp. *Investigations on Conic Invariants with Application to Object Recognition and Motion Analysis*. PhD thesis, University of Nebraska-Lincoln, August 1997.
- [6] D. R. Heisterkamp and P. Bhattacharya. Invariants of families of coplanar conics and their application to object recognition. *Journal of Mathematical Imaging and Vision*, 7(3):253–267, June 1997.
- [7] S. Hsu, P. Anandan, and S. Peleg. Accurate computation of optical flow by using layered motion representations. In *ICPR*, pages A:743–746, Jerusalem, Israel, October9–13, 1994.

Table 2: Example values of relative 3D conic projective structure and epipolar calculations

Conic	epipole	$\gamma_{\min}/\gamma_{\max}$
9	[713.4, 698.8, 1.0]	0.378
7	[561.3, 209.0, 1.0]	0.116
5	[523.4, 256.3, 1.0]	0.518
4	[199.7, 331.2, 1.0]	0.249
3	[445.6, 443.7, 1.0]	0.427
2	[319.6, 241.6, 1.0]	0.141
1	[317.6, 267.7, 1.0]	0.002

(a) Reference image conic 9 compared with all conics from view C, time 1

Time	epipole
1	[713.4, 698.8, 1.0]
2	[701.3, 682.0, 1.0]
3	[705.4, 686.9, 1.0]
4	[708.5, 685.6, 1.0]

(b) Calculated epipoles with respect to view C based on matches of conic 9.

- [8] M. Irani and P. Anandan. A unified approach to moving object detection in 2D and 3D scenes. In *ICPR*, page A8E.3, Vienna, Austria, August25–28, 1996.
- [9] M. Irani, B. Rousso, and S. Peleg. Detecting and tracking multiple moving objects using temporal integration. In *ECCV*, pages 282–287, Santa Margherita Ligure, Italy, May 1992.
- [10] M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparent motions. *IJCV*, 12(1):5–16, February 1994.
- [11] K. Kanatani and W. Liu. 3D interpretation of conics and orthogonality. *CVGIP: IU*, 58(3):286–301, November 1993.
- [12] Q. Luong, R. Deriche, O. Faugeras, T. Papadopoulo, A. Mitiche, and P. Bouthemy. Computation and analysis of image motion: A synopsis of current problems and methods. *IJCV*, 19(1):29–55, July 1996.
- [13] Q.-T. Luong and O. D. Faugeras. The fundamental matrix: Theory, algorithms and stability analysis. *IJCV*, 1(17):43–76, January 1996.
- [14] S. D. Ma. Conics-based stereo, motion estimation and pose determination. *IJCV*, 10(1):7–25, February 1993.
- [15] L. Quan. Conic reconstruction and correspondence from two views. *IEEE Trans. PAMI*, 18(2):151–160, February 1996.
- [16] H. Sawhney. 3D geometry from planar parallax. In *CVPR*, pages 929–934, Seattle, WA, June 21–23, 1994.
- [17] H. Sawhney. Simplifying motion and structure analysis using planar parallax and image warping. In *ICPR*, pages A:403–408, Jerusalem, Israel, October9–13, 1994.
- [18] H. Sawhney, S. Ayer, and M. Gorkani. Model-based 2D and 3D dominant motion estimation for mosaicing and video representation. In *ICCV*, pages 583–590, Cambridge, MA, June 20–22, 1995.
- [19] Shapiro, Zisserman, and Brady. 3D motion recovery via affine epipolar geometry. *IJCV*, (16):147–182, 1995.
- [20] A. Shashua. Projective structure from uncalibrated images: Structure-from-motion and recognition. *IEEE Trans. PAMI*, 16(8):778–790, August 1994.
- [21] G. Xu and S. Tsuji. Correspondence and segmentation of multiple rigid motions via epipolar geometry. In *ICPR*, page A74.4, Vienna, Austria, August25–28, 1996.