ESTIMATING CORRESPONDENCE BETWEEN MULTIPLE CAMERAS USING JOINT INVARIANTS

Raman Arora, Yu Hen Hu[*]

University of Wisconsin-Madison
Department of Electrical and Computer Engg.
1415 Engineering Drive, Madison, WI-53706

Charles Dyer

University of Wisconsin-Madison
Department of Computer Sciences
1210 W. Dayton St., Madison, WI-53706

ABSTRACT

The joint invariants of the projective group $\mathrm{PSL}(3,\mathbb{R})$ on $\mathbb{RP}^2$, the five-point volume cross-ratios, are studied to address the problem of correspondence in a camera network. The distribution of cross-ratios over the unit square as well as in a small local-neighbourhood of a reference point are found to have a heavy tail. No cross ratio value is unique but the collection of five point cross ratios generated by taking all possible combination of five points completely prescribes the curve. Sections of the signature submanifold that admit large enough variation of cross ratios are found to be sufficient in providing correspondence across wide perspectives. Such invariant signatures may be collected independently at cameras with different viewpoints and shared, thereby achieving the registration of objects in the image. Experimental results with license plate database are provided.

***Index Terms*—** Image registration, object recognition, computer vision.

1. INTRODUCTION

The invariant based classification schemes find utility in their ability to reduce the set of possible matches and speed up the search of similar classes or objects. These approaches have drawn a lot of attention in recent decades in the areas of computer vision and pattern recognition [1, 2]. The invariant based methods may be classified as global or local: the global invariants utilize the entire image to compute feature values whereas local invariants are computed typically from a much smaller subsets. For instance, Fourier descriptors are global whereas curvature is local. Local descriptors are more desirable due to their robustness to occlusions and noise. However, one of the fundamental problems with the use of local descriptors is that of correspondence or image-registration. The lack of correspondence (across multiple views) between the regions on which the local invariant features are computed renders any kind of classification ineffective. We describe a method in this paper that uses joint projective invariants to classify objects while simultaneously achieving correspondence between multiple views captured in a camera sensor network.

Projective (or perspective) invariants are the image descriptors that remain invariant to view-point. Projective invariants have been computed in several settings and applied to various computer vision tasks like localization [3, 4], autonomous navigation [5] and are particularly desirable for 3D scene analysis and surveillance [6]. A few researchers have focused on the probabilistic analysis of application of projective invariants. In [7] and [8], a probability distribution is derived for the four-point-cross-ratio, a classical planar projective invariant, under different assumptions on the distribution of the four points. The distribution of cross ratios is further examined in [9] as more constraints on relative distances of the four points are imposed. The performance of the cross ratios is described quantitatively in terms of probability of rejection and false alarm in [10]. Unfortunately, in all of the works mentioned above, the correspondence was assumed apriori. Without the correspondence information, the classification methodology breaks down since the cross ratios are not unique. This paper presents a simple approach to address the issue of correspondence. Although we focus on five-point volume cross-ratios that comprise the fundamental joint invariants of the projective group $\mathbb{P}\mathrm{SL}(3)$ on $\mathbb{RP}^2$ [11], the ideas presented here are applicable to other projective invariants.

The paper is organized as follows. Section 2 describes the object recognition problem in a multi-view camera network setting. The probabilistic analysis of five-point volume cross-ratios is provided in Section 3. Section 4 presents an invariant-signature based algorithm for image registration and recognition and discusses its application to the license plate database. The test dataset was generated by capturing images of license plates from various angles and distances. The images have been pre-processed to extract the binary images encoding the contour (boundary) of the alpha-numerical characters on the license plates [12].

[*]The authors have been partially supported by the National Science Foundation under Grant No. CCF-0434355, and the University of Wisconsin Graduate school Award NO. MSN119059.

2. PROBLEM FORMULATION

Let C_1, C_2 be two cameras with different perspectives of an overlapping 3D scene. Let $\{O_{1,i}\}_{i=1}^{M}, \{O_{2,j}\}_{j=1}^{N}$ be the sets of planar curves, extracted through pre-processing, in the images captured at sensors C_1 and C_2 respectively. Given a pair of curves $(O_{1,i}, O_{2,j})$ the object recognition problem is to determine if the two curves represent the same object in the scene or not. The five-point joint invariants on the planar curves are used to address the classification problem.

Let G be a Lie group acting on a manifold M. An $(n+1)-$*point joint invariant*, $I(z^0, \ldots, z^n)$, of the transformation group G is defined to be a function that is invariant to the joint action of G on the Cartesian product $M^{\times(n+1)}$ given by

$$g \cdot (z^0, \ldots, z^n) = (g \cdot z^0, \ldots, g \cdot z^n), \qquad (1)$$

where $g \in G$ and $z^0, \ldots, z^n \in M$.

The transformation group that we are interested in is the projective transformation group $\mathrm{PSL}(3, \mathbb{R})$ acting on the 2-dimensional projective space $M = \mathbb{RP}^2$. An element of the projective transformation group is described by the matrix $\begin{pmatrix} A & b \\ c^T & d \end{pmatrix} \in GL(3, \mathbb{R})$ where A is a 2×2 matrix, b, c are 2×1 vectors and d is a scalar. A point in the planar image, $z \in \mathbb{R}^2$, gets transformed to the point $w \in \mathbb{R}^2$, given by the group action

$$w = g \cdot z = \frac{Az + b}{c \cdot z + d}. \qquad (2)$$

The Geometric First Main Theorem for the Projective Group [11] states that every five-point joint invariant for the action of $\mathrm{PSL}(3, \mathbb{R})$ on $\mathbb{RP}^2$ is generated by the following cross ratios,

$$CR(0; 1, 2, 3, 4) = \frac{V(0,1,2)V(0,3,4)}{V(0,1,4)V(0,2,3)}, \qquad (3)$$

and

$$CR(1; 0, 2, 3, 4) = \frac{V(0,1,2)V(1,3,4)}{V(0,1,4)V(1,2,3)}, \qquad (4)$$

where $V(i, j, k)$ is the area of the triangle defined by z^i, z^j and z^k. The cross ratio defined in (3) is described as the ratio of the product of the areas of the non-shaded triangles in Figure 1 and the product of areas of shaded triangles.

3. DISTRIBUTION OF JOINT INVARIANTS

This section presents probabilistic analysis of the five point cross ratios as it may be applied to the computer vision problems. First, empirical probability distributions are generated for joint invariants on the unit square as well as on the test dataset. Second, we discuss that small perturbations of the points result in small excursions of cross ratios. This assures the robustness of joint invariant based method against noise. Finally, it is shown that given a cross-ratio value, the corresponding five point set on any given curve is not unique.

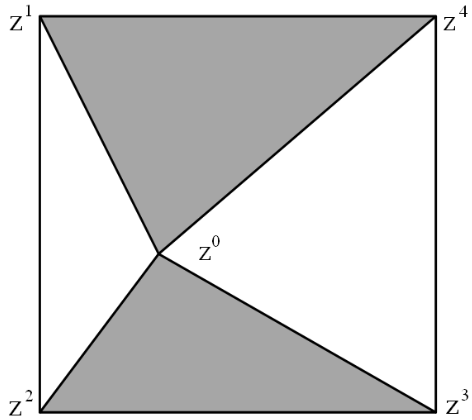

Fig. 1. Five point projective joint-invariant is the ratio of product of areas of non-shaded triangles and shaded triangles.

3.1. Probability distribution of cross ratios

The first step toward probabilistic analysis of five point cross ratio is to study its distribution. The probability distribution of the classic four point cross ratios was studied in [7, 8, 10] and a closed-form expression was given under various assumptions on the underlying distribution of points in the plane. Assuming that the points are identically and uniformly distributed over the unit square, the empirical probability distribution function (pdf) of five point volume cross ratios is plotted in Figure 2. The pdf exhibits peaks at cross ratio values equal to zero and one and is symmetric about 0.5. However, as we impose spatial separation on the points (uniformly distributed but with different means), the pdf transforms and is no longer symmetric. The distribution of cross ratios on the license-plate contour-database [12] also follows the general form as seen in Figure 2. An important observation from these plots is the heavy-tail of the pdf(s). The dotted line represents the probability that corresponding cross ratio values lie outside the interval $[-1000, 1000]$. Most of such large cross ratios are observed as five point set approaches singularity points of I in $\mathbb{R}^{5\times 2}$ (for cross ratio in equation (3), these are the points where z^0, z^1, z^4 or z^0, z^2, z^3 are almost collinear).

3.2. Local distribution of cross ratios

Next, we investigate the local distribution of the cross ratios and the effect of perturbation of points on the corresponding cross ratios. Given five points z^i with Cartesian coordinates (x_i, y_i), for $i = 0, 1, 2, 3, 4$, the area of the triangle described by z^i, z^j, z^k is given as

$$V(i, j, k) = \begin{vmatrix} x_i & x_j & x_k \\ y_i & y_j & y_k \\ 1 & 1 & 1 \end{vmatrix}. \qquad (5)$$

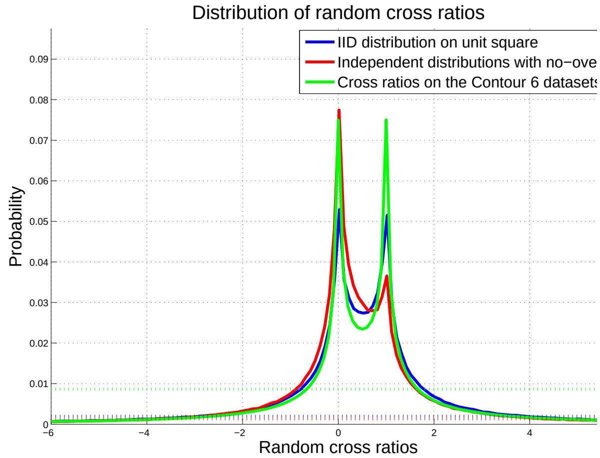

Fig. 2. Distribution of random five point volume cross ratios.

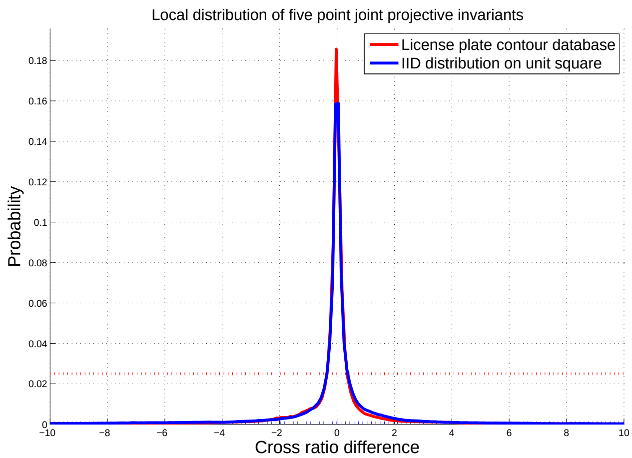

Fig. 3. Local distribution of cross ratios around a reference five point set.

Let Δz^i denote the jittered point $z^i + \Delta^i$. Then

$$
\begin{aligned}
&\Delta\mathrm{CR}_0(0;1,2,3,4)\\
&\quad = \mathrm{CR}(\Delta 0;1,2,3,4) - \mathrm{CR}(0;1,2,3,4)\\
&\quad = \frac{V(\Delta 0,1,2)\,V(\Delta 0,3,4)}{V(\Delta 0,1,4)\,V(\Delta 0,2,3)} - \frac{V(0,1,2)\,V(0,3,4)}{V(0,1,4)\,V(0,2,3)}
\end{aligned}
$$

Now,

$$
\begin{aligned}
V(\Delta 0,1,2) &= \begin{vmatrix} x_0+\Delta_x & x_1 & x_2 \\ y_0+\Delta_y & y_1 & y_2 \\ 1 & 1 & 1 \end{vmatrix}\\
&= (x_0+\Delta_x)(y_1-y_2) - x_1(y_0+\Delta_y-y_2)\\
&\qquad + x_2(y_0+\Delta_y-y_1).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
V(\Delta 0,1,2) - V(0,1,2) &= \begin{vmatrix} \Delta_x & x_1 & x_2 \\ \Delta_y & y_1 & y_2 \\ 0 & 1 & 1 \end{vmatrix}\\
&= \Delta_x(y_1-y_2) - \Delta_y(x_1-x_2)
\end{aligned}
$$

Let $M = \max\{|x_1-x_2|, |y_1-y_2|\}$. Then choosing $\epsilon_1 > 0$, such that

$$\max(|\Delta_x|, |\Delta_y|) < \epsilon_1 \cdot |V(0,1,2)|/M,$$

we get,

$$|V(\Delta 0,1,2)| < |V(0,1,2)| \cdot (1 \pm \epsilon_1). \tag{6}$$

Similar relationships hold for other volumes, yielding

$$\left|\frac{\Delta\mathrm{CR}_0(0;1,2,3,4)}{\mathrm{CR}(0;1,2,3,4)}\right| < \left|\frac{(1\pm\epsilon_1)(1\pm\epsilon_2)}{(1\pm\epsilon_3)(1\pm\epsilon_4)} - 1\right| < \epsilon. \tag{7}$$

The jitter analysis implies that the joint invariants are relatively robust to small amount of noise. Figure 3 shows the distribution of cross-ratio-differences around a given reference set of five points. The empirical pdf is in agreement with the analysis above: small jitter results in little change in cross ratios.

3.3. Uniqueness of cross-ratios

Equipped with the probabilistic analysis of the cross ratios, it is now argued that no single cross ratio is unique and should not be used for image registration or recognition. First note that on almost any continuous planar curve, there exists five-point set with cross ratio equal to (in limit) zero (by construction, simply choose three points arbitrarily close so that they are almost collinear). This corresponds to a point $u \in \mathbb{R}^{5\times 2}$. A permutation of these points will also result in the cross ratio of infinity (in limit). This is point $v \in \mathbb{R}^{5\times 2}$. The jitter analysis in the last section shows that cross ratio map is a smooth open map. Along the straight line, connecting u to v, all the intermediate cross ratios are observed. And since two lines can intersect at most at one point, but for the cross ratio value at the point of intersection, all other cross ratio values are repeated. Finally, there are an uncountable number of points like u (again by construction).

4. JOINT INVARIANT SIGNATURES

Owing to their non-uniqueness, single cross ratios are ineffective at classification. However, the signature manifold comprising of cross ratio values generated by all possible combinations of five points on the planar curve prescribes the entire curve. A 1-D slice is extracted by fixing four of the five points on the signature manifold. The representation of the original curve with this submanifold is unique up to a projective transformation [11]. Short sections of the submanifold that overlap with a singularity point provide good discrimination.

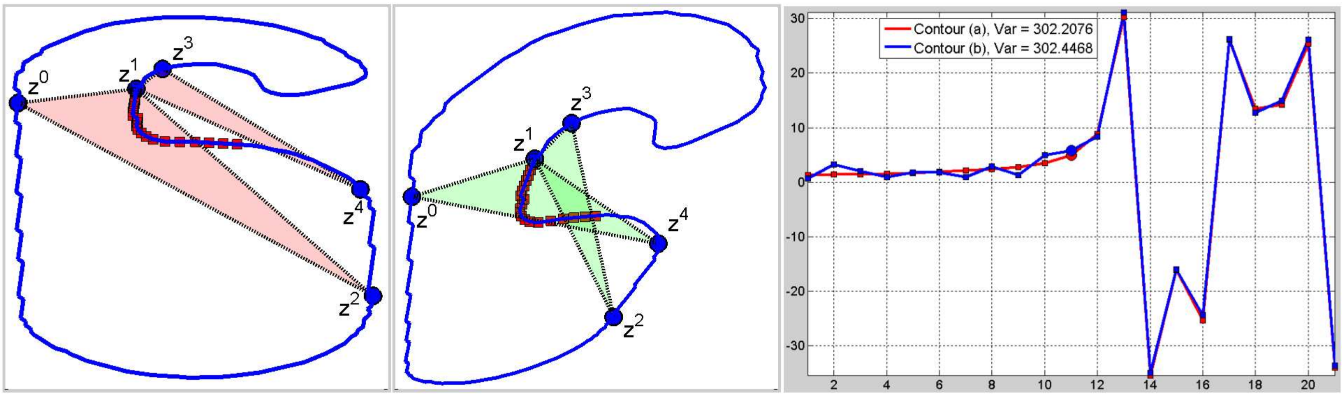

Fig. 4. (a) Contour of digit 6 extracted from license plate images at camera 1 and (b) camera 2. Shaded triangles in (a)/(b) appear in the numerator/denominator of the cross ratio, respectively. (c) Invariant joint signatures in the two contour images.

Note that matching proceeds after appropriate thresholding of cross ratios around the singularity points.

Figure 4 shows contour plots from the license plate test dataset along with invariant signatures. Figure 4(a),(b) show the contours of digit 6 (extracted from images of the license plate 67724QB) from two different viewpoints. The set of five points on contours that generated the invariant signatures (in Figure 4(c)), are highlighted with symbols. The invariant signature comprising of 21 cross ratio values (given by equation (4)) was obtained by translation of z^1 along the contour. The initial set of points is marked by circles and the jittered set is marked by squares. Note that the change of sign of the cross ratios is attributed to the flipping of the triangle defined by z^0, z^1, z^4 as z^1 is perturbed. Numerous such signatures were generated for contour (a) and the matching signatures in contour (b) were unique with probability close to one.

5. CONCLUSIONS AND FUTURE DIRECTIONS

This paper discusses the challenges of image registration or correspondence in the multiple camera setting. It is argued that five point joint invariants for projective transformations of planar curves lack the uniqueness of any single cross ratio. But with high probability the objects can be uniquely described by sections of signature submanifolds. We are currently working on simultaneous correspondence and classification based on joint invariant signatures. Future work will focus on distributed computation and matching of signatures at various nodes in a large camera network.

6. REFERENCES

[1] Joseph L. Mundy and Andrew Zisserman, Eds., *Geometric Invariance to Computer Vision*, MIT Press, 1992.

[2] Joseph L. Mundy, Andrew Zisserman, and David Forsyth, Eds., *Applications of Invariance in Computer Vision*, Springer-Verlag LNCS, 1993.

[3] Kyoung Sig Roh, Wang Heon Lee, and In So Kweon, “Obstacle detection and self-localization without camera calibration using projective invariants,” in *Proc. Intelligent Robots and Systems*, 1997, pp. 1030–1035.

[4] Bruno M. Marhic, El Mustapha Mouaddib, and Claude Pegard, “A localisation method with an omnidirectional vision sensor using projective invariant,” in *Int. conf. Intelligent Robots and Systems*, 1998, pp. 1078–1083.

[5] Vassilios S. Tsonis, Konstantinos V. Chandrinos, and Panos E. Trahanias, “Landmark-based navigation using projective invariants,” in *Proc. Int conf. Intelligent Robots and Systems*, 1998, pp. 342–347.

[6] Senem Velipasalar and Wayne Wolf, “Frame-level temporal calibration of video sequences from unsynchronized cameras by using projective invariants,” in *Proc. Advanced Video Signal-based Surveillance (AVSS)*, 2005, pp. 462–467.

[7] Stephen J. Maybank, “Probabilistic analysis of the application of the cross ratio to model based vision: Misclassification,” *International Journal of Computer Vision*, vol. 16, pp. 5–33, 1995.

[8] Kalle Astrom and Luce Morin, “Random cross ratios,” in *In Proc. 9th Scandinavian Conf. on Image Analysis*, 1995, pp. 1053–1061.

[9] D. Q. Huynh, “The cross ratio: A revisit to its probability density function,” in *The Eleventh British Machine Vision Conference (BMVC)*, Sept. 2000.

[10] Stephen J. Maybank, “Probabilistic analysis of the application of the cross ratio to model based vision: Misclassification,” *IJCV*, vol. 14, pp. 199–210, 1995.

[11] Peter J. Olver, “Joint invariant signatures,” *Foundations of computational mathematics*, vol. 1, pp. 3–67, 2001.

[12] UW Camera Network group, “License plate database,” http://www.cae.wisc.edu/∼raman/camnet/JICR5.