THE FEATURES FOR RECOGNITION OF PROJECTIVELY DEFORMED POINT SETS

Tomáš Suk and Jan Flusser

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic
E-mail: suk@utia.cas.cz, flusser@utia.cas.cz

ABSTRACT

The paper deals with features of a point set which are invariant with respect to projective transform. First, some properties of the area of a triangle under projective transform and the cross-ratio are discussed. Two types of projective triangular invariants of point sets are derived.

One of them is a quotient of two different products of the areas of triangles formed by the vertices of the polygon. It is invariant to the labeling of the vertices. The number of the features is discussed and a method for its reduction is presented.

The other type of features is expressed as functions of five-point cross-ratios and it is invariant to the choice of labeling of the points. Possibility of the normalization of these invariants is shown. The performance of the features is demonstrated by numerical experiments.

1. INTRODUCTION

One of the important tasks in image processing and computer vision is a recognition of objects on images captured under different viewing angles. If the scene is planar, then the distortion between two frames can be described by *projective transform*

$$\begin{aligned} x' &= (a_0 + a_1 x + a_2 y)/(1 + c_1 x + c_2 y) \\ y' &= (b_0 + b_1 x + b_2 y)/(1 + c_1 x + c_2 y) \ , \end{aligned} \qquad (1)$$

where x, y are the coordinates in the first frame and x', y' are the coordinates in the second one.

Feature-based recognition of such objects requires features that are invariant to projective transform. Several different approaches to this problem have been published in the recent works. One of them is based on the assumption that non-linear term of the eq. (3) is relatively slight and the projective transform can be approximated by the affine one. This assumption is true, if the distance from the object to the sensor is much greater than the size of the object. In such cases, various affine invariants can be applied such as moment invariants [1,2] or Fourier descriptors [3].

However, in some cases the distortion cannot be approximated by the affine transform and therefore the use of exact projective invariants is required. The invariants which have been developed for this purpose can be categorized into two groups: differential invariants and point-based ones [4,5].

This paper deals with point-based features for recognition of projectively deformed point sets. The points must be labeled by some way, but the correspondence between the points can be unknown and then the correct labeling of the points is also unknown. Two types of features are derived and compared. The first of them is based on the presumption that the points are vertices of a polygon. The labeling is defined by the polygon and therefore two different labelings can differ from each other only by a cyclic shift. These features are invariant to the choice of the starting point. The other type of features is derived for sets of isolated points. These features are invariant to the choice of labeling of the points, i.e. to the permutation of the points.

2. THE AREA OF A TRIANGLE AND A CROSS-RATIO

Let's denote (i, j, k) the area of the triangle with the vertices $[x_i, y_i]$, $[x_j, y_j]$ and $[x_k, y_k]$, i.e.

$$(i, j, k) = (x_i y_j - x_j y_i + x_j y_k - x_k y_j + x_k y_i - x_i y_k)/2. \quad (2)$$

If the vertices are numbered clockwise, we will consider the area negative.

If $J(x, y)$ is a Jacobian of the projective transform (1) at the point $[x, y]$ then the area of the triangle after the transform can be expressed as

$$(i, j, k)' = (i, j, k) \ \sqrt[3]{J(x_i, y_i)\, J(x_j, y_j)\, J(x_k, y_k)}. \qquad (3)$$

The proof of eq. (3) can be found in [6].

As can be seen from (3), the area of the triangle after the transform is given by the area of the original triangle multiplied by some power of Jacobian of the projective transform. The triangle area is *relative invariant* [7]. Also each product of the areas of triangles is a relative invariant to projective transform.

This property can be utilized for a construction of absolute projective invariants, i.e. invariants, which don't change values under projective transform. The quotient of two general relative triangular invariants is absolute invariant, if each vertex occurs with the same frequency in the numerator and in the denominator [6].

This work has been supported by the grant No. 102/94/1835 of the Grant Agency of the Czech Republic.

The simplest absolute invariant is the cross-ratio of five points

$$\rho = \frac{(1,2,3)(1,4,5)}{(1,2,4)(1,3,5)}. \tag{4}$$

3. INVARIANTS TO THE CHOICE OF THE STARTING POINT

To ensure the invariance to the choice of the starting point, the product of the areas of the triangles from the vertices of the polygon can be used

$$R_{k_1,k_2,k_3} = \prod_{i=1}^{n} (i, (i+k_1-1) \bmod n + 1, \\ (i+k_1+k_2-1) \bmod n + 1), \tag{5}$$

where $n = k_1 + k_2 + k_3$ and "mod" denotes modulo operation. Commas between indices will be omitted, if it doesn't lead to misunderstanding.

The invariants (5) are called relative projective triangular invariants (RPTI).

The quotient of two RPTIs defined by (5) is absolute invariant. In the denominator it is suitable to use $R_{1,1,n-2}$, which is constructed from the areas of the triangles from the adjacent vertices of the polygon. If some three adjacent vertices are collinear, the medium vertex can be omitted and the polygon has got less than n vertices. If no three adjacent vertices are collinear, $R_{1,1,n-2}$ is different from zero and it can be used for normalization of the invariants.

We will denote a normalized absolute projective triangular invariant (PTI) as I_{k_1,k_2,k_3}:

$$I_{k_1,k_2,k_3} = \frac{R_{k_1,k_2,k_3}}{R_{1,1,n-2}}. \tag{6}$$

If n is big, the values of PTIs can be high and then an extraction of the root is suitable:

$$\begin{array}{ll} I'_{k_1,k_2,k_3} = \sqrt[n]{I_{k_1,k_2,k_3}} & \text{if } I_{k_1,k_2,k_3} \geq 0 \\ I'_{k_1,k_2,k_3} = -\sqrt[n]{-I_{k_1,k_2,k_3}} & \text{if } I_{k_1,k_2,k_3} < 0. \end{array} \tag{7}$$

The number ℓ of PTIs is

$$\begin{array}{lllll} \ell & = & (n-1)(n-2)/6 - 1 & \text{if} & n \bmod 3 \neq 0, \\ \ell & = & (n-1)(n-2)/6 - 1/3 & \text{if} & n \bmod 3 = 0. \end{array} \tag{8}$$

Eqs.(8) mean that the number of PTIs grows quadraticly with n, but the polygon with n vertices has $2n$ degrees of freedom and projective transform has 8 parameters. Therefore the polygon can have only

$$m = 2n - 8 \tag{9}$$

independent invariants to projective transform. That is why it is suitable to reduce the number of invariants for polygons with more than 11 vertices.

The PTIs can be divided to groups according to the least index g ($g = \min(k_1, k_2, k_3)$). The number of these groups is $[n/3]$ (integer part). If the number of invariants should be m, then each group can consist of $2n/(n/3)$, i.e. 6 invariants maximally. Only if n is not dividable by 3, then the group with $g = 1$ can consist of 7 invariants. The PTIs in each group can be partially added up to acquire the desirable number of invariants. In the following text these invariants are denoted I_{gi}, where $i = 1, 2, \cdots$. More details can be found in [6]. The sums are divided by the number of added invariants.

4. INVARIANTS TO THE CHOICE OF LABELING

The five-point cross-ratio (4) can be used for construction of the invariants to the choice of labeling. Reiss in [1] uses $\rho + \rho^{-1}$, which is unstable near zero. If some triplet of five points in (4) is collinear the function $\rho + \rho^{-1}$ is infinite. Thus the more suitable function is $\psi = \frac{2}{\rho+\rho^{-1}} = \frac{2\rho}{\rho^2+1}$. If ρ or ρ^{-1} is zero, then the function ψ is zero.

If invariance to the choice of labeling of the points is required, some symmetric function must be used for considering all possibilities of labeling. We can use additions and multiplications

$$\begin{array}{l} F'_{+}(1,2,3,4,5) = \\ = \psi(1,2,3,4,5) + \psi(1,2,3,5,4) + \psi(1,2,4,5,3), \\ F_{\cdot}(1,2,3,4,5) = \\ = \psi(1,2,3,4,5) \cdot \psi(1,2,3,5,4) \cdot \psi(1,2,4,5,3). \end{array} \tag{10}$$

It is suitable to use e.g. the function $F_+ = 6 - \frac{80}{3(6-F'_+)}$ instead of F'_+ because of higher discriminability of the invariants. Functions F_+ and $F_\cdot$ are invariant to the choice of labeling of last four points, but the point 1 must be common at all cross-ratios. To obtain full invariance to the choice of labeling, we must alternate all five points as common ones

$$\begin{array}{l} I_{s_1,s_2}(1,2,3,4,5) = F_{s_1}(1,2,3,4,5) s_2 F_{s_1}(2,3,4,5,1) s_2 \\ s_2 F_{s_1}(3,4,5,1,2) s_2 F_{s_1}(4,5,1,2,3) s_2 F_{s_1}(5,1,2,3,4), \end{array} \tag{11}$$

where s_1 and s_2 are either sign $+$ or $\cdot$.

According to (9) only two from four invariants I_{++}, $I_{\cdot+}$, $I_{+\cdot}$ and $I_{\cdot\cdot}$ can be independent.

Similar invariants were published in [8], but the invariants F_+ and $F_\cdot$ are more stable and generalization for more than 5 points is easier.

In that case we can use the following algorithm:

$$I_{s_1,s_2,s_3,s_4,s_5,s_6,s_7} = \\ = \mathop{\mathrm{S}_7}_{i=1}^{n} \mathop{\mathrm{S}_6}_{\substack{j=1 \\ j \neq i}}^{n} \mathop{\mathrm{S}_5}_{\substack{k=1 \\ k \neq i,j}}^{n} \mathop{\mathrm{S}_4}_{\substack{l=1 \\ l \neq i,j,k}}^{n} \mathop{\mathrm{S}_3}_{\substack{m=1 \\ m \neq i,j,k,l}}^{n} I_{s_1,s_2}(i,j,k,l,m), \tag{12}$$

where S_a is either $\sum$ when s_a is $+$ or $\prod$ when s_a is $\cdot$, $a = 3, 4, \ldots, 7$.

The computing time is $n(n-1)(n-2)(n-3)(n-4)T$, what means the computing complexity is $O(n^5)$.

The number of invariants is $2^7 = 128$, independently of n, which is enough for $n \leq 68$ points, but only first m invariants are used, see (9).

The number of added and multiplied values is very high and that's why a normalization of the results is suitable. To preserve values inside acceptable intervals, we must divide each sum of n values by n and extract the nth root of each product of n values. It means that

$$\mathop{\mathrm{S}}_{i\in C}^{a} f(i) = \frac{1}{r}(\sum_{i\in C} f(i)), \tag{13}$$

when s_a is $+$, and

$$\mathop{\mathrm{S}}_{i\in C}^{a} f(i) = p \cdot \sqrt[r]{\prod_{i\in C} |f(i)|}, \tag{14}$$

when s_a is $\cdot$, p is a sign of $\prod_{i\in C} f(i)$ and r is the number of elements of the set C. The sums and products in definition of I_{s_1,s_2} (10) and (11) should be normalized in the same way.

The values of ψ always fall between -1 and 1. The values of invariants I_{s_1,s_2} are maximum, if the values of ψ in (10) are 1, 0 and 0; then $F_+ = 1$ and $F_{\cdot} = 0$. They are minimum if the values of ψ are $\frac{2}{\sqrt{5}}$, $\frac{2}{3}$ and $-\frac{2}{\sqrt{5}}$; then $F_+ = \frac{2}{3}$ and $F_{\cdot} = -\frac{8}{15}$. Thus if we use the features

$$I'_{+s_2} = 9I_{+s_2} - 2,\ I'_{\cdot s_2} = I_{\cdot s_2}(-\sqrt[3]{15})/2, \tag{15}$$

then the values of all invariants will be in range from 0 to 1 (I_{+s_2} and $I_{\cdot s_2}$ are after normalization (13) and (14)).

5. NUMERICAL EXPERIMENT

The following numerical experiment was carried out to observe the stability and discriminability of the invariants (12) to the choice of labeling (PTILs) and to compare them with the invariants (6) to the choice of the starting point.

A set of 16 points was created by a pseudorandom generator, see Fig. 1. The size of the image is 512 × 512 pixels. This image was projectively transformed and coordinates of points were rounded off to integers (Fig. 2). The numbers in the figures don't express the correspondence between points, but show labeling during computation of invariants, which was different. The least distance between the points (5 and 6 in Fig. 2) is 5 pixels.

To observe discriminability of these features, another set of points was generated, see Fig. 3. The two point sets don't differ from each other by projective transform, therefore the values of the invariants should be different.

The results are shown in Table 1. The invariants with normalization (13), (14) and (15) were used. The Euclidean distance between the sets of points in Fig. 1 and Fig. 2 in 24-dimensional feature space is 0.027 and between Fig. 1 and Fig. 3 is 0.353, what is more than 13 times greater. This fact illustrates that the discriminability is satisfactory.

The invariants to the choice of the starting point (6) with the normalization (7) were computed for the same data. The results are shown in Table 2. The values in the column Fig. 2' were computed with the correct labeling. The difference between them and the values of Fig. 1 is caused only by rounding off the coordinates. The distance between them is 0.962. The distance between the values of Fig. 1 and Fig. 2 is 17.4. It is caused by incorrect labeling. The distance between the values of Fig. 1 and Fig. 3 is 9.17. It is caused by fact these point sets don't differ from each other by projective transform. The discriminability of PTIs is similar to that of PTILs.

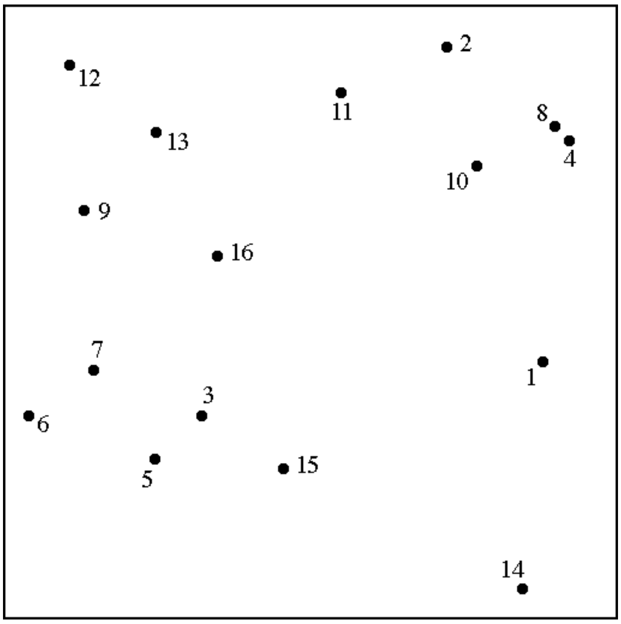

Figure 1: The set of 16 points

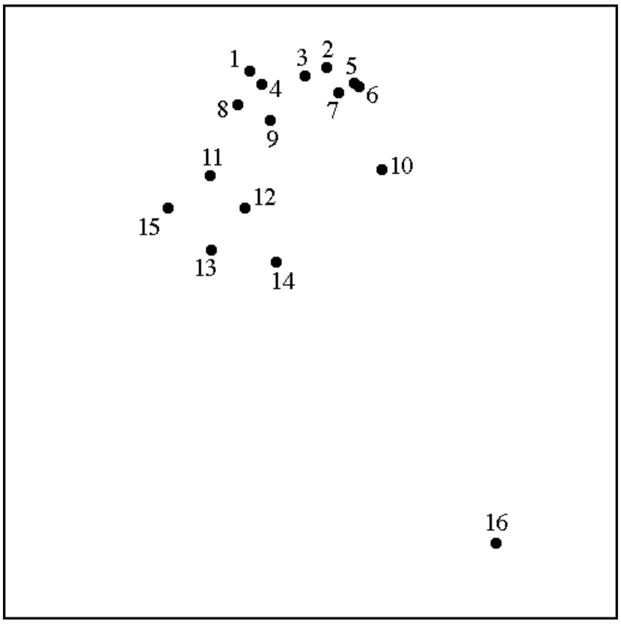

Figure 2: The projectively deformed set of 16 points

6. SUMMARY

It was shown by the numerical experiments that both types of invariants yield satisfactory results. The stability and

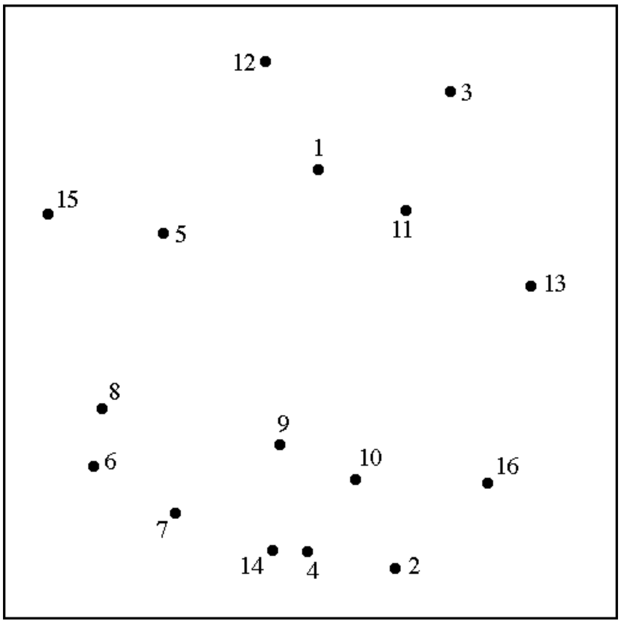

Figure 3: The other set of 16 points

the discriminability of PTIs with the normalization (7) and PTILs with the normalization (13), (14) and (15) is similar.

The disadvantage of PTILs is their high computing complexity $O(n^5)$. The computing complexity of PTIs is $O(n)$ that is much more acceptable.

	Fig.1	Fig.2'	Fig.2	Fig.3
$I_{11} = I_{1,2,13}$	979	970	1860	1122
$I_{12} = I_{1,13,2}$	1281	1319	-1548	-1105
$I_{13} = (I_{1,3,12} + I_{1,12,3})/2$	-73	-62	196	-74
$I_{14} = (I_{1,4,11} + I_{1,11,4})/2$	1594	1645	324	-1183
$I_{15} = (I_{1,5,10} + I_{1,10,5})/2$	126	135	113	1
$I_{16} = (I_{1,6,9} + I_{1,9,6})/2$	-1542	-1563	-4094	-41
$I_{17} = (I_{1,7,8} + I_{1,8,7})/2$	-1293	-1326	-3639	797
$I_{21} = I_{2,2,12}$	-904	-927	2066	-634
$I_{22} = I_{2,3,11}$	-1242	-1268	2883	-954
$I_{22} = (I_{2,11,3} + I_{2,4,10})/2$	1385	1430	2644	281
$I_{24} = (I_{2,10,4} + I_{2,5,9})/2$	369	1308	180	-152
$I_{25} = (I_{2,9,5} + I_{2,6,8})/2$	-70	-77	-23	901
$I_{26} = (I_{2,8,6} + I_{2,7,7})/2$	1535	1588	-3854	891
$I_{31} = I_{3,3,10}$	1231	1307	-4410	1275
$I_{32} = I_{3,4,9}$	1673	1702	-5023	-1165
$I_{33} = I_{3,9,4}$	1342	1381	-4061	1372
$I_{34} = I_{3,5,8}$	2741	2812	5711	-937
$I_{35} = I_{3,8,5}$	2435	2481	4360	979
$I_{36} = (I_{3,6,7} + I_{3,7,6})/2$	209	241	-328	107
$I_{41} = I_{4,4,8}$	-1379	-1416	3636	-1187
$I_{42} = I_{4,5,7}$	-1682	-1758	4481	1353
$I_{43} = I_{4,7,5}$	1849	1924	4308	-1554
$I_{44} = I_{4,6,6}$	1223	1256	3771	-697
$I_{51} = I_{5,5,6}$	2408	2468	4931	-1179

Table 2: The values of PTIs (the values were multiplied by 10^3)

	Fig. 5	Fig. 6	Fig. 7
$I_{+++++++}$	477	480	541
$I_{\cdot++++++}$	649	654	711
$I_{+\cdot+++++}$	454	457	519
$I_{\cdot\cdot+++++}$	585	592	661
$I_{++\cdot++++}$	400	406	480
$I_{\cdot+\cdot++++}$	621	627	690
$I_{+\cdot\cdot++++}$	381	386	458
$I_{\cdot\cdot\cdot++++}$	538	547	627
$I_{+++\cdot+++}$	460	464	529
$I_{\cdot++\cdot+++}$	641	646	705
$I_{+\cdot+\cdot+++}$	437	441	505
$I_{\cdot\cdot+\cdot+++}$	570	577	651
$I_{++\cdot\cdot+++}$	378	385	463
$I_{\cdot+\cdot\cdot+++}$	612	618	683
$I_{+\cdot\cdot\cdot+++}$	359	365	441
$I_{\cdot\cdot\cdot\cdot+++}$	523	533	616
$I_{++++\cdot++}$	471	475	538
$I_{\cdot+++\cdot++}$	646	651	709
$I_{+\cdot++\cdot++}$	448	451	514
$I_{\cdot\cdot++\cdot++}$	580	587	658
$I_{++\cdot+\cdot++}$	393	400	475
$I_{\cdot+\cdot+\cdot++}$	618	624	688
$I_{+\cdot\cdot+\cdot++}$	374	380	453
$I_{\cdot\cdot\cdot+\cdot++}$	533	542	623

Table 1: The values of PTILs (the values were multiplied by 10^3)

7. REFERENCES

[1] T. H. Reiss, "Recognition Planar Objects Using Invariant Image Features", Lecture Notes in Computer Science 676, Springer, 1993.

[2] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants", *Pattern Recognition* No. 26, pp. 167-174, 1993.

[3] C. C.Lin and R. Chellapa, "Classification of partial 2-D shapes using Fourier descriptors", *IEEE Trans. Pattern Anal. Mach. Intell.*, No. 9, pp. 686-690, 1987.

[4] P. Meer and I. Weiss, "Point/Line Correspondence under 2D Projective Transformation", *Proc. 11th International Conference on Pattern Recognition*, The Hague, The Netherlands, pp. 399-402, 1992.

[5] D. Forsyth, J. L. Mundy, A. Zisserman, C. Coelho, A. Heller and C. Rothwell, "Invariant Descriptors for 3-D Object Recognition and Pose", *IEEE Trans. Pattern Anal. Mach. Intell.*, No. 10, pp. 971-991, 1987.

[6] T. Suk, J. Flusser, "Vertex-based features for recognition of projectively deformed polygons", *(to appear in Pattern Recognition).*

[7] I. Shur, "Vorlesungen über Invariantentheorie", Springer, Berlin, 1968 (in German).

[8] R. Lenz, P. Meer, "Point configuration invariants under simultaneous projective and permutation transformations", *Pattern Recognition*, Vol. 27, No. 11, pp. 1523-1532, 1994.