Pattern Recognition 45 (2012) 2145–2163

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

The generalization of the R-transform for invariant pattern representation

Thai V. Hoang [a,b,*], Salvatore Tabbone [b]

[a] MICA Center, HUST – CNRS/UMI 2954 – Grenoble INP, Hanoi, Vietnam
[b] LORIA, CNRS/UMR 7503, Nancy University, 54506 Vandoeuvre-les-Nancy, France

ARTICLE INFO

Article history:
Received 9 July 2011
Received in revised form
12 October 2011
Accepted 1 November 2011
Available online 19 November 2011

Keywords:
Invariant pattern representation
Radon transform
R-transform
R-signature
Feature extraction
Dominant directions
Noise robustness

ABSTRACT

The beneficial properties of the Radon transform make it a useful intermediate representation for the extraction of invariant features from pattern images for the purpose of indexing/matching. This paper revisits the problem of Radon image utilization with a generic view on a popular Radon transform-based transform and pattern descriptor, the R-transform and R-signature, bringing in a class of transforms and descriptors spatially describing patterns at all directions and at different levels, while maintaining the beneficial properties of the conventional R-transform and R-signature. The domain of this class, which is delimited due to the existence of singularities and the effect of sampling/quantization and additive noise, is examined. Moreover, the ability of the generic R-transform to encode the dominant directions of patterns is also discussed, adding to the robustness to additive noise of the generic R-signature. The stability of dominant direction encoding by the generic R-transform and the superiority of the generic R-signature over existing invariant pattern descriptors on grayscale and binary noisy datasets have been confirmed by experiments.

1. Introduction

Many descriptors have been proposed in literature for the extraction of patterns' invariant features [1,2] using techniques that allow invariance to rotation, scaling, translation (RST) or their combinations. Translation and scaling invariance could be obtained by using the Fourier [3] and Mellin [4] transforms respectively; rotation invariance by computing the harmonic expansion [5] or performing the discrete Fourier transform on the circular slices of the pattern image in polar space [6], etc. However, the task of combining several techniques to have all RST while guaranteeing the discriminatory power of the extracted invariant features is challenging and has attracted attention of many researchers. Most of the existing methods do not allow invariance to all RST, they usually require normalization for the unavailability of any of RST. For example, methods based on the theory of moments [7] usually normalize input patterns regarding their centroid position and size: the pattern's centroid is required to coincide with the origin of the coordinate system and the longest distance between this centroid and a pattern's point is set to a fixed value. These normalizations usually introduce errors, are sensitive to noise, and thus induce inaccuracy in a later matching process.

Radon transform-based methods are different from the others in the sense that the Radon transform is used to create an intermediate representation upon which invariant features are extracted from for the purpose of indexing/matching. There are some reasons for the utilization of Radon transform. Firstly, 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. Secondly, it is a lossless transform, patterns can even be reconstructed accurately by the inverse Radon transform. Thirdly, it has low complexity, requiring $O(N \log N)$ operations for an input pattern image of N pixels [8]. And finally and more importantly, it has useful properties concerning RST transformations which have been applied on patterns. By applying the Radon transform on an RST-transformed pattern image, the transformation parameters are encoded in the radial (for translation and scaling) and angular (for rotation) slices of the obtained Radon image [9]. Current techniques thus usually exploit this encoded information to define invariant features.

A pioneer work in this direction is the R-transform, which gives rise to the R-signature [10], 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 [11] 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 normalizations. The strength of these two approaches is simplicity,

* Corresponding author at: Bureau B226, LORIA, Campus Scientifique, BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France.
Tel.: +33 3 54 95 85 60; fax: +33 3 83 27 83 19.
E-mail addresses: vanthai.hoang@loria.fr (T.V. Hoang), tabbone@loria.fr (S. Tabbone).

0031-3203/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2011.11.007

however, the obtained signatures have low discriminatory power as there is a loss of information in the compression process from the Radon image to 1D signatures. Moreover, the required normalizations for the Φ-signature prevent it from being applied to noisy images.

There was an effort to apply the 2D Fourier–Mellin transform on the Radon image [12] to get invariance to rotation and scaling. In this approach, Mellin transform and harmonic expansion are applied on the radial and angular slices of the Radon image respectively. The main weakness of this approach is the lack of translation invariance. This drawback has been overcome in [13] by replacing the Mellin transform by the 1D Fourier–Mellin transform on the radial slices. Recently, a set of spectral and structural features has also been extracted from the Radon image for pattern description [14]. In this set, the "degree of uniformity" is essentially the R-transform and the "longest line" is the information encoded in the generic R-signature described in this paper. However, and more importantly, this set of features is not invariant to rotation and consequently, in the matching step, these features need to be rotated to all possible rotating angles corresponding to potential pattern's orientations in order to compute patterns' similarity. Long matching time may prevent the application of this approach in real systems.

Another direction in using the Radon transform for pattern description is to extract pattern features directly from the Radon image, similar to the way the Hough transform [15] is used. For example, pattern primitives in edge form are detected from the Radon image and represented analytically in [16]. Moreover, their spatial relations can be made explicit [17] and this leads to a taxonomy of patterns for their characterization [18]. This approach, however, is quite limited as it requires that the edge primitives have analytical form. Generalizations of the Radon transform, called the trace and geometric transforms, have also been proposed and used for image description [19,20] by using functionals other than integral and by extending the functional domain from a line to a region delimited by a closed contour. However, their application is restricted due to high computational complexity.

Among the Radon transform-based pattern descriptors, R-signature is the most popular due to its simplicity and has been successfully applied to several applications (e.g., symbol recognition [21], activity recognition [22,23], and orientation estimation [24]). This paper provides a generic view on the R-transform and the R-signature while maintaining their beneficial properties, leading to three main theoretical contributions. The first is a better understanding of the discriminatory power of the generic R-signature, which results from the exploitation of variation in the accumulations of the pattern image along all parallel lines, leading to an increase in performance. The second is a discussion on the reasonable range of the generalization, which is limited due to the existence of singularities and the sensitivity of the generic R-transform to sampling/quantization and additive noise. The last is the ability of the generic R-transform to represent dominant directions of patterns, even in the presence of noise, resulting in the superiority of the generic R-signature on noisy datasets over comparison methods. A preliminary study of this generalization has been published in [25].

The remainder of this paper is organized as follows. Section 2 gives some background on the Radon transform along with a brief review of the conventional R-transform and R-signature. Section 3 presents the generalization of the R-transform and R-signature along with their geometric interpretation and their properties. A discussion on the meaningful domain of these generic transform/signature is carried out in Section 4. Theoretical arguments on their robustness to additive noise and their ability to encode dominant directions of patterns are presented in Section 5. Experimental results are given in Section 6 and finally conclusions are drawn in Section 7.

2. Basic material

This section provides some basics of the Radon transform, its definition and its derived beneficial properties. The inspiration for the proposal of the R-signature from these properties will also be presented afterwards.

2.1. The Radon transform

Let $f(x,y)$ be a 2D function and $L(\theta,\rho)$ be a straight line in $\mathbb{R}^2$ represented by

$$L = \{(x,y) \in \mathbb{R}^2 : x\cos\theta + y\sin\theta = \rho\},$$

where θ is the angle L makes with the y-axis and ρ is the distance from the origin to L. The Radon transform [26] of f, denoted by $\mathcal{R}_f$, is a functional defined on the space of lines $L(\theta,\rho)$ by the line integral along each line

$$\mathcal{R}_f(L) = \mathcal{R}_f(\theta,\rho) = \int_L f(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(\rho - x\cos\theta - y\sin\theta)\,dx\,dy. \quad (1)$$

In the field of shape analysis and recognition, the function $f(x,y)$ is constrained to the following particular definition:

$$f(x,y) = \begin{cases} 1, & \text{if } x \in D, \\ 0, & \text{otherwise,} \end{cases}$$

where D is the domain of the binary shape represented by $f(x,y)$. In the illustration of the Radon transform in Fig. 1, the shaded region represents the region D. The value of the line integral in Eq. (1) is equal to the length of the intersection between the line L and the shaded region.

The Radon transform has some properties that are beneficial for invariant pattern recognition problems as outlined below

P1 *linearity*: The Radon transform is linear.

$$\mathcal{R}_{(f+g)}(\theta,\rho) = \mathcal{R}_f(\theta,\rho) + \mathcal{R}_g(\theta,\rho).$$

P2 *periodicity*: The Radon transform of $f(x,y)$ is periodic in the variable θ with period 2π.

$$\mathcal{R}_f(\theta,\rho) = \mathcal{R}_f(\theta + 2k\pi,\rho), \quad \forall k \in \mathbb{Z}.$$

P3 *semi-symmetry*: The Radon transform of $f(x,y)$ is semi-symmetric.

$$\mathcal{R}_f(\theta,\rho) = \mathcal{R}_f(\theta \pm \pi, -\rho).$$

P4 *translation*: A translation of $f(x,y)$ by a vector $\vec{u} = (x_0,y_0)$ results in a shift in the variable ρ of $\mathcal{R}_f(\theta,\rho)$ by a distance

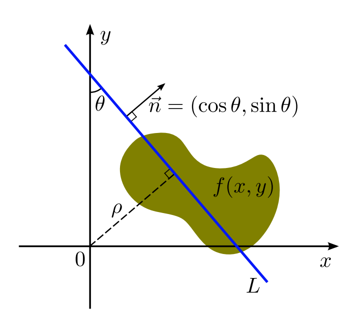

Fig. 1. Geometric illustration of the Radon transform of a function $f(x,y)$. The Radon transform is a mapping from the spatial space (x,y) to the parameter space (θ,ρ) and can be mathematically represented by a line integral of $f(x,y)$ along all the lines L parameterized by (θ,ρ) represented in the spatial space (x,y).