
Tohid Sedghi*

Department of Electrical Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran.

ABSTRACT

In CBIR, local features play a significant role in determining the similarity of images along with the shape information of the objects so this reason encourage authors to apply local features in literatures. One of the famous shapes features are one-dimensional normalized contour sequence moments which extracted from Gradient Vector Flow fields are used to compute the edge image, which will capture the object shape information. GVF fields give excellent results in determining the object boundaries irrespective of the concavities involved. In addition, Other Shape feature such as Hu Moments and edge histogram are not as efficient as GVF Fields because in mentioned shape feature extraction methods omits useful information which decrease retrieval percentage.

INTRODUCTION

It is introduced efficient and simple modified CBIR system operation and detailed study of each block will be discussed in literature [1,2,3]. We have proposed two new methods for image retrieval novel textural feature extraction in transformed domain. Then newly proposed features are extracted from these tiles. Features drawn transformed image using Spectral function from each of non-overlapping tiles computed Norm-1 energy and standard deviation in each diamond shape region of FAM algorithm, serve as texture descriptors. An integrated matching scheme based on most significant highest priority principle and adjacency matrix of a bipartite graph constructed between image tiles, is implemented for image similarity. Gradient vector flow fields are used to extract shape of objects. Invariant moments are used to describe the shape features. A combination of these texture and shape features provides a robust feature set for image retrieval.

METHODS AND MATERIALS

The SCF for a discrete-time real-valued signal $x(n)$ is defined as the Fourier-series transform of the cyclic correlation function thus $S_x^\alpha(f)$ is the cross spectrum of the pair of complex valued frequency-shifted signals $x(n) e^{-j\alpha n}$ and $x(n) e^{j\alpha n}$ where $f$ is the cross spectrum frequency variable and the parameter $\alpha$, called the cyclic frequency, is the relative frequency shift. If the signal has finite average power then there are at most a countable number of values of $\alpha$ for which $S_x^\alpha(f) \neq 0$. For $\alpha = 0$ the SCF reduces to the ordinary power spectrum:
Several computationally efficient algorithms for cyclic spectral analysis have been introduced that generally fall into two categories: those that average in frequency (frequency smoothing methods) and those that average in time (time smoothing methods). In this paper we utilize one efficient algorithm from time smoothing category namely FAM. We explain only the fundamentals of time smoothing algorithms that tends to FAM method. All time smoothing algorithms are based on the time smoothed cyclic cross periodogram:

\[
S_x^\alpha (f) = S_x^0 = \sum_k R^0(k)e^{-j2\pi fk} \quad (1)
\]

In (8), the operation < . > is the time averaging operation defined as follows:

\[
< \cdot, > = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} \cdot dt \to \infty \quad (3)
\]

and, the operation * is conjugate operator.

The physical interpretation of time smoothed cyclic periodogram is the correlation spectral components of \( x(n) \) over a time span of \( \Delta t \) seconds. The spectral components \( X_T(n, f + \alpha / 2) \) and \( X_T(n, f - \alpha / 2) \) are the complex envelopes of narrow-band, bandpass components of a signal and called complex demodulates. For the calculating of \( S_x^{\alpha}(n, f) \), a data tapering window of length \( T \) slides over the data for a time span of \( \Delta t \). At each instant the complex demodulates of the data within the window are computed. After calculation of the complex demodulates, they are correlated to produce estimates of the cyclic spectrum function. Correlation is done by time averaging of conjugate products over an interval of \( \Delta t \) seconds. To estimate the spectrum at the point \( (f_0, \alpha_0) \), demodulates separated in frequency by an amount \( \alpha_0 \) and centered about a midpoint of \( f_0 \) are correlated. The quantity \( \Delta t \) and \( \Delta f \) are referred to as the time and frequency resolutions of the point estimate. Mathematically, computation of the complex demodulates is expressed as:

\[
X_T(n, f) = \sum_{n=-N/2}^{N/2} a(r) x(n-r)e^{-j2\pi(n-r)T} \quad (4)
\]

where \( a(r) \) is a data tapering window of length \( T = N'Ts \). Then, the complex demodulates are correlated over a time span of \( \Delta t \) seconds:

\[
S_x^{\alpha_0}(n,f_0) = \sum_r X_T(r, f_1)X_T^*(r, f_2)g(n-r) \quad (5)
\]

where \( g(n) \) is a data tapering window of width \( \Delta t = NTs \); \( f_1 = f_0 + \alpha / 2 \) and \( f_2 = f_0 - \alpha / 2 \). It is shown in that the time smoothed cyclic periodogram converges to the cyclic cross spectrum in the limit, as \( \Delta t \to \infty \) followed by \( \Delta f \to 0 \), if the time windows \( a(n) \) and \( g(n) \) are properly normalized. Therefore, if \( \sum_n a^2(n) = \sum_n g(n) = 1 \)

Then,

\[
\lim_{\Delta t \to 0} \lim_{\Delta f \to \infty} S_x^{\alpha_0}(n,f_0) = S_x^{\alpha_0}(f_0) \quad (6)
\]

In the FAM algorithm, time smoothing is done by Fourier transform. If frequency is shifted from \( \alpha_0 \) to \( \alpha_0 + \Delta \alpha \), the output of the system is given by:

\[
S_x^{\alpha_0 + \Delta \alpha}(n,f_0) = \sum_r X_T(r, f_1)X_T^*(r, f_2)g(n-r)e^{-j2\pi\Delta \alpha T} \quad (7)
\]

If several values of \( \varepsilon \) are desired, evaluation of the sum can be simplified by discretizing the values of \( = q\Delta \alpha \). In this case the output of the algorithm is expressed as:
In which the sum can be evaluated with an N-point FFT. Thus, point estimates with constant cycle frequency can be computed in blocks by Fourier transforming the product sequences instead of averaging the product sequences individually. For complete coverage of the bifrequency plane a bank of bandpass filters is required to produce the necessary complex demodulates. An efficient method for producing the required complex demodulates is based on a sliding FFT. In this approach the frequencies of the filter bank are discretized to:

\[ f_k = k(f_s / N'), k = -N'/2 \ldots (N'/2) - 1 \] (9)

The locations of SCF estimation associated with the pairs of complex demodulates are \((f_j, \alpha_i)\) where the frequency coordinates are:

\[ f_j = \frac{f_k + f_i}{2} = \frac{k+1}{2} \left(\frac{f_s}{N'}\right) \] (10)

And the cycle frequency coordinates are:

\[ \alpha_i = f_k - f_i = (k-1)\left(\frac{f_s}{N'}\right) \] (11)

Fig. 5 shows tiling the bifrequency plane with the FFT accumulation method for \(N' = 8\) and the locations of SCF estimation. For an \(N'\)-point channelizer there are \((N')^2\) possible combinations of channelizer streams; hence, there are at most \((N')^2\) estimation regions (diamond regions). Due to symmetry, estimation of the cyclic spectrum of a single real signal requires only \((N')^2 / 4\) diamond region. After studying about SCF, First we partition an image into nine regions. Two, one dimensional signals are obtained from each of partition of partitioned image by ordering of pixels row by row and column by column. Then the SCF of each signal is calculated by FAM algorithm. For constructing the feature vector, the Norm-1 Energy and standard deviation are computed separately on each diamond shape region of SCFs from each signal and then feature vector is formed using these two parameter values. The basic assumption of using energy as a feature for texture discrimination is that the energy distribution in the frequency-domain identifies a textural information of an image. Resulting feature vectors from N number of regions from each partition of nine partition are as follows. where \(E_{in_j}, i = 1, 2, \ldots, N\) and \(\sigma_{in_j}, i = 1, 2, \ldots, N\), \(j = 1, 2, \ldots, N\) is the number of diamond regions at FAM algorithm and j is the number of partitions. Respectively, \(E_{in_j}\) and \(\sigma_{in_j}\) are the Norm-1 energy and standard deviation of the SCF amplitude that are calculated for the ith feature vector at the nth region namely D:

\[ E_{\text{norm}} = \frac{1}{Q} \sum_{\alpha \in \mathbb{D}} |S_{\alpha}^q(f)| \] (12)

\[ \sigma = \left( \frac{1}{Q} \sum_{\alpha \in \mathbb{D}} (S_{\alpha}^q(f) - \overline{S_{\alpha}^q(f)})^2 \right)^{1/2} \] (13)

Where Q and \(\overline{S_{\alpha}}(f)\) are the number of estimate points and the mean of spectral correlation function at corresponding D, respectively. At last for producing the main feature vector, we combine nine feature vector of each partition as follows to get the main feature:

**EXPERIMENTAL RESULTS**

In this section, the performance of system is evaluated using image set. There are two sets of images \(1,2, \ldots, N\) and \(\sigma_{in_j}\) are calculated for the corresponding D, respectively. At last for producing the main feature vector, we combine nine feature vector of each partition as follows to get the main feature.
opposite to $I_i^q$. Table 1 shows Comparison of average precision (%) with proposed method and other standard retrieval systems. If $I_i^q$ exists among the L database images, we say the system has correctly found the expected image. Otherwise, the system has failed to find the expected image. In the following, the accuracy rate of replying a query will be explained with accuracy (ACC, %). Table 2 shows comparison of accuracy (ACC%) of retrieved image

The experiment is to compare the retrieval accuracy of our system approaches.

Table 1: Comparison of average precision (%) with proposed method and other standard retrieval systems

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>48%</td>
<td>30%</td>
<td>55%</td>
</tr>
<tr>
<td>Beaches</td>
<td>32%</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td>Building</td>
<td>35%</td>
<td>25%</td>
<td>40%</td>
</tr>
<tr>
<td>Bus</td>
<td>36%</td>
<td>26%</td>
<td>68%</td>
</tr>
<tr>
<td>Dinosaur</td>
<td>95%</td>
<td>90%</td>
<td>99%</td>
</tr>
<tr>
<td>Elephant</td>
<td>38%</td>
<td>36%</td>
<td>62%</td>
</tr>
<tr>
<td>Flower</td>
<td>42%</td>
<td>40%</td>
<td>69%</td>
</tr>
<tr>
<td>Horses</td>
<td>72%</td>
<td>38%</td>
<td>77%</td>
</tr>
<tr>
<td>Mountain</td>
<td>35%</td>
<td>25%</td>
<td>45%</td>
</tr>
<tr>
<td>Food</td>
<td>38%</td>
<td>20%</td>
<td>55%</td>
</tr>
<tr>
<td>Mean</td>
<td>47.1%</td>
<td>36%</td>
<td>60.7%</td>
</tr>
</tbody>
</table>

Table 2: Comparison of accuracy (ACC%) of retrieved image

<table>
<thead>
<tr>
<th>L</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method [5]</td>
<td>62.4</td>
<td>70.7</td>
<td>74.8</td>
<td>76.6</td>
<td>79.0</td>
<td>84.0</td>
<td>87.7</td>
<td>90.2</td>
<td>92.3</td>
</tr>
<tr>
<td>Method [6]</td>
<td>85.5</td>
<td>90.5</td>
<td>92.3</td>
<td>93.2</td>
<td>93.6</td>
<td>95.1</td>
<td>96.7</td>
<td>97.7</td>
<td>98.9</td>
</tr>
<tr>
<td>Present Method</td>
<td>88.3</td>
<td>92</td>
<td>93.9</td>
<td>94.4</td>
<td>94.6</td>
<td>95.7</td>
<td>97</td>
<td>98.1</td>
<td>99</td>
</tr>
</tbody>
</table>

CONCLUSION

Local information is captured for two resolutions and two grid layouts that provide different details of the same image. Shape information is captured in terms of edge images computed using Gradient Vector Flow fields. Invariant moments are then used to record the shape features. The combination of the texture features between image and the shape features provide a robust feature set for image retrieval. It is this that makes texture a particularly interesting facet of images and results in a plethora of ways of extracting texture features. To enable us to explore different approaches for computing texture feature.

REFERENCES