INTERNATIONAL JOURNAL OF COMPUTER ENGINEERING & TECHNOLOGY (IJCET)

ISSN 0976 – 6367(Print)
ISSN 0976 – 6375(Online)
Volume 6, Issue 2, February (2015), pp. 29-37
© IAEME: www.iaeme.com/IJCET.asp
Journal Impact Factor (2015): 8.9958 (Calculated by GISI)
www.jifactor.com

IJCET

©IAEME

IMAGE SIMILARITY USING FOURIER TRANSFORM

Siddharth Narayanan[1], P K Thirivikraman[2]

[1]Research Center Imarat, DRDO Hyderabad, India 500069
[2]Birla Institute of Technology & Science Hyderabad, India 500078

ABSTRACT

In this paper, a similarity measure for images based on values from their respective Fourier Transforms is proposed for image registration. The approach uses image content to generate signatures and is not based on image annotation and therefore does not require human assistance. It uses both, the real and complex components of the FFT to compute the final rank for measuring similarity. Any robust approach must accurately represent all objects in an image and depending on the size of the image data set, diverse techniques may need to be followed. This paper discusses implementation of a similarity rating scheme through the Open CV library and introduces a metric for comparison, carried out by considering Intersection bounds of a covariance matrix of two compared images with normalized values of the Magnitude and Phase spectrum. Sample results on a test collection are given along with data using existing methods of image histogram comparison. Results have shown that this method is particularly advantageous in images with varying degrees of lighting.

Keywords: Discrete Fourier Transform (DFT), Fast Fourier Transform (FFT), Feature extraction, Histogram Intersection, Image Signature

1. INTRODUCTION

Several approaches have been proposed to the problem of identifying image features that could be used as objective indicators of their contents. One set of approaches require explicit image annotation by humans to mark objects, regions and shapes. The other set contains methods for automatic extraction of physical properties like color and contour. These are computed through values and distributions that are used in image classification. Automatic feature extraction in many cases provides good performance in large databases and for applications where cost of image processing must be kept low. A specific area of application is retrieval by similarity in thematic databases. For retrieval by similarity, a query is made of one or more parts of an image. The process to procure similar images is based on some selected visual properties according to an appropriate metric. The content of a thematic database is specific to a certain domain. In this context, it is assumed that the thematic nature allows queries to be targeted to visual aspects rather than to content interpretation.

Two major approaches can be followed in extracting information about textures and lines direction in an image. The first is based on images segmented by an edge finding procedure while the second uses full colors or b/w multilevel images. This paper proposes an approach to the problem of finding the distribution of image lines direction by analyzing its Fourier transform. As noted by several authors, the 2D Fourier power spectrum preserves direction information of an image [1]. Once the Fourier transform is computed, its frequency domain representation can be scanned and required values generated.

This paper is organized as follows. In Section 2 a brief overview of the current literature is given. Section 3 discusses the use of the FFT algorithm as a means for matching similar images. Section 4 describes the implementation technique, and Section 5 presents the results of the retrieval technique based on similarity. Conclusions are given in Section 6 and references follow.

2. RESEARCH ELABORATIONS

2.1 IMAGE ANALYSIS

The large objects in an image are usually relevant for interpreting the image. These can be assumed to be most frequently localized in central areas like foreground objects or spanning an entire image like in the case of recurrent shapes or landscapes. These objects contribute to the low frequency information of the image [2]. Higher frequencies are mostly located in small details and fine-grain textures. In the frequency domain, components with continuous and low frequency tend to be located at the center. Higher frequencies that are related to fast variations in the image are located close to the border. Most of data is located near the origin. Image information is largely related to larger objects, rather than to details and continuous background.

2.2 RETRIEVAL SYSTEMS

Comparing two images, or an image and a model, is the fundamental operation for any retrieval systems [3].Retrieval by similarity refers to retrieving images which are similar to a previous image. This can also be called retrieval by example, or to a model or schema. Retrieval by similarity requires specific definitions of what similar means. Representation of features of images - like color, texture, shape, motion, etc. - is a fundamental problem in visual information retrieval. Algorithms for pattern recognition and image analysis provide the means to extract numeric descriptors thus giving us a quantitative measure of these features. Computer vision enables object and motion identification by comparing extracted patterns with predefined models. Coupling it with localization techniques would allow autonomous robots to globally localize them selves to reliably track their positions and recover from localization failures. Such a system would also be robust against distortion and occlusions.

Today, similarity queries arise naturally in a large variety of applications like electronic catalogues, Medical databases for ECG, X-ray and TAC, Weather prediction and Criminal investigations. The drawbacks of traditional approaches can be overcome through Similarity search by using numerical features computed by direct analysis of the information content. Content-Base Image Retrieval (CBIR) were proposed in the early 1990’s that use visual features to represent the image content. This approach allows features to be computed automatically, and the information used during the retrieval process is always consistent as it does not depend on human interpretation. A query image, or select a prototype image is taken by the user, searching for something similar. The result is a list of images sorted by increasing or decreasing values of similarity relative to the query image depending on the method of comparison. Retrieval is immediate, based on an appropriate predefined similarity criterion, and able to measure the grade of similarity between two images using only low level image properties (i.e. no human experts should provide additional information). Moreover, an efficient way to obtain the most similar DB images to the query has to be defined. At query time, the feature vector of the query image is computed and the DB is searched for the most similar feature vectors.

2.3 SIGNAL DISCRETIZATION

The time signal squared;$x^2(t)$ represents how the energy contained in the signal distributes over time, while its spectrum squared;$X^2(f)$ represents how the energy distributes over frequency (therefore the term power density spectrum). Obviously, the same amount of energy is contained in either time or frequency domain, as indicated by Parseval's formula (1).

$$\int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{-\infty}^{\infty} |X(f)|^2 \, df \qquad (1)$$

Signal discretization is used to process a given continuous physical signal and in Time/Frequency domain computer processing, the signal is digitized and stored in a digital computer to be later processed by any desired processing algorithm with maximum flexibility. It is therefore not a real time process. The processing can either be in the time domain or the frequency domain. For the latter, the Fast Fourier Transform algorithm is used to transform the data between the time and frequency domains. The physical signal is truncated and sampled before it can be further analyzed and processed numerically by a digital computer.

The DFT for an N by N matrix Wcan be written as (2).

$$X[n] = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} x[m] e^{-\frac{j2\pi mn}{N}} = \sum_{m=0}^{N-1} w_N^{mn} \, x[m] = \qquad (2)$$

$$\sum_{m=0}^{N-1} w[n,m] x[m] \quad (n = 0,1, \dots, N-1)$$

Similarly, (3) represents the inverse transform.

$$x[m] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X[n] e^{\frac{j2\pi mn}{N}} = \qquad (3)$$

$$\sum_{m=0}^{N-1} w^*[n,m] x[m] \quad (m = 0,1, \dots, N-1)$$

The computational complexity of DFT is (N^2). The Fast Fourier Transform algorithm allows the complexity to be much improved based onequations (4) to (6).

$$w_N^{kN} = e^{-\frac{j2kN\pi}{N}} = e^{-2kj\pi} \equiv 1 \qquad (4)$$

$$w_{2N}^{2k} = e^{-\frac{j2k2\pi}{2N}} = e^{-\frac{2kj\pi}{N}} \equiv w_N^k (5)$$

$$w_{2N}^{N} = e^{-\frac{j2N\pi}{2N}} = e^{-j\pi} \equiv -1 \qquad (6)$$

The complexity is herefore reduced from $(N^2)$to $(Nlog_2N)$. This major improvement of computational makes Fourier transform practically possible in many applications.

C code for FFT:

```
for (i=0; i<N; ++i) { // bit reversal
j=0;
for (k=0; k<m; ++k)
j=(j << 1) | (1 & (i >> k));
if (j < i)
{ SWAP(xr[i],xr[j]); SWAP(xi[i],xi[j]); }
}
for (i=0; i<m; i++) { // for log N stages
n=pow(2.0,(float)i); w=Pi/n; if (inverse) w=-w;
k=0;
while (k<N-1) { // for N components
for (j=0; j<n; j++) { // for each section
 c=cos(-j*w); s=sin(-j*w);
 j1=k+j;
tempr=xr[j1+n]*c-xi[j1+n]*s;
tempi=xi[j1+n]*c+xr[j1+n]*s;
                xr[j1+n]=xr[j1]-tempr;
                xi[j1+n]=xi[j1]-tempi;
                xr[j1]=xr[j1]+tempr;
xi[j1]=xi[j1]+tempi;
}
k+=2*n;
}
}
```

3. PROPOSED METHOD

Many methods for comparison currently exist like Correlation, Chi-Square, Bhattacharya Distance and Intersection of images to provide numerical parameters that express how well two image histograms match with each other. This paper proposes a new metric using their respective FFT to compare to given images for similarity. A function is used to get a numerical parameter that express how well two image histograms match with each other. This approach is useful also because it allows cancelling out and negation of values. When considering both the real and complex matrices of the FFT, negative components help to keep values low and makes for a much simpler metric system.

$$\frac{(\sum_{i=1}^{N} F_{1i}F_{2i} - N\overline{F}_1\overline{F}_2)^2}{(\sum_{i=1}^{N} |F_{1i}|^2 - N\overline{F_1}^2)(\sum_{i=1}^{N} |F_{2i}|^2 - N\overline{F_2}^2)} \qquad (7)$$

The metric for comparison is computed using the formula in (7). For each frequency, an intensity value is obtained from the real and complex parts of the Fourier Transform. A suitable image size $(512px \times 512px)$ is taken to maintain uniformity for generating required results.$F_{1i}$is the intensity value of i^{th} pixel of the first image while F_{2i} is the intensity value of i^{th} pixel of the second image. $\overline{F_1}$and$\overline{F_2}$ represent the average frequency values of each image. The input base image is loaded along with Images to be compared with it and all images are transformed into gray scale format. Convert the images to the CV_32F format to capture intensity values in float data type for accurate computation. The float values have a much wider range from 0 to 1. Histogram for all the images is calculated and normalized in order to compare them. The histogram of the base image is compared with respect to all other test histograms &numerical and matching parameters are obtained.

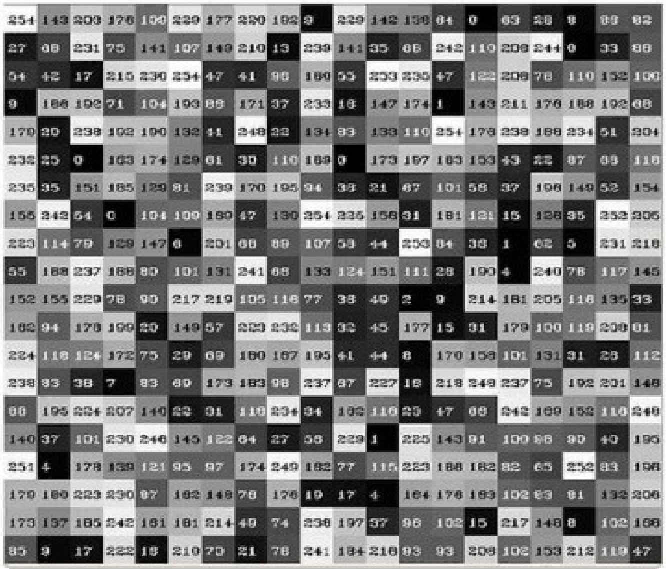

Fig 1: Values ranging from 0 to 255 for a 20X20 pixel image in CV_8U format. A similar mapping is generated in CV_32F.

4. IMPLEMENTATION

The proposed technique has been implemented using the Open CV library and the Microsoft Visual C++ Suite. The current code base has been configured to support gray scale images from a sample set of 300 images. The CV_8U format of the Open CV libraryis useful for displaying the final FFT and filtered results. Figure1 shows the intensity mapping for a sample image in this format [4]. While performing calculations during the study however, the restricted set of values of this format (i.e. intensity in range 0 - 255) did not provide very accurate results. The CV_32F format was then employed due to its wider range of allowed values.

The proposed metric is also compared to results of existing metrics to compute the matching of image histograms. The three techniques used for analysis in this study are –

Correlation:

$$d(P_1, P_2) = \frac{\sum_I (P_1(I)-\overline{P_1})(P_2(I)-\overline{P_2})}{\sqrt{\sum_I (P_1(I)-\overline{P_1})^2 \sum_I (P_2(I)-\overline{P_2})^2}} \quad (8)$$

Intersection:

$$d(P_1, P_2) = \sum_I Min(P_1(I), P_2(I)) \quad (9)$$

Bhattacharyya distance:

$$d(P_{1,}P_2) = \sqrt{1 - \frac{1}{\sqrt{\bar{P_1}\bar{P_2}N^2}} \sum_I \sqrt{P_1(I).P_2(I)}} \qquad (10)$$

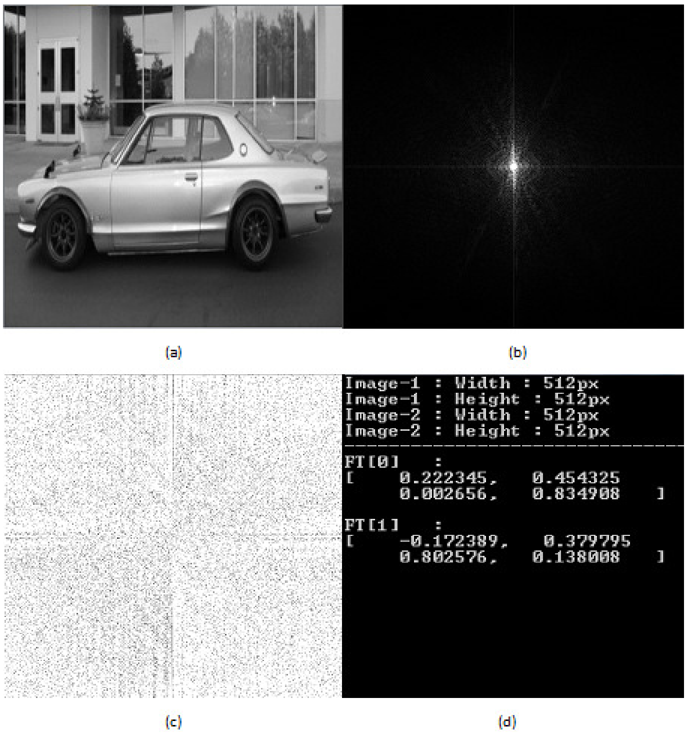

Fig2: (a) The resized image (b) The FFT function in OpenCV allows us to view the resulting Magnitude and (c) Phase spectrum images. (d) Output values for 2X2 matrix for real and complex components of the image

4.1 APPLICATION OF FILTERS

In the frequency domain all image components are represented as the sum of periodic functions characterized by different periods, centered on the zero frequency components. They therefore retain information only about the direction of image components. Also, image mirroring corresponds to mirroring of the spectrum, while image rotation rotates the spectrum by the same angle. In all three cases, noise and aliasing may be introduced due to the discrete nature of the image pixels. This in turn may add or exclude certain objects thus modifying the overall content of the image. Noise and aliasing can be eliminated through the application of filters before FFT computation.

4.2 COMPUTATION OF FOURIER TRANSFORM

The procedure discussed in this paper is based on the FFT computation and its interpretation. The computation of FFT is a basic operation in image and signal processing [5]. The total number of floating point operations needed for FFT computation of an N-point sequence is $Nlog_2N$.

Thus, for a$M \times N$ pixel image, the total number of operations required is $M \times N\ (log_2M + log_2N)$. Applying the same in the case of this study yields:

$512 \times 512\ (log_2 512 + log_2 512) = 4718592$ Operations

International Journal of Computer Engineering and Technology (IJCET), ISSN 0976-6367(Print), ISSN 0976 - 6375(Online), Volume 6, Issue 2, February (2015), pp. 29-37 © IAEME

5. RESULTS

The following images are taken as input images to test the proposed metric and the results are compared to existing methods of comparison using the OpenCV library. A perfect match is displayed on comparing the base image histogram with itself. Also, for the other two test images do not provide very good matches since they have very different lighting conditions. The numerical results of the study are presented in the following table

Table 1: The results of the comparative study are presented for all the four approaches implemented using C++ on OpenCV

Image	Correlation	Intersection	Bhattacharyya distance	**Rank - FFT**
Car 1	1	37.348704	0	**0.995642**
Car 2	0.065667	8.238745	0.782008	**0.455956**
Car 3	0.031694	1.203044	0.831347	**0.410324**
Car 4	0.009159	1.111309	0.896255	**-0.22937**
Car 5	-0.017561	2.113258	0.896092	**-2.1178**
Car 6	-0.0116	0.785085	0.927125	**-5.32439**
Car 7	0.011672	0.918962	0.914182	**0.264427**
Car 8	0.108535	4.70819	0.705022	**0.643458**
Car 9	0.016492	0.57914	0.908365	**0.32495**

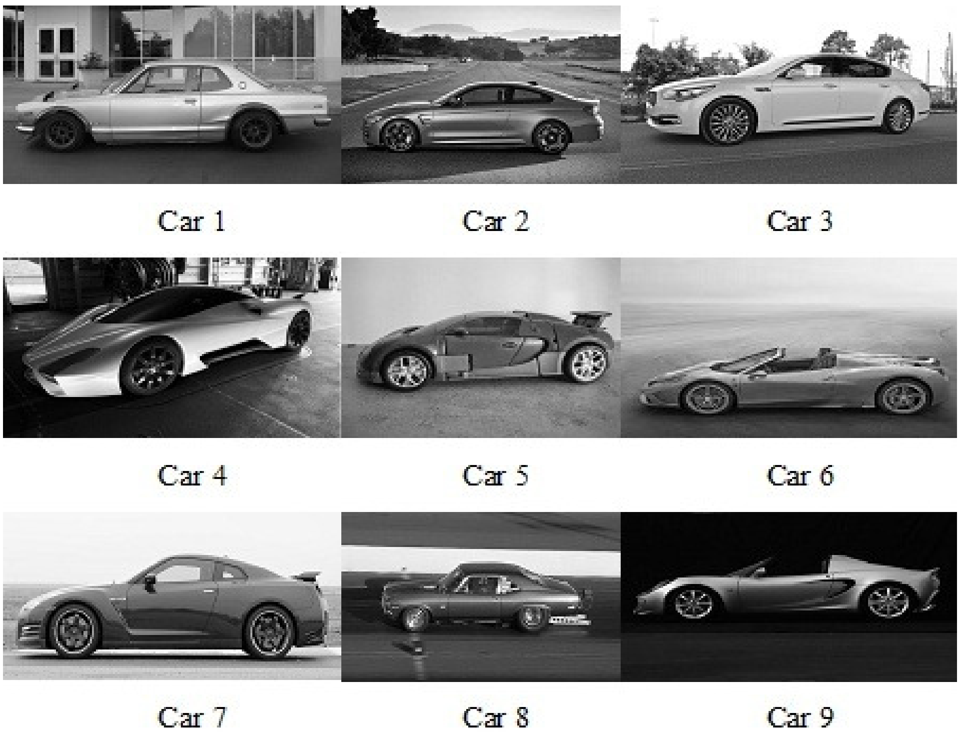

Fig3: Car 1 is selected as the base image and should yield the highest similarity measure with itself in all approaches.

For the Correlation and Intersection methods, the higher the metric, the more accurate the match while for the Bhattacharyya distance metric, the lesser the result, the better the match [6].As expected, the match of the base-base image is the highest and the matches of other images with respect to the base are comparatively worse. When collating the data for all images, this method was found especially robust in images with similar image components and differences in lighting effects. Example images presented in this study use cars in a particular orientation to highlight this feature. These images are first resized and padded for $(512px \times 512px)$ image size and provided as input parameters to the implemented FFT rank function.

5.1 TRANSLATION

In the frequency domain all the image components are represented as the sum of periodic functions characterized by different periods centered about the zero frequency component. Being centered in the origin, the image components do not retain any information about the original position, but only information about their direction. Image translation can be taken into account by calculating it in terms of the angular spectrum computed on the Fourier transform.

5.2 ROTATION AND MIRRORING

Image rotation corresponds to a horizontal shift of the spectrum, and image mirroring corresponds to mirroring of the spectrum. It is important to note that the discrete nature of the image pixels may introduce noise and aliasing, and that translation and rotations may modify the overall content of the image by adding or excluding some objects. Like Image translation, rotation and mirroring can also be accounted for by interpreting them in terms of the angular spectrum computed on the Fourier transform.

6. CONCLUSION

This paper discusses the use of intersection of normalized values of the Magnitude and Phase spectrum from FFT of two compared images as a means of comparing image content. The comparison is performed between a reference image, assumed as the query, and a collection of images. In order to be effective, the collection must be semantically homogeneous so that visual similarity can be a surrogate for content match. The database population can be executed in acceptable time using almost all programming environments. Results based on a sample test collection are provided along with a comparison with more assessed techniques. The approach is suitable for image retrieval application where the graphical content or layout bear most of the information. It was also found to be useful in applications which have different lighting. While, initial tests in specific domain like the analysis of human signatures and fingerprints are promising, it has to be compared with the safer and more assessed techniques to further refine the process.

REFERENCES

1. S. K. Mitra, J. F. Kaiser, Handbook for Digital Signal Processing, John Wiley & Sons, 1993Wachs, J.P., Kölsch, M., Stern, H., Edan, Y., Vision-Based Hand-Gesture Applications. Communications of the acm, 54 (2) (2011) 60-71.
2. Similarity Matching in Computer Vision and Multimedia NicuSebe (2008) Computer Vision and Image Understanding, online at www.sciencedirect.com [Accessed 22 January 2014]
3. A FFT based technique for image signature generation Augusto Celentanoa and 2014.aUniversitàCa' Foscari di Venezia,

4. OpenCV 2.4.9.0 documentation » OpenCV Tutorials » imgproc module. Image Processing » Histogram Calculation http://docs.opencv.org/doc/tutorials/imgproc/histograms/histogram_calculation/histogram_calculation.html [Accessed 4 November 2014].
5. fourier.eng.hmc.edu/e101/lectures/image_processing/node2.html [Accessed 14 November 2014].
6. OpenCV 2.4.9.0 documentation » OpenCV Tutorials » imgproc module. http://docs.opencv.org/doc/tutorials/imgproc/histograms/histogram_comparison/histogram_comparison.html [Accessed 4 December 2014].
7. C. Basavaraju and Dr. Chandrakanth.H.G, "FFT Based Spectrum Analysis Model For An Efficient Spectrum Sensing" International Journal of Advanced Research in Engineering & Technology (IJARET), Volume 5, Issue 12, 2014, pp. 87 - 96, ISSN Print: 0976-6480, ISSN Online: 0976-6499.
8. Dr. Hanaa m. A. Salman, "Information Hiding In Edge Location of Video Using Amalgamate Fft and Cubic Spline" International journal of Computer Engineering & Technology (IJCET), Volume 4, Issue 4, 2013, pp. 240 - 247, ISSN Print: 0976 – 6367, ISSN Online: 0976 – 6375.
9. Priyanka Chauhan and Girish Chandra Thakur, "Efficient Way of Image Encryption Using Generalized Weighted Fractional Fourier Transform with Double Random Phase Encoding" International Journal of Advanced Research in Engineering & Technology (IJARET), Volume 5, Issue 6, 2014, pp. 45 - 52, ISSN Print: 0976-6480, ISSN Online: 0976-6499.
10. Prof. Maher K. Mahmood and Jinan N. Shehab, "Image Encryption and Compression Based on Compressive Sensing and Chaos", International journal of Computer Engineering & Technology (IJCET), Volume 5, Issue 1, 2014, pp. 68 - 84, ISSN Print: 0976 – 6367, ISSN Online: 0976 – 6375.