DTW-Radon-based Shape Descriptor for Pattern Recognition

Santosh K.C., Bart Lamiroy, Laurent Wendling

HAL Id: hal-00823961
https://hal.inria.fr/hal-00823961

Submitted on 20 May 2013

DTW–Radon-based Shape Descriptor for Pattern Recognition

K.C. Santosh⋆, Bart Lamiroy†

⋆INRIA Nancy Grand Est, †Université de Lorraine – LORIA
BP 239 – 54506 Vandoeuvre-lés-Nancy Cedex, France
Santosh.KC@inria.fr, Bart.Lamiroy@loria.fr

Laurent Wendling‡

‡SIP – LIPADE, Université Paris Descartes (Paris V)
45, rue des Saints-Pères,
75270 Paris Cedex 06, France
Laurent.Wendling@parisdescartes.fr

Abstract

In this paper, we present a pattern recognition method that uses dynamic programming (DP) for the alignment of Radon features. The key characteristic of the method is to use dynamic time warping (DTW) to match corresponding pairs of the Radon features for all possible projections. Thanks to DTW, we avoid compressing the feature matrix into a single vector which would otherwise miss information. To reduce the possible number of matchings, we rely on a initial normalisation based on the pattern orientation. A comprehensive study is made using major state-of-the-art shape descriptors over several public datasets of shapes such as graphical symbols (both printed and hand-drawn), handwritten characters and footwear prints. In all tests, the method proves its generic behaviour by providing better recognition performance. Overall, we validate that our method is robust to deformed shape due to distortion, degradation and occlusion.

Keywords

The Radon Transform; Dynamic Programming; Shape Descriptor; Pattern Recognition.

◆

1 INTRODUCTION

1.1 Context

Automatic recognition, description, classification and grouping of patterns are important issues in many disciplines such as biology, computer vision, artificial intelligence or remote sensing where shape analysis plays an important role Loncaric (1998); Zhang and Lu (2004). Patterns for instance, can be graphical elements like symbols in documents or other visual structure of any kind including cursive characters, footwear and finger prints Watanabe (1985).

Under the purview of pattern recognition through shape analysis, in this paper, two major issues will be considered:

1) shape representation and
2) matching.

Global shape representation is widely used because of its implementation simplicity since it does not necessarily require extra pre-processing and segmentation, in contrast to local pattern representation Zhang and Lu (2004). Shape representation must be sufficiently rich such that important internal content is not missed. Therefore, in order to be useful in a broad range of contexts, the choice of vector space to represent it is important. The feature vector space representing shape distribution of the studied pattern determines what matching techniques will be employed to achieve optimal performance. For instance, normalising pattern information into a fixed size single vector[1] (let's say the normalised $1D$ signature as in global signal based descriptors Zhang and Lu (2004) and $\mathcal{R}$-signature Tabbone *et al.* (2006) in particular) provides fast matching, at the cost of not offering complete shape information.

In this paper, our goal is to develop a method, covering different applications with optimal performance without integrating any ad-hoc techniques. We take advantage of the Radon transform Deans (1983), which is a particular

---

1. Fixed size vector, here refers to the set-up feature vector space. For example, the feature vector size in $\mathcal{R}$-signature Tabbone *et al.* (2006) is 180 i.e., a single compressed projection value at each projecting angle in the range $[0, \pi[$, and in generic Fourier descriptor (GFD) Zhang and Lu (2002), it has a size of $m \times n$, where m and n respectively refers to radial and angular frequencies.

case of the Trace transform Kadyrov and Petrou (2001) to represent patterns. Pattern matching is done between the corresponding Radon histograms by using dynamic programming (DP). The idea is simple, parameter free and can be used in a global context. The work is primarily based on the works reported in Fränti *et al.* (2000); Leavers (2000); Coetzer *et al.* (2004); Boulgouris and Chi (2007). For encoding, we use the whole shape instead of only contour-based information. Since the Radon feature at every projecting angle varies according to image size, we propose to employ the DTW algorithm Kruskall and Liberman (1983); Bellman and Kalaba (1959) to absorb varying histogram sizes resulting from image signal variations. The use of DTW thus can exploit complete shape information of the studied pattern and as a consequence, it can outperform the shape descriptors that basically provide normalised $1D$ signatures of fixed size (for instance, $\mathcal{R}$-signature Tabbone *et al.* (2006) where the Radon information at every projecting angle is compressed into a single value).

The rest of the paper is organised as follows. We start with discussing the pertinent state-of-the-art in Section 1.2, followed by the motivation and contribution outline in Section 1.3. In Section 2, we give a detailed explanation of the proposed method, which mainly includes pattern representation, matching and recognition and finally retrieval. Section 3 describes experimental set-up. It introduces the principal datasets, benchmarking methods and the evaluation metrics used. Experimental results are reported in Section 4. In addition, we analyse and discuss performances in Section 5. Furthermore, a case study will be reported in Section 6. Section 7 concludes the paper.

1.2   Related work on shape analysis

Shape features are often categorised into two classes Zhang and Lu (2004), which are

1) contour-based descriptors and
2) region-based descriptors.

1.2.1   Contour-based descriptors

There are several approaches to describe boundary contours based on a small set of features, and the particular choice is basically driven by requirements such as robustness to noise and tolerance to small distortions. In contour-based descriptors, one of the most commonly used techniques is the *Fourier descriptor* (FD) Persoon and Fu (1986); Zhang and Lu (2005); El-ghazal *et al.* (2009). Others are based on polygonal primitives from contour information and have been used for matching via dynamic programming Maes (1991); Attalla and Siy (2005). A polygonal approximation of any shape induces a loss of information (for instance, the degree of ellipticity is not well suited for recognition of polygons Maes (1991)). In Gerdes *et al.* (1995), another approach has been presented based on a contour-oriented $2D$ object where recognition shows robustness towards polygonal approximation inconsistency. The major drawback however, is its high time complexity. The generalised Hough transform Ballard (1981); Fränti *et al.* (2000) can be taken as a useful technique but high computational time and storage requirements, are its substantial issues. From a practical point of view, the advancement such as the work presented in Kassim *et al.* (1999) does not provide a faster technique. Curvature approaches Pelillo *et al.* (1999); Mokhtarian and Abbasi (2002); Bandera and Sandoval (2002); Bernier and Landry (2003) basically describe shapes in scale space. In Pelillo *et al.* (1999); Mokhtarian and Abbasi (2002), a shape is described in a scale space by the maximum of the curvature feature vector considering object's boundary contours. The similarity between shapes can be measured by the distance between their corresponding scale space representations. For instance, *shape context* (SC) Belongie *et al.* (2002) is robust to small perturbations while it does not guarantee scale-invariance. Skeleton extraction has also been used for pattern matching Kimia *et al.* (1995); Zhu and Yuille (1996) where shapes can be defined as a variant of contour-based approaches. In these approaches, as reported in Tabbone *et al.* (2006), each instance of an object is represented by a graph built on the medial axis of the shape silhouette. Improved versions of the graph have been proposed Sharvit *et al.* (1998); Siddiqi *et al.* (1999) with the shock graphs (locus of singularities) and several works have focused on methods to efficiently match two graphs Pelillo *et al.* (1999); Zhu and Yuille (1996). These methods are highly effective since they are based on global optimizations. However, they have a high complexity and comparative studies show that they are less tolerant to scale variation than polar curvature methods Bernier and Landry (2003). In short, all these contour-based descriptors are

1) appropriate for silhouette shapes, and
2) limited to certain applications since they cannot capture interior contents, disconnected shapes or shapes with multiple holes where boundary information is not available.

1.2.2   Region-based descriptors

Region-based descriptors take into account all pattern pixels. Common methods are based on the moment theory Teh and Chin (1988); Belkasim *et al.* (1991); Prokop and Reeves (1992); Bailey and Srinath (1996) including *geometric*, *Legendre*, *Zernike*, and *Pseudo-Zernike* moments. Comparative studies Bailey and Srinath (1996); Teh and Chin (1988) have demonstrated the interest of improving invariance properties Kim and Kim (2000); Kamila *et al.* (2005) and

reducing execution time of the *Zernike* moments Chong *et al.* (2003). Zhang and Lu Zhang and Lu (2002); Zhang (2002) have proposed a region-based *generic Fourier descriptor* (GFD) to overcome the limitations of contour-based FD in one hand. On the other hand, since this approach uses the 2D *Fourier transform* (FT) on a polar-raster sampled shape, it can avoid the problem of rotation in the Fourier spectra. Experimental results show that GFD outperforms common contour-based (classical Fourier and curvature approaches) and region-based (*Zernike* moments) shape descriptors Zhang and Lu (2002). For a thorough study, one can refer to Zhang (2002). On the whole, we can point out the following.

1) In region-based descriptors, a normalisation procedure (representing global shape by a single vector, for instance) is used to satisfy common geometric invariance properties by using centroid position, re-sampling and re-quantisation, for instance.

2) Such a single feature vector on the other hand, may not always capture all shape information, and eventually becomes less discriminant. However, discrimination power and robustness strongly depend on the selection of an optimal set of features as well as proper fusion of classifiers Kittler *et al.* (1998); Kudo and Sklansky (2000); Ruta and Gabrys (2000); Duda *et al.* (2001). For instance, in the field of symbol recognition, to increase performance, descriptors have been combined with several classifiers according to Tombre *et al.* (1998).

1.3   Motivation and contribution outline

Under the purview of shape analysis in pattern recognition, the Radon transform-based methods have been widely used because

1) it is a rich transform with one-to-many mapping, each pattern's point lies on a set of lines in the spatial domain and contributes a curve to the Radon image;

2) it is a lossless transform, patterns can even be reconstructed accurately; and

3) it has useful properties concerning rotation, scaling and rotation transformations, and thus possible to use in the recognition framework.

A very prominent work based on the Radon transform that yields $\mathcal{R}$-signature Tabbone *et al.* (2006) is obtained by using an integral function and then the discrete Fourier transform on the radial and angular slices of the Radon image respectively. Similarly, the ϕ-signature Nacereddine *et al.* (2010) is computed by using an integral function on the angular slices of the Radon image to get rotation invariance. Invariance to translation and scaling is made possible by normalisations. The strength of these two approaches is their simplicity, however, they are less discriminant since there is a loss of information in the compression process from the Radon image to 1D signatures (*cf.* Section 1.1). Moreover, the required normalisations in the signatures prevent it from being applied to noisy images. We note that, among the Radon transform-based pattern descriptors, $\mathcal{R}$-signature (due to its simplicity, without a surprise) has been successfully applied to several applications (e.g., symbol recognition Terrades *et al.* (2009) and activity recognition Souvenir and Parrigan (2009); Khan and Sohn (2013)). Very recently, the generalisation of the $\mathcal{R}$-transform has been reported Hoang and Tabbone (2012a): it basically aims to provide more insight to theoretical geometric interpretations as well as their properties, while addressing exponent values variation when computing integral function used in the $\mathcal{R}$-transform instead of just relying on just a single value i.e., 2 as in Tabbone *et al.* (2006). Such a generalisation however, does not really tune a single exponent value that can generally be applied. The results vary as soon as exponent values change – it happens more specifically when noisy samples are concerned. For instance, as reported in Hoang and Tabbone (2012a), high exponent value may cause high variation in the $\mathcal{R}$-transform and thus it is different from the ideal analytical values. The heavily deformed $\mathcal{R}$-transform with different exponent values (except 1) due to noise will make the representation inappropriate for recognition or matching. Besides, as in Tabbone *et al.* (2006); Nacereddine *et al.* (2010), since feature vector computation concerns the compression of the Radon image, it possesses similar limitations. Similarly, other work Hoang and Tabbone (2012b) does not exploit complete information since it yields a completely transformation invariant feature vector i.e., 1D signature via normalisation. Rather than using just a single $1D$ signature, in Tabbone *et al.* (2008), the histograms of radon transform (HRT) shows its usefulness for shape retrieval. In contrast to Tabbone *et al.* (2006) in particular, the HRT is more efficient since it yields a $2D$ histogram representing the shape length at each orientation. More importantly, this set of histograms is not invariant to rotation and consequently, in the matching step, these histograms need to be rotated to all possible rotating angles corresponding to potential pattern's orientations in order to compute similarity between two shapes. In this context, it is important to notice that to make histograms i.e., projections fixed sizes, images are scaled into a fixed size window. To maintain the quality of the image, loss-less resizing/scaling techniques need to be integrated, otherwise important shape information of degraded patterns may be missed Charrier *et al.* (2007). Therefore, it is not wise to pay heavy computational time since it does not really discriminate shape deformation due to degradations, distortions, etc. In Hasegawa and Tabbone (2011), authors extend the work by using logarithmic of HRT (LHRT) in order just to normalise the histogram, and employing phase-only-correlation (POC) technique to make it rotation invariant. The principle drawback however, remains

(a) Projection example

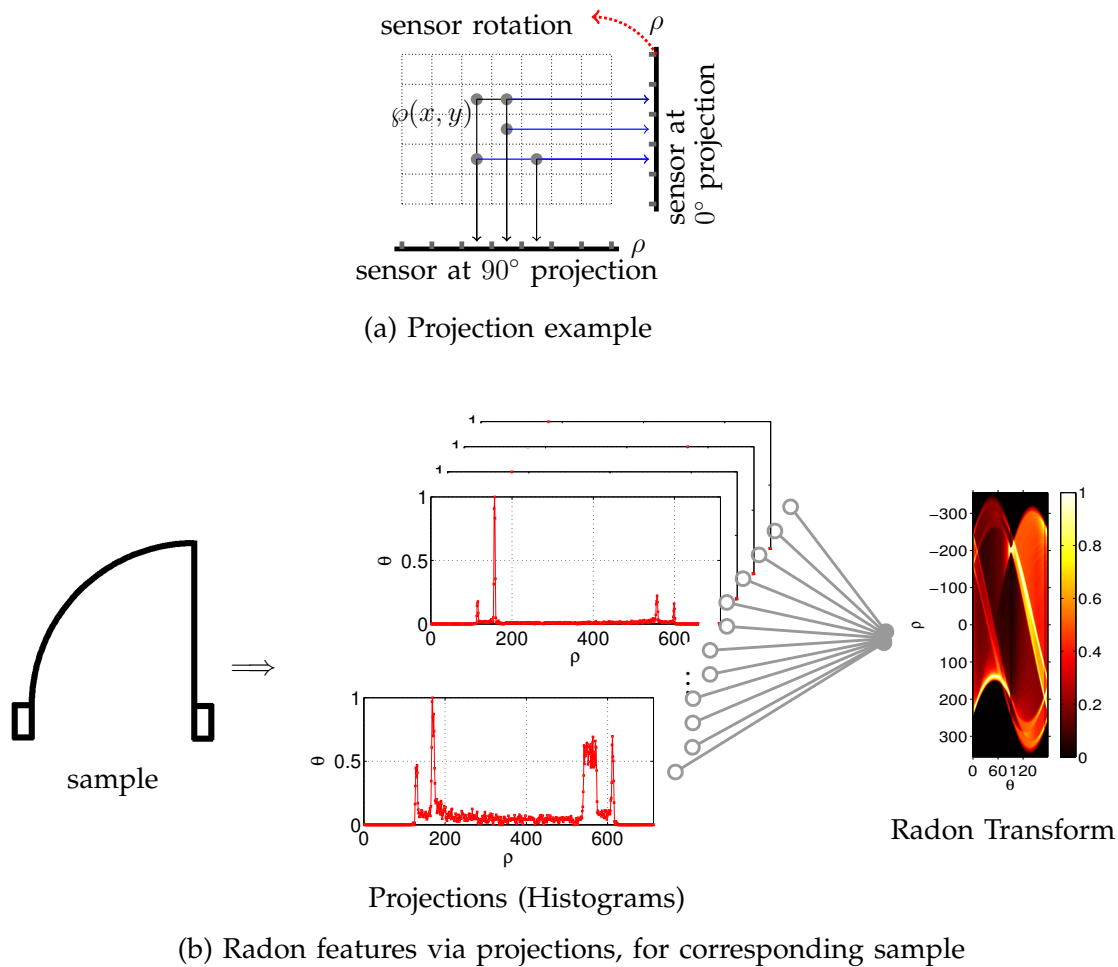(b) Radon features via projections, for corresponding sample

Fig. 1. The Radon transform – a complete illustration: (a) pattern projection using the Radon transform theory and (b) the Radon features for all possible projections over $[0, \pi[$ and the complete Radon transform i.e., a collection of all the Radon histograms or features.

the same, in addition to the limitation of POC where many peaks can be possible in the output when periodic images like chess-board are taken. On the whole, these methods do not really work on deformed shapes due to degradations and distortions, occlusions, including non-uniform scaling like stretching.

We note that the Radon transform is essentially a set of parametrised histograms or features. In order to fully exploit information of the Radon image, resizing an image is neither a good choice nor compressing the Radon histogram into a fixed size $1D$ signature. Based on this, in this paper, we consider the following.

1) We employ DTW to absorb varying histogram sizes resulting from image signal variations.

2) Besides, unlike the previous works Coetzer (2005); Santosh *et al.* (2011); Santosh (2011), another interesting part of the paper is the optimal selection of number of bins (or histograms) rather than using just the straightforward discrete Radon transform. This aims to optimise the computational time.

3) Furthermore, to reduce number of DTW matchings, we have integrated pattern rotation beforehand to make the Radon transform rotation invariant. It eventually provides one-to-one features alignment.

4) In order to establish the interest of the proposed method in the broader perspectives, several different datasets having different levels of noise, distortion and occlusion, are employed. They are publicly available shape datasets, electrical and architectural symbol (both printed and hand-drawn) datasets, character and footwear print datasets.

The experimental results confirm the proposed method can be applied to many pattern analysis and recognition or retrieval applications.

2 METHOD

In this work, we use the Radon transform to represent patterns Deans (1983). Then we apply the DTW algorithm to align every corresponding histogram for each projecting angle since representation size varies according to image
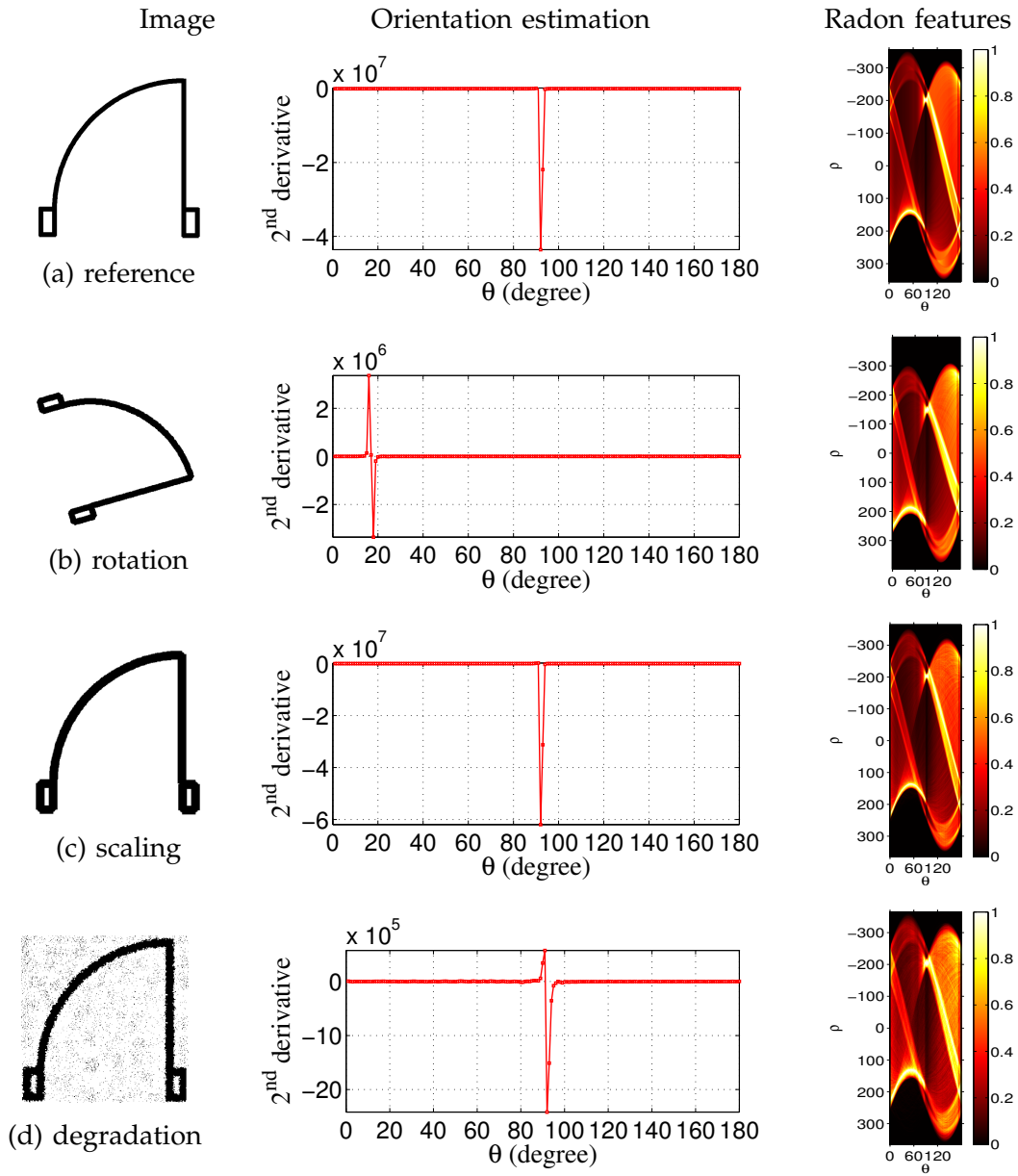
Fig. 2. Images, their corresponding orientation estimation and the Radon features. Samples are taken from graphical symbol dataset GREC (2003). Orientation angle α has been estimated as follows: $90°$ for reference, scale and degradation images while $17°$ for rotation image.

size. DTW makes it possible to avoid compressing pattern representation, which may lead to a loss of information.

2.1   Pattern representation

As shown in Fig. 1(a), a collection of projections of the pattern at different angles refers to the Radon transform Deans (1983). In other words, the Radon transform for an image pattern $\mathcal{P}(x, y)$ and for a given set of angles can be thought of the projection of all non-zero points. The resulting projection is the sum of the non-zero points for any pattern $\mathcal{P}(x, y)$ in each direction, which eventually form a matrix. Therefore the integral of $\mathcal{P}$ over a line $\mathcal{L}(\rho, \theta)$ defined by $\rho = x \cos \theta + y \sin \theta$ can formally be expressed as,

$$R(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{P}(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy \tag{1}$$

Image	Radon features	Image	Radon features

(a) rotation 1

(b) rotation 2

(c) rotation 3
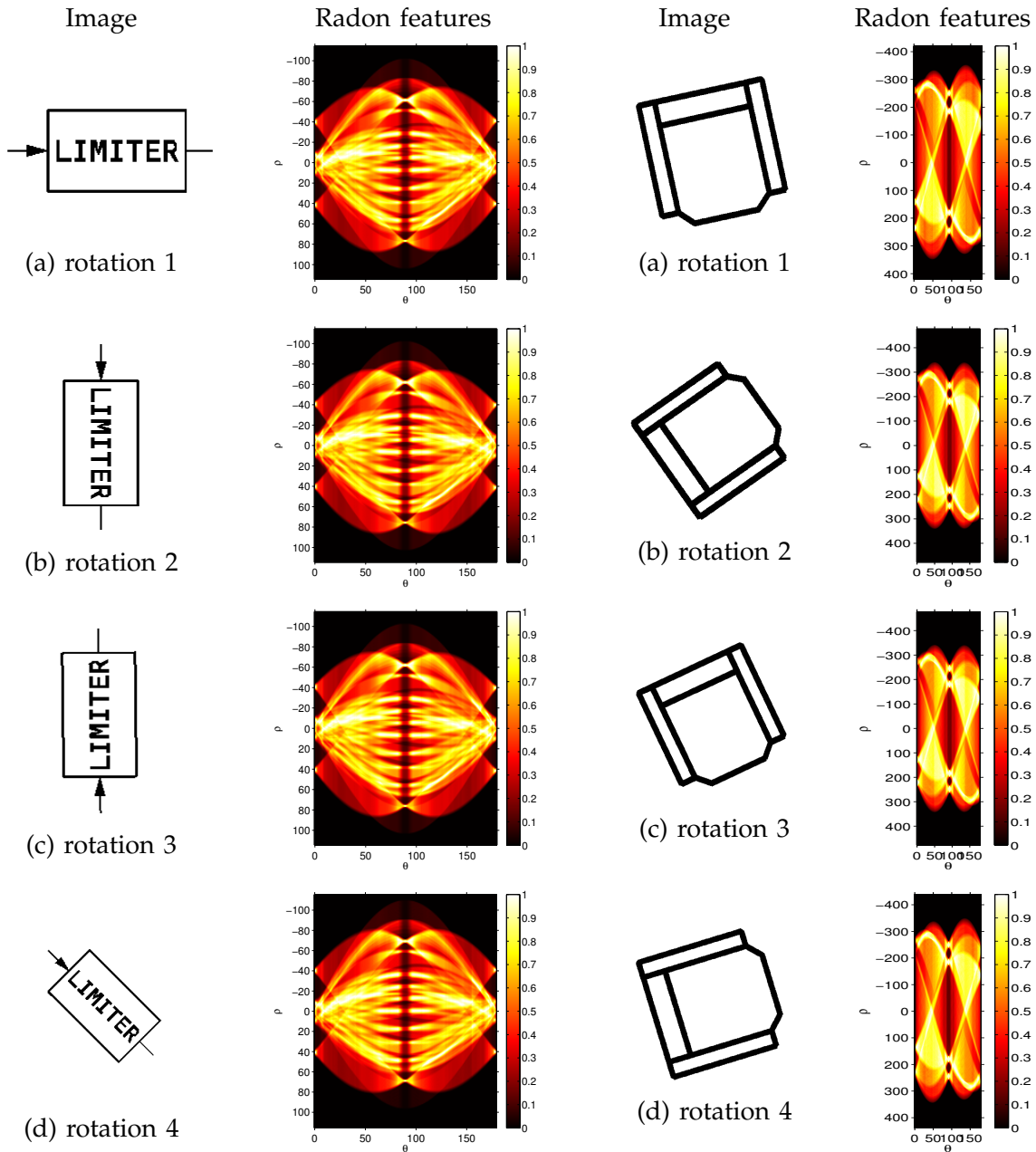
(d) rotation 4

Fig. 3. Rotational images GREC (2003); Tooley and Wyatt (2008) and the corresponding Radon features.

where $\delta(.)$ is the Dirac function,

$$\delta(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

with $\theta \in [0, \pi[$ and $\rho \in]-\infty, \infty[$. For the Radon transform, $\mathcal{L}_i$ be in normal form (ρ_i, θ_i). For all θ_i, the Radon transform now can be described as the length of intersections of all lines $\mathcal{L}_i$. Note that the range of ρ i.e., $-\rho_{min} < \rho \leq \rho_{max}$ is entirely based on the size of pattern. A complete illustration is provided in Fig. 1(b).

2.1.1 Affine transformation properties

The following transformations are required in order to make the Radon transform invariant to well-known affine properties Toft (1996).

1) *Translation* – To handle translation, we use the image centroid (x_c, y_c) such that translation vector is $\vec{u} = (x_c, y_c)$. In that case, the Radon transform is $R(\rho - x_c \cos \theta - y_c \sin \theta, \theta)$. Therefore, translation results in the shift of its transform in ρ by a distance equal to the projection of translation vector of the line $\mathcal{L}$.

2) *Scaling* – For scaling, histograms are normalised into $[0, 1]$ at every projecting angle.
3) *Rotation* – For rotation, the orientation angle can be estimated as in Jafari-Khouzani and Soltanian-Zadeh (2005),

$$\alpha = arg \left[\min_{\theta} \frac{d^2 \sigma_{\theta}^2}{d\theta^2} \right] \qquad (2)$$

where $\sigma_{\theta}^2 = \frac{1}{P} \sum_{\rho} (R(\rho, \theta) - \mu_{\theta})^2$ is the variance of projection at θ, $\mu_{\theta} = \frac{1}{P} \sum_{\rho} R(\rho, \theta)$ and P, the number of samples. If angle of rotation is α, then $R^{\alpha}(\rho, \theta) = R(\rho, \theta + \alpha)$. This simply implies a circular shift of the histograms.

As reported in Jafari-Khouzani and Soltanian-Zadeh (2005), the Radon transform can be used to detect linear trends in images. An object's principal direction can be roughly defined as the direction along which there exists more pixels. This means that the Radon transform along this direction usually has larger variations. Therefore, the variance of the projection at this direction is locally maximum. It is important to notice that taking the derivative removes the low frequency components of the variance, and thus the technique is robust. In case of multiple local maxima, computing the second derivative of the variance would provide an accurate principal object's shape orientation.

Fig. 2 shows the Radon features for reference, rotation, scaling, as well as degradation instances from a known class of graphical symbol GREC (2003). In this illustration, the orientation angle estimation, the Radon histograms (features) from their corresponding sample images are provided. In all cases, the Radon histograms are similar to each other except the magnitude difference. Similarly, more examples are provided in Fig. 3.

2.1.2 Number of bins

As said before, the Radon transform is essentially a set of parametrised histograms or features since projecting angle extends over $[0, \pi[$. Each projecting angle represents a bin. This means that every bin yields a single Radon histogram as a feature (*cf.* Fig. 1). Therefore, in its generic form, a complete set of the Radon features $R(\rho, b)$ can be expressed as,

$$\mathcal{F} = \{\mathbf{F}_b\}_{b=1,\dots,B}, \qquad (3)$$

where the total number bins B can be formulated as,

$$B = \frac{180}{\Theta} = \begin{cases} 180 & \text{when } \Theta = 1° \\ 90 & \text{when } \Theta = 2° \\ 60 & \text{when } \Theta = 3° \\ 36 & \text{when } \Theta = 5° \\ \dots & \text{so on,} \end{cases}$$

with Θ as the projection angle-range. A single Radon feature at bin b is $\mathbf{F}_b$ is the collection of histograms at every discrete projection angle. If $\Theta = 1°$, it is called the discrete Radon transform where there are 180 bins. While, for instance if $\Theta = 5°$, all histograms within it are averaged to form a single Radon feature and altogether, there are 36 bins.

Performances of using different numbers of bins will be reported in Section 4. Since we have used different numbers of bins starting from $B = 2$ to 180, different performances i.e., outputs are obtained. This means that the output can be expressed as $\mathcal{O}(B)$ where an optimal number of bins is selected by taking $\arg \max_B \mathcal{O}(B)$. Without a surprise, the higher the value of B, the better the recognition performance. Its value represents the trade-off between the recognition performance and time/space requirements. Therefore, instead of using all possible projections i.e. $B = 180$, we can select smaller value of B by using $\arg \min_B (\arg \max_B \mathcal{O}(B))$. In detail, execution time will be discussed in Section 2.2.2.

2.2 Pattern matching

As explained in Section 2.1, we have a collection of sets of features $\mathcal{F}$ in a specified number of bins B, to represent any pattern $\mathcal{P}$. Given two patterns: query $\mathcal{P}^q$ and database $\mathcal{P}^d$, matching can be calculated between corresponding features (*cf.* eq. (3)) from the complete set of $\mathcal{F}^q$ and $\mathcal{F}^d$. The Radon transform generates different ρ sizes depending on the image contents' size. In order to be able to adapt to these differences in size, we use the DTW algorithm Kruskall and Liberman (1983). In what follows, we explain matching computation first and then derive the matching score between two patterns.
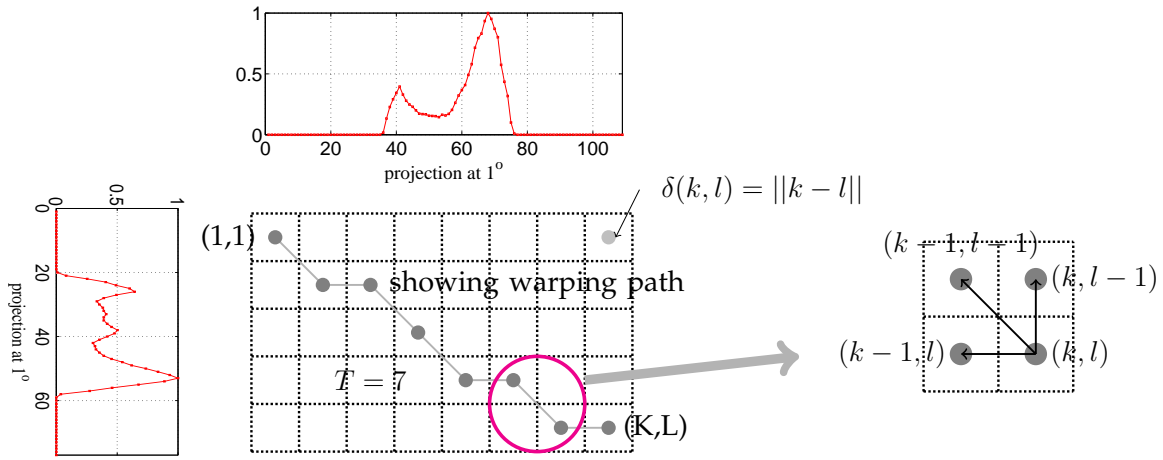
Fig. 4. DTW algorithm – illustration.

2.2.1 Dynamic time warping algorithm

Let us consider two feature sequences $\mathbf{X} = \{\mathbf{x}_k\}_{k=1,\ldots,K}$ and $\mathbf{Y} = \{\mathbf{y}_l\}_{l=1,\ldots,L}$ of size K and L, respectively. The aim of the algorithm is to provide the optimal alignment between both sequences potentially having different lengths. At first, a matrix M of size $K \times L$ is constructed. Then for each element in matrix M, local distance metric $\delta(k,l)$ between the events e_k and e_l is computed. Let $D(k,l)$ be the global distance up to (k,l),

$$D(k,l) = \min \begin{bmatrix} D(k-1,l-1), \\ D(k-1,l), \\ D(k,l-1) \end{bmatrix} + \delta(k,l) \tag{4}$$

with an initial condition $D(1,1) = \delta(1,1)$ such that it allows warping path going diagonally from starting node $(1,1)$ to end (K,L). The main aim is to find the path for which the least cost is associated. The warping path therefore provides the difference cost between the compared signatures. Formally, the warping path is, $\mathcal{W} = \{w_t\}_{t=1\ldots T}$, where $\max(k,l) \leq T < k+l-1$ and t^{th} element of $\mathcal{W}$ is $w(k,l)_t \in [1:K] \times [1:L]$ for $t \in [1:T]$. The optimised warping path $\mathcal{W}$ satisfies the following three conditions: (c1) boundary condition, (c2) monotonicity condition and (c3) continuity condition:

c1. $w_1 = (1,1)$ and $w_T = (K,L)$;
c2. $k_1 \leq k_2 \leq \cdots \leq k_K$ and $l_1 \leq l_2 \leq \cdots \leq l_L$; and
c3. $w_{t+1} - w_t \in \{(1,1)(0,1),(1,0)\}$ for $t \in [1:T-1]$.

c1 conveys that the path starts from $(1,1)$ to (K,L), aligning all elements to each other. **c2** forces the path advances one step at a time. **c3** restricts allowable steps in the warping path to adjacent cells, never be back. Note that **c3** implies **c2**.

We then define the global distance between $\mathbf{X}$ and $\mathbf{Y}$ as,

$$\text{Dist.}(\mathbf{X},\mathbf{Y}) = \frac{D(K,L)}{T}. \tag{5}$$

The last element of the $K \times L$ matrix gives the DTW-distance between $\mathbf{X}$ and $\mathbf{Y}$, which is normalised by T i.e., the number of discrete warping steps along the diagonal DTW-matrix. The warping path can be handled via backtracking following the minimum cost index pairs (k,l) starting from (K,L) by using DP:

$$w_{t-1} = \begin{cases} (1,l-1) & \text{if } k=1 \\ (k-1,1) & \text{if } l=1 \\ \text{argmin} \begin{Bmatrix} D(k-1,l-1), \\ D(k-1,l), \\ D(k,l-1) \end{Bmatrix} & \text{otherwise,} \end{cases}$$

where we take the lexicographically smallest pair in case the 'argmin' is not unique. The overall process is illustrated with the help of an example in Fig. 4 where two non-linear sequences having different lengths are employed.
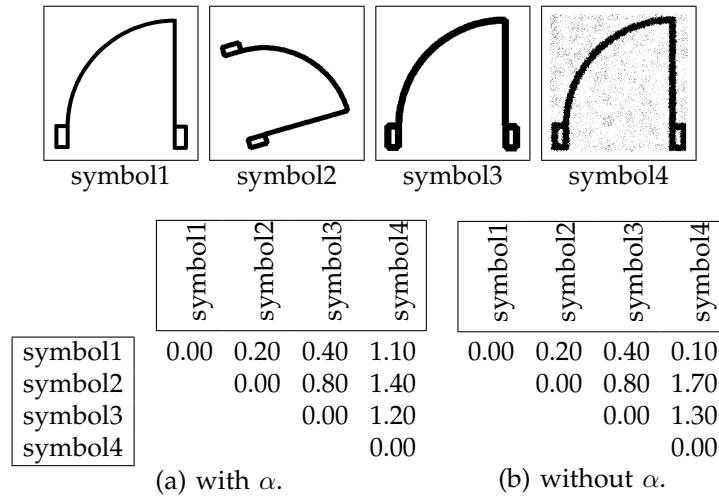
(a) with α.

	symbol1	symbol2	symbol3	symbol4
symbol1	0.00	0.20	0.40	1.10
symbol2		0.00	0.80	1.40
symbol3			0.00	1.20
symbol4				0.00

(b) without α.

	symbol1	symbol2	symbol3	symbol4
symbol1	0.00	0.20	0.40	0.10
symbol2		0.00	0.80	1.70
symbol3			0.00	1.30
symbol4				0.00

Fig. 5. Matching scores (in 10^{-3}) between reference, rotation, scaling and degradation sample images. A reference sample image is taken from GREC dataset GREC (2003). It follows Fig. 2.
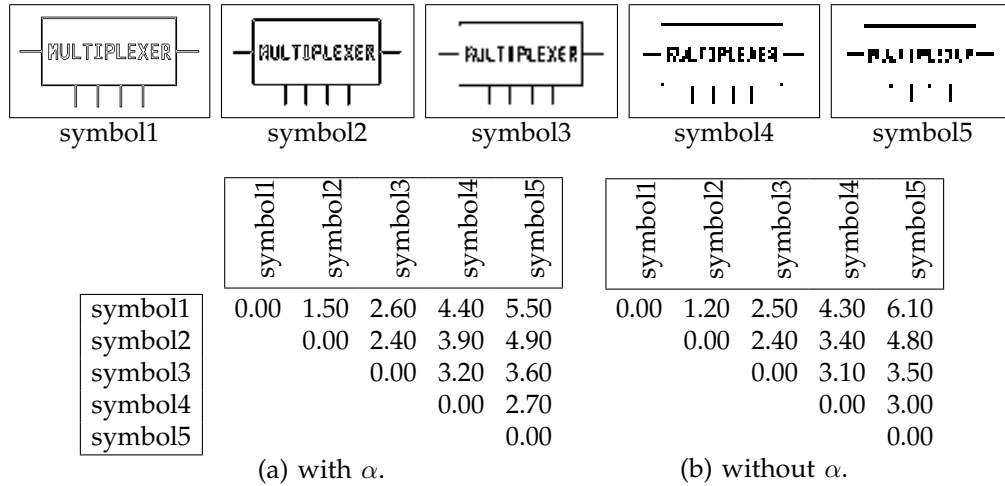

(a) with α.

	symbol1	symbol2	symbol3	symbol4	symbol5
symbol1	0.00	1.50	2.60	4.40	5.50
symbol2		0.00	2.40	3.90	4.90
symbol3			0.00	3.20	3.60
symbol4				0.00	2.70
symbol5					0.00

(b) without α.

	symbol1	symbol2	symbol3	symbol4	symbol5
symbol1	0.00	1.20	2.50	4.30	6.10
symbol2		0.00	2.40	3.40	4.80
symbol3			0.00	3.10	3.50
symbol4				0.00	3.00
symbol5					0.00

Fig. 6. Matching scores (in 10^{-3}) between distortion as well as deformed images. A reference sample image is taken from FRESH dataset Tooley and Wyatt (2008).

2.2.2 Matching score

Aggregating distances between the Radon features in all corresponding bins b between the two patterns $\mathcal{P}^q$ and $\mathcal{P}^d$ (cf. eq. (5)) yields a global pattern matching score,

$$m.\text{score}(\mathcal{P}^q, \mathcal{P}^d) \quad = \quad \sum_{b \in B} \text{Dist.}(\mathbf{F}_b^q, \mathbf{F}_b^d). \tag{6}$$

Overall, since we have employed estimated orientation α, the Radon feature alignment can go one-to-one basis. In case where α is not integrated, it is also possible to treat the method as rotation invariant via the use of the Radon feature alignments for all i.e., one-to-all in order to select the best of all.

The difference between the two different ways of aligning the Radon features lies in the number of matchings to compute $m.\text{score}(,)$ i.e.,

$$\text{possible number of matchings in } m.\text{score}(\mathcal{P}^q, \mathcal{P}^d) = \begin{cases} O(B) & \text{one-to-one alignment, and} \\ O(B^2) & \text{one-to-all alignment.} \end{cases}$$

Therefore, the overall execution time depends on how many bins are taken. Its value determines the number of matchings associated with it. In Section 5, execution time is reported based on an optimal number of bins especially determined to provide better performance.

Let us take a few examples to see the difference between the two DTW matching scores. Fig. 5 and 6 provide matching score matrices between rotated, scaled, degraded, distorted as well as deformed images. In both cases, we use the Radon features with 180 bins. These illustrations provide similar behaviour between the two distance

computation methods: 'with' and 'without α'. Therefore, in our experiments, considering time complexity issue, the Radon features alignment will be made with α.

2.3 Recognition and retrieval

We can now use the previously described approach as a global pattern matching score (*cf.* eq. (6)). This score expresses the similarity between database patterns and query,

$$\text{Similarity}(\mathcal{P}^q, \mathcal{P}^d) = 1 - \overline{m.\text{score}}(\mathcal{P}^q, \mathcal{P}^d) = \left\{ \begin{array}{ll} 1 & \text{for the closest candidate and} \\ 0 & \text{for the farthest candidate,} \end{array} \right. \tag{7}$$

where matching scores are normalised into $[0, 1]$ by,

$$\overline{m.\text{score}}(.) = \frac{m.\text{score}(.) - m.\text{score}^{min.}(.)}{m.\text{score}^{max.}(.) - m.\text{score}^{min.}(.)}.$$

Ranking can therefore be expressed in the decreasing order of similarity. In our experiments, we will distinguish "recognition" (i.e., search for the closest candidate) from "retrieval" (i.e., the closest candidates are retrieved for a given short-list).

3 EXPERIMENTAL SET-UP

3.1 Datasets

We work on several different datasets in different contexts: recognition and retrieval. Before coming to applications, we first test shape retrieval. Our applications will then include

1) symbol recognition and retrieval,
2) character recognition and
3) footwear impression evidence verification.

We primarily focus on deformed shape via distortion, degradation, noise as well as non-uniform scaling. Besides, stroke thickness in samples (in graphical symbol datasets) is varied from a few pixels to hundreds of pixels. In this situation, graph-based pattern shape representation would be inconvenient since we cannot extract sufficient primitives like points, lines and arcs nor do we decompose into meaningful parts like circles and triangles. Furthermore, skeletons are always disturbed and unifying step is not so trivial. Therefore, as reported in Valveny *et al.* (2007), we take statistical methods into account for comparison.

3.2 Benchmarking methods

We confront $\mathcal{D}-$Radon[2] for several different number of bins B, with well-known descriptors:

1) $\mathcal{R}$-signature Tabbone *et al.* (2006),
2) GFD Zhang and Lu (2002),
3) SC Belongie *et al.* (2002) and
4) ZM Kim and Kim (2000).

For those descriptors, it is important to fit the best parameters. For the Radon transform, projecting range is $[0, \pi[$. In case of GFD, one needs to tune the radial $(4 : 12)$ and angular $(6 : 20)$ frequency parameters to get the best combinations. Note that such a best combination of radial and angular frequencies can be different from one dataset to another. For SC, we use 100 sample points as reported in Belongie *et al.* (2002). In case of ZM, we have used 36 *zernike* functions of order less than or equal to 7.

3.3 Evaluation criteria

Within the recognition and retrieval framework described in Section 2.3, different evaluation metrics will be used according to the appropriateness of the datasets. Therefore, before providing results, we shortly mention the evaluation metrics once each dataset is introduced.

4 EXPERIMENTAL RESULTS

Throughout the experiments, we will determine how many bins for $\mathcal{D}-$Radon are the most appropriate for recognition. This means that we first compare the state-of-the-art of shape descriptors before confronting 'the best' performer with $\mathcal{D}-$Radon for different numbers of bins.

For ease of reading, the compared results are highlighted. For benchmarking methods, the best performer is highlighted first. Then in case of $\mathcal{D}-$Radon, the highlighted numeric figures from different number of bins will be identified as optimal selection of B. Moreover, bold-faced numeric figures are the highest values in every test.

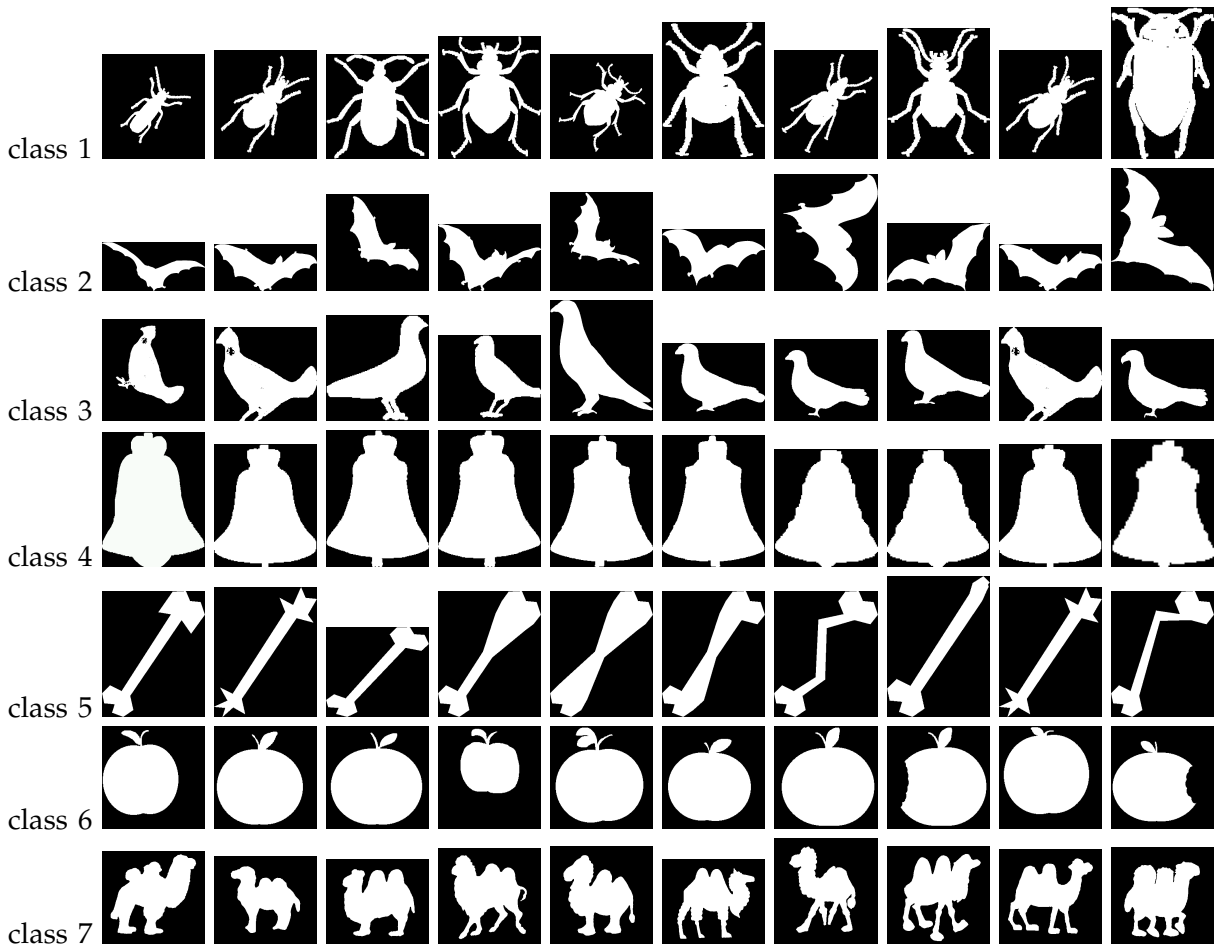2. For simplicity, DTW–Radon is named as $\mathcal{D}-$Radon.

Fig. 7.   A few samples (of some classes) from MPEG-7 CE-1 shape dataset.

TABLE 1
Average retrieval rate (in %) for MPEG-7 CE-1 shape dataset.

Requested List	GFD	ZM	SC	$\mathcal{R}$-sign.	$\mathcal{D}-$Radon for different values of B						
					180	90	60	36	18	09	02
top-4	97	84	**98**	89	**98**	86	83	82	78	72	65
top-8	84	67	**88**	73	85	75	70	62	55	54	49
top-12	72	55	**75**	63	73	62	55	53	47	44	41
top-16	61	46	**68**	53	65	55	47	43	39	38	34
top-20	55	39	**61**	46	51	43	36	34	31	29	26

4.1   Shape retrieval

Dataset. In this test, we have taken the MPEG-7 CE-1 shape silhouette dataset as described in Jeannin and Bober (1999). The dataset consists of 70 classes $\times$ 20 samples. Fig. 7 shows a few samples of it. **Evaluation criterion.** To

evaluate the performance of the methods, each sample is used as a query, and the number of similar images (which belong to the same class) is counted for a given short-list. In this test, retrieval goes with the step of +4 until total number of samples in each class.

Observations. Table 1 shows the experimental results. Compared to ZM and $\mathcal{R}-$signature, $\mathcal{D}-$Radon outperforms. On the other side, the performance can be compared with GFD and SC where there exists marginal difference.

4.2   Symbol recognition and retrieval

We have used both computer-printed and hand-drawn symbols. It is composed of symbols from architectural and electrical wiring diagrams in separate datasets. A comprehensive study is made over three different datasets. They
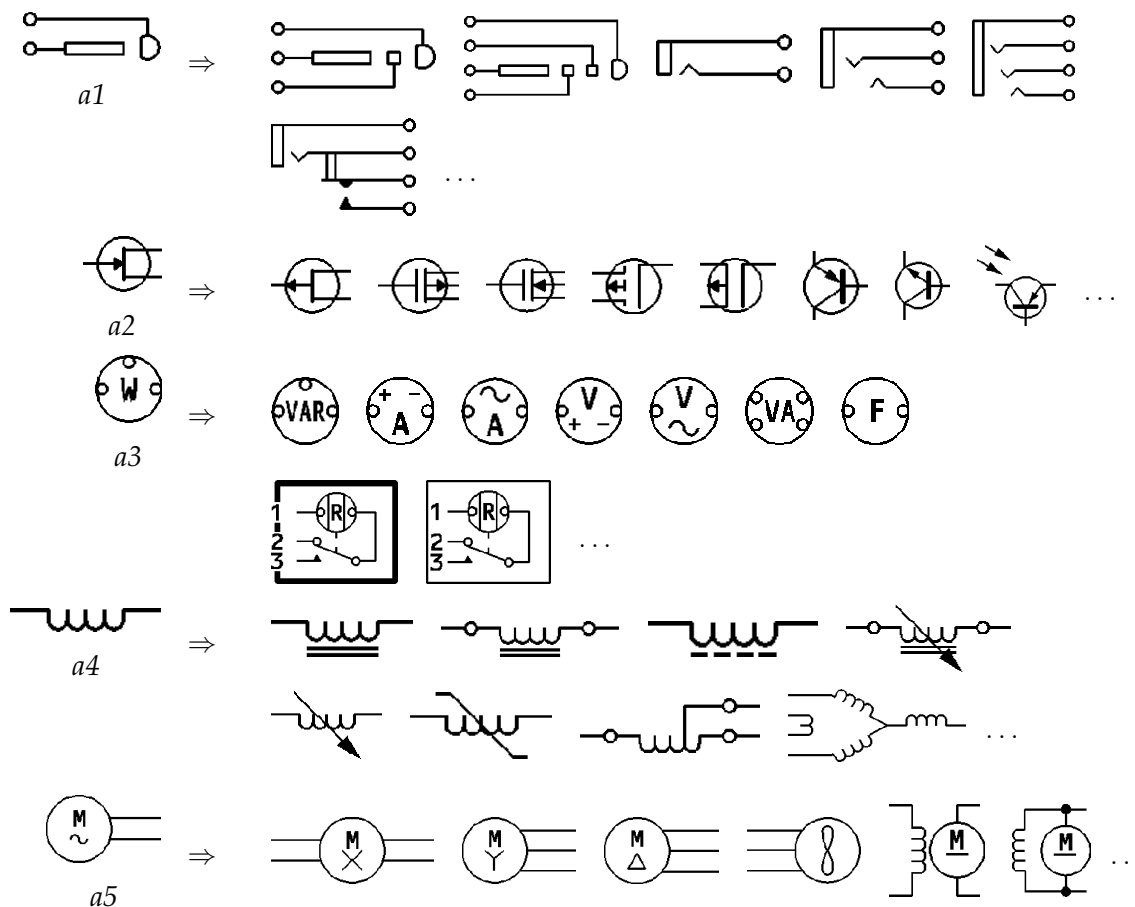
Fig. 8.   A sample of few electrical symbols and their similar sample images Tooley and Wyatt (2008). For every test symbol: *a1* to *a5*, a few relevant symbols are enlisted based on human evaluation.

are named as

1) FRESH,
2) GREC and
3) CVC dataset.

4.2.1   FRESH dataset

Dataset. We work on a real world industrial problem to identify a set of different known symbols in aircraft electrical wiring diagrams extracted from Tooley and Wyatt (2008). Fig. 8 gives some examples of symbols in the database. Symbols may either be very similar in shape – and only differ by slight details – or either be completely different from a visual point of view. Symbols may also be composed of other known and significant symbols and need not necessary be connected. It is composed of roughly 500 different known symbols.

Ground-truth formation. Since there is no absolute ground-truth associated to our dataset, we have proceeded by using human validation, but by taking care of eliminating subjective bias. In order to achieve this we have asked six volunteers to manually select what they consider as "similar" symbols, for all queries executed in this section. Human evaluators have chosen the candidates which have similar visual overall appearance or which have significantly similar parts with respect to the chosen query. The evaluators were not required to provide any ranking order nor degree of visual resemblance. In more formal terms, as in our testing protocol established in Santosh *et al.* (2012), for each query the "ground-truth" is considered to be the set of symbols formed by the union of all human selected sets. Fig. 8 provides a few examples.

Evaluation criterion. Our aim is not only limited to distinguishing symbols but also extends to ranking them. It is important to notice that the number of ground-truths varied from one query to another. Therefore traditional precision and recall cannot be used. We use retrieval efficiency Kankanhalli *et al.* (1995) as the evaluation metric.

TABLE 2
Retrieval efficiency (in %) over 30 queries for FRESH dataset.

Requested List	GFD	ZM	SC	$\mathcal{R}$-sign.	$\mathcal{D}-$Radon for different values of B						
					180	90	60	36	18	09	02
top-2	91	88	87	84	**92**	**92**	91	87	78	77	75
top-4	80	72	72	71	**83**	82	81	76	66	62	60
top-6	74	65	63	60	**77**	75	74	69	57	53	50
top-8	71	60	59	51	**76**	75	73	64	49	45	45
top-10	69	56	54	49	**73**	71	69	58	44	42	41

For every chosen query, retrieval efficiency for a given short-list of size $\mathcal{K}$ is expressed as,

$$\eta_{\mathcal{K}} = \begin{cases} n/\mathcal{N} & \text{if } \mathcal{N} \le \mathcal{K} \\ n/\mathcal{K} & \text{otherwise,} \end{cases} \qquad (8)$$

where n is the number of returned relevant symbols, $\mathcal{N}$ the total number of relevant symbols and $\mathcal{K}$ the number of ranked symbols requested. Note that $\eta_{\mathcal{K}}$ computes the traditional recall if $\mathcal{N} \le \mathcal{K}$ and computes precision otherwise. The main advantage of this is that the average retrieval efficiency curve is not biased even with different ground-truths for different queries while it happens for precision measure when $\mathcal{N} < \mathcal{K}$.

Observations. Average retrieval efficiency, for $\mathcal{K}$ ranging from 1 to 10 is shown in Table 2. From the set of benchmarking shape descriptors, GFD performs better, followed by SC and ZM. In contrast, $\mathcal{D}-$Radon outperforms for different values of bins like $B = 180, 90, 60$. The performance of $\mathcal{D}-$Radon using 36 number of bins can also be compared since it provides marginal difference with 'the best' performer i.e., GFD.

4.2.2   GREC dataset

Dataset. In this dataset[3] GREC (2003), we have used the following different categories: ideal, rotation, scaling, distortion as well as degradation.

Altogether, there are 50 different model symbols. Those symbols are grouped into 3 sets, containing 5, 20 and 50 model symbols. Each model symbol has 5 test images in every category except the ideal one. Ideal test images are directly taken from the set of model symbols and therefore the test is to evaluate the ability of simple shape discrimination, as the number of symbols increases. Since vectorial distortion works only for symbols with straight lines, and not arcs, it is applied to a subset of 15 model symbols. Furthermore, there are 9 models of degradation, aiming to evaluate the robustness to the scalability with degradation. Fig. 9 shows a few samples of GREC dataset.

Evaluation criterion. To evaluate the method, each test image is matched with the model symbols to get the closest model.

Observations. Experimental results for all types of aforementioned categories of datasets are shown in Table 3. More precisely, it provides a results for ideal, rotation, scaling, combination of rotation and scaling, distortion and degradation categories.

Following Table 3, we observe the following.

- Based on the results from ideal test images, every shape descriptor provides similar recognition performance. In case of $\mathcal{D}-$Radon, it is interesting to notice that it gives 100% recognition rate from all provided number of bins. For rotated images, GFD performs better, proving a marginal difference with SC and $\mathcal{D}-$Radon for 180 bins. For scaled images, $\mathcal{D}-$Radon outperforms all, where $B = 180, 90, 60$ and 36 provide 100% recognition rates while not offering substantial difference with the benchmarking descriptors. For those test images where rotation and scale are combined, $\mathcal{D}-$Radon for 180, GFD and SC provide similar results. Overall, not a single 'the best' performer from a major set of state-of-the-art has been observed. Besides, one cannot judge the superiority of the methods. Only execution time comparison would be an alternative.
- Results from test images with vectorial distortions show the worst case scenario for $\mathcal{R}$-signature while for others, there exists similar behaviour. In case of $\mathcal{D}-$Radon, 36 bins can also be compared with 'the best' performer from the state-of-the-art.
- However, we receive notable differences between $\mathcal{D}-$Radon and GFD, in case of binary degradations. In case of $\mathcal{D}-$Radon, 60 bins can provide competitive results.

---

3. International symbol recognition contest, 2003

(a) ideal   (b) rotation   (c) scaling   (d) distortion

m1   m2   m3   m4   m5   m6   m7   m8   m9
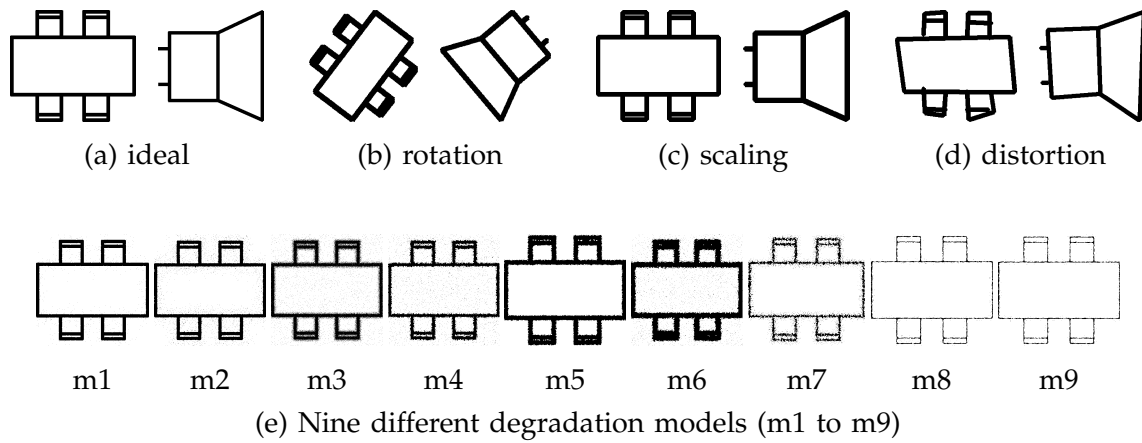
(e) Nine different degradation models (m1 to m9)

Fig. 9.   GREC2003 samples – graphical symbol.

TABLE 3
Average recognition rates (in %) for six different categories: ideal, rotation, scaling, rotation + scaling, distortion and degradation of GREC dataset.

Images Set	GFD	ZM	SC	$\mathcal{R}$-sign.	$\mathcal{D}-$Radon for different values of B						
					180	90	60	36	18	09	02
ideal	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
rotate	**98**	94	97	94	97	94	88	73	82	77	71
scale	99	98	99	96	**100**	**100**	**100**	**100**	84	74	57
rotate+scale	**98**	93	**98**	92	98	97	94	82	79	73	62
distort	**100**	94	**100**	92	**100**	**100**	**100**	**100**	85	72	47
degrade	91	79	78	76	**99**	98	95	84	67	47	35

Overall, 'the best' performer (from the comparison set) can be changed from time to time (depending on the dataset) while $\mathcal{D}-$Radon provides consistent results. $\mathcal{D}-$Radon in particular, shows an interesting behaviour for different values of bins.

4.2.3   CVC dataset

Dataset. This dataset comes from Fornés *et al.* (2010), we have tested 10×300 sample images i.e., 10 different known classes of hand-drawn architectural symbols with 300 instances in each. Symbols have been drawn by 10 different people. Fig. 10 shows a few samples of the dataset. Samples provide distortions, gaps, overlapping as well as missing parts within the shapes.

Evaluation criterion. To validate the methods, each test image is matched with all images and number of correct matches is reported over the requested list. In this test, since we have 300 samples per class, the size of the requested list is 300, aiming to retrieve all similar images of every class including itself. To handle in-depth experimental results analysis, we provide retrieval rate in every 20 increasing step. Table 4 shows the average retrieval rate for all requested short-lists (e.g., top-20, top-40 and so on).
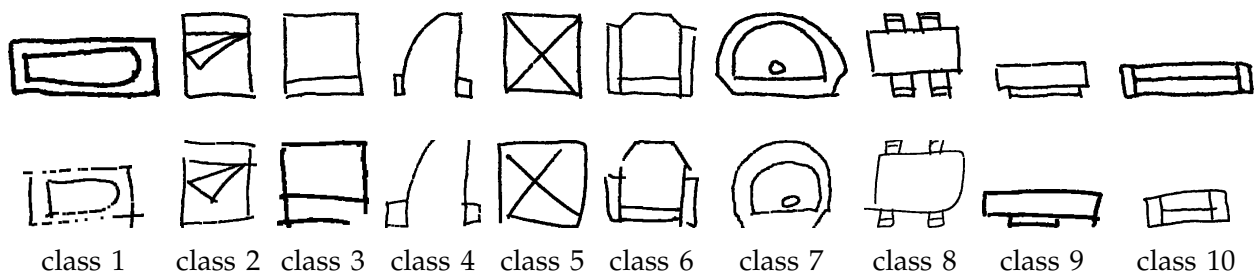

class 1   class 2 class 3 class 4 class 5 class 6   class 7   class 8   class 9   class 10

Fig. 10.   Two hand-drawn samples from 10 different known classes.

TABLE 4
Average retrieval rate (in %) for CVC dataset.

Requested List	GFD	ZM	SC	$\mathcal{R}$-sign.	$\mathcal{D}-$Radon for different values of B						
					180	90	60	36	18	09	02
top-20	96	68	98	82	**99**	**99**	**99**	98	94	92	83
top-40	93	62	95	75	**99**	**99**	98	97	92	88	79
top-60	90	59	95	69	**97**	**97**	97	96	90	86	67
top-80	88	57	92	65	**97**	**97**	97	94	88	81	64
top-100	85	55	91	62	**97**	**97**	96	94	84	80	61
top-120	83	54	88	59	**97**	96	96	93	80	74	60
top-140	81	52	87	56	**95**	**95**	95	92	76	67	58
top-160	78	50	85	54	**95**	**95**	94	91	73	64	57
top-180	76	50	83	51	**93**	**93**	94	89	67	61	54
top-200	73	48	81	49	**93**	92	91	86	62	56	53
top-220	71	44	78	48	**93**	92	91	86	59	53	50
top-240	68	41	78	46	**92**	90	88	84	55	51	49
top-260	66	39	75	45	**91**	89	86	83	52	44	46
top-280	63	37	73	43	**88**	87	85	81	48	42	45
top-300	61	36	70	42	**86**	**86**	84	78	47	39	44

Observations. Unlike the previous dataset, SC provides the best performance among benchmarking descriptors. In contrast, $\mathcal{D}-$Radon outperforms 'the best' performer.

Up to top-60, one cannot decide which method performs well since there exist no notable retrieval rate differences among them. It is only determined after top-60. The aim of the test is to evaluate the retrieval stability of the methods. $\mathcal{D}-$Radon provides the rate of more than 16% difference with SC. SC lags GFD by approximately 9%. $\mathcal{R}-$signature provides an average results compared to ZM.

Overall, recognition rates from $\mathcal{D}-$Radon for $B = 180, 90, 60$ and 36 provide better results than GFD, SC, $\mathcal{R}$-signature and ZM.

4.3   Character recognition

Datasets. Several different datasets from different scripts[4] are tested. They are Roman, Bangla, Devanagari and Oriya Bhowmik *et al.* (2006); Bhattacharya and Chaudhuri (2009). Fig. 11 shows a couple of sample images in each class. We have 10 classes for all scripts. In the case of Roman and Oriya scripts, there are 100 samples for a single class, while there are 300 and 400 samples for Devanagari and Bangla, respectively.

As shown in Fig. 11, we can highlight a few major challenges associated with the datasets such as multi-class similarity, symmetric shape similarity, missing parts, and stroke length or size variation.

Evaluation criterion. While experimenting, every test sample is matched with training candidates and the closest one is reported. The closest candidate corresponds to the labelled class, which we call 'character recognition'.

To evaluate the performance of the methods, $\mathbb{K}$-fold cross-validation (CV) has been implemented unlike traditional dichotomous classification. In $\mathbb{K}$-fold CV, the original sample for every class is randomly partitioned into $\mathbb{K}$ sub-samples. Of the $\mathbb{K}$ sub-samples, a single sub-sample is used for validation, and the remaining $\mathbb{K} - 1$ sub-samples are used for training. This process is then repeated $\mathbb{K}$ times, with each of the $\mathbb{K}$ sub-samples used exactly once. Finally, a single value results from averaging all. The aim of the use of such a series of rigorous tests is to avoid the biasing of the samples that can be possible in conventional dichotomous classification.

Observations. Table 5 shows the average recognition rates for all datasets using $\mathbb{K}$-fold CV. In the Roman dataset, all shape descriptors provide comparable recognition performances. The significant differences between them exist in case of Devanagari and Bangla datasets. SC yields consistent recognition rates for all, while others do not follow such a characteristic. GFD however, comes closer to SC. $\mathcal{D}-$Radon can now be compared with SC and sometimes with GFD. In this category, there exists no surprising differences between them i.e., approximately 1–2%. However, it is found that number of bins $B = 90$ for $\mathcal{D}-$Radon, provides better results and $B = 60$ can be compared with.

4. ISI character datasets for Indian scripts, CVPR unit, India
   ETL3 Roman dataset, AIST, Japan

Fig. 11.   Two numeral samples from four different known scripts for 10 $(0-9)$ classes.

TABLE 5

Average recognition rate in $\%$ using $\mathbb{K}$-fold CV (where $\mathbb{K} = 5$) for different numeral datasets.

Dataset	GFD	ZM	SC	$\mathcal{R}-$sign.	$\mathcal{D}-$Radon for different values of B						
					180	90	60	36	18	09	02
Roman	97	83	98	78	**100**	**100**	**100**	88	79	76	71
Oriya	98	44	98	58	**100**	98	92	81	67	60	51
Devanagari	86	40	96	55	**99**	97	89	79	71	62	52
Bangla	73	47	**95**	48	**95**	**95**	84	72	60	51	41

5 DISCUSSIONS

We analyse the behaviour of the methods based on the key characteristics as well as major challenges found in datasets. In general, we focus on those samples which are distorted, embedded with different levels of noise and even degraded. We also take into account those samples with missing parts including severe vectorial distortions in hand-drawn symbol dataset in addition to a significant size variation as well as multi-class similarity between the classes. Within this framework, based on the observations in all experimental tests, we discuss the methods in two different issues:

1) recognition rate and
2) computational complexity.

Recognition performance is related to how discriminative the shape descriptor is.

All descriptors perform well except for deformed shapes due to distortion and degradation as well as non-uniform scaling. In particular, $\mathcal{R}$-signature is found to be severely disturbed. SC shows almost similar behaviour where the level of noise is high since it takes those pixels into account while sampling. In addition, it does not provide optimal performance when image size changes. For ZM, we have observed the effect in case of degradation models. GFD provides fairly satisfying results. However, a single shape descriptor cannot be always 'the best' performer for all types of sample images. $\mathcal{D}-$Radon in contrast, shows discriminative power.

On the other hand, $\mathcal{D}-$Radon is constrained by how many number of bins are used. As said previously, the larger the number of bins, the higher the discrimination power and vice versa. The statement is validated with the help of a series of tests. In all experimental test results, $\mathcal{D}-$Radon provides competitive recognition and retrieval performances even when the number of bins is reduced upto 60. In some cases, 36 number of bins can also be compared with benchmarking shape descriptors. This means that the performance largely depends on how complex

TABLE 6

Average execution time (in sec) for a pattern representation and recognition single pair.

Shape descriptors	Time (in sec)
1. $\mathcal{R}$-signature Tabbone *et al.* (2006)	01
2. ZM Kim and Kim (2000)	04
3. GFD Zhang and Lu (2002)	03
4. Shape Context Belongie *et al.* (2002)	22
5. $\mathcal{D}-$Radon	56

the dataset is. For instance, only 2 bins (i.e., vertical and horizontal projections) provides 100% recognition rate for ideal sample set in GREC dataset (*cf.* Table 3). Considering the situation, since execution time is related according to how many number of bins are used, we can possibly reduced the delay.

Execution time processing is usually high since it uses DTW for matching[5]. It also largely depends on how big the image is. As far as computational cost is concerned, the observed average execution time for all methods is given in Table 6. In case of $\mathcal{D}-$Radon, an average value is computed when an optimal recognition performance is achieved. In our tests, 180, 90 and 60 bins are taken. Overall, it is important to notice that state-of-the-art of shape descriptors provide different results according to the behaviour of the dataset i.e., not a single shape descriptor can be 'the best' performer. As a consequence, for comparison, execution time will be an alternative. In contrast, $\mathcal{D}-$Radon outperforms of all 'the best' performers in all types of dataset while paying more execution time.

6 CASE STUDY

In order to establish the interest as well as extensibility of the method, we employ footwear print dataset. It is yet another application to see the behaviour of the methods when pattern complexity increases.

In this dataset[6], we have 64 footwear impression or footwear print classes, each with height instances. Footwear impression images are categorised into full print and partial print. Fig. 12 shows sample images for a couple of classes. As shown in Fig. 12, there are four images in the full print category. These images may include rotation for each class. The remaining ones are partial prints. Partial prints contain toe and heel portions.

Given a footwear impression at a crime scene, a key task is to find the closest match in database so as to determine footwear brand and model. Every test image is matched with 64 templates to find the best match which corresponds to the target footwear print.

The global signal based descriptors (used before in our bench-marking methods) do not provide interesting results for footwear impression evidence verification. This is due to insufficient extraction of local information. An average recognition rate is provided in Table 7. In contrast to the benchmarking methods, $\mathcal{D}-$Radon yields a much higher recognition rates for most number of bins.

7 CONCLUSIONS

In this paper, we have presented a method based on the Radon transform for representing patterns and the DTW algorithm for matching. The key characteristic of the method is to use the DTW algorithm to match corresponding histograms at every projecting angle. We have validated the method over several different datasets in different contexts like recognition and retrieval. The robustness of the method has been tested by using samples subject to distortion, degradation, noise, occlusion, missing parts as well as size variation. We have established the discriminative power of the method even when complexity increases like in our case study using footwear prints. The main result of our findings is that $\mathcal{D}-$Radon provides a coherent and constant behaviour over a wide range of datasets in comparison to the state-of-the-art shape descriptors.

Computing the Radon transform is very simple and immediate. But the execution time for matching is high when using the standard DTW algorithm. However, it could be substantially reduced with the use of the advanced DTW Keogh (2002); Salvador and Chan (2007). Graphical processing units (GPU) will definitely be a possible solution for commercial use. Further investigation following other transform like the trace transform Kadyrov and Petrou (2001) would be worth-considering.

---

5. It is important to note that we have used basic DTW algorithm in order to attest the discrimination power of the shape descriptor. However, execution time could be substantially reduced by using the advanced DTW algorithm Keogh (2002); Salvador and Chan (2007).

6. Thanks to Signal Processing Laboratory, Kinki University, Japan for providing this dataset.
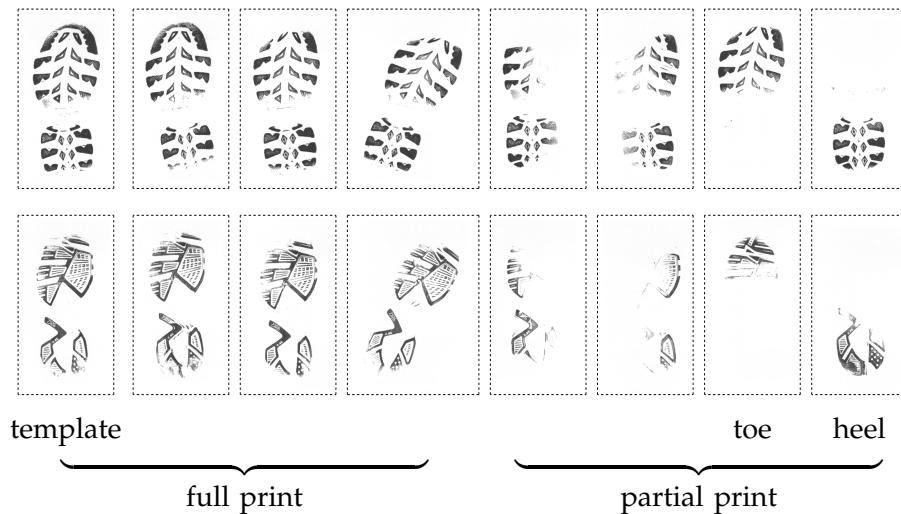
Fig. 12.   Eight samples of 2 different known footwear print classes.

TABLE 7
Average recognition rate (in $\%$) for full print as well as partial footwear print.

Images Set		GFD	ZM	SC	$\mathcal{R}$-sign.	$\mathcal{D}-$Radon for different values of B						
						180	90	60	36	18	09	02
Full print	64×4	65	45	55	46	**81**	78	74	73	61	55	54
Partial print												
heel	64×1	18	19	17	03	**44**	43	42	35	29	27	24
toe	64×1	09	08	03	03	**35**	34	33	32	22	17	12
other	64×2	11	17	04	07	**56**	52	52	41	31	29	22

REFERENCES

Attalla, E. and Siy, P. (2005).   Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching. *Pattern Recognition*, **38**(12), 2229–2241.

Bailey, R. R. and Srinath, M. (1996).   Orthogonal moment features for use with parametric and non-parametric classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**(4), 389–399.

Ballard, D. (1981).   Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, **13**(2), 111 – 122.

Bandera, C. U. A. and Sandoval, F. (2002). Non-parametric planar shape representation based on adaptive curvature functions. *Pattern Recognition*, **35**, 43–53.

Belkasim, S. O., Shridar, M., and Ahmadi, M. (1991). Pattern Recognition with Moment Invariants: A Comparative Study and New Results. *Pattern Recognition*, **24**, 1117–1138.

Bellman, R. and Kalaba, R. (1959). On adaptive control processes. *Automatic Control*, **4**(2), 1–9.

Belongie, S., Malik, J., and Puzicha, J. (2002).   Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(4), 509–522.

Bernier, T. and Landry, J.-A. (2003). A new method for representing and matching shapes of natural objects. *Pattern Recognition*, **36**(8), 1711–1723.

Bhattacharya, U. and Chaudhuri, B. B. (2009).   Handwritten numeral databases of indian scripts and multistage recognition of mixed numerals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**(3), 444–457.

Bhowmik, T. K., Parui, S. K., Bhattacharya, U., and Shaw, B. (2006).   An hmm based recognition scheme for handwritten oriya numerals. In *Proceedings of International Conference on Information Technology*, pages 105–110.

Boulgouris, N. V. and Chi, Z. X. (2007). Gait recognition using radon transform and linear discriminant analysis. *IEEE Transactions on Image Processing*, **16**(3), 731–740.

Charrier, C., Maloney, L. T., Cherifi, H., Knoblauch, K., and Lô, S. (2007). Maximum likelihood difference scaling of image quality in compression-degraded images. *Journal of the Optical Society of America A Optics, Image Science and Vision*, **24**(11), 3418–3426.

Chong, C., Raveendran, P., and Mukudan, R. (2003). A comparative analysis of algorithms for fast computation of zernike moments. *Pattern Recognition*, **36**, 731–742.

Coetzer, J. (2005). *Off-line Signature Verification*. Phd thesis, Department of Applied Mathematics, University of Stellenbosch.

Coetzer, J., Herbst, B., and du Preez, J. (2004). Offline signature verification using the discrete radon transform and a hidden markov model. *EURASIP Applied Signal Processing*, **4**, 559–571.

Deans, S. R. (1983). *Applications of the Radon Transform*. Wiley Interscience Publications, New York.

Duda, R., Hart, P., and Stork, D. (2001). *Pattern Classification - 2me dition*. Wiley Interscience Publications, New York.

El-ghazal, A., Basir, O., and Belkasim, S. (2009). Farthest point distance: A new shape signature for fourier descriptors. *Signal Processing: Image Communication*, **24**(7), 572–586.

Fornés, A., Lladós, J., Sánchez, G., and Karatzas, D. (2010). Rotation invariant hand-drawn symbol recognition based on a dynamic time warping model. *International Journal on Document Analysis and Recognition*, **13**(3), 229–241.

Fränti, P., Mednonogov, A., Kyrki, V., and Kälviäinen, H. (2000). Content-based matching of line-drawing images using the hough transform. *International Journal on Document Analysis and Recognition*, **3**(2), 117–124.

Gerdes, R., Otterbach, R., and Kammüller, R. (1995). Fast and robust recognition and localization of 2-d objects. *Machine Vision and Applications*, **8**(6), 365–374.

GREC (2003). International symbol recognition contest at grec2003.

Hasegawa, M. and Tabbone, S. (2011). A shape descriptor combining logarithmic-scale histogram of radon transform and phase-only correlation function. In *Proceedings of IAPR International Conference on Document Analysis and Recognition*, pages 182–186.

Hoang, T. V. and Tabbone, S. (2012a). The generalization of the r-transform for invariant pattern representation. *Pattern Recognition*, **45**(6), 2145–2163.

Hoang, T. V. and Tabbone, S. (2012b). Invariant pattern recognition using the rfm descriptor. *Pattern Recognition*, **45**(1), 271–284.

Jafari-Khouzani, K. and Soltanian-Zadeh, H. (2005). Radon transform orientation estimation for rotation invariant texture analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**(6), 1004–1008.

Jeannin, S. and Bober, M. (1999). Description of core experiments for mpeg-7 motion/shape. In *Technical Report ISO/IEC JTC 1/SC 29/WG 11 MPEG99/N2690, MPEG-7*, Seoul.

Kadyrov, A. and Petrou, M. (2001). The trace transform and its applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**(8), 811–828.

Kamila, N. K., Mahapatra, S., and Nanda, S. (2005). Invariance image analysis using modified zernike moments. *Pattern Recognition Letters*, **26**(6), 747–753.

Kankanhalli, M. S., Mehtre, B. M., and Wu, J. K. (1995). Cluster-based color matching for image retrieval. *Pattern Recognition*, **29**, 701–708.

Kassim, A. A., Tan, T., and Tan, K. H. (1999). A Comparative Study of Efficient Generalised Hough Transform Techniques. *Image and Vision Computing*, **17**, 737–748.

Keogh, E. J. (2002). Exact indexing of dynamic time warping. In *Proceedings of 28th International Conference on Very Large Data Bases*, pages 406–417. Morgan Kaufmann.

Khan, Z. A. and Sohn, W. (2013). A hierarchical abnormal human activity recognition system based on r-transform and kernel discriminant analysis for elderly health care. *Computing*, **95**(2), 109–127.

Kim, W.-Y. and Kim, Y.-S. (2000). A region-based shape descriptor using zernike moments. *Signal Processing: Image Communication*, **16**(1-2), 95 – 102.

Kimia, B. B., Tannenbaum, A., and Zucker, S. W. (1995). Shapes, shocks, and deformations i: The components of two-dimensional shape and the reaction-diffusion space. *International Journal of Computer Vision*, **15**(3), 189–224.

Kittler, J., Hatef, M., Duin, R., and Matas, J. (1998). On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**(3), 226–239.

Kruskall, J. B. and Liberman, M. (1983). The symmetric time warping algorithm: From continuous to discrete. In *Time Warps, String Edits and Macromolecules: The Theory and Practice of String Comparison*, pages 125–161. Addison-Wesley.

Kudo, M. and Sklansky, J. (2000). Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition*, **33**(1), 25–41.

Leavers, V. (2000). Use of the two-dimensional radon transform to generate a taxonomy of shape for the characterization of abrasive powder particles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(12), 1411–1423.

Loncaric, S. (1998). A Survey of Shape Analysis Techniques. *Pattern Recognition*, **31**(8), 983–1001.

Maes, M. (1991). Polygonal shape recognition using string-matching techniques. *Pattern Recognition*, **24**(5), 433–440.

Mokhtarian, F. and Abbasi, S. (2002). Shape similarity retrieval under affine transforms. *Pattern Recognition*, **35**(1), 31–41.

Nacereddine, N., Tabbone, S., Ziou, D., and Hamami, L. (2010). Shape-based image retrieval using a new descriptor based on the radon and wavelet transforms. In *Proceedings of IAPR International Conference on Pattern Recognition*,

pages 1997–2000.

Pelillo, M., Siddiqi, K., and Zucker, S. W. (1999). Matching hierarchical structures using association graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **21**(11), 1105–1120.

Persoon, E. and Fu, K.-S. (1986). Shape discrimination using fourier descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8**(3), 388 –397.

Prokop, R. J. and Reeves, A. P. (1992). A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graph Models Image Processing*, **54**(5), 438–460.

Ruta, D. and Gabrys, B. (2000). An overview of classifier fusion methods. *Computing and Info. Sys.*, **7**(1), 1–10.

Salvador, S. and Chan, P. (2007). Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, **11**(5), 561–580.

Santosh, K. C. (2011). Character recognition based on dtw-radon. In *Proceedings of IAPR International Conference on Document Analysis and Recognition*, pages 264–268.

Santosh, K. C., Lamiroy, B., and Wendling, L. (2011). Dtw for matching radon features: A pattern recognition and retrieval method. In J. Blanc-Talon, R. P. Kleihorst, W. Philips, D. C. Popescu, and P. Scheunders, editors, *Proceedings of International Conference on Advances Concepts for Intelligent Vision Systems*, volume 6915 of *Lecture Notes in Computer Science*, pages 249–260. Springer-Verlag.

Santosh, K. C., Lamiroy, B., and Wendling, L. (2012). Symbol recognition using spatial relations. *Pattern Recognition Letters*, **33**(3), 331–341.

Sharvit, D., Chan, J., Tek, H., and Kimia, B. B. (1998). Symmetry-based indexing of image databases. *Journal of Visual Communication and Image Representation*, **9**, 366–380.

Siddiqi, K., Shokoufandeh, A., Dickinson, S. J., and Zucker, S. W. (1999). Shock graphs and shape matching. *International Journal of Computer Vision*, **35**(1), 13–32.

Souvenir, R. and Parrigan, K. (2009). Viewpoint manifolds for action recognition. *EURASIP Journal on Image & Video Processing*, **2009**, 1–13.

Tabbone, S., Wendling, L., and Salmon, J.-P. (2006). A new shape descriptor defined on the radon transform. *Computer Vision and Image Understanding*, **102**(1), 42–51.

Tabbone, S., Terrades, O. R., and Barrat, S. (2008). Histogram of radon transform. a useful descriptor for shape retrieval. In *Proceedings of IAPR International Conference on Pattern Recognition*, pages 1–4.

Teh, C.-H. and Chin, R. T. (1988). On image analysis by the methods of moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **10**(4), 496–513.

Terrades, O. R., Valveny, E., and Tabbone, S. (2009). Optimal classifier fusion in a non-bayesian probabilistic framework. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**(9), 1630–1644.

Toft, P. (1996). *The Radon Transform - Theory and Implementation*. PhD Thesis, Department of Mathematical Modelling, Technical University of Denmark.

Tombre, K., Ah-Soon, C., Dosch, P., Habed, A., and Masini, G. (1998). Stable, robust and off-the-shelf methods for graphics recognition. In *Proceedings of IAPR International Conference on Pattern Recognition*, volume 1, page 406.

Tooley, M. and Wyatt, D. (2008). *Aircraft electrical and electronic systems: principles, operation and maintenance*. Aircraft engineering principles and practice. Butterworth-Heinemann, ISBN: 9780750686952.

Valveny, E., Dosch, P., Winstanley, A. C., Zhou, Y., Yang, S., Yan, L., Liu, W., Elliman, D., Delalandre, M., Trupin, É., Adam, S., and Ogier, J.-M. (2007). A general framework for the evaluation of symbol recognition methods. *International Journal on Document Analysis and Recognition*, **9**(1), 59–74.

Watanabe, S. (1985). *Pattern Recognition: Human and Mechanical*. John Wiley & Sons, ISBN: 0471808156, New York.

Zhang, D. and Lu, G. (2002). Shape-based image retrieval using generic fourier descriptor. *Signal Processing: Image Communication*, **17**, 825–848.

Zhang, D. and Lu, G. (2004). Review of shape representation and description techniques. *Pattern Recognition*, **37**(1), 1–19.

Zhang, D. and Lu, G. (2005). Study and evaluation of different fourier methods for image retrieval. *Image and Vision Computing*, **23**(1), 33–49.

Zhang, Z. (2002). *Image Retrieval Based on Shape*. Thèse de doctorat, Faculty of Information Technology, Monash University.

Zhu, S. C. and Yuille, A. L. (1996). FORMS: A Flexible Object Recognition and Modelling System. *International Journal of Computer Vision*, **20**(3), 187–212.

K.C. Santosh was graduated from Pokhara University, Nepal in 2003 with an Engineering degree in Electronics and Communication. He received a M.S. degree by research and thesis in Information Technology from SIIT, Thammasat University, Thailand in 2007 under ADB-JSP program. In 2011, he received a Ph.D. in Computer Science from Institut National Polytechnique de Lorraine (INPL) (now Université de Lorraine), France under INRIA-CORDI research grant. Currently, he is working as a post-doctoral researcher at INRIA Nancy Grand Est Research Centre, LORIA - Univesité de Lorraine, France. His research interests include handwriting analysis and recognition, graphics recognition, table extraction, sketch recognition, document information content exploitation and forensic image analysis like footwear impression evidence verification. He serves on the international conference on document analysis and recognition (ICDAR) 2013, IAPR as program committee member.

Bart Lamiroy is a permanent faculty member at the Université de Lorraine, in Nancy, France, and member of the associated LORIA research lab since 2000. He was a visiting scientist at Lehigh University from January 2010 to July 2011. He has a broad experience in Machine Perception. Over the years, his research topics have ranged from Content Based Image Retrieval over Visual Servoing to Document Analysis.
From 2007 to 2009, he was head of the Computer Science and IT Department at the École des Mines de Nancy, France, and has been a permanent faculty member there since 2000. Before that he was a research contractor at INRIA, after having obtained his Ph.D. in computer vision at the Institut National Polytechnique de Grenoble, France in 1998. He received his bachelor's degree in applied mathematics in 1993. He also serves on the International Association for Pattern Recognition TC-10 Committee as Dataset Curator, and on the Publicity and Publications Committee.

Laurent Wendling was born in Lunéville, France, in 1969. He received the Ph.D. degree in computer science from the University of Paul Sabatier, Toulouse, France, in 1997. He received the "HDR" degree in 2006. From 1993 to 1999, he was with the IRIT in the field of pattern recognition. From 1999 to 2009, he was an Assistant Professor at the ESIAL Nancy, and a member of LORIA in the field of symbol recognition. His current research topics are spatial relation, feature selection, and image segmentation. He is currently a full professor at the Paris Descartes University, in the field of computer science. He is also a member of the SIP team, LIPADE.