Visualizing Projective Shape Space

John T. Kent

University of Leeds, Leeds LS2 9JT, UK
j.t.kent@leeds.ac.uk

Abstract. Projective shape consists of the information in a configuration of points invariant under projective transformations. It is usually studied through projective invariants, the most familiar example being the cross ratio for four collinear points. In this paper a standardized representation of the configuration is investigated which is better suited for quantitative comparisons between different projective shapes.

1 Introduction

Kent and Mardia [1] gave a new approach to the study of projective shape through the use of "Tyler standardized" configurations. In this paper we shall give some further details about the implications of this standardization. We start with a review of the projective shape of k landmarks in $\mathbb{R}^m$, $m \geq 1$, and describe the relevance to machine vision. The classical approach to projective shape is through the use of projective invariants. However, the use of Tyler standardization is better suited for the construction of metrics to compare quantitatively different projective shapes. A detailed examination is given for the 1-dimensional case, $m = 1$, especially for $k = 4$ landmarks. Sections 1–3 are largely review; Section 4 is mainly new material.

Start with a configuration $X_0(k \times m)$ of k points or landmarks in m dimensional space. The projective shape of X_0 consists of the information in X_0 that is invariant under projective transformations. The easiest way to deal with projective transformations is to introduce homogeneous coordinates. Thus introduce an augmented matrix

$$X = \left[X_0\ 1\right],$$

where 1 is a k-vector of ones. Then X is a $k \times p$ matrix, where throughout the paper we set $p = m + 1$. We can write X in terms of its rows and columns as

$$X \begin{bmatrix} x_1^T \\ \vdots \\ x_k^T \end{bmatrix} = \left[x_{(1)}, \ldots, x_{(p)}\right],$$

where x_i^T denotes the ith row (so that x_i without the transpose denotes a column vector), and $x_{(j)}$ denotes the jth column, $j = 1, \ldots, p$.

From the point of view of homogeneous coordinates each row x_i of X is well-defined only up to a scalar multiple. Indeed the i row of X_0 can be recovered from

F. Nielsen and F. Barbaresco (Eds.): GSI 2013, LNCS 8085, pp. 335–342, 2013.

the multiple $\lambda_i x_i$ ($\lambda_i \neq 0$) through simple division, $(\lambda_i x_{i1}, \ldots, \lambda_i x_{im})/(\lambda_i x_{ip})$. Note that the coordinate system for X is slightly more general than that for X_0 because the ith row of X can have a vanishing pth entry, $x_{ip} = 0$, corresponding to the ith row of X_0 lying "at ∞".

In homogeneous coordinates X represents an equivalence class of matrices,

$$X \equiv \{DX : D = \text{diag}(d_i) \text{ is } k \times k \text{ diagonal with nonzero diagonal entries}\}.$$

A projective transformation on X_0 corresponds simply to a linear transformation on X, i.e. $X \to XB^T$, where B is $p \times p$ nonsingular. If

$$B = \begin{bmatrix} B_0 & \gamma \\ 0_m & 1 \end{bmatrix}$$

then B represents an affine transformation of X_0, $x_i^{(0)} \to B_0 x_i^{(0)} + \gamma$. If B is replaced by cB, $c \neq 0$, then the projective transformation is unchanged.

Projective shape consists of the information in X_0 or X that is unchanged under projective transformations. Algebraically, the projective shape of X can be viewed as the equivalence class of configurations

$$[X] = \{DXB^T : D(k \times k) \text{ diagonal nonsingular}, B(p \times p) \text{ nonsingular}\}.$$

Projective shape is important in computer vision when using a camera to take a film image in $m = p - 1$ dimensions of a scene in p dimensions containing k points lying in an m-dimensional hyperplane. The most important cases in practice are $k \geq 4$ collinear points ($m = 1$) and $k \geq 5$ coplanar points ($m = 2$). The film image depends on the focal point of the camera and on the position of the film. However, the projective shape is invariant under these choices and can be recovered from the film image. For more information about the use projective geometry in computer vision, see e.g. [2] or [3].

2 Representations of Projective Shape

The description of projective shape in terms of equivalence classes of configurations is elegant mathematically; however, it is difficult to use this approach to make quantitative comparisons between different projective shapes. For this purpose we need to construct a metric on projective shape space.

In this section we describe two methods that have been used to give more explicit representations of projective shape. The first method involves projective invariants. Although these invariants are a classic tool in differential geometry, they cannot be used directly to construct of a metric on projective shape space. The second method is Tyler standardization, which does lead naturally to several choices of metric.

2.1 Projective Invariants

Projective invariants of $[X]$ are usually defined as ratios of products of determinants of $p \times p$ matrices. For example, if $p = 3$ and $k \geq 5$, a typical projective invariant is given by

$$\frac{|\,[x_1\ x_2\ x_3]\,|\ |\,[x_3\ x_4\ x_5]\,|}{|\,[x_1\ x_3\ x_4]\,|\ |\,[x_2\ x_3\ x_5]\,|}. \tag{1}$$

The key properties are (i) each index should appear at most once in each determinnt, and (ii) each index should appear as many times in the numerator as in the denominator. Under these conditions it is easy to check that (1) is unchanged if X is replaced by DXB^T.

The simplest case is $k = 4, p = 2$ when there is essentially just one projective invariant, the cross ratio, one definition of which is

$$\tau = \frac{|\,[x_1\ x_2]\,|\ |\,[x_3\ x_4]\,|}{|\,[x_1\ x_3]\,|\ |\,[x_2\ x_4]\,|}. \tag{2}$$

Any other ordering of the indices leads to an invariant related to τ by a one-to-one transformation. There are 6 choices in all,

$$\tau,\ 1-\tau,\ 1/(1-\tau),\ 1/\tau,\ -(1-\tau)/\tau,\ -\tau/(1-\tau). \tag{3}$$

Projective invariants provide a concrete coordinate system to represent projective shape. But unfortunately they are not suitable for metric comparisons.

2.2 Tyler Standardization

Using results of Tyler ([4] and [5]), it was shown in [1] that up to mild regularity conditions on X, it is possible to find a member of the equivalence class such that

$$X^T X = (k/p)I_p, \quad x_i^T x_i = 1,\ i = 1, \ldots, k. \tag{4}$$

That is, up to a scaling constant the columns of X are orthonormal, and the rows of X are unit vectors. Then X is unique up to (i) multiplication on the right by a $p \times p$ orthogonal matrix and (ii) the choice of sign of each row.

The arbitrary choice of orthogonal transformation can by removed by looking at the inner product matrix $M = XX^T$, with elements,

$$M = (m_{ij}), \quad m_{ij} = x_i^T x_j,\ i, j = 1, \ldots, k. \tag{5}$$

The arbitrary sign for each row of X, or equivalently each row and column of M, can be removed by squaring each element. Define a matrix N by

$$N = (n_{ij}), \quad n_{ij} = m_{ij}^2. \tag{6}$$

Then N can be viewed as a matrix-valued projective invariant.

If $p = 2$, it can be shown that N determines the Tyler standardized configuration X, up to an orthogonal transformation on the right and a possible sign change of each row. Work is in progress to determine an analogous set of summary features which determine the projective shape when $p > 2$. Hence for the remainder of the paper, we restrict attention to the case $p = 2$.

3 The General Case $p = 2$

If X is Tyler standardized, then each row of X is a unit vector and takes the form $x_i^T = v(\theta_i)^T,\ i = 1, \ldots, k$, for some collection of angles $\{\theta_i\}$, where $v(\theta)^T = (\cos\theta, \sin\theta)$. The various indeterminacies in X can be summarized as follows.

(a) The sign change from x_i to $-x_i$ corresponds to replacing θ_i by $\theta_i + \pi$.
(b) A rotation of X corresponds to replacing θ_i by $\theta_i - \psi$ for all $i = 1, \ldots, k$, for some angle ψ.
(c) A reflection of X about the direction $\theta = 0$ correponds to replacing θ_i by $-\theta_i, i = 1, \ldots, k$.

The indeterminacies can be removed as follows:

(a) Angle doubling. Replace θ_i by $\phi_i = 2\theta_i$, and set $y_i = v(\phi_i),\ i = 1, \ldots, k$. Let $Y(k \times 2)$ denote the corresponding angle-doubled configuration.
(b,c) Inner products. The inner products for the angle-doubled landmarks take the form

$$y_i^T y_j = \cos(\phi_i - \phi_j), = 2\cos^2(\theta_i - \theta_j) - 1 = 2n_{ij} - 1.$$

Thus the inner products for Y contain the same information as N, and do not depend on the indeterminacies in X.

Several proposals were given in [1] for metrics on projective shape space, two of them based on the embedding defined by N. One choice is given by Euclidean distance between the N matrices for two projective shapes. The other choice is given by Euclidean distance between the abs(M) matrices. For a Tyler standardized configuration, the matrix abs(M) is defined by taking elementwise absolute values in M; that is abs(M) = (abs(m_{ij})) has elements

$$\text{abs}(m_{ij}) = |m_{ij}| = n_{ij}^{1/2}.$$

4 The Special Case $p = 2,\ k = 4$

In this case the matrix X takes a particularly simple form. For later convenience, label the landmarks by the letters A, B, C, D. Up to the labelling of the landmarks, the sign of each row, and rotation, it is shown in [1] that the θ angles take the form

$$(\theta_A, \theta_B, \theta_C, \theta_D) = (-\delta/2, \delta/2, \pi/2 - \delta/2, \pi/2 + \delta/2). \tag{7}$$

When the 4 landmarks in X are distinct, the angle δ lies in the interval $(0, \pi/2)$. Note that landmarks A and C are orthogonal, as are landmarks B and D. The corresponding inner product matrix M takes the form

$$M = \begin{bmatrix} 1 & \cos\delta & 0 & -\sin\delta \\ \cos\delta & 1 & \sin\delta & 0 \\ 0 & \sin\delta & 1 & \cos\delta \\ -\sin\delta & 0 & \cos\delta & 1 \end{bmatrix}. \tag{8}$$

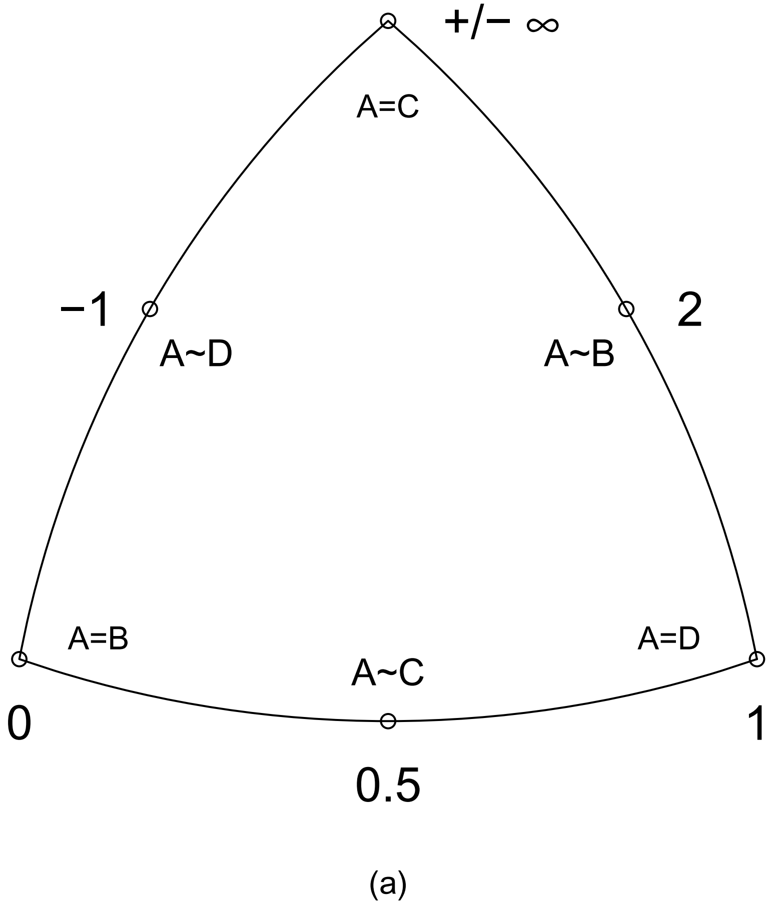

Fig. 1. Representation of projective shape space, $p = 2,\ k = 4$ as spherical triangle, by embedding abs(M) in $\mathbb{R}^3$. The vertex labels, e.g. $A = B$ describe the landmark coincidences, the edge labels, e.g. $A \sim B$, describe the landmark separation, and the outer labels give the corresponding values of the cross ratio. Figure adapted from [1].

Ignoring the diagonal element, each row is a unit vector of length 3 with a structural zero. Under relabelling of the landmarks, the structural zero can lie in one of three places. After taking absolute values, all 4 rows contain the same information. Hence $\frac{1}{2}$abs(M) can be isometrically embedded as a spherical triangle on the unit sphere in $\mathbb{R}^3$; see Figure 1.

It is worth emphasing several features of this representation.

(a) *edge labels.* The edge labels indicate the "separation" properties of the 4 landmarks. For example, if the landmarks are ordered $ABCD$, then A is separated from C (and also B is separated from D); this separation is written concisely as $A \sim C$.

(b) *vertex labels.* The vertex labels indicate the landmark coincidences, which always occur in pairs under Tyler standardization. For example the coincidence $A = B$ also implies $C = D$, and is written concisely as $A = B$.
(c) *angle* δ. Each edge of the spherical triangle is indexed by the angle δ. As δ moves through the interval $(0, \pi/2)$, the projective shape moves from one vertex to the other, in a counterclockwise direction, say. More information on visualizing and interpreting δ is given in Section 4.
(d) *relationship to cross ratio.* Each value of the cross ratio τ in the extended real line corresponds to one point on this triangle. Figure 1 shows the correspondences for selected values of the cross ratio. The effects on the cross ratio in (3) from relabelling the landmarks correspond to one of six possible reflections and rotations of this figure.

Thus every projective shape can by described by an angle δ and a discrete label giving one of the 3 edges. Each edge includes $24!/3 = 8$ orderings of the 4 landmarks.

4.1 Why Does the Triangle for Projective Shape Space Have Corners?

The spherical triangle representation of projective shape space has 3 distinct corners or vertices. However, the reason for such corners is not clear when looking at the cross ratio.

Represent the landmarks by four real numbers A, B, C, D. The cross ratio in (2) can be written as $\tau = \{A - B)(C - D)\}/\{A - C)(B - D)\}$. Hold $A < C < D$ fixed and let B vary. If we allow B to vary through the extended real line, then the cross ratio varies in a bijective fashion through the extended real line. If we avoid the singularity at $B = D$, then the cross ratio is an infinitely differentiable function of B. In particular, there is no hint of a singularity as B passes through A and C. At these points the cross ratio takes the values 0 and 1, respectively, corresponding to two of the vertices in projective shape space. Hence it is natural to ask where the singularities (i.e. vertices or corners) come from in our representation for projective shape shape? This question can be addressed from several perspectives.

(a) The first answer is that when B approaches one of the other three landmarks, Tyler standardization forces the other two landmarks to come together as well. Thus the three situations where B matches one of the other landmarks are single-pair singularities in the simple cross ratio description, but are actually double-pair singularities in the Tyler-standardized description.
(b) Further, there are two distinct ways to move away from a singularity in terms of the ordering of the landmarks, where for the moment we allow all four landmarks to move. Suppose the singularity corresponds to $A = B < C = D$. After breaking the singularity, there are two choices. The first choice involves two subchoices,

$$A < B << C < D \quad \text{or} \quad B < A << D < C,$$

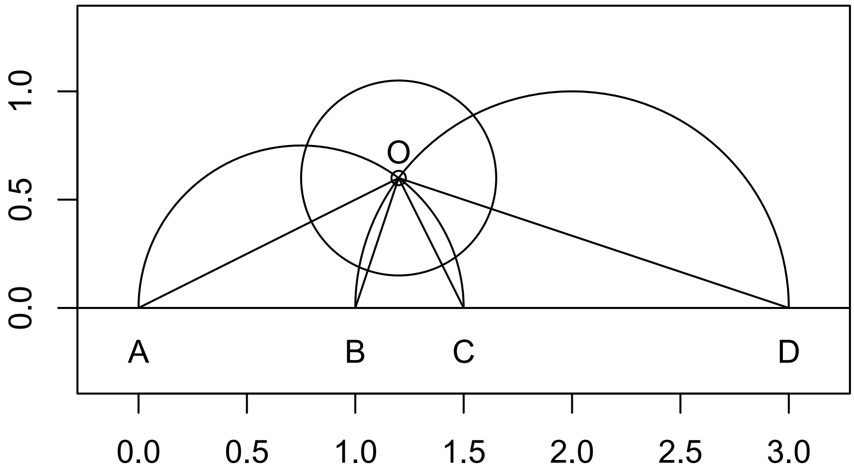

Fig. 2. Geometric construction of the focal point which yields Tyler standardized coordinates, starting from four collinear points

both of which correspond to moving from the vertex into one particular edge of the triangle (labelled $A \sim C$ in Figure 1). The notation indicates that two landmarks in each pair are close together, relative to the distance between the two pairs.

The other choice involves the two subchoices

$$A < B << D < C \quad \text{or} \quad B < A << C < D,$$

which corresponds to moving into the other edge (labelled $A \sim D$ in Figure 1) at that vertex.

(c) The phrase "mild regularity conditions" in the construction of a Tyler standardized version of X hides some subtleties. Although the regularity condition always holds without any problems for projective shapes lying in the interior of each edge, it holds only in a limiting sense at the vertices.

(d) Another way of looking at the lack of Tyler regularity at the vertices can be given in terms of the double angle configuration Y. This matrix has rank 2 for a configuration lying on the interior of an edge, but it only has rank 1 at the vertices. For a projective shape lying at a vertex, two of the $\{y_i\}$ describe the same direction on the circle; the other two lie in the opposite direction. Hence all four of the $\{y_i\}$ lie on the same line through the origin.

4.2 Does the Angle δ Have a More Direct Geometric Interpretation?

The answer is yes. Consider four landmarks on the line in increasing order, $A < B < C < D$, and construct two semicircles in the upper half plane, the first having A and C at opposite sides of a diameter, and the other having B and D at opposite sides of a diameter; see Figure 2. These two semicircles intersect at a point O, say. This point is precisely the choice of focal point which is needed to produce Tyler standardized coordinates. Figure 2 shows a circular "film" about O. Treating the focal point as the origin and projecting the data onto the film yields a Tyler-standardized configuration. For this figure the cross ratio is $\tau = 0.5$ with corresponding angle $\delta = \pi/4$.

To understand why, note that the angle AOC is a right angle since the corresponding triangle is inscribed in a semicircle; so is BOD. Thus the four lines through O form two orthogonal frames. Let δ denote the angle AOB. If the landmarks A, C, D are held fixed and B is allowed to vary between A and C, then δ lies in the range $(0, \pi/2)$ and the corresponding projective shape lies on the bottom edge of projective shape space ($A \sim C$) in Figure 1. Further, if the coordinate system on the circular film is chosen to that the zero direction lies midway between the rays OA and OC, then the angular coordinates of the four landmarks take the form given in (7).

In many ways projective shape space for the case of four collinear landmarks ($k = 4,\ m = 1$) is very special. It is possible to visualize projective shape space completely in this case, and the singular points are very distinctive. However, many of the general principles extend to higher values of k and m. In particular, the Procrustes approach allows questions about similarities and differences between different projective shapes to be tackled independently of a particular view of a set of k landmarks in an m-dimensional image of a scene in $\mathbb{R}^p,\ p = m + 1$, provided the landmarks lie in an m-dimensional hyperplane.

Acknowledgement. This work benefited from a visit to the Statistical and Applied Mathematics Sciences Institute (SAMSI), North Carolina and a workshop at the Mathematical Biosciences Institute (MBI), Ohio, especially through helpful discussions with Ian Dryden, Thomas Hotz, Huiling Le, Stephan Huckemann, Kanti Mardia, Ezra Miller, and Vic Patrangenaru.

References

1. Kent, J.T., Mardia, K.V.: A geometric approach to projective shape and the cross ratio. Biometrika 99, 833–849 (2012)
2. Faugeras, O., Luong, Q.-T.: The Geometry of Multiple Images. MIT Press, Cambridge (2001)
3. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge Univ. Press, Cambridge (2000)
4. Tyler, D.E.: A distribution-free M-estimator of multivariate scatter. Ann. Stat. 15, 234–251 (1987)
5. Tyler, D.E.: Statistical analysis for the angular central Gaussian distribution on the sphere. Biometrika 74, 579–589 (1987)