# The Recipe for Some Invariant Numbers and for a New Projective Invariant Feature Descriptor

Raphael dos S. Evangelista[a] and Leandro A. F. Fernandes[b]

*Instituto de Computação, Universidade Federal Fluminense (UFF),*
*Av. Gal. Milton Tavares de Souza, Niterói, Brazil*

Keywords: Visual Feature, Junction, Invariant Theory, Similarity Map, Affine Map, Projective Map.

Abstract: The Computer Vision literature provides a range of techniques designed to detect and describe local features in images. The applicability of these techniques in visual tasks is directly related to the invariance of each kind of descriptor to a group of geometric transformations. To the best of our knowledge, there is no local feature descriptor solely based on single intensity images that are invariant to projective transformations. We present how to use existing monomials invariant to similarity, affine, and projective transformations to compute invariant numbers from junctions' geometry. In addition, we present a new junction-based invariant number and use it to propose a new local feature descriptor invariant to projective transformations in digital images.

## 1 INTRODUCTION

The ability to recognize visual patterns is essential to complete our most common daily tasks. Computational vision systems were developed and applied to a countless number of problems and, regardless of the visual challenge, the most basic task is the identification, recognition, and classification of objects or parts of them. Researches have demonstrated that contours and junctions play an essential role in the recognition process (Barrow and Tenenbaum, 1981). This affirmative does not put texture properties aside on recognition tasks but emphasizes that geometrical features are of the utmost importance.

The exclusion of junctions from an edge image affects the human visual recognition system more negatively than the elimination of continuous edges, as exposed by Biederman (1985). Therefore, junctions and their surrounded regions are assumed to be highly discriminative. Besides, the perception of this kind of feature by the human visual system is robust to changes in viewpoint.

Computer Vision researchers have been challenged to find ways to detect and describe these particular regions in such a way that the description becomes invariant to the viewpoint of the scene, the lighting conditions, and other variables. Geomet-

[a] https://orcid.org/0000-0002-2808-5865
[b] https://orcid.org/0000-0001-8491-793X

rical transformations are classified into groups, depending on its characteristic. The group of linear projective transformations describes operations that map primitives between projective planes, after a finite series of projections. The projective group encompasses the affine group (*e.g.,* shearing, and non-uniform scaling), the similarity group (*e.g.,* uniform scaling), and the isometric group (*e.g.,* rotations, reflexions, and translations). However, most existing image feature descriptors ignore structural geometry, preventing the invariance from reaching more general groups of transformations, *e.g.,* the projective group.

The word *invariant* means that a value does not change even when the application of a mapping transforms the element. It is not the same as to affirm that something has the *same behaviour*, as usually occurs with regions and textures (Tuytelaars and Mikolajczyk, 2008). Some researchers have been studying how to extract geometrical information from images (*e.g.,* lines, ellipses, and general conic sections) in such a way that we could acquire actual invariants (Quan, 1995; Luo et al., 2013; Jia et al., 2016). Many invariant might be used if the extraction succeeds. For instance, the cross-ratio is a well-known projective invariant (Hartley and Zisserman, 2004).

In this paper, we show how to use the geometry of junctions and monomials from the invariant theory of $n \times n$ matrices (Cayley, 1858) to compute a set of invariant numbers to similarity and affine mappings. Also, we show how to use the cross-ratio to calcu-

late projective invariant values for a subclass of junctions. Finally, we present a new junction-based number, which is invariant to general homographies. Despite the unsuccessful of practical attempts, we hope that future works might explore the background theory presented in this paper.

## 2 RELATED WORK

Many techniques address the problem of finding and describing local discriminative visual features in intensity images. The Scale-Invariant Feature Transform (SIFT) (Lowe, 1999) uses the Difference of Gaussians operator to achieve invariance to similarity transformations and partial invariance to illumination changes. The Speed-Up Robust Features (SURF) (Bay et al., 2008) implements the Hessian Operator, which seems to be more robust against (but not completely invariant to) affine transformations (Lindeberg, 2012). The Binary Robust Independent Elementary Features (BRIEF) (Calonder et al., 2012) was the first binary feature descriptor aiming to be computationally efficient. It uses the FAST corner detector (Rosten and Drummond, 2006; Rosten et al., 2010) and strings of bits to encode the features. All these techniques and their extensions (*e.g.,* Bosch et al. (2008); Ke and Sukthankar (2004); Yu and Morel (2011); Aldana-Iuit et al. (2016)) take the textured region around discontinued edges and corners to build local descriptors from gradient signatures, random difference of intensities and other relations (Tuytelaars and Mikolajczyk, 2008). When using only texture data, the invariance to each group of transformations is treated as an independent problem, which makes the formulation of descriptors more complicated and less intuitive. The Affine-SIFT (ASIFT) (Yu and Morel, 2011), for example, consider the SIFT descriptors extracted from many random synthetic projections of the input image to retrieve affine invariant feature descriptors.

A different approach is to describe interesting contours from the silhouette of objects by using expected shapes like rectangles, circles, and ellipses (Gdalyahu and Weinshall, 1999; Alajlan et al., 2007, 2006). These techniques are strongly based on invariant geometric relations among the shapes. But they don't work well with natural images and usually require controlled environments, with limited illumination and points of view. Recent works (Luo et al., 2013; Fan et al., 2014; Jia et al., 2016; Li et al., 2019) achieved good results at describing the entire contours but, despite its limited contour of interest (*i.e.,* the shape), they are still susceptible to noisy and occlu-

sion. The more localized the feature, the more robust it may be under projective mappings.

There are several recent solutions based on convolutional neural networks (Balntas et al., 2016; Tian et al., 2017; Mishchuk et al., 2017; Tian et al., 2019) which outperform the standard unsupervised feature detection-description-matching algorithms. Despite their effectiveness, there are still computational systems with limited resources, that might be interested in standard and old-fashion algorithms. Thus, the classical approach still plays an import role in Computer Vision. In this sense, we use projective invariant geometric relations of the branches of junctions to describe local discriminative information in images.

## 3 INVARIANT OF MATRICES

In this section, we review the concept of invariance when applied to $3 \times 3$ symmetric matrices, as we can observe their behavior as if they encode conic sections. In other words, given any symmetric matrix, they would be transformed as a primal conic section or a dual conic section, depending on the situation.

Let M be an invertible $n \times n$ matrix. The Cayley-Hamilton theorem states that every matrix M satisfies its characteristic polynomial, derived from:

$$\det(M - kI) = k^n - c_1 k^{n-1} + \cdots + (-1)^n c_n,$$

where $I$ is an identity $n \times n$ matrix, $k$ is a unknown linear coefficient and $c_1, c_2, \cdots, c_n$ are the characteristic coefficients of M. As a generalization, M may be a linear combination of $n$ matrices

$$\det(k_1 I M_1 + k_2 I M_2 + \cdots + k_n I M_n) = f_1 c_1 + \cdots + \underbrace{k_1^{p_{i1}} k_{i2}^{p_{i2}} \cdots k_n^{p_{in}}}_{f_i} c_i + \cdots + f_n c_n,$$

where $f_i$ denotes the combination of the combinatory factors $k_1, k_2, \cdots, k_n$ and $p_{i1}, p_{i2}, \cdots, p_{in}$ are scalar numbers, to elucidate the general formula.

Assuming that M is symmetric, it encodes a conic section. The characteristic coefficients are naturally invariant under the isometric group of transformations. The characteristic coefficients usually arise as a function of the determinant, trace, or the combination of both. One achieves invariance under the similarity group of transformations (excluding translations) considering the ratio (Semple and Kneebone, 1952):

$$\gamma = \frac{\text{dsc}(M)}{\text{tr}(M)^2}, \tag{1}$$

where $\text{tr}(M)$ is the trace and $\text{dsc}(M)$ is the second discriminant, defined as the determinant of the leading

$2 \times 2$ submatrix of M. The discriminant determines the shape of the curve. When the conic section is degenerated, the discriminant is *zero*.

Nevertheless, it is possible to assemble invariant monomials in a more systematic way, because any polynomial raised from the combination of characteristic coefficients is also a characteristic polynomial (Forsyth et al., 1991; Quan et al., 1991). Therefore, conveniently, we can consider the product of all coefficients $\prod_{i=0}^{n} (f_i c_i)^{a_i}$ where all $a_i \in \mathbb{Q}$ are variable exponents related to the characteristic coefficients $c_i$ and its combinatorial factors $f_i$. We can vary the value of $a_i$ as much as we want to build different characteristic polynomials. But it is important to keep in mind that not all of them are invariant due to the presence of the factors $f_i$. We are interested in the polynomials that can neutralize the $f_i$ effects. So, what we need is:

$$\prod_{i=0}^{n} (f_i c_i)^{a_i} = \prod_{i=0}^{n} \left( \underbrace{k_1^{p_{i1}} k_2^{p_{i2}} \cdots k_n^{p_{in}}}_{f_i} c_i \right)^{a_i} = \prod_{i=0}^{n} c_i^{a_i},$$

where the above equality is possible if, and only if, each combinatorial factor $f_i = 1$, which leaves the characteristic coefficients $c_i$ untouched. Thus, we want the solution of the homogeneous system $P\,a = 0$, where P is a square $(n+1) \times (n+1)$ matrix composed by scalar powers $p_{ij}$, and $a = (a_0, \cdots, a_n)^\top$. Due to space restrictions, see Quan (1995) for details.

The product operator allowed us to find a solution where $\sum_{i=0}^{n} p_{ij} = 0$, for any given factor $k_j$. With this powerful tool in mind, we can derive as many geometrical quantities of conic sections as intended and combine them to achieve invariance. Quan (1995) proved that a true projective invariant quantity is only reachable with the combination of two curves, so

$$\beta_{i,j} = \frac{\operatorname{tr}\left(M_i^{-1} M_j\right) \det\left(M_i\right)}{\operatorname{tr}^2\left(M_j^{-1} M_i\right) \det\left(M_j\right)} \qquad (2)$$

where $M_i$ and $M_j$ are symmetric matrices related to any two conic sections.

Section 4 presents our main contribution, which is how to encode junctions as dual conic sections and compute invariant numbers for them.

# 4 THE PROPOSED INVARIANT MONOMIALS FOR JUNCTIONS

For a given bidimensional intensity image $I : \mathbb{R}^2 \mapsto \mathbb{R}^+$, where the function $I(x, y)$ returns the gray level related to point $p = (x, y) \in \mathbb{R}^2$, the location of a (continuous) junction correspond to the intersection of linear ridges in the gradient of $I$, while the branches of a junction are given by the line segments defined by the ridges. Thus, a continuous junction $\mathcal{J}_n$ can be defined as the set

$$\mathcal{J}_n = \{p \mid p \in \overline{p_0 p_i}, \exists\, i \in \{1, 2, \ldots, n\}\}, \qquad (3)$$

where $n \in \{2, 3, \ldots\}$ is the number of branches or *degree of the junction*, $p_0$ denotes the *central vertex* (*i.e.,* the point shared by the branches), and $p_i = (x_i, y_i)$ and $\overline{p_0 p_i}$ are, respectively, the *delimiter vertex* and the line segment defining the $i$-th branch, with $p_i \notin \overline{p_0 p_j}$ for all $i \neq j$ and $i, j \in \{1, 2, \ldots, n\}$. We assume that branches are taken counterclockwise. In digital images, the branches of digital junctions having $n \geq 5$ overlap. Thus, $n \in \{2, 3, 4\}$ in this case.

By definition, the central vertex $p_0$ of a junction is always a proper point (*i.e.,* a point with finite location). When the delimiter vertices $p_i$ are also proper points, we say that we have a *proper junction*. However, when $p_i$ are ideal points (*i.e.,* points at infinity or directions), we say that they define an *ideal junction*.

Section 4.1 presents a similarity invariant monomial adapted for 2-junctions. Affine invariance is achieved by the monomial used with 3-junctions in Section 4.2. Sections 4.3 and 4.4 describe projective invariant numbers for 4-junctions. As shown in Section 4.4, only junctions with four (or more) branches could provide the required input for the projective invariant, despite the ambiguous visual connotation related to them in practice (*e.g.,* occlusions).

## 4.1 Similarity Invariant for Proper and Ideal 2-junctions

The key insight in this work was to build symmetric matrices from the geometry of junctions and use the $\gamma$ number defined in (1) as invariant for 2-junctions under similarity transformations. More specifically, we explore the duality between points in projective space $\mathbb{P}^2$ and dual conic sections to represent point vectors $x = (x, y, w)^\top \in \mathbb{P}^2$ as symmetric matrices with the form (Perwass and Forstner, 2006):

$$D = x\,x^\top = \begin{pmatrix} x^2 & xy & xw \\ yx & y^2 & yw \\ wx & wy & w^2 \end{pmatrix}, \qquad (4)$$

where $\square^\top$ denotes matrix transposition.

It is important to notice that $\operatorname{rank}(D) = 1$. Therefore, D cannot be used directly in (1) since $\det(D) = 0$. But a net of conics (Semple and Kneebone, 1952) computed from at least three dual conic sections related to linearly independent point vectors $x \in \mathbb{P}^2$ produces a symmetric $3 \times 3$ matrix with full

rank. In our work, we use the following net of conics to compute invariant $\gamma$ numbers for 2-junctions:

$$N = D_0 + D_1 + D_2. \tag{5}$$

In (5), we assume that the central vertex $p_0$ is the origin of the coordinate system. Therefore, (4) relates the dual conic section matrix $D_k$ to the point vector $x_k = (x_k - x_0, y_k - y_0, 1)^\top$ if $p_k$ is a proper point, or to $x_k = (x_k, y_k, 0)^\top$ if the vertex is an ideal point.

*Proof.* Under a point transformation $x' = Tx$, a dual conic matrix D transforms to:

$$D' = TDT^\top. \tag{6}$$

Using (6), the transformed net of conics (5) is:

$$N' = TD_0 T^\top + TD_1 T^\top + TD_2 T^\top = TNT^\top.$$

Recall that mapping the central vertex $p_0$ to the origin (see (5)) cancel out the translation, leaving only uniform scale and rotation in $T$. We prove the invariance of $\gamma$ (1) to similarity transformations $S$ applied to $\mathcal{J}_2$ by replacing M by N' in (1):

$$\gamma = \frac{\mathrm{dsc}\,(N')}{\mathrm{tr}^2\,(N')} = \frac{\mathrm{dsc}\,(SNS^\top)}{\mathrm{tr}^2\,(SNS^\top)}$$

$$= \frac{\mathrm{dsc}\,(S)\,\mathrm{dsc}\,(N)\,\mathrm{dsc}\,(S^\top)}{\mathrm{tr}^2\,(NS^\top S)} = \frac{s^4\,\mathrm{dsc}\,(N)}{s^4\,\mathrm{tr}^2\,(N)} = \gamma,$$

where $s$ is the uniform scale factor in $S$. □

## 4.2 Affine Invariant for Proper and Ideal 3-junctions

The three vertices of a 2-junction do not provide enough information for covering the six degrees of freedom of an affine mapping. Therefore, it is not possible to define affine invariant values from these junctions. Using one additional vertex, the proposed affine invariants for proper 3-junctions (3) are $\beta_{1,2}$, $\beta_{1,3}$, and $\beta_{3,1}$, from (2). It is important to note that (2) is proved to be projective-invariant (Quan, 1995), but that property is lost as soon as the junction is projected on the image plane after a transformation. Despite that, the affine-invariance is preserved.

We define the matrices $Q_k$ used in (2) from the coordinates of vertices of a proper 3-junction as:

$$Q_k = v_i v_i^\top + v_j v_j^\top, \tag{7}$$

where $v_l = (x_l - x_0, y_l - y_0)^\top$, for $i \neq j \neq k$ and $i, j, k \in \{1, 2, 3\}$. By doing so, $\beta_{1,2} = \beta_{3,2}$, $\beta_{1,3} = \beta_{2,3}$, $\beta_{3,1} = \beta_{2,1}$, and $\beta_{1,2} + \beta_{1,3} + \beta_{3,1} = 1$. As a result, the number of invariant values can be reduced to two, since $\beta_{i,j} = 1 - (\beta_{j,k} + \beta_{j,i})$.

*Proof.* Let $A$ be an affine transformation whose translation was canceled out by mapping $p_0$ to the origin of the space (see $v_l$ in (7)). Let $Q'$ be the symmetric matrix in (7) after the 3-junction $\mathcal{J}_3$ be transformed by $A$:

$$Q'_k = \dot{A}v_i v_i^\top \dot{A}^\top + \dot{A}v_j v_j^\top \dot{A}^\top = \dot{A}Q_k \dot{A}^\top,$$

where $\dot{A}$ is the leading $2 \times 2$ submatrix of $A$. The invariance of $\beta_{i,j}$ to affine mapping is proved by:

$$\beta'_{i,j} = \frac{\mathrm{tr}\left(Q'^{-1}_i Q'_j\right)\det\left(Q'_i\right)}{\mathrm{tr}^2\left(Q'^{-1}_j Q'_i\right)\det\left(Q'_j\right)}$$

$$= \frac{\mathrm{tr}\left(\dot{A}^{-\top}Q_i^{-1}Q_j\dot{A}^\top\right)\det\left(\dot{A}Q_i\dot{A}^\top\right)}{\mathrm{tr}^2\left(\dot{A}^{-\top}Q_j^{-1}Q_i\dot{A}^\top\right)\det\left(\dot{A}Q_j\dot{A}^\top\right)}$$

$$= \frac{\mathrm{tr}\left(Q_i^{-1}Q_j\dot{A}^\top\dot{A}^{-\top}\right)\det^2\left(\dot{A}\right)\det\left(Q_i\right)}{\mathrm{tr}^2\left(Q_j^{-1}Q_i\dot{A}^\top\dot{A}^{-\top}\right)\det^2\left(\dot{A}\right)\det\left(Q_j\right)}$$

$$= \frac{\mathrm{tr}\left(Q_i^{-1}Q_j\right)\det\left(Q_i\right)}{\mathrm{tr}^2\left(Q_j^{-1}Q_i\right)\det\left(Q_j\right)} = \beta_{i,j}.$$

□

## 4.3 Projective Invariant for Ideal 4-junctions

The cross-ratio (Hartley and Zisserman, 2004) is a well-known projective invariant number related to a list of four collinear points $q_i$, for $i \in \{1, 2, 3, 4\}$:

$$\lambda = (q_1, q_2; q_3, q_4) = \frac{\|\overline{q_1 q_3}\|}{\|\overline{q_2 q_4}\|}\frac{\|\overline{q_2 q_3}\|}{\|\overline{q_1 q_4}\|}, \tag{8}$$

where $\|\square\|$ is the signed length of a line segment in Euclidean space. The permutation of the points alter the cross-ratio, resulting in up to six different values:

$$\left\{\lambda, \quad \frac{1}{\lambda}, \quad \frac{1}{1-\lambda}, \quad 1-\lambda, \quad \frac{\lambda}{\lambda-1}, \quad \frac{\lambda-1}{\lambda}\right\} \tag{9}$$

For ideal 4-junctions (3), the central vertex $p_0$ is a proper point, while delimiter vertices $p_i = (x_i, y_i)$ are directions, *i.e.,* $x_0 = (x_0, y_0, 1)^\top$ and $x_i = (x_i, y_i, 0)^\top$ in projective space $\mathbb{P}^2$, for $i \in \{1, 2, 3, 4\}$. Let $r_i = x_0 \times x_i$ be the straight line defined by the central vertex of the junction and the direction of its $i$-th branch. Here, $\times$ denotes the cross product. Let l be a straight line not parallel to any line $r_i$. We use the cross-ratio (8) of points $q_i = l \times r_i$ shared by l and $r_i$ to compute projective invariant values for ideal 4-junctions. Among all possible results for $\lambda$ (9), there is only one in the $[-1, 0]$ range. By taking its modulus, we have a projective invariant number in the $[0, 1]$ range that does not depend on the branches' order.

*Proof.* The proof for the invariance of $\lambda$ is presented by Hartley and Zisserman (2004). □

## 4.4 Projective Invariant for Proper 4-junctions

Similarity and affine transformations do not affect the homogeneous coordinate $w$ under a point transformation $x' = Tx$. Thus, for the invariant numbers presented in Section 4.1 and 4.2, one doesn't have to worry about normalizing the homogeneous coordinate of $x'$ in order to retrieve the actual location of the mapped point. Projective transformations, on the other hand, may change the homogeneous coordinate of mapped points. The normalization turns the definition of truly projective invariants a challenging task.

The proposed projective invariant number is robust to the application of homographies to 4-junctions and to the subsequent normalization of the homogeneous coordinates of vertices. It is computed as:

$$\alpha = \frac{\det(PP)}{\operatorname{tr}^3(PP)}, \quad (10)$$

where

$$P = \sum_{f \in \mathcal{F}} \left( D_{f_1} C D_{f_2} C D_{f_3} C D_{f_4} C \right). \quad (11)$$

In (11), $\mathcal{F}$ is a set including the 8 cyclic permutations of the indices of branches in $\mathcal{J}_4$ (3) taken clockwise and counterclockwise, C is the matrix representation of the (primal) conic section which passes through the vertices of $\mathcal{J}_4$, and $D_{f_i}$ is the dual conic section matrix related to vertex $p_{f_i}$ by (4). We have used PP instead of just P in (10) to force $\alpha \geq 0$.

The reason for multiplying matrices $D_{f_i}C$ in (11) is twofold: (i) $\alpha$ must be independent of the selection of the first branch in $\mathcal{J}_4$; and (ii) the multiplication allows us to collect from the summation the scalar factors multiplying $D_{f_i}$ and C, making $\alpha$ robust against the normalization of homogeneous coordinates.

*Proof.* Recall that under the projective transformations of points, *i.e.*, $x' = Tx$, a primal conic section matrix C transforms to:

$$C' = T^{-\top} C T^{-1}. \quad (12)$$

Also, recall that $x \equiv \delta x$, $C \equiv \delta C$, and $D \equiv \delta D$ in projective space $\mathbb{P}^2$ for any real value $\delta \neq 0$, where x, C, and D denote, respectively, a point vector, a primal conic section matrix, and a dual conic section matrix.

Using (6) and (12), let $D''_{f_i} = \delta_{f_i} D'_{f_i} = \delta_{f_i} H D_{f_i} H^\top$ and $C'' = \delta_C C' = \delta_C H^{-\top} C H^{-1}$ be, respectively, the dual conic related to vertex $p_{f_i}$ and the primal conic defined by the vertices of a 4-junction under a projective transformation $H$, for $\delta_\square \neq 0$ being any scalar factor multiplying these matrices, including the normalization factor of the homogeneous coordinate.

The transformed P matrix is given by:

$$\begin{aligned}
P' &= \sum_{f \in \mathcal{F}} \left( D''_{f_1} C'' D''_{f_2} C'' D''_{f_3} C'' D''_{f_4} C'' \right) \\
&= \sum_{f \in \mathcal{F}} \left( \delta_f H D_{f_1} C D_{f_2} C D_{f_3} C D_{f_4} C H^{-1} \right) \\
&= \delta_{\mathcal{F}} H \left( \sum_{f \in \mathcal{F}} \left( D_{f_1} C D_{f_2} C D_{f_3} C D_{f_4} C \right) \right) H^{-1} \\
&= \delta_{\mathcal{F}} H P H^{-1},
\end{aligned}$$

where $\delta_{\mathcal{F}} = \delta_f = \delta_C^4 \prod_{i=1}^4 \delta_{f_i}$ for all $f \in \mathcal{F}$. We prove the invariance of $\alpha$ by replacing P by P' in (10):

$$\begin{aligned}
\alpha' &= \frac{\det(P'P')}{\operatorname{tr}^3(P'P')} = \frac{\det\left( \delta_{\mathcal{F}}^2 H P P H^{-1} \right)}{\operatorname{tr}^3\left( \delta_{\mathcal{F}}^2 H P P H^{-1} \right)} \\
&= \frac{\delta_{\mathcal{F}}^6 \det(H) \det(PP) \det\left(H^{-1}\right)}{\delta_{\mathcal{F}}^6 \operatorname{tr}^3\left( PP H^{-1} H \right)} = \frac{\det(PP)}{\operatorname{tr}^3(PP)} = \alpha.
\end{aligned}$$

$\square$

We performed controlled experiments with continuous junctions and their transformed counterparts (see Fig. 1). As expected, each proposed number is invariant to a group of transformations.

To demonstrate the distinctiveness of the given projective invariant monomial, we built continuous 4-junctions and their projective-transformed counterparts (see Fig. 2). For each digital version of each junction, we took its central and delimiter pixels and produced 100,000 random continuous junctions by sampling points in the pixels' area. Then, we computed the histogram of $\alpha$ numbers related to the set of random samples and compared the histograms of original and transformed junctions. Fig. 2 (top) illustrates junctions used as reference and histograms of invariant values produced using the random samples. Notice that there is a well-defined peak of votes close to the actual $\alpha$ number computed for the reference junction (the dashed line). The distribution of random $\alpha$ samples produced for an unrelated junction generates a peak of votes in a different $\alpha$ value (Fig. 2, middle), while the peak of votes generated for the transformed version of the reference junction preserves the signature of the distribution (Fig. 2, bottom). Such behavior suggests that the histogram of samples may be used to characterize well-defined discrete junctions related by projective transformations.

## 5 THE PROPOSED PROJECTIVE INVARIANT DESCRIPTOR

The invariant numbers presented in Section 4 are defined for continuous junctions ($\mathcal{J}_n$) in bidimensional

| Reference | Similarity Transformation | Affine Transformation | Projective Transformation |
|---|---|---|---|

$\gamma = 0.14821$    $\gamma = 0.14821$    $\gamma = 0.238821$    $\gamma = 0.143078$

$\beta = \begin{pmatrix} 0.15 & 0.16 & 0.00 \\ 0.13 & 0.13 & 0.00 \end{pmatrix}$    $\beta = \begin{pmatrix} 0.15 & 0.16 & 0.00 \\ 0.13 & 0.13 & 0.00 \end{pmatrix}$    $\beta = \begin{pmatrix} 0.15 & 0.16 & 0.00 \\ 0.13 & 0.13 & 0.00 \end{pmatrix}$    $\beta = \begin{pmatrix} 0.02 & 0.22 & 0.01 \\ 0.01 & 0.22 & 0.01 \end{pmatrix}$

$\alpha = 0.496638$    $\alpha = 0.496638$    $\alpha = 0.496638$    $\alpha = 0.496638$
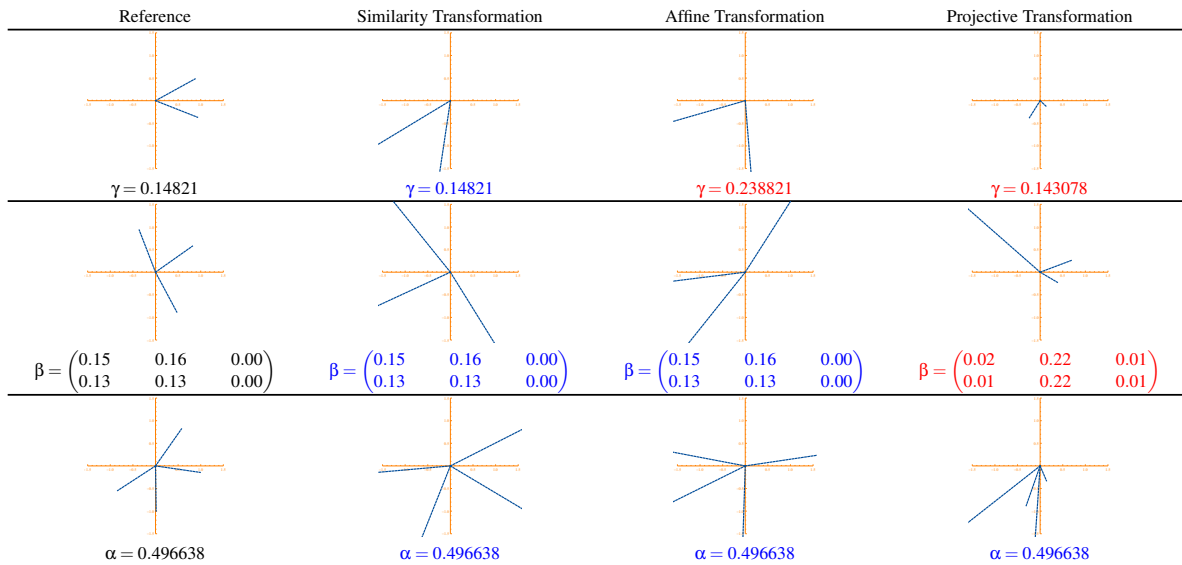
Figure 1: Each row presents one type of junction transformed by similarity, affine, and projective transformations (columns). The respective invariant value, $\gamma$, $\beta$, and $\alpha$ is at the bottom of each sample junction. Blue and red texts denote when the value remains the same or changed after the transformation of the reference junctions.

intensity images ($I$). For a given bidimensional intensity digital image $\hat{I} : \mathbb{Z}^2 \mapsto \mathbb{R}^+$, the discrete domain of $\hat{I}$ leads to the definition of digital junctions as:

$$\hat{\jmath}_n = \left\{ \hat{p} \mid \hat{p} \in \hat{\mathcal{B}}_i, \exists i \in \{1, 2, \ldots, n\} \right\},$$

where $n \in \{2, 3, 4\}$ is the degree of the junction, and $\hat{\mathcal{B}}_i$ denotes the set of pixels $\hat{p} = (u, v) \in \mathbb{Z}^2$ defining the $i$-th branch of the junction having the central pixel $\hat{p}_0 = \hat{\mathcal{B}}_i \cap \hat{\mathcal{B}}_j$ as the only pixel shared by the branches, for all $i \neq j$ and $i, j \in \{1, 2, \ldots, n\}$. By definition, the branches $\hat{\mathcal{B}}_i$ are digital straight line segments ranging from the central pixel $\hat{p}_0$ to the delimiter pixel $\hat{p}_i$.

The challenges for defining a projective invariant junction descriptor include: (i) detect well-defined 4-junctions in images (Section 5.1); (ii) compute descriptors for the detected junctions (Section 5.2); and (iii) define a matching procedure (Section 5.3).

## 5.1 Detecting Junctions in Digital Images

It is quite difficult to perform accurate detection of well-defined 4-junctions in digital images of natural scenes using computationally cheap procedures. Our proposed strategy to accomplish this task is to find image corners using the Minimum Eigenvalue Corner Detector (Shi and Tomasi, 1994) and, for each detected corner, define a region of interest (ROI) from which the analysis of the gradient leads to the identification of candidate branches leaving the corner. Candidate 4-junctions are produced from the combination
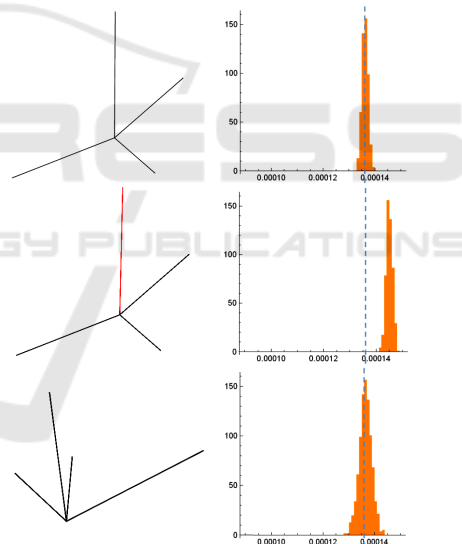
Figure 2: The distributions of $\alpha$ numbers (right) computed for random junctions samples produced from a digital junctions (left). The junction in the bottom was created by applying a homography to the junction in the top. The junction in the middle is not related to the others by a homography.

of those branches. Thus, instead of retrieve one digital junction for each detected corner, we use each corner as the central vertex of a set of junctions representing the image gradient around it. Section 5.2 shows how the $\alpha$ numbers related to these candidate junctions are combined to produce a descriptor for the ROI.

## 5.2 Computing the Feature Descriptors

Our feature descriptor is a histogram having *m* bins, which count the (weighted) frequency of invariant numbers computed for the set of digital junctions. However, our experience shows that α numbers close zero are more frequent (see Fig. 3, left). As a result, a histogram on the observations of α having a small *m* might be unrepresentative (*i.e.,* all resultant histograms would look the same). We have overcome this issue by performing a non-linear order-preserving bijective mapping $\alpha \mapsto \hat{\alpha}$ based on the theoretical cumulative distribution of α. Fig. 3 (right) shows the frequency distribution of $\hat{\alpha}$ numbers computed from the α values assumed in Fig. 3 (left). Note that frequencies are better distributed along with the $\hat{\alpha}$ axis.

Without loss of generality, let

$$\alpha = \frac{\det(PP)}{\operatorname{tr}^3(PP)} \equiv \frac{\mu_2^2 \mu_3^2}{\left(1 + \mu_2^2 + \mu_3^2\right)^3} \quad (13)$$

be the invariant number defined in (10) rewritten in terms of $1 \geq \mu_2^2 \geq \mu_3^2 \geq 0$, where 1, $\mu_2^2$, and $\mu_3^2$ are the eigenvalues of the symmetric matrix PP after being divided by the eigenvalue with the largest magnitude (recall that scaling PP does not affect α). Using (13), it is possible to conclude that $\alpha \in [0, 1/27]$, since $\max_{\mu_2, \mu_3} \alpha = 1/27$. Also, (13) allows us to define the theoretical cumulative distribution function of α as:

$$F(\alpha) = 1 - 2 \int_{\dot{\mu}_2(\alpha)}^{1} (\mu_2 - \dot{\mu}_3(\mu_2, \alpha)) \, d\mu_2,$$

where functions $\dot{\mu}_2$ and $\dot{\mu}_3$ are obtained by solving (13) for $\mu_2 = \mu_3$ and $\mu_2 = 1$, respectively.

For a given α number, we map it to $\hat{\alpha}$ using:

$$\hat{\alpha} = F(\alpha), \quad (14)$$

where $\hat{\alpha} \in [0, 1]$. We use numerical integration to evaluate *F* and its inverse, and keep a lookup list with *m* entries to perform fast mapping between α and $\hat{\alpha}$.

Each histogram assigned to a given corner pixel $\hat{c}$ is computed as follows: (i) each digital junction $\hat{\mathcal{J}}_4$ is approximated by the continuous junction $\mathcal{J}_4$ whose vertices correspond to the center of the endpixels of $\hat{\mathcal{J}}_4$; (ii) the α value of each $\mathcal{J}_4$ is computed using (10) and mapped to $\hat{\alpha}$ using (14); (iii) the $\hat{\alpha}$ values contribute to the final histogram according to the weight of their respective digital junctions, which is computed as the product of the strength of the branches.

## 5.3 Matching Procedure

Given two sets of invariant histograms, namely $\mathcal{A}$ and $\mathcal{B}$, we use the *k*-Nearest Neighbors algorithm (*k*-NN) (Friedman et al., 1977) to find the best


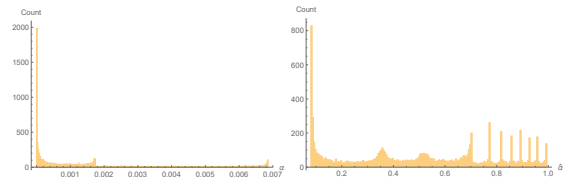
Figure 3: Distribution of α numbers before (left) and after (right) applying the non-linear mapping $\alpha \mapsto \hat{\alpha}$. The frequencies concentrate on the left side of the α domain. After mapping, they distribute better in the $[0, 1]$ domain of $\hat{\alpha}$.

match in $\mathcal{A}$ for each descriptor in $\mathcal{B}$. In our experiments, we have assumed $k = 1$ and the Earth Mover's Distance (EMD) (Rubner et al., 1998) to compare two histograms. Alternatively, any other algorithms could be used to match our feature histogram, as long as an appropriate distance metric is adopted (*e.g.,* Kullback-Leibler Divergence and Chi-Squared Test).

# 6 CONCLUSIONS

In this paper, we presented how to compute symmetric matrices from the geometry of junctions having degree two or three, and how to use these matrices with monomials invariant to, respectively, similarity and affine transformations. Using the cross-ratio, we showed how to compute a projective invariant number for ideal junctions having degree four. Also, we presented the derivation of a new projective invariant number for 4-junctions, which is independent of the cyclic order assumed for the branches of the junction and to the normalization of the homogeneous coordinate in projective space. Using the latter invariant, we defined a new procedure for detecting, describing, and matching local projective invariant junction-based features in digital images.

Results show that our descriptor is promising since it is capable of representing discriminative information. Unfortunately, the practical attempts to implement the projective descriptor in natural images failed due to several reasons, most of them related to finding digital 4-junctions, which is an open research topic, and that's why we did not perform comparisons with state-of-the-art techniques. We are currently investigating ways to improve the detection of junctions' vertices in digital images. We hope that by showing connections between the junction's geometry and existing invariant monomials and by presenting new invariant numbers and a new descriptor, this work may stimulate further researches.

## ACKNOWLEDGEMENTS

## REFERENCES

Alajlan, N., El Rube, I., Kamel, M. S., and Freeman, G. (2007). Shape retrieval using triangle-area representation and dynamic space warping. *Pattern Recogn.*, 40(7):1911–1920.

Alajlan, N., Kamel, M. S., and Freeman, G. (2006). Multi-object image retrieval based on shape and topology. *Image Commun.*, 21(10):904–918.

Aldana-Iuit, J., Mishkin, D., Chum, O., and Matas, J. (2016). In the saddle: Chasing fast and repeatable features. In *ICPR*, pages 675–680.

Balntas, V., Riba, E., Ponsa, D., and Mikolajczyk, K. (2016). Learning local feature descriptors with triplets and shallow convolutional neural networks. In *BMVC*.

Barrow, H. G. and Tenenbaum, J. M. (1981). Interpreting line drawings as three-dimensional surfaces. *Artif. Intell.*, 17(1-3):75–116.

Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (SURF). *Comput. Vis. Image Underst.*, 110(3):346–359.

Biederman, I. (1985). Human image understanding: recent research and a theory. *Computer Vision, Graphics, and Image Processing*, 32:29–73.

Bosch, A., Zisserman, A., and Muñoz, X. (2008). Scene classification using a hybrid generative/discriminative approach. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(4):712–727.

Calonder, M., Lepetit, V., Ozuysal, M., Trzcinski, T., Strecha, C., and Fua, P. (2012). BRIEF: Computing a local binary descriptor very fast. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(7):1281–1298.

Cayley, A. (1858). A memoir on the theory of matrices. *Phil. Trans.*, 148:17–37.

Fan, X., Luo, Z., Zhang, J., Zhou, X., Jia, Q., and Luo, D. (2014). Characteristic number: Theory and its application to shape analysis. *Axioms*, 3:202–221.

Forsyth, D., Mundy, J. L., and Zisserman, A. (1991). Invariant descriptors for 3D object recognition and pose. *IEEE Trans. Pattern Anal. Mach. Intell.*, pages 971–991.

Friedman, J. H., Bentely, J., and Finkel, R. A. (1977). An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Math. Softw.*, 3(3):209–226.

Gdalyahu, Y. and Weinshall, D. (1999). Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(12):1312–1328.

Hartley, R. I. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition.

Jia, Q., Fan, X., Liu, Y., Li, H., Luo, Z., and Guo, H. (2016). Hierarchical projective invariant contexts for shape recognition. *Pattern Recognit.*, 52:358–374.

Ke, Y. and Sukthankar, R. (2004). PCA-SIFT: A more distinctive representation for local image descriptors. In *CVPR*, pages 506–513.

Li, E., Mo, H., Xu, D., and Li, H. (2019). Image projective invariants. *IEEE Trans. Pattern Anal. Mach. Intell.*, 41(5):1144–1157.

Lindeberg, T. (2012). Scale selection properties of generalized scale-space interest point detectors. *J. Math. Imaging Vision*, 46(2):177–210.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *ICCV*, pages 1150–1157.

Luo, Z., Luo, D., Fan, X., Zhou, X., and Jia, Q. (2013). A shape descriptor based on new projective invariants. In *ICIP*, pages 2862–2866.

Mishchuk, A., Mishkin, D., Radenović, F., and Matas, J. (2017). Working hard to know your neighbor's margins: Local descriptor learning loss. In *NIPS*, pages 4829–4840.

Perwass, C. and Forstner, W. (2006). *Uncertain Geometry with Circles, Spheres and Conics*, pages 23–41. Springer Netherlands.

Quan, L. (1995). Invariant of a pair of non-coplanar conics in space: definition, geometric interpretation and computation. In *ICCV*, pages 926–931.

Quan, L., Gros, P., and Mohr, R. (1991). Invariants of a pair of conics revisited. In *BMVC*, pages 71–77.

Rosten, E. and Drummond, T. (2006). Machine learning for high-speed corner detection. In *ECCV*, pages 430–443.

Rosten, E., Porter, R., and Drummond, T. (2010). Faster and better: a machine learning approach to corner detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32:105–119.

Rubner, Y., Tomasi, C., and Guibas, L. J. (1998). A metric for distributions with applications to image databases. In *ICCV*, pages 59–66.

Semple, J. G. and Kneebone, G. T. (1952). *Algebraic Projective Geometry*. Oxford Science Publication.

Shi, J. and Tomasi, C. (1994). Good features to track. In *CVPR*, pages 593–600.

Tian, Y., Fan, B., and Wu, F. a. (2017). L2-Net: deep learning of discriminative patch descriptor in Euclidean space. In *CVPR*, pages 6128–6136.

Tian, Y., Yu, X., Fan, B., Wu, F., Heijnen, H., and Balntas, V. (2019). SOSNet: second order similarity regularization for local descriptor learning. In *CVPR*, pages 11016–11025.

Tuytelaars, T. and Mikolajczyk, K. (2008). Local invariant feature detectors: A survey. *Found. Trends. Comput. Graph. Vis.*, 3(3):177–280.

Yu, G. and Morel, J. (2011). ASIFT: An algorithm for fully affine invariant comparison. *IPOL*, 1:11–38.