A Rotation Invariant Image Descriptor based on Radon Transform

Yudong Zhang, Lenan Wu

(School of Information Science and Engineering, Southeast University, Nanjing, China)
zhangyudongmu@gmail.com, wuln@seu.edu.cn
doi:10.4156/jdcta.vol5. issue4.26

Abstract

In this paper, we proposed a rotation invariant image descriptor based on Radon transform (RT) and energy operator. Radon transform captures the directional features of the pattern image by projecting the pattern onto different orientation slices, and its most attractive ability is to transform rotational components to circular shift components. Meanwhile, the energy operator can remove the circular shift components. Therefore, the proposed descriptor is invariant with orientations. Moreover, the time complexity of traditional RT is $O(N^3)$. In order to hasten the procedures, an acceleration strategy was introduced based on projection slice theorem, so that a RT only needs $O(N^2 \log N)$ multiplications. The experiments demonstrate the anti-noising and rotation invariant abilities of the proposed descriptor, and a classification test on 20 distinct natural textures selected from Brodatz’s album shows that the proposed descriptor is superior to wavelet packet analysis, combined invariant feature, and radon transform plus multiscale analysis in respect to classification accuracy and computation time.

Keywords: Radon Transform, Rotation Invariant, Fourier Transform, Projection Slice Theorem, Feature Extraction, Pattern Recognition

1. Introduction

Feature extraction is a crucial step in invariant pattern recognition. Good features must satisfy the following requirements [1]: I) intraclass variance must be small; II) interclass separation should be large; III) features should be independent of the size, orientation, and location of the pattern, which means the features should be scale-, rotation-, and translation-invariant. The translation invariance can be achieved by moving the centroid of the pattern to the center of the pattern image. Also, the scale invariance can be done by scaling the pattern to a predefined image size. So the key task is to develop a rotation-invariant descriptor [2].

Tsai et al. [3] proposed a wavelet decomposition approach for rotation-invariant template matching. In the matching process, they first decomposed an input image into different multi-resolution levels in the wavelet-transformed domain, and used only the pixels with high wavelet coefficients in the decomposed detail subimage at a lower resolution level to compute the normalized correlation between two compared patterns. Sookhanaphibarn et al. [4] proposed a novel rotation-invariant pattern based on rotational direction algorithm (RDA) and scaling-invariant competitive learning algorithm (SIC-LA). Lahajnar et al. [5] first obtained energy-normalized features by multiscale and multichannel decomposition using Gabor and Gaussian filters, and then the rotation invariance is achieved by the Fourier expansion of these features with respect to orientation. Arivazhagan et al. [6] presented a new approach for rotation invariant texture classification using Gabor wavelets. Lo et al. [7] proposed a new scale and rotation invariant, texture-based segmentation algorithm, that performs feature extraction using the Dual-Tree Complex Wavelet Transform (DT-CWT).

Those methods are quite effective in extracting rotation-invariant features, however, most of them have the shortcoming in the way the number of features is limited, which will decrease the classification accuracy [8]. The Radon transform (RT) is an alternative representation that allows one to derive more number of features from an image, and in this study a fast RT was introduced of which the sequential computation complexity is $O(N^2 \log N)$ for a $N$-by-$N$ image [9] compared to the $O(N^3)$ of traditional RT [10].

The paper is organized as follows: next section 2 introduced in the basic principles of Radon transform; section 3 proposed a novel image descriptor based on RT and energy operator, discussed the anti-noising and rotation invariant abilities of the proposed descriptor, and introduced in an acceleration algorithm of
calculating RT. Experiments in Section 4 demonstrated the two aforementioned abilities of the proposed
descriptor, and applied it the texture classification. The results showed our proposed descriptor performs
best considering both the classification accuracy and the computation time. Final section 0 concluded the
paper and gave future research direction.

2. Radon Transform

The RT is based on the parameterization of straight lines and evaluation of integrals of an image along
these lines. Due to inherent properties of RT, it is a useful tool to capture the directional features of an
image [11]. The RT of a 2-D image function \( f(x,y) \) is denoted as \( R(\theta, r) \), which is defined as follows

\[
R(\theta, r) = \Re \{ f(x,y) \} = \int \int f(x,y) \delta(r - x \cos \theta - y \sin \theta) \, dx dy
\]

where \( \delta(.) \) is the Dirac function. We can image RT as the line integrals from multiple sources along
parallel paths called beams as shown in Fig 1, where \( \theta \in [0, \pi) \) denoting the angle between the beam and
\( x \)-axis, and \( r \in (-\infty, \infty) \) is the perpendicular distance from the beam crossing the origin.

![Image of the geometric illustration of the RT](image)

Fig 1. The geometric illustration of the RT

3. Rotation Invariant Feature Extraction

3.1 Robust to Noise

The fist advantage of RT is its robustness to zero mean white noise. Suppose an image is represented
as

\[
g(x,y) = f(x,y) + \sigma \times \epsilon(x,y)
\]

where \( \epsilon(x,y) \) is the white noise with zero mean and unit variance, and \( \sigma \) is the variance of the noise. It is
easy to know its RT is

\[
\Re \{ g(x,y) \} = \Re \{ f(x,y) \} + \sigma \times \Re \{ \epsilon(x,y) \}
\]

For the continuous case, the line integral of the RT of white noise is constant for all of the points and
all the directions, and the result is equal to its mean value which is assumed to zero. Therefore,

\[
\Re \{ g(x,y) \} = \Re \{ f(x,y) \}
\]

However, for the digital case, the above conclusion is not true because the image is composed of a
finite number of pixels. In this case, the noise immunity is expressed in terms of signal-to-noise ratio of
Radon projection (SNRp) and that of an image (SNR) given as follows [12]:
\[ SNR_p = SNR_i + 1.7 \times N \times SNR_i \]  

So SNR of projections is much higher than that of an image if the size of the image \((N)\) is large enough. Thus, the RT is nearly immune to zero mean white noise.

### 3.2. Rotation Invariance

The second advantage of RT is that the RT of the rotation of \(f(x,y)\) by angle \(\phi\) leads to a circular shift of the RT of original \(f(x,y)\) in the variable \(\theta\), viz.

\[ \Re[f(x,y)] = R(\theta,r) \Rightarrow \Re[f_{\phi}(x,y)] = R(\theta + \phi, r) \]  

Here \(f_{\phi}(x,y)\) denotes the rotation of \(f(x,y)\) by angle \(\phi\) and \(R(\theta + \phi, r)\) denotes the circularly shift by angle \(\phi\) along the \(\theta\) dimension. We use the circuit image as an example, rotate it by increasing angles from 0 to 90, and perform the RT of the rotated image. **Tab 1** 1 shows the result, where the horizontal axis of RT denotes the angle \(\theta\) and vertical axis of RT denotes the distance \(r\).
3.3. Feature Extraction

Since RT can transfer the rotation change to circular translation change, it is easy to take a energy
operation (summation) along the $\theta$ direction due to the energy operator is a circular-translation invariant
descriptor.

Therefore, our strategy to extract rotation-invariant features is a two-step process: First is to perform a
RT on the rotated pattern to transform the rotation component to circular translation component, followed
by an energy operator along the angle direction to transform the circular translation component to the
invariant features. The detailed steps are described as follows:

Step 1 Normalize the pattern of size $N \times N$ so that it is translation- and scale-invariant;
Step 2 Discard all those pixels that are outside the surrounding circle with center ($N/2$, $N/2$) and radius
$N/2$;
Step 3 Project the pattern in $2N$ different orientations to get the RT coefficients;
Step 4 Conduct energy operator on the radon coefficients along the $\theta$ direction;
Step 5 Save the features into the feature database.

3.4. Fast RT Calculation

The projection-slice theorem [13] of two-dimensional condition states that the following two
calculations are equal: (1) Project an image into one-dimensional line, and do a Fourier Transform (FT)
of that projection; (2) Do a two-dimensional FT first, followed by extracting the slice through the origin,
which is parallel to the projection line. In operator terms, for an image $f$,

\[ F_1 P_1 (f) = S_1 (F) \]  (7)

where $F_1$ is the 1-dimensional FT operators, $P_1$ is the projection operator which projects an image onto a
1-D line, and is a slice operator which extracts a 1-D central slice from an image, and $F$ is the 2D FT of
original image $f$.

![An illustration of projection slice theorem](image)

Fig 2. An illustration of projection slice theorem

Fig 2 shows a graphical illustration of projection-slice theorem. The $f$ and $F$ are 2-D FT pairs. The
projection of $f(x,y)$ onto the $x$-axis is the integral of $f$ along lines parallel to the $y$-axis and is labeled $P_1 (f)$.
Meanwhile, the slice through $F(k_x, k_y)$ is on the $k_x$-axis and is labeled $S_1 (F)$. The projection slice theorem
states that $P_1 (f)$ and $S_1 (F)$ are 1D FT pairs. Therefore, a fast RT can reduce the time complexity to
$O(N^2 \log N)$ with the help of projection slice theorem. The detailed procedures can be seen in Ref. [9].

4. Experiments

The experiments were performed with an IBM P4 with 2GHz processor and 1GB memory and by the
Windows XP operation system. The image processing toolbox and neural network toolbox of Matlab
2010b were employed for programming. For testing purpose, 20 classes of natural texture images are selected from the Brodatz’s texture album (http://www.ux.uis.no/~tranden/brodatz.html). Original size is 640-by-640, while we resize it to only 201-by-201 for simple. Row 1 contains D1, D3, D4, D6, and D16; Row 2 contains D19, D20, D21, D24, D31; Row 3 contains D34, D35, D37, D46, D47; Row 4 contains D50, D52, D53, D55, D56.

4.1. Rotation-invariant Property

In this case, we study the rotation-invariant property of the proposed feature. D1, D16, and D24 images are chosen with different orientations (0-179° with 20° intervals). Afterwards, we plot the extracted features in Fig 4, indicating that the features keep steady with the variation of the image orientations.
4.2. Anti-noising Property

In this case, we study the effect of anti-noising ability of the proposed feature. D20, D47, and D55 images are chosen with white Gaussian noises of different variances (0.01 to 0.1 with 0.01 intervals). The results are shown in Fig 5, indicating that the features changes within a relative small range with the variation of variances of the noises from 0.01 to 0.1.
4.3. Texture Classification

In the final case, the total 20 classes with different orientations (0-350 with 10 intervals) and different variances of noises (0-0.08 with 0.02 intervals) are generated. In this way, a database of 20*36*5=3600 images is created for the purpose of classification.

The procedures are depicted as follows: Firstly, features are extracted by different descriptors; then followed by a principle component analysis (PCA) [14] to reduce the number of features; then followed by a multi-layer neural network, of which the $k$-fold cross validation ($k=10$ in this study) was utilized to enhance the generalization ability. The whole procedure is repeated 10 times to reduce the randomness errors. Finally, the average classification accuracy and the computation time were calculated as the comparative indicators. We compare our descriptor with popular rotation invariant descriptors including the wavelet packet analysis (WPA) [15], combined invariant feature (CIF) [16], and radon transform plus multiscale analysis (RT-MSA) [17].

Tab 2 shows the average recognition results of three different descriptors. It indicates that traditional wavelet-based descriptors such as WPA and CIF can obtain good recognition rates as 92.25% and 92.39%, respectively, but they still need improvements. RT-based methods such as RT-MSA get larger recognition rate as 93.62%, and our method (RT-Energy) acquires the largest recognition rates (93.75%) among all descriptors. For computation time, wavelet-based descriptors cost very little time. The computation time of WPA and CIF are 0.87s and 0.91s, respectively. However, their classification accuracies are not satisfying. The RT-MSA performs well in classification but requires much more time as 3.21s. Our
method employs a fast RT algorithm so its computation time is only 1.15s.

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Average Classification Accuracy (%)</th>
<th>Computation Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WPA [15]</td>
<td>92.25</td>
<td>0.87</td>
</tr>
<tr>
<td>CIF [16]</td>
<td>92.39</td>
<td>0.91</td>
</tr>
<tr>
<td>RT-MSA [17]</td>
<td>93.62</td>
<td>3.21</td>
</tr>
<tr>
<td>RT-Energy (Our)</td>
<td>93.75</td>
<td>1.15</td>
</tr>
</tbody>
</table>

5. Conclusions

In this paper, a novel rotation invariant image descriptor was proposed based on RT and energy operator. RT can transform rotational components to circular shift components, and the energy operator eliminates the circular shift components, so the descriptor is invariant with orientations of given images. Besides, a fast RT calculation method was introduced in, the time complexity of which is \(O(N^2 \log N)\) compared to the one of traditional RT of \(O(N^3)\).

The future works focus on applying the proposed image descriptor on various industrial applications such as image segmentation [18], protein prediction [19], face recognition [20]. Moreover, it is worthy to test the classification performance of the combination of the proposed descriptor and more powerful classifiers such as support vector machine.

6. References

A Rotation Invariant Image Descriptor based on Radon Transform
Yudong Zhang, Lenan Wu


