Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computer Vision and Image Understanding 102 (2006) 42–51
Computer Vision and Image Understanding
www.elsevier.com/locate/cviu

A new shape descriptor defined on the Radon transform

S. Tabbone *, L. Wendling, J.-P. Salmon

LORIA, Campus Scientifique, BP 239, 54 506 Vandœuvre-les-Nancy Cedex, France

Received 13 November 2003; accepted 10 June 2005
Available online 9 November 2005

Abstract

This paper presents a novel approach to identify complex shapes based on the Radon transform. We propose an adaptation of Radon transform called $\mathcal{R}$-transform, which is invariant to common geometrical transformations. Moreover, to improve the uniqueness of the approach, a binary shape is projected into the Radon space for different levels of the Chamfer distance transform. The accuracy and the efficiency of the proposed algorithm in the presence of a variety of transformations are evaluated within a shape recognition process.

Keywords: Radon transform; Distance transform; Shape representation; Shape recognition

1. Introduction

Shape representation for object recognition has been the subject of much research and extensive surveys of shape analysis can be found in [15,21]. Many approaches have been proposed to describe the boundary contours from a small set of features. The choice of a particular representation scheme is usually driven by the need to cope with requirements such as robustness against noise, stability with respect to small distortions, invariance to common geometrical transformations or tolerance to occlusions. Two classes of feature descriptors are often encountered: those that work on a shape as a whole (called region-based descriptors) and those that work on the contours of the shape (called contour-based descriptors).

Usual contour-based descriptors include Fourier descriptors [17,24,26,38] which have been widely used and modified versions [1,26] have been proposed to compute the affine transformation between one shape to another. Curvature approaches [6,22,33] have also been used. In [22,33] a shape is described in a scale space by the maximum of the curvature. The similarity of two shapes are determined by measuring the distance between their corresponding scale space representations. The similarity may be computed only at high scale or at all scales. The shape context [5] is a robust descriptor to small perturbations of parts of the shape. The main drawback of this approach relies in the lack of guarantee with scale-invariance. Skeletal methods [18,35] can be defined as a variant of contour-based approaches. In these approaches, 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 [30,31] with the shock graphs (locus of singularities) and several works have focused on methods to efficiently match two graphs [23,28,29,31]. 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 [6].

Since contour descriptors are based on the boundary of a shape, they cannot capture the internal structure of a shape. Furthermore, these methods are not suited to disjoint shapes or shapes with holes because the boundary information is not available. Consequently, they are limited to certain kinds of applications.

Region-based descriptors work on a shape as a whole taking into account all the pixels within a shape. Common methods are based on moment theory [2–4,25,32,39] including geometric, Legendre, Zernike, and pseudo-Zernike

* Corresponding author. Fax: +33 3 83 41 30 79.
E-mail address: tabbone@loria.fr (S. Tabbone).

1077-3142/$ - see front matter
doi:10.1016/j.cviu.2005.06.005

moments. Comparative studies [2,39] have demonstrated the interest of the Zernike moments, and a lot of work has been focused on improving the invariance properties [10,16] and speeding up the fast computation of the Zernike moments [10]. To overcome the drawbacks of contour-based Fourier descriptors, Zhang and Lu [37] have proposed a region-based generic Fourier descriptor. To avoid the problem of rotation in the Fourier spectra, the two-dimensional Fourier transform is applied on a polar-raster sampled shape image. Experimental results show that this approach outperforms common contour-based (classical Fourier and curvature approaches) and region-based (Zernike moments) shape descriptors [37].

Region-based methods are more suited to general applications. However, they are more computationally intense and most approaches need to normalize (centroid position, re-sampling, and re-quantization) the image to achieve common geometrical invariances. These normalizations introduce errors, sensitivity to noise, and thus inaccuracy in the recognition process.

In this paper, we propose a new region-based method for the recognition of complex shapes based on the Radon transform. Earlier works [13,19,20] on the 2D Radon transform were dedicated to finding high-valued coefficients in the transformed domain to detect specific shape primitives like straight lines or arcs of conics. In all these approaches the encoded information is contour-based allowing for the characterization of simple shapes. Furthermore, this kind of representation is not suited for a recognition task because it needs to be normalized with respect to geometric parameters (translation, rotation, and scaling). Indeed, it is difficult to recover all the geometric parameters of the transformation between two objects using directly the Radon transform. We propose an original adaptation of the Radon transform to overcome this problem. We define a new representation (called $\mathcal{R}$-transform) which has a low time complexity and a nice behavior with respect to common geometrical transformations. Furthermore to improve the uniqueness [1] of the approach, a binary shape is projected in the Radon space for different levels of the Chamfer distance transform. Thus, the link between the internal structure and the boundaries of the shape is taken into account. In this manner, we provide an approach which is shape independent. Moreover, the transform is centroid independent and consequently the approach is less sensitive to noise. Experimental results show that the proposed method outperforms common shape descriptors (contour and region-based approaches).

The remainder of the paper is organized as follows. The definition of the Radon transform is recalled in Section 2. The current method and digital considerations are described in Sections 3 and 4. The measure of similarity between two shapes is described in Section 5 and experimental results are given in Section 6. Finally, the conclusion and future investigations are discussed in Section 7.

[1] We mean by uniqueness a representation that specify a shape uniquely.

2. The Radon transform

By definition the Radon transform of an image is determined by a set of projections of the image along lines taken at different angles. For discrete binary image data, each non-zero image point is projected into a Radon matrix. Let $f(x,y)$ be an image. Its Radon transform is defined by [12]:

$$T_{R^f}(\rho,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - \rho)\,\mathrm{d}x\,\mathrm{d}y, \tag{1}$$

where $\delta(\cdot)$ is the Dirac delta-function ($\delta(x)=1$ if $x=0$ and 0 elsewhere), $\theta \in [0,\pi[$ and $\rho \in]-\infty,\infty[$. In other words, T_{R^f} is the integral of f over the line $L_{(\rho,\theta)}$ defined by $\rho = x\cos\theta + y\sin\theta$.

In shape recognition field it is of particular interest to consider the case where the general function f is

$$f_D(x,y) = \begin{cases} 1 & \text{if}\,(x,y)\in D \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where D is the domain of the binary shape (see Fig. 1).

In the plane, let L_i be in normal form (ρ_i, θ_i). The Radon transform $T_{R^f}(\rho,\theta)$ describes the length intersection of all the line L_i with the function f_D for all $\theta_i \in [0,\pi[$ and $-\rho_{\min} < \rho \leqslant \rho_{\max}$. $\rho_{\min}$ and $\rho_{\max}$ are set to finite values which rely on the image size.

Since the Radon transform is linear by definition, geometric properties like straight lines or curves can be made explicitly by the Radon transform which concentrates energies (loci of intersection of several sinusoidal curves) from the image in few high-valued coefficients in the transformed domain. These remarks are illustrated in Fig. 2 where white pixels correspond to high energies loci.

The Radon transform has several useful properties. Some of them are relevant for shape representation [12]:

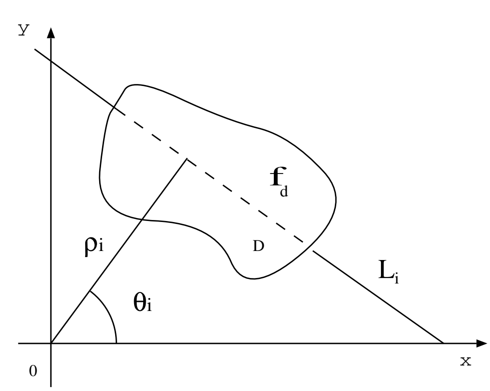

Fig. 1. Definition of the Radon transform.