International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013

Performance Comparison of Cosine, Haar, Walsh-Hadamard, Fourier and Wavelet Transform for shape based image retrieval using Fuzzy Similarity Measure

Alina Banerjee\textsuperscript{a, \ast}, Ambar Dutta\textsuperscript{b, \ast}
\textsuperscript{a}Regent Education and Research Foundation, Barrackpore, Kolkata 700121
\textsuperscript{b}Birla Institute of Technology, Mesra, Kolkata Campus, 4\textsuperscript{th} floor, 1582 Rajdanga Main Road, Kolkata 700107

Abstract

Shape is one of the most important features in Content Based Image Retrieval (CBIR). When a shape is used as feature, edge detection might be the first step of feature extraction. Invariance to the different transformations like translation, rotation, and scale is required by a good shape representation. In this paper a performance comparison is done on various image transforms like Wavelet transform, Fourier transform, Haar transform, Walsh-Hadamard transform and discrete cosine transform using a fuzzy similarity measure. It is seen that according to retrieval performance Wavelet transform gives the best result among the other mentioned transforms. It has higher recall and precision values and higher crossover point.

© 2013 The Authors. Published by Elsevier Ltd.
Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering

Keywords: Fuzzy similarity measure; Wavelet transform; Fourier transform; Haar transform; Walsh-Hadamard transform; discrete cosine transform

1. Introduction

Content based image retrieval came into existence in the early 90s making use of the low level features of images like texture, color, shape and structure. Extensive research has been done to extract the features of images and measure the similarity between images so that relevant image can be retrieved. In CBIR, images are indexed by features extracted from the image itself. Here manual annotations are not needed. Information retrieved from images is rather low level like color, texture, shape, structures and their combinations. The main focus of CBIR is the...
extraction of the image features and the calculation of similarity measure between the features. The representation of
the image contents can be compiled into a feature vector of \( d \) dimensions. It can be represented as follows:

\[
\mathbf{f}^l = (f_1^l, f_2^l, f_3^l, \ldots, f_d^l) \tag{I}
\]

The dimensionality \( d \) directly affects the query image processing. For non indexed query processing it takes
\( O((Nd)^2) \) number of computations to compare each element of the vector for \( N \) number of images. A good feature
should have sufficient power to discriminate between similar and dissimilar images. It should also be invariant to
spatial transformations such as rotation, translation and minor changes due to image capturing environment.

2. Related Works

Shape based image retrieval has been a very interesting area of research. The most important part of shape based
retrieval is the extraction of the shape descriptors which are hence compared for image retrieval. Zhang et. al [1]
gave a comparative study on shape matching with various signatures of Fourier transform. They concluded saying
that centroid distance Fourier descriptors gave the best performance. Zhang et. al. [2] gave a review of shape
representation and description techniques. Xiaojun Qi [3] used the concept of wavelet transform modulus maxima
for extracting the contour descriptors. Guru et. al [4, 5] described a new similarity measure for matching two
dimensional shapes. Banerjee et.al [6] described a fuzzy matching scheme on Fourier descriptors. This literature

3. Image Transform and Fuzzy Similarity Measure

A transform maps image data into a different mathematical space via a transformation equation. One example of
transform is the transform from one color space to another color space-RGB to SCT (spherical coordinate
transform), RGB to HSL (hue/saturation/lightness). However, the transform from one color space to another color
space has a one-to-one correspondence between a pixel in the input and the output. But the transforms used in this
literature map the image data from the spatial domain to frequency domain (spectral domain). All the pixels in the
input (spatial domain) contribute to each value in the output (frequency domain). Discrete transforms are performed
based on specific functions, which are called the basis functions. These functions are typically sinusoidal or
rectangular. The discrete versions of 1-D basis function are called basis vectors. The discrete versions of 2-D basis
function are called basis images (or basis matrices). The process of transforming the image data into another domain
involves projecting the image onto the basis images.

Given a query image \( Q \) the boundary points are calculated and an image transform is applied on the boundary
points. The absolute values and the phase values are found out. It is observed that the obtained feature values of \( Q \)
are simply a vector of real values instead of a vector of interval values as we have only one instance of \( Q \).

\[
\mathbf{F}_Q = [f_{Q1}, f_{Q2}, f_{Q3}, f_{Q4}, \ldots, f_{Qn}] \tag{II}
\]

where the \( k^{th} \) feature is given as follows:

\[
f_{Qk} = \{\eta_{Qk}, \mu_{Qk}\} \tag{III}
\]

Here it can be seen that the \( k^{th} \) feature is multi-valued rather than multi-interval valued [4, 5]. The feature vector of
the model images in the database is given as \( \mathbf{F}_M = [f_{M1}, f_{M2}, f_{M3}, f_{M4}, \ldots, f_{Mn}] \) where the \( k^{th} \) feature is given
as \( f_{Mk} = [\eta_{Mk}, \mu_{Mk}, \mu_{Mk}^{-}] \) which is of multi-interval valued type [4, 5]. The degree of similarity between \( Q \)
and \( M \) is defined as the average of the degree of similarities between their respective features.

The similarity between query image \( Q \) and model image \( M \) is given as follows [4,5]:

\[
\text{Sim}(f_{Qk}, f_{Mk}) = \frac{1}{l} \sum_{i=1}^{l} V_i \tag{IV}
\]

where \( V_i \) is given as follows [4, 5]:
\[
V_i = \max \left\{ \frac{1}{1 + |q_{i} - m_{i}|^\beta}, \frac{1}{1 + |m_{i} - q_{i}|^\beta} \right\} \quad \text{if} \quad m_{i} \leq q_{i} \leq m_{i}^* \\
1 \quad \text{otherwise}
\] (V)

It is seen that if \[\left[ \mu_{M}^{k-}, \mu_{M}^{k+} \right] \] interval contains \[\left[ \mu_{Q}^{k}, \mu_{Q}^{k} \right] \] then the similarity value is evaluated as 1. But if it does not lie in that interval the value is checked to which extent it is closer to the upper bound or the lower bound. Also it is difficult to say that which association of the boundary points is the best. So to get best association, cyclic shifts are carried out on the model shapes and the maximum value of similarity is selected from them.

4. Experimental Results

In this section, we compared the performance of shape based image retrieval using different image transforms. The experiments were conducted on a machine with Pentium 4 processor with 512 MB RAM and clock speed of 1.83 GHz.

The image database comprises of 5 types of images of 5 categories. The shape images are taken from [5]. The images in the database are shown in Fig 1. The average precision and recall graphs are plotted against number of images retrieved. The graphs are shown in Fig 2 and Fig 3.

![Representative images of the image database](image.png)
The average recall and average precision graphs given in Fig 2 and Fig 3 show that wavelet transform has higher recall and precision values than the other transforms. The crossover points plotted in Fig 4, a concept utilized by Kekre et.al. [7] also shows higher value for the wavelet transform depicted by the intersection of the pink and dark blue lines in the graph. It is observed that wavelet transform has the highest crossover point (0.59). Then in descending order of crossover points we have Walsh-Hadamard transform (0.54), Haar transform (0.5122), Fourier transform (0.49) and Discrete Cosine transform (0.44).

Fig 4 Crossover points

5. Conclusion

In this paper, a comparison is done between various transforms on a fuzzy similarity measure for retrieval of 2 dimensional shapes. It is seen that according to retrieval performance Wavelet transform gives the best result with respect to average precision and recall values among the transforms used in comparison. It has higher recall and precision values and higher crossover point. In future, the method could further be improved so as to achieve higher precision and recall values. This method is applicable to just 2D shapes. So it can be thought to incorporate it in 3D shape matching with some modification.

References