Invariant Signatures for Planar Shape Recognition under Partial Occlusion

Alfred M. Bruckstein  Robert J. Holt  Arun N. Netravali  Thomas J. Richardson

Technion - IIT Haifa, Israel  AT&T Bell Laboratories, Murray Hill, New Jersey

Abstract

A planar shape distorted by a projective viewing transformation can be recognized under partial occlusion if an invariant description of its boundary is available. Recent research in this area has provided a theory for invariant boundary descriptions based on an interplay of differential, local, and global invariants. Differential invariants require high-order derivatives. The use of global invariants and point matches on the distorting transformations enables one to reduce the order. Trade-offs between the highest order derivatives required and the quantity of additional information constraining the distorting viewing transformations are made explicit.

1 Introduction

In [1–13] the problem of recognizing planar shapes with smooth or piecewise smooth boundaries when these shapes are distorted by a projective-type viewing transformation and perhaps only partially visible has been considered. A statement of the problem is, see e.g. [4]: given a planar curve described by \( P(t) = [x(t), y(t)] \) with an arbitrary parametrization and a distorting viewing transformation \( T_\psi : \mathbb{R}^2 \to \mathbb{R}^2 \) described by a vector of parameters \( \psi \) can we efficiently test whether a curve segment \( Q(\tilde{t}) \) is a portion of \( T_\psi[P(t(t))] \) for some \( \psi \) and some reparametrization \( t(\tilde{t}) \)? In the sequel we show that for the common viewing transformations there exist signature functions that can be associated to suitably reparametrized versions of planar curves, so that if \( Q(\tilde{t}) \) is indeed a portion of a distorted and reparametrized \( P(t) \) then the respective signature functions will match over the corresponding intervals. Thus the recognition problem is reduced to partial function matching.

This paper is organized as follows. Sections 2, 3, and 4 deal with viewing transformations of increasing complexity: Euclidean (rigid) motion, affine transformations, and projective transformations. Each section has a similar form. We first address the issue of finding invariant reparametrizations. Sometimes linearly scaled reparametrizations are easier to find, we consider these also. After reparametrization, a signature function, \( \rho \), is determined for \( P(\tau) \) which is invariant under \( T_\psi \), i.e. \( \rho(\tilde{\tau}) = \rho(\tau + \tau_0) \) when \( P = T_\psi[P] \).

To accomplish this one needs either very precise descriptions of the local behavior of the curves \( P(t) \) (enabling the extraction of higher derivatives, see e.g. [4, 5, 12, 13]) or some further information constraining \( T_\psi : \mathbb{R}^2 \to \mathbb{R}^2 \), in the form of point matches or line matches or explicit knowledge of some of the parameters \( \psi \), see e.g. [1, 2, 6–9]. For each transformation we consider how point matches can be incorporated into the invariants to reduce the order of derivatives required.

If an invariant reparametrization is found we can also use it in conjunction with global invariants of the transformation \( T_\psi \) to obtain local signatures based not on derivatives but rather on locally applied geometric invariants. This approach was used in [3] in the case of similarity transformations. We also consider such signatures here.

For sake of compactness we define the transformations here. In a projective transformation, a point \( u \in \mathbb{R}^2 \) is mapped according to

\[
T_\psi : u \rightarrow \tilde{u} = \frac{1}{z(u)} (Au + v)
\]

where \( z(u) = w \cdot u + 1 \) for some \( w \in \mathbb{R}^2 \) and \( A \) is a general invertible \( 2 \times 2 \) matrix. This is the most general transformation. Affine transformations are defined by restricting \( z \equiv 1 \). Euclidean transformations are defined by the further restriction that \( A = U_\omega \) where \( U_\omega \) is a rotation matrix. Thus in this case \( \psi \) consists of an angle of rotation \( \omega \) and the two components of a translation vector \( v \).
2 Euclidean plane transformations

This section describes straightforward results; however, by it we shall introduce and fix much of our notation. If a curve \( P(t) = [x(t), y(t)] \) is transformed into another \( \hat{P}(t) = T_\omega(P(t)) \) then, assuming that the curve \( P(t) \) is smooth, we have

\[
\frac{d}{dt} \begin{bmatrix} \hat{x}(t) \\ \hat{y}(t) \end{bmatrix} = U_\omega \frac{d}{dt} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} dt.
\]

Since \( U_\omega \) is unitary, i.e., \( U_\omega^T U_\omega = I \), we deduce that an invariant reparametrization is obtained by setting

\[
d\tau = |\dot{\hat{P}}(t)| dt \quad \text{and} \quad d\hat{\tau} = |\dot{\hat{P}}(\hat{\tau})| d\hat{\tau}.
\]

This is not surprising: we know that reparametrizing both curves by the arc length should indeed be the first step towards an invariant signature function. After thus reparametrizing we can write

\[
\frac{d^n}{d\hat{\tau}^n} \begin{bmatrix} \hat{x}(\hat{\tau}) \\ \hat{y}(\hat{\tau}) \end{bmatrix} = U_\omega \frac{d^n}{d\tau^n} \begin{bmatrix} x(\tau) \\ y(\tau) \end{bmatrix} \quad \text{for all } n.
\]

Since \( \det(U_\omega) = 1 \), for \( n \neq m, (n, m \geq 1) \), we get

\[
K^{n,m}[\hat{x}, \hat{y} | \hat{\tau}] = \hat{z}^{(n)} \hat{y}^{(m)} - \hat{z}^{(m)} \hat{y}^{(n)} = K^{n,m}[x, y | \tau].
\]

Thus, computed relative to the arc length parametrization, all forms \( K^{n,m}[x, y | \tau] \) are invariant. The lowest degree nontrivial one is \( K^{1,2}[x, y | \tau] \), better known as the curvature, \( k(\tau) \), of \( P(\tau) \). The computation of the curvature requires second derivatives of the reparametrized \( P(\tau) \), and it is quite obvious that we cannot get an invariant signature using lower derivatives in this way.

A point match, i.e., knowing \( T_\omega(P_1) = \hat{P}_1 \) for some point pair, can be used, by making an appropriate translation of coordinates, to set \( \nu = 0 \). We thereby obtain

\[
\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = U_\omega \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} \begin{bmatrix} dt/d\hat{\tau} & 0 \\ 0 & 1 \end{bmatrix},
\]

and by taking determinants we find that

\[
d\tau^* = |K^{-1}[x, y | t]| dt \quad \text{and} \quad d\hat{\tau}^* = |K^{-1}[\hat{x}, \hat{y} | \hat{\tau}]| d\hat{\tau}.
\]

This gives an invariant reparametrization:

Using either this reparametrization or the classical arc length one we can then define, for example, the distance from \( P_1 \) to \( P(\tau^*) \) (or \( P(\tau^*) \)) as an invariant signature. The message is now clear: using one point correspondence enables us to derive invariant signatures with lower order derivatives. This "exchange principle" was introduced and discussed in [1, 2, 6–10]. Catalogues containing invariants derived following this principle can be found in [6] for the affine case and in [1, 2, 7–9] for the projective case. In this paper we use invariant reparametrizations based on this principle, extending and partially reiterating these catalogues.

We also consider the incorporation of global invariants into these schemes, following ideas put forward in [3]. Note that

\[
\begin{bmatrix} \hat{x}(\hat{\tau} + s) - \hat{x}(\hat{\tau}) \\ \hat{y}(\hat{\tau} + s) - \hat{y}(\hat{\tau}) \end{bmatrix} = U_\omega \begin{bmatrix} x(\tau + s) - x(\tau) \\ y(\tau + s) - y(\tau) \end{bmatrix}.
\]

Hence, for example, for a fixed \( s \) the quantity \( \delta[P(\tau + s), P(\tau)] = \delta[P(\tau) + s, P(\tau)] \), i.e., the distance between the points, is a valid invariant signature function. This is a local application of the global invariance of distances under Euclidean transformations. The parameter, \( s \), can be regarded as a "locality" parameter. The smaller \( s \) is the more "local" the calculation of the invariant signature \( \delta[P(\tau + s), P(\tau)] \) becomes.

3 Affine transformations

For the affine case it is necessary to consider higher order derivatives. Here we have

\[
\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} = A \begin{bmatrix} \dddot{x} \\ \dddot{y} \end{bmatrix} \begin{bmatrix} (dt/d\hat{\tau})^2 & 0 \\ dt^2/d\hat{\tau}^2 & dt/d\hat{\tau} \end{bmatrix}.
\]

From this we derive

\[
K^{1,2}[x, y | t] = (\det A)^{1/3}[x, y | t] \left( \frac{dt}{d\hat{\tau}} \right)^3.
\]

If we reparametrize using \( d\hat{\tau}^* = |K^{1,2}[x, y | t]|^{1/3} dt \) then we have \( d\hat{\tau}^* = |\det A|^{1/3} d\tau^* \) and thus

\[
|K^{n,m}[x, y | \tau^*]|^{3/(3-n-m)} = |\det A|^{|K^{n,m}[x, y | \tau^*]|^{3/(3-n-m)}},
\]

for any \( n + m \neq 3 \). Therefore ratios of these forms will be independent of \( (\det A) \), i.e., absolute invariants. In particular we have

\[
\begin{align*}
\frac{|K^{2,4}[x, y | \tau^*]|^{-1}}{|K^{3,3}[x, y | \tau^*]|^{-3/2}} &= \frac{|K^{2,4}[x, y | \tau^*]|^{-1}}{|K^{3,3}[x, y | \tau^*]|^{-3/2}}.
\end{align*}
\]

This shows how one obtains an invariant signature versus a scaled "arc length" parameter using up
to fourth derivatives with respect to \( \tau^* \) and \( \bar{\tau}^* \), or fifth derivatives with respect to \( t \) and \( \bar{t} \). For absolute reparametrizations we could use the absolute invariant functions (2) to further reparametrize the curves via

\[
d\tau = \left| \frac{d}{d\tau^*} \left( \frac{|K^{2,4}[x, y | \tau^*]|}{|K^{2,3}[x, y | \tau^*]|^{3/2}} \right) \right| d\tau^*
\]

which implies that \( \tau \) will be simply a monotized version of the invariant function that we have found. However, we must find another independent absolute invariant. We could now use, say

\[
\frac{K^{1,2}[x, y | \tau]}{K^{2,3}[x, y | \tau]} = \frac{K^{1,2}[z, \bar{y} | \bar{\tau}]}{K^{2,3}[z, \bar{y} | \bar{\tau}]}
\]

as a candidate for the invariant signature: see [4, 5].

As before, a point match allows us to set \( \nu = 0 \). Proceeding in the usual way we obtain

\[
K^{0,1}[z, \bar{y} | \bar{t}] = (\det A) K^{0,1}[x, y | t] \frac{dt}{d\bar{t}}
\]

Combining this with (1) we find an invariant reparametrization via

\[
d\tau = \left( \frac{K^{1,2}[x, y | t]}{|K^{0,1}[x, y | t]|} \right)^{1/2} dt
\]

(3)

With this parametrization it follows that any ratios of the forms \( K^{n,m} \) will be an absolute invariant and, hence, a valid signature function. Clearly the lowest derivatives should be used, yielding

\[
\frac{K^{1,2}[z, \bar{y} | \bar{\tau}]}{K^{2,3}[z, \bar{y} | \bar{\tau}]} = \frac{K^{1,2}[x, y | \tau]}{K^{2,3}[x, y | \tau]}
\]

as an invariant signature obtainable with up to third order derivatives of \( P(\tau) \) and \( \bar{P}(\bar{\tau}) \) and hence fourth order derivatives of \( P(t) \) and \( \bar{P}(\bar{t}) \).

Suppose now that we have two point matches, i.e.

\[
T_0[P_1] = \bar{P}_1 \quad \text{and} \quad T_0[P_2] = \bar{P}_2
\]

In this case we have the following absolute invariant

\[
\frac{\det \begin{pmatrix} \bar{P}_1 - \bar{P}_2 \end{pmatrix}}{\det \begin{pmatrix} P_1 - P_2 \end{pmatrix}} = \frac{\det \begin{pmatrix} (P_1 - P_2) \end{pmatrix}}{\det \begin{pmatrix} (P - P_2) \end{pmatrix}}
\]

Thus combined with reparametrization via (3) we obtain an invariant signature function requiring only second order derivatives. If we had three point correspondences we could determine the complete transformation.

An absolute invariant based on the global invariance of ratios of areas is the following. Let \( T(Q) \) represent the line tangent to the curve \( P(t) \) at the point \( Q \). Let \( P_r \) be an arbitrary point on the curve. For convenience we assume that the curve is convex near \( P_r \).

Let \( P_6 \) be a point on curve near \( P_r \), and let \( Q_0 \) denote the point of intersection between \( T(P_6) \) and \( T(P_r) \). Define \( A^+(P_0) \) to be the area bounded by the curve and the line segment \( P_r P_6 \), and define \( A^-(P_0) \) to be the area bounded by the curve and line segments \( P_0 Q_0 \) and \( P_r Q_0 \).

Now we can choose two numbers \( k_F \) and \( k_B \), and define the points \( P_F = P(\tau + \tau_F) \) and \( P_B = P(\tau - \tau_B) \) by requiring that the ratios \( A^+(P_F)/A^-(P_F) \) and \( A^+(P_B)/A^-(P_B) \) equal \( k_F \) and \( k_B \) respectively, with minimal \( |\tau_F| \) and \( |\tau_B| \). Since area ratios are invariant the points \( \tau_F \) and \( \tau_B \) will be invariant with respect to the affine transformation. The points \( Q_F \) and \( Q_B \) will also be invariant, therefore the ratio \( |P_F Q_F|/|P_B Q_B| \) (where \( |\cdot| \) denotes length) is invariant and can be used as a signature function. There is a case in which this procedure breaks down: when the area ratios are constant. This behavior should be readily detectable, and in this case both curves are affine transformations of a parabola.

4 Projective transformations

The projective transformation is considerably more problematic than the affine since it involves a nonlinear scaling. However, invariant signatures can still be constructed using matrix and determinant methods [1, 2, 7–10], similar in spirit to the methods used for affine transformations, [4, 11]. In this section we survey the approaches proposed above and present several new results.

By introducing the matrix

\[
B = \begin{bmatrix} A & v \\ w^T & 1 \end{bmatrix}
\]

the projective transformation takes the form

\[
\begin{bmatrix} T_0(u) \\ 1 \end{bmatrix} = \frac{1}{z(u)} B \begin{bmatrix} u \\ 1 \end{bmatrix}
\]

Proceeding much as in the derivation of (1) we obtain

\[
K^{2,1}[z, \bar{y} | \bar{t}] = \left( \det B \right) \left( \frac{1}{z} \right)^3 \left( \frac{dt}{d\bar{t}} \right)^3 K^{2,1}[x, y | t]
\]

(4)
Unfortunately the $(1/z)^3$ factor makes this relation inadequate for reparametrization. Work on affine invariants for the projective map, by Halphen, Lane, and Wilczynski, is available in the mathematics literature, see e.g. [14–16]. To illustrate the ideas we present a new result for the case when a point match is available, derived by a method which recalls the approach of Wilczynski (developed in 1905) [12–16]. A more complete development can be found in [17].

By exploiting one point match we can set $v = 0$. This provides us with the valuable relation

\[ K^{1,0}[\tilde{z}, \tilde{y} | \tilde{t}] = \left( \det A \right)^{1/2} K^{1,0}[x, y | t] \frac{dt}{d\tilde{t}}. \]

Combining this with (4), noting $\det B = \det A$ since $v = 0$, and reparametrizing (as introduced and discussed in detail in [8, 9]) via

\[ d\tilde{t} = (\det A)^{-1/3} d\tau, \]

ensures

\[ d\tilde{t} = (\det A)^{-1/3} d\tau. \]

This is achieved using up to second order derivatives of the curve description. We need an invariant signature too.

The vector valued function $\mathcal{X}(t) \triangleq [x(t), y(t)]^T$ obeys a second order differential equation

\[ \dot{\mathcal{X}} + p_1 \mathcal{X} + p_2 \mathcal{X} = 0. \]

where, generically, $p_1$ and $p_2$ are uniquely determined. Let $\xi_{\mathcal{X}}(t) = [1 \ p_1 \ p_2]^T$. Scaling $\mathcal{X}$ to $\mu \mathcal{X}$ so that $\xi_{\mu \mathcal{X}}(t) = [1 \ 0 \ P_2(t)]$, we find that $\mu$ is uniquely determined modulo a multiplicative constant, and hence $P_2(t)$ is invariant under scalings. Solving for $P_2(t)$ we obtain

\[ P_2(t) = \frac{1}{2} K^{3,0}[x, y | t] + \frac{3}{2} K^{2,1}[x, y | t] - \frac{3}{2} \left( \frac{K^{2,1}[x, y | t]}{K^{1,0}[x, y | t]} \right)^2. \]

Assume $\mathcal{X}(t)$ is a canonically scaled representation obeying $[\dot{\mathcal{X}} \ \mathcal{X} \ \mathcal{X}]^T[1 \ 0 \ P_2] = 0$. Then to bring a reparametrized version $\hat{\mathcal{X}}$ to this form requires multiplication by a scale factor $\lambda$ satisfying

\[ \left( \frac{dt}{d\tilde{t}} \right)^2 = c = \text{Constant}. \]

From this (setting $c = 1$) one can derive the relation

\[ P_3 = \lambda^3 \hat{\lambda} + \lambda^4 \hat{P}_2. \]

Clearly we need $\hat{\lambda} = 0$ to get an invariant. But we already know that by reparametrizing using (5) we can have $\tau \propto \tilde{t}$ and hence $\lambda = \text{Constant}$. Therefore, after a scaled reparametrization, we shall have

\[ \sqrt{P_2} d\tau = \sqrt{\hat{P}_2} d\tilde{t}. \]

This relation can be used to obtain a non-scaled reparametrization, with the used of third order derivatives with respect to $\tau$ and $\tilde{t}$, and hence fourth order derivatives with respect to $t$ and $\tilde{t}$.

Suppose now that we have one additional pair of matching points, i.e. that besides $(0, 0) - (0, 0)$, we have that $T_{\Psi}(P_2) = \hat{P}_2$. In this case it can be shown that

\[ \hat{Q} \triangleq \left\{ \begin{array}{cc} \det \hat{\mathcal{P}} & \det \hat{\mathcal{P}} \\ \det \mathcal{P} & \det (\hat{\mathcal{P}} - \hat{\mathcal{P}}_2) \end{array} \right\} \]

\[ = \left\{ \begin{array}{cc} \det \hat{\mathcal{P}} & \det \mathcal{P} \\ \det \hat{\mathcal{P}} & \det (\mathcal{P} - \mathcal{P}_2) \end{array} \right\} \left( \begin{array}{c} dt \\ d\tilde{t} \end{array} \right) \]

\[ \triangleq Q \frac{dt}{d\tilde{t}}. \]

An invariant reparametrization is thus achieved via

\[ d\tilde{t} = Q dt \quad d\tilde{t} = Q d\tilde{t}. \]

This invariant parametrization has also appeared in [10] and has already been used for object recognition and symmetry detection, along with several projective invariants of a similar type [7–9].

Once an invariant reparametrization is found we can also use locally applied global invariants for the generation of local signatures. Such signatures can be found in [17].

5 Conclusions

This work is based on results reported in [3, 4] and on some recent ideas put forward in [1, 7, 8, 13].

The ideas of [3] for the recognition of planar curves distorted by similarity transformations showed that on curves for which an invariant reparametrization was first found one can use global geometric invariants of the transformation employed locally. The motivation, that of reducing the order of derivatives required, led [1, 7, 8] to suggest the use of point matches as additional sources of information in the recognition tasks. Point matches effectively reduce the parameter space
of projective transformations by imposing various relationships among those parameters. This paper shows, in a unified way, how to integrate such ideas in order to carry out the program proposed in [4] of determining invariant reparametrizations and associated signature functions using lowest possible derivatives of curve representations. Our paper fills several gaps left by the above mentioned work, and, more importantly, proposes the use of local (non semi-differential) invariant signatures based on employing locally used global invariants of the viewing transformation.

In closing, we note that the use of algebraic and global invariants in vision has also attracted a lot of attention recently, as exemplified by the papers [23–30].

6 Acknowledgments

The authors thank I. Weiss, M. Brill, L. Van Gool and J. Lagarias for many interesting discussions.

References


