

Affine Shape Descriptor Algorithm for Image Matching and Object Recognition in a Digital Image

Algoritmo para clasificación de formas invariante a transformaciones afines para objetos rígidos en una imagen digital

Algoritmo de classificação para as formas invariantes afins para objetos rígidos em uma imagem digital

DOI: <http://dx.doi.org/10.23913/reci.v6i11.57>

Jacob Morales G.

Centro Universitario de Ciencias Exactas e Ingenierías, Universidad de Guadalajara, México

Jacobinho86@gmail.com

Nancy Arana D.

Centro Universitario de Ciencias Exactas e Ingenierías, Universidad de Guadalajara, México

nancyaranad@gmail.com

Alberto A. Gallegos

Centro Universitario de Ciencias Exactas e Ingenierías, Universidad de Guadalajara, México

gallegos.alberto.a@gmail.com

Resumen

El uso de la forma como criterio discriminante entre clases de objetos extraídos de una imagen digital, es uno de los roles más importantes en la visión por computadora. El uso de la forma ha sido estudiado extensivamente en décadas recientes, porque el borde de los objetos guarda suficiente información para su correcta clasificación, adicionalmente, la cantidad de memoria utilizada para guardar un borde, es mucho menor en comparación a la cantidad de información contenida en la región que delimita. En este artículo, se propone un nuevo descriptor de forma. Se demuestra que el descriptor es invariante a la perspectiva, es estable en la presencia de ruido, y puede diferenciar entre diferentes clases de objetos. Un análisis comparativo es incluido para mostrar el rendimiento de nuestra propuesta con respecto a los algoritmos del estado del arte.

Palabras Clave: Reconocimiento de Patrones, Visión por Computadora, Descriptor de Forma, Transformación por Afinidad.

Abstract

The use of shape, as a mean to discriminate between object classes extracted from a digital image, is one of the major roles in machine vision. The use of shape has been studied extensively in recent decades, because the shape of the object holds enough information for its correct classification; additionally, the quantity of memory used to store a border is much less than that of the whole region within it. In this paper, a novel shape descriptor is proposed. The algorithm demonstrates that it has useful properties such as: invariance to affine transformations that are applied to the border (e.g., scales, skews, displacements and rotations), stability in the presence of noise, and good differentiability between different object classes. A comparative analysis is included to show the performance of our proposal with respect to the state of the art algorithms.

Key words: Pattern Recognition, Machine Vision, Shape Descriptor, Affine Transformation.

Resumo

Usando critério discriminante do caminho entre as classes extraídas de uma imagem digital, objetos é um dos papéis mais importantes na visão de computador. A utilização da forma tem sido extensivamente estudado nas últimas décadas, porque a borda de objectos guardados suficientes relatórios-ing de classificação correcta, para além disso, a quantidade de memória usada para armazenar uma aresta, é muito mais baixa em comparação com a música quantidade de informação na região que delimita. Neste artigo, um novo descritor forma é proposto. Mostra-se que o descritor é invariante em perspectiva, que é estável na presença de ruído, e é possível diferenciar entre diferentes tipos de objectos. Uma análise comparativa é in-concluiu para mostrar o desempenho da nossa proposta sobre o estado dos algoritmos de arte.

Palavras-chave: Reconhecimento de Padrões, visão computacional, descritor de formulário, Affinity transformação.

Fecha recepción: Julio 2016

Fecha aceptación: Diciembre 2016

Introduction

A shape descriptor is a qualitative or quantitative abstraction of the features that are present in the border of an object (Gonzalez & Woods, 2008). The shape of an object is a powerful low-level descriptor because it provides plenty of information about the object because the dynamic of the form is embedded in the border and, can be captured in a digital image. Working with only the border is faster computationally given that the number of points in the border is smaller than the number of points in the region within it, because the shape is used as a discriminating feature in object recognition applications. The shapes present in the image can be affected by various factors, such as: illumination, occlusion or a perspective change.

The change in perspective is of special interest because it can be modeled with functions including: rotation, translation, scaling and skewing; all of these functions share the same algebraic representation in the Euclidean plane. They preserve points, straight lines, collinearity, planes and ratios of distances; also, they can be reversed (Artzy, 2008). These types of functions are classified as affine transformations or geometrical distortions. Thus, it is said that if two different shape descriptors are connected by affine transformations, then one can be obtained from the other, i.e., both come from the same object class. Many affine invariant descriptors have been proposed in the literature (Kazmi, You, & Zhang, 2013; Pillai, 2013; Krig, 2014), that extract boundary or region information. One way to classify boundary descriptors is by the type of information that is captured by it; if it is global information, the descriptor is robust against noise but the border dynamic is lost in the process (shape number descriptors are an example (Gonzalez & Woods, 2008)), in contrast, if the information captured by the boundary descriptor is local information, then the border dynamic is preserved but is very unsteady in the presence of noise or geometrical distortions (e.g., Fourier border descriptors (Gonzalez & Woods, 2008)).

It is worth mentioning that an effective descriptor ideally has the following properties:

Uniqueness, which means that every class of object has a unique descriptor representation.

Differentiability, which means that every different class of object has a different representation.

Stability, which means that minimal change in the object induces minimal change in the representation.

Other criteria such as robustness to illumination, incompleteness or resolution require specific techniques and considerations (Krig, 2014); for brevity; these will not be addressed here.

Recent developments in shape descriptors seek to clean the points of interest from any affine transformation before processing them in other stages of the application. Statistical functions are used to reverse all affine transformations that could have been applied to the shape. The statistical process has the very attractive property that there is no need to search or infer parameters. In the work presented in (Mei & Androustos, 2008), ICA (independent component analysis) and PCA (principal component analysis) are the methods proposed for reducing the shape into a canonical form; both statistical mechanisms demonstrated their merit by reversing arbitrary geometrical distortions without the need to calculate the specific parameters of their inverse functions. However, the resulting canonical form is not unique, such indetermination is a result of the nature of both processes; specifically, for every affine transformed class of borders that lie in the two dimensional Euclidean space, two canonical forms exist. One form is the 180-degree rotation of the other. In order to deal with such indetermination, the authors proposed computing the Fourier transform of the sampled signature curve of the border. Later, in (Mei & Androustos, 2009), the authors used the CSS (curvature scale space) maxima as the affine invariant descriptor. The proposed descriptors share invariance to 180° rotations. In contrast in (Guney & Ertuzun, 2006), an exhaustive search that maximizes correlations between the original and transformed object boundaries is done to correct the sign ambiguities; similarly, in (Huang, Wang, & Zhang, 2005), a maximization is conducted over the negentropy parameter. Also in (Sener & Unel, 2006), the authors use higher p - q order moments of normalized shapes to solve the indetermination. In this work, the proposed method uses a similar approach; first, the borders extracted are normalized with PCA and then the skew of the resulting shapes is calculated. Depending on its sign, the border will either be rotated 180° or not, leaving a unique canonical form. Note

that only the sign of the skew is of interest, and that the calculation will be simplified because of the properties of the transformed points.

The sections of this paper are organized as follows. Section "State of the Art" presents a review of the shape descriptors in the state of the art. In section "Preliminaries", all the mathematical background and border processing algorithms that are needed are briefly described. Section "Affine Invariant Shape Matching Classification" explains in detail the theoretical development of our proposal and presents the practical development. Next, section "Experiments and Analysis" explains the applied experimental tests and shows the obtained results, highlighting the pros and cons of our approach with respect to the proposed state of the art algorithms, and finally, in section "Conclusions", the results are interpreted.

State of the Art

In the following subsections, some state of the art algorithms are described; these were chosen because they have a large body of work behind them, they are largely used in industry and academia and they have been proven to be effective (Kazmi et al., 2013; Pillai, 2013).

Fourier Descriptor

The Fourier shape descriptor is one of the most applied shape descriptors (Rajput & Mali, 2010; Ekombo, Ennahahi, Oumsis, & Meknassi, 2009). The descriptor is obtained after the application of the Fourier transform to the coefficients of the shape signature. The four most utilized shape signatures are the centroid distance, complex coordinates, curvature function and cumulative angular function; among them, the centroid distance is the one that performs best (Kazmi et al., 2013).

Before taking the border signature, the border has to be normalized to translations by subtracting the centroid of the shape; then, every point is divided by the standard deviation of each component to make it invariant to scale (Gonzalez & Woods, 2008). Next, the signature

is sampled. Then, the Fourier transform is applied to each coefficients of the sampled signature; see Equation 1.

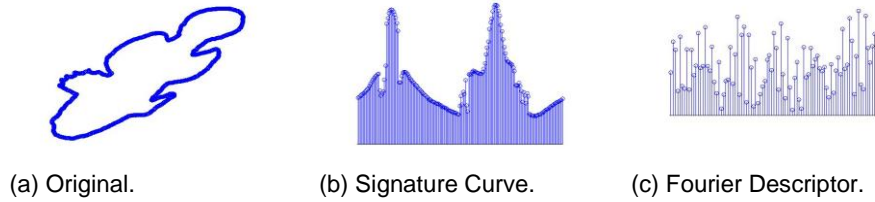


Figure 1. Fourier Descriptor.

$$u_n = \frac{1}{N} \sum_{t=0}^{N-1} s(t) \exp\left(\frac{-j2\pi nt}{N}\right), n = 0, 1, \dots, N - 1 \quad (1)$$

The descriptor is completed by taking the module of every u_n coefficients; Figure 1 depicts an example.

Wavelet Descriptor

The wavelet descriptor of closed planar curves, is a multi-resolution approach that decomposes the shape into components of several scales and is widely used in shape recognition applications (Ming et al., 2009). The elements in higher resolutions contain the global information, and the components in the finer resolutions contain detailed local information (Kazmi et al., 2013). The reduction of forms as stated in (Chuang & Kuo, 1996) consist of the decomposition of every dimension of the parametrized curve into approximation coefficients $(x_a(t), y_a(t))$ at scale M and detailed coefficients $(x_d(t), y_d(t))$ at scale M by means of the wavelet transform (with a bi-orthogonal basis); see Equation 2.

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} x_a^M(t) \\ y_b^M(t) \end{bmatrix} + \sum_{M-m_0}^M \begin{bmatrix} x_d^M(t) \\ y_d^M(t) \end{bmatrix} \quad (2)$$

Next, these coefficients have to be normalized to scale, rotation and shift (Chuang & Kuo, 1996). With the normalized coefficients, the shape is recovered with the inverse transform and the signature is taken, because this method requires a fixed evaluation starting point (Chuang & Kuo, 1996). Finally, the signature curve of the normalized shape is sampled; see Figure 2.

CSS Descriptor

This is one of the mostly widely used algorithms in content-based image retrieval (CBIR) and general shape analysis (Frejlichowski, 2012). This method divides the shape into convex and concave segments, identifying a set of points where the curvature changes sign. The CSS algorithm involves calculating the curvature of the contour while progressively smoothing the curve. The smoothing process is done until the point when there are no zero curvature points left and the whole contour becomes convex. The CSS image is a curve on a two-dimensional plane that plots all the zero crossing points. The horizontal.

Axis contains the position of the points on the contour while the vertical axis contains the level of smoothing performed on the contour (Kazmi et al., 2013). As stated in (Abbasi, Mokhtarian, & Kittler, 1999), given a parametrized curve in space $l(t) = (x(t), y(t))$, its points of inflection at σ level of smoothing can be found in the zero crossings of Equation 5.

$$X_t(t, \sigma) = x(t) * g_t(t, \sigma), \quad Y_t(t, \sigma) = y(t) * g_t(t, \sigma) \quad (3)$$

$$X_{tt}(t, \sigma) = x(t) * g_{tt}(t, \sigma), \quad Y_{tt}(t, \sigma) = y(t) * g_{tt}(t, \sigma) \quad (4)$$

$$k_t(t, \sigma) = \frac{X_t(t, \sigma)Y_{tt}(t, \sigma) - X_{tt}(t, \sigma)Y_t(t, \sigma)}{(X_t(t, \sigma)^2 + Y_t(t, \sigma)^2)^{3/2}} \quad (5)$$

The descriptor consists of the peaks of every curve in the CSS image, with a height bigger than one sixth of the biggest peak in the image; see Figure 3. Before starting the process, the shape points have to be equally sampled so every shape has the same perimeter; shifting the shape in space does not change the descriptor because of the nature of the descriptor. The nature of the descriptor is independent of a coordinate system, and rotation is attacked by a circularly shifting model and input descriptors to find the best match, i. e., the version of the minimum distance classifier in the CSS algorithm presented in (Abbasi et al., 1999); see Figure 3.

Other Algorithms

There are other the state of the art algorithms that exploit features other than shape; for completeness, it is worth mentioning them briefly. The local binary pattern descriptor (Pietikainen, Hadid, Zhao, & Ahonen, 2011) is a local texture codification that is quick to compute and capture the features in a tight manner. There are several variations of the LBP descriptor where the uniform, rotational invariant and symmetric stand out from the others and are widely used in recognition applications (Nanni, Lumini, & Brahnam, 2010). The binary robust invariant scalable keypoints (Leutenegger, Chli, & Siegwart, 2011), and the fast retina keypoint (Ortiz, 2012) descriptors are also local neighborhood descriptors that also include a keypoint detector. They were designed for quick computation, and are widely used in real time recognition applications. The speed-up robust features descriptor (Bay, Ess, Tuytelaars, & Van Gool, 2008), is also a local texture descriptor with many variations, such as Gauge SURF (Alcantarilla, Bergasa, & Davison, 2013). The SURF family of descriptors is regarded as one of the standard descriptors in industrial applications.

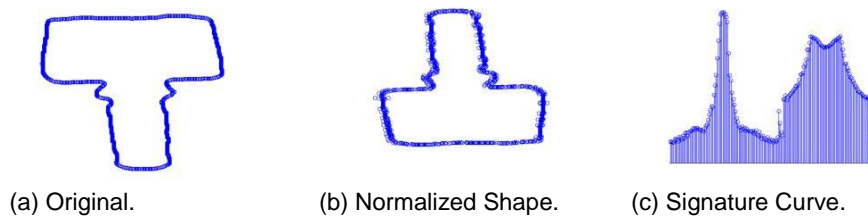


Figure 2 . Wavelet Descriptor.

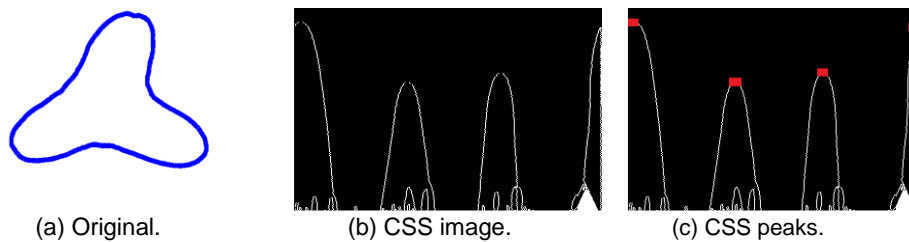


Figure 3 . CSS Descriptor.

Every descriptor mentioned before was tested with the battery constructed for this work and was noted that the shape descriptors gave the best results. Confirming that shape descriptors are a better choice when the information in the frame has low variance, i.e., the intensity values in the frame are all clustered in regions, the extreme case is a black frame with white regions; in that case, there are no relevant keypoints to be detected. Thus, the present work focuses on shape descriptors.

Preliminaries

Border

The borders obtained from objects in a digital image, can compose a set of two dimensional points, that is, every border can be defined as a finite and countable subset of the \mathbb{R}^2 Euclidean space with the property that every point is connected with the others by a path that, is defined by a criterion of a neighborhood (e.g., four neighbors or eight neighbors) and every point is a frontier point, meaning that at least one neighbor in the neighborhood of every point is not part of the object in question (Haralick & Shapiro, 1992; Gonzalez & Woods, 2008). If we denote an arbitrary border as β , then its definition is as follows:

$$\beta = \{p \mid p \in \mathbb{R}^2\} \quad (6)$$

Affine transformations

An affine transformation can be decomposed in a linear transformation followed by a translation, where such a lineal application must preserve the collinearity between points and the ratio of vectors along a line. An affine transformation is invertible if and only if the linear transformation is invertible. The invertible affine transformations (of a space into itself) form the affine group. Affine transformations in two real dimensions include the following:

- Pure translations.
- Scaling in a given direction, with respect to a line in another direction, combining negative signs and translations. This application includes projection, reflection and glide reflection.

- Rotation.
- Shear mapping.
- Squeeze mapping.

Equation 7 shows the algebraic representation of an affine transformed point in the plane, with original coordinates (x, y) .

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (7)$$

Or in a more compact algebraic notation as follows.

$$\bar{x} = Ax + B \quad (8)$$

Where the non-singular matrix with real entries A can be viewed as a composition of scale, skew and rotation functions in the plane as seen in Equation 9.

$$\begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (9)$$

Vector B represents a constant shift function. Two different borders, β_1 and β_2 are affine-connected if one is the result of applying the same affine transformation to every point of the other border.

$$\beta_2 = \{p_2 \mid p_2 = Ap_1 + B\}, \forall p_1 \in \beta_1 \quad (10)$$

More detailed information of geometrical transformations can be found in (Artzy, 2008).

Principal Component Analysis

We have defined a border β as a finite countable set of points p where $p = [x, y]^T$ is a 2×1 real entry matrix. The following auxiliary quantities are defined over the points in

the border, i.e., the mean vector and the covariance matrix.

$$M_p = E\{p\}, \forall p \in \beta \quad (11)$$

$$C_p = E\{(p - M_p)(p - M_p)^T\} \quad (12)$$

Where $E\{\bullet\}$ denotes the expected value of the argument. For K points in a border, the mean vector can be approximated with the averaging expression.

$$M_p = \frac{1}{K} \sum_{i=1}^K p_i \quad (13)$$

Similarly, the covariance matrix could be approximated from the points in the border.

$$C_p = \frac{1}{k} \sum_{i=1}^K p_i p_i^T - M_p M_p^T \quad (14)$$

Because the points are two-dimensional the vector M_p is also two-dimensional and C_p is a 2X2 matrix. Given that C_p is a real symmetric matrix, finding a set of two orthonormal eigenvectors is always possible (Noble & Daniel, 1988). Let e_1 and e_2 denote the eigenvectors and let λ_1 and λ_2 denote the eigenvalues associated with each eigenvector, respectively, arranged in descending order so that $\lambda_1 > \lambda_2$. Now let S denote a matrix where its rows are the eigenvectors that were found earlier ordered so that the first row corresponds to the vector associated with the largest eigenvalue. Now let us use S as a transformation matrix and define a mapping as seen in Equation 15.

$$\bar{p} = S(p - M_p) \quad (15)$$

The last relation is known as the Hotelling transform or the principal component transformation. These new transformed points have several useful properties; the most notorious is that these p^- points have zero mean. Other property come from linear algebra, i.e., the covariance matrix C_p of the original points is similar to the covariance matrix C_{p^-} of the transformed points via the similitude matrix S (Strang, 2003).

$$C_{\bar{p}} = S C_p S^T \quad (16)$$

The resulting matrix C_{p^-} in Equation 16 is a diagonal matrix whose elements along the main diagonal are the eigenvalues of C_p (for more information about similitude relations refer to (Strang, 2003).

$$C_{\bar{p}} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad (17)$$

Equation 17 gives two of the properties that will be used extensively later; first, the variance along each dimension and the statistical independence of the dimensions.

Lagrange Polynomial

Lagrange polynomial interpolation is a reformulation of the Newton polynomial avoiding the calculations of divided differences. The precise definition for a $n - 1$ grade polynomial that passes through n points in the plane ($p_i = [x_i, y_i]$, for $i = 0, \dots, n$) can be

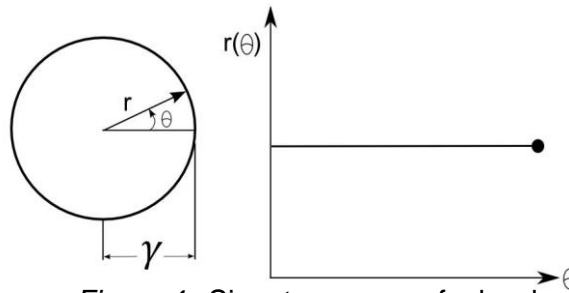


Figure 4 . Signature curve of a border.

Seen in Equation 18.

$$P_{n-1} = \sum_{i=0}^n L_i(x)f(x) \quad (18)$$

where:

$$L_i(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j}, j \neq i \quad (19)$$

Between two points in the plane ($p_0 = [x_0, y_0]$ and $p_1 = [x_1, y_1]$) exists only one line that passes through both (Press, Teukolsky, Vetterling, & Flannery, 2007); any value within such a line can be interpolated by a Lagrange polynomial of degree one with the following formula.

$$l(x) = \frac{x - x_1}{x_0 - x_1}y_0 + \frac{x - x_0}{x_1 - x_0}y_1 \quad (20)$$

Border Signature

The border signature is a one-dimensional curve that captures the border variation with respect to one parameter. The most utilized parametrization is to compute the distance to the centroid of every point in the border as a function of the angle that is formed between the point and the positive x axis (Gonzalez & Woods, 2008; Haralick & Shapiro, 1992). Regardless of how the signature is parametrized, the goal is to achieve a reduction in the

dimensionality of a two-dimension problem to a one-dimension problem that, is apparently easier to manage. Figure 4 shows a border and the associated signature curve.

Affine Invariant Shape Matching Classification (AISMC)

Given an arbitrary cloud of points that is extracted from the border of an object in a digital image as input, the proposed AISMC algorithm consists of the following steps:

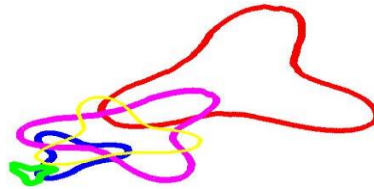
1. Compute the mean vector of the points using Equation 13.
2. Compute the covariance matrix of the points using Equation 14.
3. Obtain the eigenvectors and eigenvalues of the covariance matrix.
4. Build the transformation matrix S , using the eigenvectors ordered according to their associated eigenvalues.
5. Apply Equation 15 to every point.
6. Scale the points obtained in the last step with the standard deviation of each dimension.
7. Compute the skew of one dimension and multiply it by -1 if necessary.
8. Get the norm and angle of every point in the shape.
9. Sort the points by angle.
10. Build the sampled signature curve of the shape (using Equation 19 to interpolate when necessary).
11. Obtain the difference vector between the signature that was computed in the last step and every signature curve in the knowledge set.
12. Get the norm of every difference vector in the last step.
13. Find the minimum norm.
14. Assign the class label to the class related to the vector that returned the minimum norm.

Steps 1 to 10 comprise the canonical shape normalization phase of the algorithm, in which any arbitrary composition of affine transformations is reversed, reducing the shape to a canonical form. The next steps encompass the classification stage by using a minimum distance policy of classification.

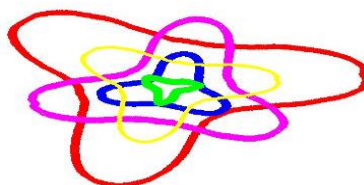
Canonical Form Normalization

Start with a collection of affine-connected shapes as seen in Figure 5a. Each border contains an arbitrary number of points. Given an arbitrary shape in the collection, first, the mean vector has to be computed (step 1) in addition to the covariance matrix (step 2), over the points in the border. Next, build the transformation matrix S (the matrix whose rows are the eigenvectors of the covariance matrix arranged by descending eigenvalue) and apply the PCA transformation to every point in the border (steps 3 to 5). The first evident result is that all the borders are centered at a point (see Figure 5b), but there is still a sign ambiguity that is not as obvious (that represents a 180° rotation between the same shapes) and a loud magnitude discrepancy. It has been shown that the standard deviation that is used as a scaling factor gives good results (Haralick & Shapiro, 1992; Gonzalez & Woods, 2008; Pratt, 2001); this, is very convenient when computing Equation 17, where we obtain the variance along every dimension of the new shape. Next, (step 6) normalizes each component of every point in the shape with the standard deviation of the corresponding direction (the square root of the eigenvalue associated with that dimension); in short, every point in the shape should have the following form:

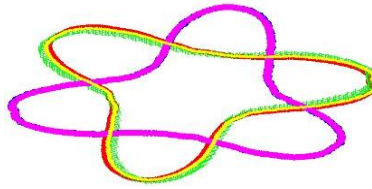
$$\bar{p} = S \begin{pmatrix} \frac{p_x - m_x}{\sqrt{\lambda_1}} & \frac{p_y - m_y}{\sqrt{\lambda_2}} \end{pmatrix} \quad (21)$$



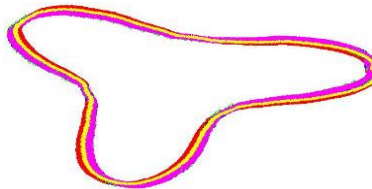
(a) Affine connected shapes.



(b) PCA mapped points.



(c) Whitened shapes.



(d) Canonic form.



(e) Signature Curve.

Figure 5. AISMC form normalization.

Equation 21 is known as whitening because the points have zero mean and unitary variance, such as white noise (Mei & Androutsos, 2009); the result is depicted in Figure 5c. Now that the inconsistency of the sign is very evident, let us discuss that inconsistency for a moment. The results showed that for any affine connected border, two valid canonical forms exist from which the transformed shape might have been obtained (breaking the uniqueness property). One way to think about this indetermination is to take a look back at the relation that the eigenvectors and eigenvalues of the covariance matrix C_p must meet.

$$(C_p - \lambda_i I)e_i = 0, \text{ for } i = 1, 2 \quad (22)$$

If vector e_i is multiplied by -1, the result would be the same; therefore there is a choice between two vectors with the same magnitude but opposite directions, and Figure 5c shows two classes of shapes with opposite directions. Specifically, one class is rotated 180° with respect to the other. Now it becomes clear that the process so far is intrinsically not deterministic. In general, it is not possible to make a bijective relation between the arbitrary rotations and the resulting eigenvectors of the points; therefore, it is necessary to find

alternative ways to work around this problem. Let us propose the following: compute the skew of the x component of the points, and use it as a measure of orientation.

$$\varepsilon = \frac{1}{K} \sum_i^K x_i^3, x_i \in \bar{p}, \forall \bar{p}_i \in \beta \quad (23)$$

The y dimension of the points is neglected because it is a statistically independent set of two-dimensional points (see Equation 17). Because the magnitude of the skew is not of interest, the sign is sufficient to determine the orientation of the shape. Taking the positive direction (arbitrarily), it is decided that every cloud of points with a negative skew is going to be multiplied by -1 (see that number as a 180° rotation), as seen in Equation 24.

$$\bar{p} = \begin{cases} -\bar{p}, & \text{if } \varepsilon < 0 \\ \bar{p}, & \text{otherwise} \end{cases} \quad (24)$$

The results can be seen in Figure 5d. With this we can find one canonical form for a class of affine connected shapes. Next (step 8) is to take the k points of the border and compute the distance to the origin and the angle that the point makes with the positive x axis. Now a sampling rate must be selected to sample the border; if the situation occurs when the difference between angles of consecutive points is too large to accommodate the sampling step, then the value can be interpolated using Equation 19. Otherwise, if the difference between angles of consecutive points is too small to accommodate the sampling step then the next point must be considered. The result is the border signature of the sampled border as seen in Figure 5e.

It should be noted that the descriptor rendered by the process was designed to render a unique representation, i.e., it has the uniqueness property. It also has the differentiability property because different border dynamics will have different signature curves. If minimal changes are applied to the border, such as Gaussian noise, then the points in the border will be minimally shifted from the centroid; therefore, the border signature will be shifted proportionally, so the descriptor has the stability property.

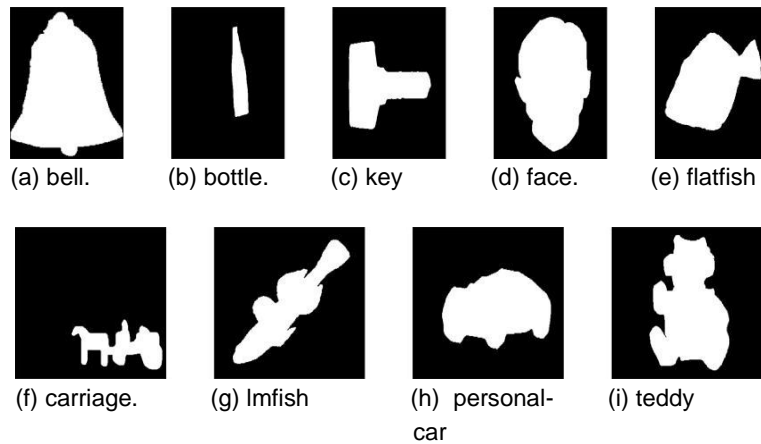


Figure 6. Repository Subset.

It should be noted that the mean vector is a key factor in the calculation; hence, this descriptor is sensitive to anything that changes it. For example occlusion, where partial incompleteness of the shape would render a different mean vector to the one obtained from the same non occluded shape resulting in two different descriptors. However, same shapes that take on different illumination conditions can effectively change the descriptors. Illumination is the source of all images, so it is the most important factor to consider, it also represents its own problematic (Krig, 2014).

Minimum Distance Classification

The classification phase needs a knowledge set that is a collection of sampled border signatures denoted by Λ . Every element in the collection represents a prototype of the canonical form of a known object; let us use ω_i to denote an arbitrary signature set in the collection. Taking these border signatures as descriptors, the classification step is defined by taking the difference of the unknown signature δ and every prototype signature ω_i in the knowledge set. The class label is assigned to the vector that renders the minimum norm, as follows:

$$class = i, for \min \|\delta - \omega_i\|, \forall \omega_i \in \Lambda \quad (25)$$

It should be noted that this equation is a very low complexity calculation with respect to the other classification algorithms, but it will be used to measure the descriptor properties.

Experiments and Analysis

In this section, the test that was used to demonstrate the performance of the algorithms is explained in Section. Then, the methodology used to pick the sample rate parameter is presented in Section and finally, the experiments and results are described in Section.

Test Set

The test set was extracted from two different sources: the first subset test was taken from the image repository (manual in preparation, 2016) and is shown in Figure 6, a second subset was selected because of the similarity of the shapes in order to measure the differentiability of the descriptors and consists of six pollen particles (Figure 7); taken as models; test patterns were constructed by applying the affine transformation shown in Equation 26.

$$SRI_i, \text{ for } i = 1,2,3,4,5,6 \quad (26)$$

Where S represents a resolution transformation, the magnitude of the original model was transformed by a factor taken randomly from the interval [10%,200%], R represents a rotation transformation, the angle of the rotation is a number taken randomly from the interval $[0^\circ,360^\circ]$, and I represents the original image or original shape. The displacement that completes the affine transformation and the scale transformation is embedded in the resolution factor, and the skew is a result of the composition of the resolution and rotation functions. Then, another set was constructed by adding noise to each image of the object that in general represents non affine transformations and non-reversible maps, specifically as follows:

Blur filter, which is a 5x5 Gaussian filter with $\sigma = 20$.

Pepper noise, which is 25% of the pixels in the image, taken randomly, where changed to 0.

Half information reduction, which reduces the image by taking one of every two rows and columns.

Quarter information reduction, which reduces the image by taking one of every four rows and columns.

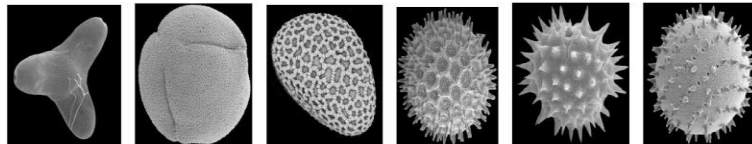


Figure 7. Synthetic Subset.

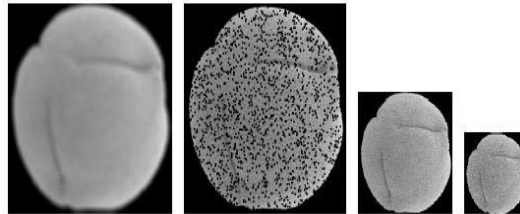


Figure 8. Adding noise to the synthetic set.

<u>Sampling Rate</u>	<u>Degrees</u>
Fourier	12.5
Wavelet	19.47
Canonical Shape	21.6

Figure 9. Sampling Rate.

The results of the above processes are depicted in Figure 8. All of the above processes comprised the set that was used to test the algorithms; in total, 385 pattern tests were applied.

Sample Rate

As mentioned in sections: , and , at some point in the algorithms pipeline, a sampling rate needs to be established, but there is no deterministic way of doing that because there is no wrong answer; any value could deliver a valid answer, so the problem becomes a search for the sampling rate that provides the best performance. In other words, the sampling rate has to be optimized.

For that, the real domain version of the univariate marginal distribution algorithm was used to find such a value (Simon, 2013). It was executed with a search domain defined in the interval [0.1, 360] and a set of 10 candidate solutions, from which a subset of 6 solutions is taken to create the next generation with a maximum of 100 generations; the algorithm was executed 10 times. The results can be seen in Figure 9.

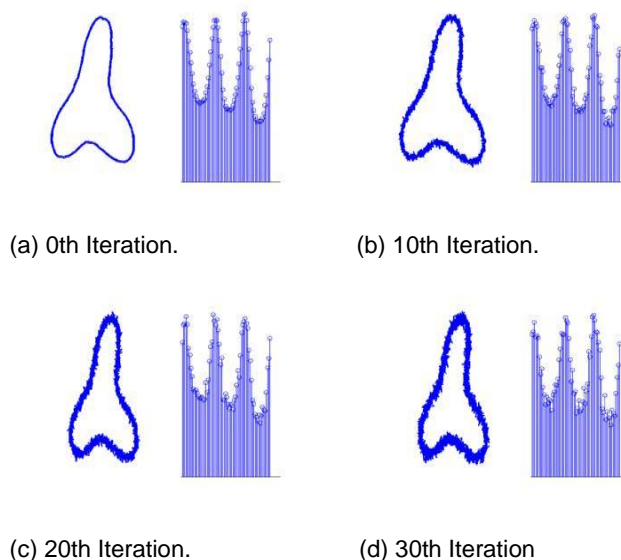
Results

The same data set was used with each algorithm (Section 2). The performance was measured with the minimum distance classifier, excluding the CSS algorithm that has defined its own classification step (Abbasi et al., 1999), the implementations were made in Matlab 2012b, and test were performed in an acer M5 computer.. The results are listed in Figure 10, as a measure of differentiability. In order to test the stability property of the proposed descriptor, a class prototype was chosen and noise was added in an iterative manner. In each iteration, each point in the shape was randomly moved in a ten-unit ratio, then the canonical shape was taken and the stability was measured by taking the medium quadratic error (MSE) of each point in the modified shape with respect to its corresponding point in the initial prototype, as noted in Figure 11. The obtained descriptor reflects the changes induced to the initial shape in a proportional way, providing enough stability to the recognition process of similar shapes. Finally, the same methodology of computing the MSE

Algorithm	Classification Rate(%)
Canonical Shape	74
Fourier	51
Wavelet	40.5
CSS	63.5

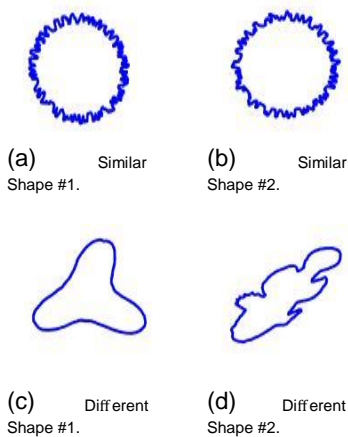
Figure 10. Differentiability Results.

Magnitude was used to measure uniqueness between some prototypes in the knowledge set; specifically, the error was measured between a couple of different shapes in the set and with similar shapes, as depicted in Figure 12. The results are measured in error %. As the reader can see, similar shapes as shown in the Figure have a lower percentage of uniqueness; meanwhile, very differentiable shapes have a higher MSE, thus complying with the property that similar shapes must have similar descriptors and different shapes must have different descriptors.



Iteration	Error Rate(%)
0th Iteration	0
10th Iteration	0.24
20th Iteration	0.6
30th Iteration	0.8

Figure 11. Stability Results.



Iteration	Error Rate(%)
Similar	1.34
Different	26.5

Figure 12. Uniqueness Results.

Conclusions

The method proposed to reverse affine transformations in a deterministic manner have an easy formulation, and its implementation results in a mostly linear cost algorithm. Because the computational load depends directly on the number of points, it has to be noted that obtaining eigenvectors and eigenvalues of a matrix are a computationally heavy operation, but in this context, all the matrices are two by two, making it reasonable to always take the worst case as a constant. It was also shown that our descriptor meets the properties of uniqueness, differentiability and stability. However, it has to be said that it is not robust against partial occlusion, because it does not capture local information; this method is a global shape descriptor. The results were promising; the canonical shape signature captures the shape dynamic in a manner comparable with the state of the art descriptors.

Bibliography

- Abbasi, S., Mokhtarian, F., & Kittler, J. (1999, November). Curvature scale space image in shape similarity retrieval. *Multimedia Syst.*, 7 (6), 467–476. Retrieved from <http://dx.doi.org/10.1007/s005300050147> doi: 10.1007/s005300050147
- Alcantarilla, P. F., Bergasa, L. M., & Davison, A. J. (2013, January). Gauge-surf descriptors. *Image Vision Comput.*, 31 (1), 103–116. Retrieved from <http://dx.doi.org/10.1016/j.imavis.2012.11.001> doi: 10.1016/j.imavis.2012.11.001
- Artzy, R. (2008). *Linear geometry* (first ed.). Dover.
- Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008, June). Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110 (3), 346–359. Retrieved from <http://dx.doi.org/10.1016/j.cviu.2007.09.014> doi: 10.1016/j.cviu.2007.09.014
- Chuang, G.-H., & Kuo, C.-C. (1996, Jan). Wavelet descriptor of planar curves: theory and applications. *Image Processing, IEEE Transactions on*, 5 (1), 56-70. doi: 10.1109/83.481671

Ekombo, P. L. E., Ennahahi, N., Oumsis, M., & Meknassi, M. (2009). Application of affine invariant fourier descriptor to shape based image retrieval. *International Journal of Computer Science and Network Security*, 9 (7), 240–247.

Frejlichowski, D. (2012). Application of the curvature scale space descriptor to the problem of general shape analysis. *Przegląd Elektrotechniczny*, 88 (10b), 209–212.

Gonzalez, R. C., & Woods, R. E. (2008). *Digital image processing* (third ed.). Prentice Hall.

Guney, N., & Ertuzun, A. (2006). Undoing the affine transformation using blind source separation.

In J. Rosca, D. Erdogmus, J. Prncipe, & S. Haykin (Eds.), *Independent component analysis and blind signal separation* (Vol. 3889, p. 360-367). Springer Berlin Heidelberg.

Haralick, R. M., & Shapiro, L. G. (1992). *Computer and robot vision* (first ed., Vol. 1). Addison Wesleyl.

Huang, X., Wang, B., & Zhang, L. (2005, July). A new scheme for extraction of affine invariant descriptor and affine motion estimation based on independent component analysis. *Pattern Recogn. Lett.*, 26 (9), 1244–1255. Retrieved from <http://dx.doi.org/10.1016/j.patrec>

.2004.11.006 doi: 10.1016/j.patrec.2004.11.006

Kazmi, I., You, L., & Zhang, J. J. (2013, Aug). A survey of 2d and 3d shape descriptors. In *Computer graphics, imaging and visualization (cgiv), 2013 10th international conference* (p. 1-10). doi: 10.1109/CGIV.2013.11

Krig, S. (2014). *Computer vision metrics* (first ed.). Apress. manual in preparation, P. (2016). *Mpeg7 data set*. http://rduin.nl/manual/PRTTools_9.html. (Accesed: 2016-04-20)

- Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011, Nov). Brisk: Binary robust invariant scalable keypoints. In *2011 international conference on computer vision* (p. 2548-2555). doi: 10.1109/ICCV.2011.6126542
- Mei, Y., & Androustos, D. (2008, Dec). Affine invariant shape descriptors: The ica-fourier descriptor and the pca-fourier descriptor. In *Pattern recognition, 2008. icpr 2008. 19th international conference on* (p. 1-4). doi: 10.1109/ICPR.2008.4761381
- Mei, Y., & Androustos, D. (2009, May). Pca-whitening css shape descriptor for affine invariant image retrieval. In *Electrical and computer engineering, 2009. ccece '09. canadian conference on* (p. 235-238). doi: 10.1109/CCECE.2009.5090127
- Ming, D., Zhang, C., Bai, Y., Wan, B., Hu, Y., & Luk, K. D. K. (2009, May). Gait recognition based on multiple views fusion of wavelet descriptor and human skeleton model. In *2009 ieee inter-national conference on virtual environments, human-computer interfaces and measurements systems* (p. 246-249). doi: 10.1109/VECIMS.2009.5068902
- Nanni, L., Lumini, A., & Brahmam, S. (2010, June). Local binary patterns variants as texture descriptors for medical image analysis. *Artif. Intell. Med.*, 49 (2), 117–125. Retrieved from <http://dx.doi.org/10.1016/j.artmed.2010.02.006> doi: 10.1016/j.artmed.2010.02.006
- Noble, B., & Daniel, J. W. (1988). *Applied linear algebra* (third ed.). Prentice Hall.
- Ortiz, R. (2012). Freak: Fast retina keypoint. In *Proceedings of the 2012 ieee conference on computer vision and pattern recognition (cvpr)* (pp. 510–517). Washington, DC, USA: IEEE Computer Society. Retrieved from <http://dl.acm.org/citation.cfm?id=2354409.2354903>
- Pietikainen, M., Hadid, A., Zhao, G., & Ahonen, T. (2011). *Computer vision using local binary pat-terns*. London: Springer Verlag. Retrieved from <http://opac.inria.fr/record=b1133835>

Pillai, C. (2013). A survey of shape descriptors for digital image processing. *IRACST-International Journal of Computer Science and Information Technology and Security*, 3 (1).

Pratt, W. K. (2001). *Digital image processing, piks inside* (third ed.). Wiley.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes* (third ed.). Cambridge.

Rajput, G., & Mali, S. (2010). Fourier descriptor based isolated marathi handwritten numeral recognition. *International Journal of Computer Applications*, 3 (4), 9–13.

Sener, S., & Unel, M. (2006). A new affine invariant curve normalization technique using independent component analysis. In *Pattern recognition, 2006. icpr 2006. 18th international conference on*

(Vol. 2, p. 48-48). doi: 10.1109/ICPR.2006.111

Simon, D. (2013). *Evolutionary optimization algorithms* (first ed.). Wiley.

Strang, G. (2003). *Introduction to linear algebra* (third ed.). Wellesley Cambridge.