Content Based Image Retrieval Using Improved Radon Transform under Various Queries

S. Singaravelan and D. Murugan

Department of Computer Science and Engineering, PSR Engineering College, Sivakasi, Tamilnadu, India
Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu, India

Abstract: Image Retrieval is very one of the biggest task in the recent years. It is widely used in many real time databases to retrieve related images in various fields like medical, military, online shopping etc. This paper offers with using radon transform followed by PCA and LDA techniques for image retrieval is called as Combined Radon Space Features Set (CRSFS). Caltech 101 database image sets used in this paper. The radon transform used in FFT based slice theorem. The correct direction is select means the computation time and complexity of operation is less to achieve good retrieval rate. Here PCA is used to reduce the dimensionality in feature vectors produced by radon transform and LDA is used to find the set of basic vectors which maximizes the ratio between-class scatter and within-class scatter. In order to verify our method to various data set in different condition like rotation, scaled and noisy query environments like additive, Gaussian, salt and pepper. And our proposed method was achieving better retrieval rate.

Key words: Image Retrieval (IR) · Combined Radon Space Features Set (CRSFS) · Radon Transform (RT) · Fast Fourier Transform (FFT) · Principal Component Analysis (PCA) · Linear Discriminant Analysis (LDA) · Noise Reduction

INTRODUCTION

Content based image retrieval (CBIR) is a one of the popular techniques which is visual contents to search and browsing digital images from image databases according to user’s interests [1-22]. The CBIR systems is used many real time applications and databases in medical, military, human identification & security systems and online searching & shopping [23]. The term ‘content based’ means the actual contents of the image rather than keywords, tags or descriptions associated with the image. The CBIR classified in two ways first one is text based retrieval and second one in content based retrieval. In our proposed method mainly concentrate in content based image retrieval. The two mode of CBIR is analyzed first one is low level features like color, shape and texture value to classified and retrieved the images. Low level features future classified in to two ways like global features and local features. Global features mean its measure the color, shape and texture information on entire image. Local features mean its measure the color, shape and texture information on some specific region and condition based to analyze the image [1].

Existing Work: Ho young [2] is used spatial color descriptor for image retrieval and video segmentation using the color histogram and angle histogram and spatial distribution of color pairs at color edges in an image. The error rate is 28.6%. The classical image database used it contains 5000 images.

Gwenole [3] is introduced adaptive non separable wavelet transform to CBIR its extracted image signature is derived from this new method. The error rate is 4.6%. The medical database DRD, DDSM & nonmedical database FD,VisTex is used.

Corresponding Author: S. Singaravelan, Department of Computer Science and Engineering, PSR Engineering College, Sivakasi, Tamilnadu, India.
Manesh kokare [4] is used new rotated complex wavelet filters to measure the texture values based on image retrieval. The error rate is 17.2%. The texture image database like Brodatz and VisTex is used.

Hui hui [5] is proposed a novel techniques for object spatial relationships semantic extraction and low level features extraction values was integrated. The error rate is 8.3%. The classical image database used.

Te –wei [6] is used multi resolution wavelet features of interest is apply to CBIR in discrete wavelet transform. The error rate is 23.7%. The classical image database used it contains 1000 images.

Gwenole [7] is integrated to proposed estimate query image in two function wavelet basic and regression function. The error rate is 3.1%. The Caltech 101, MOSSIDOR, face database and VisTex database used.

Yue Gao [8] is proposed an approach that simultaneously utilize both visual and textual information to estimate the relevance of user tagged images using hypergraph learning approach. The error rate is 2.4%. The Flicker datasets used it contains 104,000 images and 83,999 unique tags is used.

Qianni [9] is used to automatically combine heterogeneous visual features for histology image retrieval. The approach is multi objective learning methods. The error rate is 24%. The BiMed image database used it contains 20000 images.

Kekre [10] is proposed innovative idea of sectorization of full haar wavelet transformed images for extracting features. The paper proposed two phases to sector the image plane is forward and backward plane. The error rate is 24%. The classical image databases used.

Yi-chen [11] is using radon transform to examine in-plane rotation and scale invariant clustering in cbir objects. The error rate is 6.4%. The Smithsonian isolated leaf and Kimia - shape based database & Brodatz - texture based database used.

**Proposed Method:** The method used in this paper for content based image retrieval if Fig (1). It has a query phase and another one is data/database registered phase. In the data/database registered phase the set of images are transformed in to radon space using radon transform. Here radon transform performed based on 2D-FFT slice theorem is applied in this method. The data generated by radon transform are reduced using principal component analysis (PCA). From those reduced dataset (ie. the mainly valuable components) a set of basic vector which maximize the radio between-class scatter and within-class scatter using linear discriminant analysis(LDA). One basic vector for each image is derived and stored in the database. So for image database there are set of basic vectors are stored regardless of the number of images used for each dataset in the training set. The length of the basic vector depends on the number of components selected by the PCA stage. The query phase the input image is transformed into the radon space using radon transform based on 2D-FFT slice theorem. The PCA reduction is carried out as same as in the data/database registered phase. The resultant vector is calculated using distance measure comparison methods was used to check the query phase weight vector values W was compared to data/database registration phase weight vectors values. The most similar weight vectors values images are listed in retrieval results. In this work two types of distance measure used, L1 Norm and L2 Norm. L1 Norm measures the absolute distance between the pixels of the original and the corrected image. L2 Norm measures the Euclidian distance between the pixels of the original and the corrected image. The method which is stated in this paper was evaluated using the Caltech 101 database datasets which contains classical images & texture images. There are 8 different datasets used in this work. The images are color with a resolution of 120x120. For some of the datasets, the image were taken at different angle, scaling and noise added and apply for our experiments. The subsequent part discussion of each main part of the method process is presented.
Radon Transform: The radon transform is used many applications like categories the visual objects [12], rotation invariants in objects identification [13, 15], face recognitions [14], image recognitions [16]. The radon transform is the projection of the image intensity along a radial line oriented at a specific angle. The radial coordinates are the values along the x'-axis, which is oriented at degree counter clockwise from the x-axis. The origin of the both axis is the center pixel of the image. Radon transform of an image representation as a matrix, called sinogram.

Horizontal and Vertical Projection: The popular form expresses lines is

\[ t = x \cdot \cos(\theta) + y \cdot \sin(\theta) \]  

(1)

\( \theta \) – angle & t – smallest distance to the origin of the coordinate system

Projection can be computed along any angle \( \theta \), by use general equation of the radon transform

\[ g(t, \theta) = \int_{-\infty}^{\infty} f(x, y) \delta(t - x \cdot \cos(\theta) - y \cdot \sin(\theta)) \, dx \, dy \]  

(2)

The very strong property of the radon transform is the ability to extract lines (curves in general) from very noisy images.

Rotation Invariance: The radon transform of a two variable function \( x \) is defined as

\[ R_\theta x(t) = \int_{-\infty}^{\infty} x(t \cdot \cos \theta - s \cdot \sin \theta, t \cdot \sin \theta + s \cdot \cos \theta) \, ds \]  

(3)

where \((t, \theta) \in (-\infty, \infty) \times (0, \pi) \) \( \forall s \in \mathbb{R} \)

To compute the value at any given point \((t, \theta)\) in the radon domain by integrating along the line:

\[ (u, v) = (t \cdot \cos \theta - s \cdot \sin \theta + t \cdot \cos \theta) \]  

(4)

If \( x' \) is a rotated copy of \( x \) by on angle \( \theta \)

Simple proof shows \( R_\theta \tau(t) = R_{\theta + \theta} \tau(t) \) \( \forall t \)

Let \( \sigma_\theta = \text{var}_t [R_\theta x(t)] \) denotes the variance with respect to \( t \) of the RT at \( \theta \).

\( \sigma_\theta \) - Useful in estimating the principal orientation in an image.

Estimate \( \hat{\theta} \)

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

(5)

Scale Invariance: Let \( x' \) is the scaled copy of \( x \) with scaling factor \( \xi \) such that

\[ \tau(u, v) = x(\xi u, \xi v) \]  

(6)

Relate this to RT,

\[ R_\theta x(t) = \int_{-\infty}^{\infty} \tau(t \cdot \cos \theta - s \cdot \sin \theta + s \cdot \cos \theta) \, ds \]

\[ = \int_{-\infty}^{\infty} x(\xi t \cdot \cos \theta - \xi \cdot \sin \theta, \xi t \cdot \sin \theta + \xi \cdot \cos \theta) \, ds \]

\[ = \frac{1}{\xi} R_\theta x(\xi t) \]

From this observation, scale invariance can be achieved through the normalization in the RT.

Principal Component Analysis (PCA): Principal Components Analysis (PCA) is one of several statistical tools available for reducing the dimensionality of a data set [19, 20 and 24]. Its relative simplicity and both computational and in terms of understanding what’s happening and make it a particularly popular tool. The major goal of principal components analysis is to reveal hidden structure in a data set. In so doing, we may be able to

- Identify how different variables work together to create the dynamics of the system

Fig. 2: Radon Transform Horizontal & Vertical Projection View


1954
Fig. 4: FFT based Radon Transform Sinogram for an image

- Reduce the dimensionality of the data
- Decrease redundancy in the data
- Filter some of the noise in the data
- Compress the data

The training set D and testing set Q on basic vectors:

**Step 1:** Calculate mean value \( m \) for training set and subtract it from the training set

\[
m = \frac{1}{n} \sum_{i=1}^{n} D_i
\]  

All training set \( D_i = D_i - m \)

**Step 2:** Calculate the eigenvectors and eigenvalues of the training set covariance matrix \( C_i \)

\[
C_i = D_i \cdot D_i^T
\]  

The eigenvectors corresponding \( N \) largest eigenvalues of \( C_i \). These construct the principal component matrix (V).

**Step 3:** The basic vector matrix (E) is defined by:

\[
E = V \cdot D_i^T
\]  

**Step 4:** The reduced feature vectors for training and testing data calculated by:

\[
WD = D \cdot E
\]

\[
WQ = Q \cdot E
\]

**Linear Discriminant Analysis (LDA):** The LDA is used to find the set of basic vectors which maximizes the radio between -class scatter and within-class scatter [19, 21, 22-24].

The between class scatter matrix is defined by:

\[
S_B = \sum_{i=1}^{c} N_i (m_i - m_t)(m_i - m_t)^T
\]  

Here, \( C \) – Classes, \( C_i \)- ith Class, \( N_i \) – number of samples, \( M_t \) – mean of the hole dataset, \( M_i \) – mean of the ith class.

The within-class scatter matrix is defined by:

\[
S_W = \sum_{i=1}^{c} \sum_{mk \in ci} (mk - m_i)(mk - m_i)^T
\]  

The basic vector is defined by:

\[
W = \arg \max \left\{ \frac{W^T S_B W}{W^T S_W W} \right\}
\]

Solving Equations (11) to produces a matrix W whose columns are the eigenvectors corresponding to the largest eigenvalues are calculated. These columns are the linear discriminant function related with classes as one of the function for each data. These functions are stored in the database to perform the classification.

**Noises in Image:** Image noise [17] is random (not present in the object imaged) variation of brightness or color information in images and is usually an aspect of electronic noise. It can be produced by the sensor and circuitry of a scanner or digital camera. Image noise can also originate in film grain and in the unavoidable shot noise of an ideal photon detector. Image noise is an undesirable by-product of image capture that adds spurious and extraneous information.

Fat-tail distributed or "impulsive" noise is sometimes called salt-and-pepper noise or spike noise. An image containing salt-and-pepper noise will have dark pixels in bright regions and bright pixels in dark regions. This type of noise can be caused by analog-to-digital converter errors, bit errors in transmission, etc. It can be mostly eliminated by using dark frame subtraction and interpolating around dark/bright pixels.

In this equation, if \( f_a \) or \( f_b \) is zero, we have unipolar impulse noise. If both are nonzero and almost equal, it is called salt-and-pepper noise. Impulsive noises can be
positive and / or negative. It is often very large and can
go out of the range of the image. It appears as black and
white dots, or saturated peaks.

Amplifier noise is a major part of the "read noise" of
an image sensor, that is, of the constant noise level in
dark areas of the image. In color cameras where more
amplification is used in the blue color channel than in the
green or red channel, there can be more noise in the blue
channel

\[ I_p(x, y) = I(x, y) + V(x, y) \]  

(16)

where Iv is the observed image with noise, I is the true
signal (image) and V is the noise component. Many
additive noise models exist and the following are some
common additive noise models with their Probability
Density Function (PDF) [18]. Gaussian noise is most
commonly used as additive white noise to yield additive
white Gaussian noise.

\[ f(x) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]  

(17)

where the parameter \( \mu \) is called the mean and it determines
the location of the peak of the density function, parameter
\( \sigma \) is called standard deviation and \( \sigma^2 \) is variance of the
distribution.

Uniform noise is not often encountered in real-world
imaging systems, but provides a useful comparison with
Gaussian noise. The PDF of uniform distribution is given
by

\[ f(x) = \begin{cases} 
\frac{1}{\sigma^2 \sqrt{3}} & \text{for } |x| \leq \sigma \sqrt{3} \\
0 & \text{else} 
\end{cases} \]  

(18)

Distance Measures:

**L1 & L2 Norm:**

L2 Norm measures the Euclidian distance between the
weight vector values of the query and the database
images.

Formula for the L2 norm is:

\[ W_2 = (\text{abs(img_query(w) - img_database(w)))}^2 \]

Performance Measure: In the CBIR performance measure
is calculated in this work was universal methods Precision
& Recall.

\[ \text{Precision} = \frac{\text{No of relevant images retrieved}}{\text{Total no of images retrieved}} \]

\[ \text{Recall} = \frac{\text{No of relevant images retrieved}}{\text{Total no of images in the database}} \]

**Experiments:** The proposed image retrieval method
(CRSFS) was applied Caltech 101 database datasets
images in the size 120x120 pixel resolution. Totally four
experiments conducted. Each with different query
environments like normal, rotated, scaled and noisy and
performance was evaluated and represent in Tables
(1, 2, 3, 4). The two types of distance measure function
was used in our proposed method is L1 & L2 performance
was evaluated and represent in Table (5). The proposed
method overall performance was compared to other
existing traditional methods in Fig 7.

**Situation-1 :** In this situation performance measured in
different rotated angle in query image. It’s applied in 8
datasets combined with classical & texture images.
The selected Query rotated angles was 30°, 60°, 90° and
120°. See Table (1). It obtained percentage is around
97.8%.

**Situation - 2:** In this situation performance measured in
different scaling factor in query image. It’s applied in 8
datasets combined with classical & texture images.
The selected Query scaling factor was 30x30, 60x60 and
120x120. See Table (2). It obtained percentage is around
98.3%.

**Situation-3:** In this situation performance measured in
different noise information added to query image like salt
& pepper, additive and uniform noise in query image. It’s
applied in 8 datasets combined with classical & texture
images. See Table (3). It obtained percentage is around
98.5%.

**Situation-4:** In this situation performance measured in
standard image used as query image. It’s applied in 8
datasets combined with classical & texture
images. See Table (4). It obtained percentage is around
98.8%.
Fig. 3: Caltech 101 database Images

Table 1: Recognition Rate for Query in Different Angles

<table>
<thead>
<tr>
<th>Image Set</th>
<th>Angle 30°</th>
<th>Angle 60°</th>
<th>Angle 90°</th>
<th>Angle 120°</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical Image</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set 1</td>
<td>91.5</td>
<td>93.3</td>
<td>97.8</td>
<td>94.7</td>
</tr>
<tr>
<td>Set 2</td>
<td>96.4</td>
<td>90.4</td>
<td>88.5</td>
<td>94.5</td>
</tr>
<tr>
<td>Set 3</td>
<td>95.2</td>
<td>92.4</td>
<td>90.9</td>
<td>96.7</td>
</tr>
<tr>
<td>Set 4</td>
<td>89.6</td>
<td>95.3</td>
<td>92.7</td>
<td>88.3</td>
</tr>
<tr>
<td>Set 5</td>
<td>95.6</td>
<td>91.5</td>
<td>97.2</td>
<td>89.9</td>
</tr>
<tr>
<td>Texture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set 6</td>
<td>97.7</td>
<td>88.4</td>
<td>93.7</td>
<td>95.2</td>
</tr>
<tr>
<td>Set 7</td>
<td>96.8</td>
<td>95.2</td>
<td>92.7</td>
<td>94.1</td>
</tr>
<tr>
<td>Set 8</td>
<td>89.5</td>
<td>96.2</td>
<td>93.8</td>
<td>97.1</td>
</tr>
</tbody>
</table>

Table 2: Recognition Rate for Query in Different Scaling Factor

<table>
<thead>
<tr>
<th>Image Set</th>
<th>Scaling Factor 30x30</th>
<th>Scaling Factor 60x60</th>
<th>Scaling Factor 90x90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical Image</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set 1</td>
<td>95.4</td>
<td>91.5</td>
<td>98.1</td>
</tr>
<tr>
<td>Set 2</td>
<td>93.2</td>
<td>97.5</td>
<td>96.3</td>
</tr>
<tr>
<td>Set 3</td>
<td>97.8</td>
<td>88.3</td>
<td>89.4</td>
</tr>
<tr>
<td>Set 4</td>
<td>96.4</td>
<td>92.8</td>
<td>98.2</td>
</tr>
<tr>
<td>Set 5</td>
<td>97.1</td>
<td>91.3</td>
<td>94.7</td>
</tr>
<tr>
<td>Texture</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set 6</td>
<td><strong>98.3</strong></td>
<td>95.5</td>
<td>88.9</td>
</tr>
<tr>
<td>Set 7</td>
<td>88.5</td>
<td>96.7</td>
<td>91.9</td>
</tr>
<tr>
<td>Set 8</td>
<td>89.4</td>
<td>87.2</td>
<td>96.2</td>
</tr>
</tbody>
</table>

Table 3: Recognition Rate for Query in Normal

<table>
<thead>
<tr>
<th>Image Sets</th>
<th>Recognition Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1</td>
<td>98.3</td>
</tr>
<tr>
<td>Set 2</td>
<td>97.8</td>
</tr>
<tr>
<td>Set 3</td>
<td>98.1</td>
</tr>
<tr>
<td>Set 4</td>
<td>98.6</td>
</tr>
<tr>
<td>Set 5</td>
<td>97.3</td>
</tr>
<tr>
<td>Set 6</td>
<td><strong>98.8</strong></td>
</tr>
<tr>
<td>Set 7</td>
<td>97.4</td>
</tr>
<tr>
<td>Set 8</td>
<td>97.8</td>
</tr>
</tbody>
</table>

DISCUSSIONS

Image retrieval completed using our proposed method Combined Radon Space Features Set. This proposed methods recognition rate is compared with the mention literatures & references [2, 4, 6, 9, 10]. This is represent our proposed combined feature space matrix is very effectively represent the image features. From the different situation conducted in different query environments it appears that the performance is improved with the increased number of related images retrieved followed by query image. Table [1–4] represented the different query types used and achieved the better recognition rate.
Table 4: Recognition Rate for Query in Different Noises

<table>
<thead>
<tr>
<th>Noise Types</th>
<th>CRSFS</th>
<th>PCA</th>
<th>RT</th>
<th>LDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salt &amp; Pepper Noise (Variance = 0.05)</td>
<td>98.5</td>
<td>66.9</td>
<td>96.45</td>
<td>94.725</td>
</tr>
<tr>
<td>Salt &amp; Pepper Noise (Variance = 0.1)</td>
<td>94.45</td>
<td>64.7</td>
<td>92.538</td>
<td>90.7</td>
</tr>
<tr>
<td>Salt &amp; Pepper Noise (Variance = 0.15)</td>
<td>95.25</td>
<td>60.025</td>
<td>96.34</td>
<td>94.725</td>
</tr>
<tr>
<td>Salt &amp; Pepper Noise (Variance = 0.2)</td>
<td>85.525</td>
<td>56.775</td>
<td>81.67</td>
<td>76.725</td>
</tr>
<tr>
<td>Additive Noise (mean = 0.05, variance =0.05)</td>
<td>91.5</td>
<td>60.2</td>
<td>86.45</td>
<td>84.45</td>
</tr>
<tr>
<td>Additive Noise (mean = 0.1, variance =0.1)</td>
<td>68.258</td>
<td>44.75</td>
<td>61.56</td>
<td>55.575</td>
</tr>
<tr>
<td>Additive Noise (mean = 0.15, variance =0.15)</td>
<td>46.725</td>
<td>34.375</td>
<td>40.34</td>
<td>43</td>
</tr>
<tr>
<td>Additive Noise (mean = 0.2, variance =0.2)</td>
<td>33.58</td>
<td>22.625</td>
<td>30.25</td>
<td>26.95</td>
</tr>
<tr>
<td>Uniform Noise (Speckle Variance =0.05)</td>
<td>98.625</td>
<td>68.5</td>
<td>97.5</td>
<td>96.925</td>
</tr>
<tr>
<td>Uniform Noise (Speckle Variance =0.1)</td>
<td>97.924</td>
<td>66.125</td>
<td>96.125</td>
<td>94.175</td>
</tr>
<tr>
<td>Uniform Noise (Speckle Variance =0.15)</td>
<td>92.350</td>
<td>63.725</td>
<td>89.55</td>
<td>90.2</td>
</tr>
<tr>
<td>Uniform Noise (Speckle Variance = 0.2)</td>
<td>86.540</td>
<td>61.575</td>
<td>84.875</td>
<td>83.625</td>
</tr>
</tbody>
</table>

Table 5: Recognition Rate for L1 & L2 Distance Measures

<table>
<thead>
<tr>
<th>Image Set</th>
<th>Distance Measures</th>
<th>L1</th>
<th>L2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical Image</td>
<td>Set 1</td>
<td>0.25</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>Set 2</td>
<td>0.31</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>Set 3</td>
<td>0.18</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Set 4</td>
<td>0.29</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Set 5</td>
<td>0.45</td>
<td>0.88</td>
</tr>
<tr>
<td>Texture</td>
<td>Set 6</td>
<td>0.13</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>Set 7</td>
<td>0.17</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>Set 8</td>
<td>0.21</td>
<td>0.73</td>
</tr>
</tbody>
</table>

The enhanced Radon Transform achieved in this method better than related works stated in the literature review section 2.

REFERENCES


