



HAL
open science

A Shape Descriptor combining Logarithmic-scale Histogram of Radon Transform and Phase-only Correlation Function

Makoto Hasegawa, Salvatore Tabbone

► **To cite this version:**

Makoto Hasegawa, Salvatore Tabbone. A Shape Descriptor combining Logarithmic-scale Histogram of Radon Transform and Phase-only Correlation Function. 10th International Conference on Document Analysis and Recognition - ICDAR 2011, Sep 2011, Pekin, China. pp.182-186, 10.1109/ICDAR.2011.45 . hal-00657641

HAL Id: hal-00657641

<https://hal.science/hal-00657641>

Submitted on 7 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Shape Descriptor Combining Logarithmic-scale Histogram of Radon Transform and Phase-only Correlation Function

Makoto Hasegawa
LORIA, UMR 7503,
54506 Vandoeuvre-lès-Nancy, France
hasegawa.makoto@m.ieice.org

Salvatore Tabbone
LORIA, UMR 7503, Université Nancy 2,
54506 Vandoeuvre-lès-Nancy, France
tabbone@loria.fr

Abstract—A shape descriptor combining the histogram of the Radon transform, the logarithmic-scale histogram, and the phase-only correlation function is proposed. Applying a logarithmic-scale to the Radon transform, the shape scaling and rotation become two-dimensional translation in our descriptor without any normalization. The geometric invariance to translation, when we match two shapes, are kept using the phase-only correlation function. In addition, we can determine with this function the rotation angle and the scale parameter between two shapes. Our descriptor is robust to shape occlusion and noise also.

Keywords—The Radon transform; logarithmic scale; histogram; phase-only correlation;

I. INTRODUCTION

Shape description plays an important role in document analysis for character, symbol or logo recognition. Suitable shape descriptors should have invariant properties to geometric transformations (translation, rotation, and scaling), robustness to shape occlusion and noise, high recognition performance, and high processing speed. Many approaches have been proposed in the literature. D. Zhang and G. Lu [1] propose a method using the Fourier transform called the generic Fourier descriptor (GFD). The GFD is invariant toward shape rotation. The shape translation is normalized by a shape centroid shifted to the origin. The shape scaling is normalized using a magnitude control with the area of shapes. Subsequently, Q. Chen et al. [2] propose a method using the Fourier-Mellin transform (FMT). It is an enhancement of the Fourier transform using the log-polar mapping. The FMT is invariant toward such geometric transformations without any normalization. Moreover, a compact descriptor using the Radon transform called the \mathcal{R} -transform is proposed by S. Tabbone et al [3]. The \mathcal{R} -transform is invariant toward shape translation and rotation, however it needs a scaling normalization by a magnitude control. These normalizations are weak for shape occlusion and noise.

The histogram of Radon transform (HRT) proposed recently by S. Tabbone et al.[4] is a useful descriptor for shape retrieval. The HRT descriptor uses the Radon transform, which is efficient to the dominant lines in an image, and provides a 2D histogram representing the shape length at

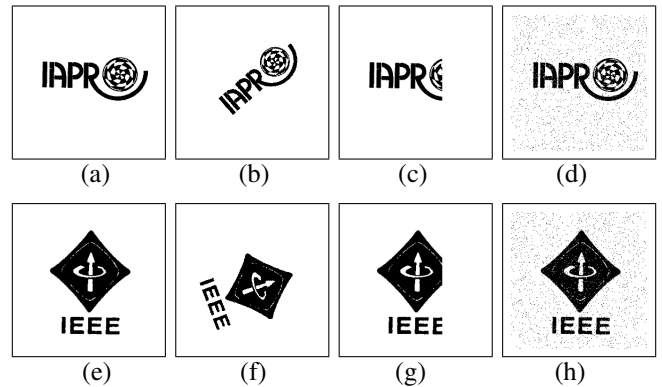
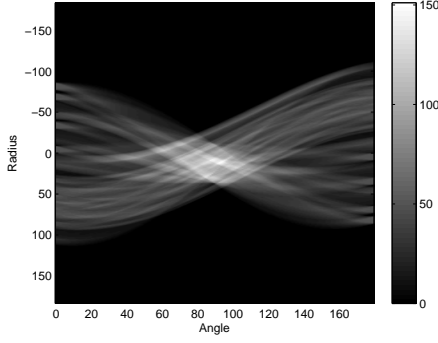


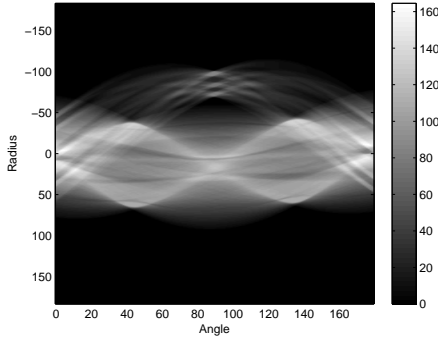
Figure 1. (a) Original image "IAPR." (b) Geometric transformations with translation, rotation, and scaling. (c) Shape occlusion: the area of 30% is occluded. (d) salt and pepper noise: 5% of the pixels flipped by the noise. (e)-(h) are the case of "IEEE."

each orientation. It is robust to the shape translation and rotation. In [4], the Radon image and the histogram are normalized with the area of shapes to achieve its scaling invariance. Another advantage of this descriptor is the compactness compared to other well know descriptors. However, if the original image has any occlusion or any noise, the area of shapes for the normalization is changed and affects negatively.

In this paper, we propose an extension of HRT descriptor combining the logarithmic conversion and the phase-only correlation function without any normalizations (LHRT). Applying the logarithmic conversion to the Radon transform, the shape scaling and the shape rotation become a two-dimensional translation in our descriptor. The geometric invariance to translation kept using the phase-only correlation function [5], which permits to recover the rotation angle and the scaling parameter. Our method does not need any normalization with the area of shapes. Consequently, the LHRT is more robust to the shape occlusion and the noise. Moreover, our descriptor can reduce the noise so that the noise-band is separated from the shape component in our descriptor. Excluding the noise-band provides better performances. In Section II, we recall the classical HRT



(a)



(b)

Figure 2. The Radon image of "IAPR" and "IEEE."

descriptor. After, we describe our method in Section III. Experimental results are shown in Section IV, and we draw conclusions in Section V.

II. HRT DESCRIPTOR

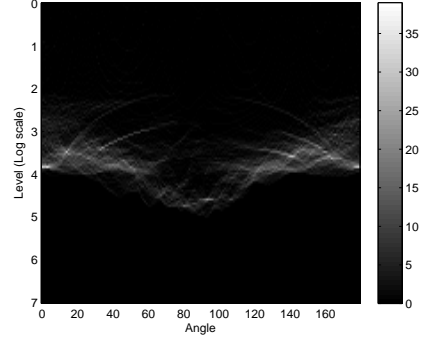
Let $f(x, y)$ be an original image of binary shape. It is described as:

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

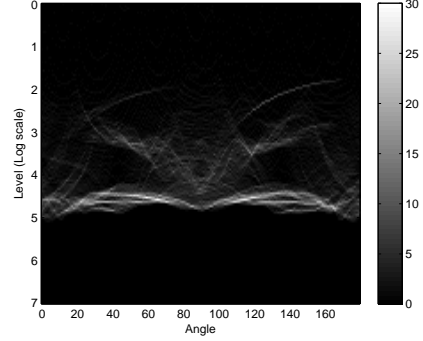
where X is the domain of the binary shape. The Radon transform is defined as:

$$R_f(\theta, \rho) = \int \int f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy, \quad (2)$$

where $\delta(\cdot)$ is the Dirac delta-function ($\delta(x) = 1$ if $x = 0$ and 0 elsewhere); $\theta \in [0, \pi)$ and $\rho \in [-A/2, A/2]$; A is the size of the image diagonal. In other words, the Radon transform is the integral of f over the line $L_{(\rho, \theta)}$ defined by $\rho = x \cos \theta + y \sin \theta$. The Radon images of Figs.1.a and 1.e are shown in Fig.2. In the case of geometric transformations as Fig.1, the shape translation becomes a sinusoidal shift relative to the radial ρ ; the shape rotation becomes a translation relative to the angle θ horizontally; a



(a)



(b)

Figure 3. The logarithmic-scale histogram of "IAPR" and "IEEE."

shape scaling becomes a 1-D scaling relative to the radial ρ vertically (see Figs.2.a and 4.a).

The HRT descriptor is defined as a matrix of frequencies computed on the Radon image aggregated by the angle parameter of the Radon transform. The HRT descriptor $D(\theta, y)$ of $\dot{R}_f(\theta, \rho)$ for each orientation θ is:

$$D(\theta, y) = H(\dot{R}_f(\theta, \cdot))(y), \quad (3)$$

where $H(f)(y)$ is a histogram of f as:

$$H(f)(y) = \frac{\#\{x \in X | y = f(x)\}}{|X|}, \quad (4)$$

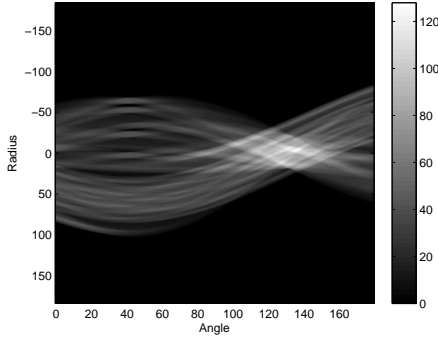
where $\#$ is the cardinality of a set, and $|X|$ means the area of shapes. The histogram need a normalization by $|X|$ for the scale invariance. Moreover, $\dot{R}_f(\rho, \theta)$ is:

$$\dot{R}_f(\theta, \rho) = \frac{R_f(\theta, \rho)}{\sqrt{|X|}}, \quad (5)$$

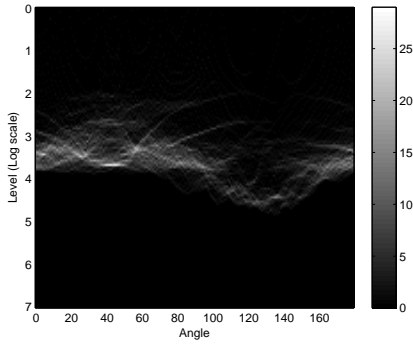
it also needs a normalization by $|X|$ for the scale invariance. The area of shapes $|X|$ can be calculated as:

$$|X| = \int R_f(\theta, \rho) d\rho \quad (6)$$

using any θ . The HRT descriptor is invariant toward the shape translation (see [4]). The shape rotation becomes a



(a)



(b)

Figure 4. (a) The Radon image of Fig.1.b and (b) the logarithmic-scale histogram.

translation relative to the angle θ . To match two images, a rotation invariant measure based on Euclidean distance $\|\cdot\|_2$ is applied as:

$$\text{Sim}(D_A(\theta, y), D_B(\theta, y)) = \min_{\alpha \in [0, \pi]} \|D_A(\theta, y) - D_B(\theta, y)\|_2. \quad (7)$$

We can remark that in the HRT descriptor if the shape has any occlusion or any noise, these normalizations with the area of shapes affect the whole Radon domain and the whole histogram, negatively. In the next section we propose an extension to solve this drawback.

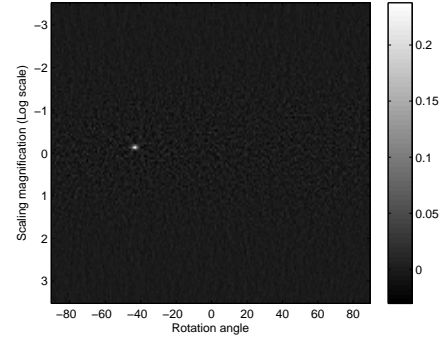
III. PROPOSED SCHEME

First, we calculate the Radon transform of an original binary image $f(x, y)$ with Eq.(2). After, we apply a logarithm conversion to the Radon image as:

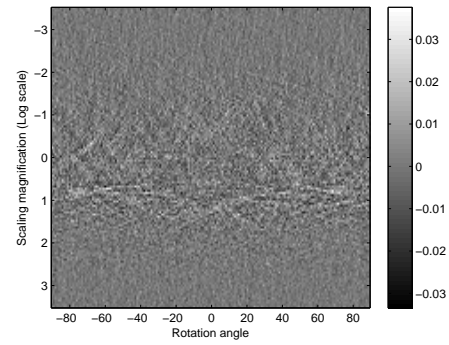
$$\tilde{R}_f(\theta, \rho) = \ln R_f(\theta, \rho). \quad (8)$$

Finally, the logarithmic-scale histogram of Radon transform (LHRT) is calculated as:

$$D(\theta, y) = H(\tilde{R}_f(\theta, \cdot))(y), \quad (9)$$



(a)



(b)

Figure 5. (a) The phase-only correlation with Figs.1.a and 1.b; (b) the case with Figs.1.a and 1.e.

where

$$H(f)(y) = \#\{x \in X | y = f(x)\}. \quad (10)$$

In our histogram, we do not normalize with the area of shapes $|X|$ (see Eq.(4)). The LHRT descriptors of Figs.1.a and 1.e are shown in Fig.3.

When the shape in f is scaled by α factor, the Radon image is scaled by α relative to the radius parameter ρ also, that is:

$$R_g(\theta, \rho) = \alpha R_f(\theta, \rho/\alpha), \quad (11)$$

where $R_f(\theta, \rho)$ is for the original shape, and $R_g(\theta, \rho/\alpha)$ is for the scaled shape. Subsequently,

$$\begin{aligned} \tilde{R}_g(\theta, \rho) &= \ln R_g(\theta, \rho/\alpha) + \ln \alpha \\ \tilde{R}_g(\theta, \rho) &= \tilde{R}_f(\theta, \rho/\alpha) + \ln \alpha. \end{aligned} \quad (12)$$

Consequently, LHRT descriptor is given by:

$$D_g(\theta, y) = \alpha D_f(\theta, y - \ln \alpha). \quad (13)$$

The shape scaling becomes a translation relative to the coordinate y (a vertical translation). Moreover, the shape rotation becomes, as for the HRT descriptor (see [4]), a translation relative to the θ coordinate (a horizontal translation). The



Figure 6. The UMD logo dataset.

descriptor is invariant toward the shape translation as the HRT descriptor does.

The Radon transform of Fig.1.b is defined Fig.4.a, and its LHRT descriptor in Fig.4.b. We can remark that the differences between the two shapes (Figs.1.a and 1.b) are -42° degree clockwise rotation and scaling by 0.87. Therefore, the corresponding Radon image of Fig.1.b is translated horizontally -42° and expanded vertically by 0.87 (see Figs.2.a and 4.a). However, our descriptor (Fig.4.b) is merely translated in both horizontal and vertical directions (42° in the angle coordinate horizontally, and $-0.15 = \ln 0.87$ in the level coordinate vertically; see Figs.3.a and 4.b).

To match two images, we apply the phase-only correlation function (POC) [5] as:

$$C(\theta, y) = \mathcal{F}^{-1} \left\{ \frac{G_1(u, v)G_2^*(u, v)}{|G_1(u, v)G_2^*(u, v)|} \right\}, \quad (14)$$

where $G(u, v)$ is a Fourier transform of $D(\theta, y)$, G^* is the complex conjugate of G , and \mathcal{F}^{-1} means the inverse Fourier transform. The matching score between two images is defined as the maximum value in $C(\theta, y)$. When the two shapes are same, there is a sharp peak in the POC (see Fig.5.a), and the matching score is relatively high (0.24 in this example). Moreover, the peak position denotes the rotation angle (clockwise -42° degree) and scaling magnification ($\exp -0.15 = 0.86$). When the two shapes are different, POC function is fuzzy (see Fig.5.b) and the matching score is low (0.04 in this case).

IV. EXPERIMENTAL RESULTS

Performances of our method have been evaluated using a trademark dataset provided by the Epeires project¹. Our dataset (see Fig.6) is composed of 32 categories, and each category has 11 shapes geometrically transformed (352 logos in total).

We create 3 datasets (a geometric transformed dataset; a geometric transformed and occluded dataset; a geometric transformed and noisy dataset). Each shape is matched against all the other shapes, so that 61,776 comparisons

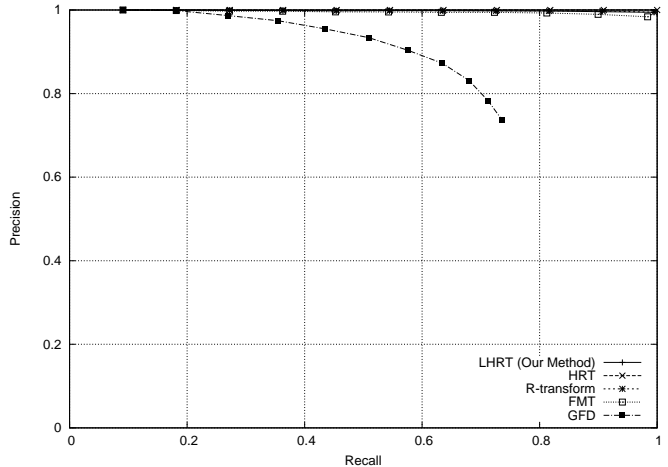


Figure 7. Precision recall curves on the geometric transformed dataset.

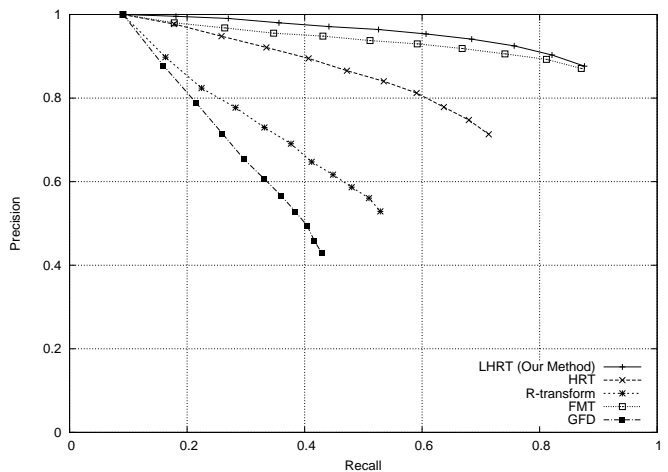


Figure 8. Precision recall curves on the geometric transformed and occluded dataset. The area of 30% is occluded.

are made in each dataset. Each shape is translated, rotated, and scaled randomly. For the occlusion, the area of 30% is reasonably occluded. In the noisy dataset, we add "salt & pepper" noise; 5% of the pixels flipped by the noise.

Our method (LHRT) is compared with the generic Fourier descriptor (GFD) [1], the Fourier-Mellin transform (FMT) [2], the \mathcal{R} -transform [3], and the histogram of Radon transform (HRT) [4].

The performances of these methods are evaluated using the precision-recall curves defined as:

$$\begin{aligned} \text{Precision} &= \frac{|\text{Relevant images} \cap \text{Retrieved images}|}{|\text{Retrieved images}|} \\ \text{Recall} &= \frac{|\text{Relevant images} \cap \text{Retrieved images}|}{|\text{Relevant images}|}, \end{aligned} \quad (15)$$

where $|\cdot|$ means the number of images. The precision and

¹<http://www.epeires.org>

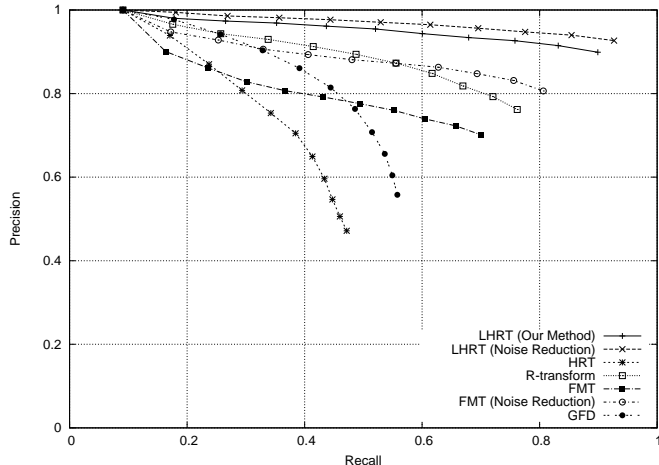


Figure 9. Precision recall curves on the geometric transformed and noisy dataset. The "salt & pepper" noise is added, and 5% of the pixels flipped by the noise.

the recall are inversely related each other. However, the descriptor whose precision and recall are both higher is a better one.

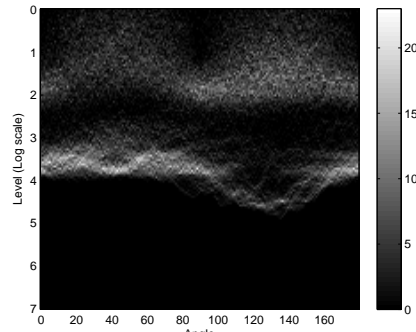
Fig.7 shows the precision recall curves on the geometric transformed dataset. Our method (LHRT), HRT, \mathcal{R} -transform, and FMT have a nearly perfect performance, so that it is observed that they are robust to the geometric transformations. However, the performance of GFD is poor for the geometric transformations.

In the case of occlusion, these performances are shown as Fig.8. Our method (LHRT) and the FMT have comparable and good performance; our method is better than the FMT. The HRT has the normalization with $|X|$ in Eq.(4), and the normalization decreases its performance. Our proposition without the normalization improves the performances remarkably.

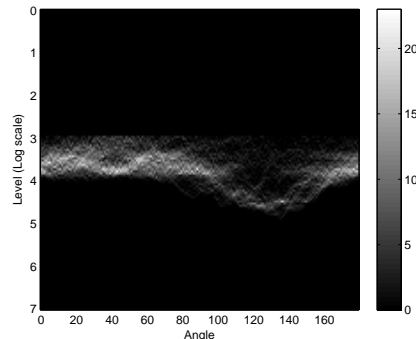
With the noisy dataset, as Fig.9, our method (LHRT) is the better. The performance of HRT decreases by its normalization. Our proposition without the normalization also improves the performance remarkably. Moreover, our descriptor can reduce its noise so that the noise-band is separated from the shape component (see Fig.10.a). Exclusion of the noise-band (the log-scale level is under 3.0) provides the better performance (LHRT in Fig.9). The FMT also can execute noise reduction (exclusion of the noise-band, its log-scale frequency is over 4.5), however our performance is better than the FMT.

V. CONCLUSION

Applying the logarithmic-scale to the HRT, the shape scaling and rotation are projected into two-dimensional translation. Therefore, the geometric invariance pattern matching can be executed using the phase-only correlation function, and we can detect the rotation angle and the scaling mag-



(a)



(b)

Figure 10. (a) The LHRT descriptor of noisy image and (b) our noise reduction.

nification. Our descriptor is robust to shape occlusion and noise without any normalization. Moreover, we can reduce noise excluding the noise-band.

REFERENCES

- [1] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," *Signal Processing: Image Communication*, vol. 17, no. 10, pp. 825-848, 2002.
- [2] Q. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transform for image registration and recognition," *IEEE Trans. PAMI*, vol. 16, no. 12, pp.1156-1168, Dec. 1994.
- [3] S. Tabbone, L. Wendling, and J. -P. Salmon, "A new shape descriptor defined on the Radon transform," *CVIU*, 102(1), pp.42-51, Apr. 2006.
- [4] S. Tabbone, O. Ramos Terrades, and S. Barrat, "Histogram of Radon transform. A useful descriptor for shape retrieval," *ICPR*, pp.1-4, Tampa, 2008.
- [5] C. Kuglin and D. Hines, "The phase correlation image alignment method," *Proc. Int. Conf. on Cybernetics and Society*, pp.163-165, 1975 .