

Research and Reviews: Journal of Engineering and Technology

A Machine Learning Approach to No-Reference Objective Video Quality Assessment for High Definition Resources.

Ruhbakhsh Yahoo*.

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

Research Article

Received: 13/07/2014
Revised: 12/08/2014
Accepted: 16/08/2014

*For Correspondence

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

Keywords: machine, image, High definition

ABSTRACT

Content based image retrieval, the problem of finding images from data base according to their content, has been the subject of a significant amount of research in the last decade. Image retrieval based on region is one of the most promising and active research directions in recent years. As literature prove that region segmentation will produce better results. Human visual perception is more effective than any machine vision systems for extracting semantic information from image, hitherto no specific system has been suggested with the ability of extracting object individually. We introduce a new idea, hence object detection has been obtained as main contribution of this paper and a new feature extraction based on spectral analysis is presented.

INTRODUCTION

Expectation Maximization (EM) algorithm is utilized to segment image into different regions. A new image representation which provides a transformation from the raw pixel data to a small set of image regions which are coherent in color and texture space is presented. In addition the EM algorithm performs automatic segmentation based on image features [1-6]. EM iteratively models the joint distribution of color and texture with a mixture of Gaussians. The resulting pixel cluster memberships provide a segmentation of the image. After the image is segmented into regions, system select the region where contain main object. More over a description of chosen region, based on novel feature extraction is produced [7]. The other option of the proposed system is that the user can access the regions directly in order to see the segmentation of the query image and specify which aspects of the image are important to the query. The deficiency of traditional retrieval systems is due to either both image representation and method of accessing those representations to find images, while users generally want to find images containing particular objects [15]. Most existing image retrieval systems represent images based only on their low-level features, with little regard for the spatial organization of those features. Systems based on user querying are often unintuitive and offer little help in understanding why certain images were returned and how to refine the query [13-15]. Often the user knows only that he has submitted a query for, say, a horse and retrieved very few pictures of horses in return. For general image collections, there are currently no systems that can automatically classify images or recognize the objects they contain. In particular, this letter demonstrates how the segmentation and new feature extraction can considerably enhance object based retrieval system.

METHODOLOGY

The EM algorithm is used for finding maximum likelihood parameter estimates when there is missing or incomplete data [1]. In our case, the missing data is the region to which the points in the feature space belong. We estimate values to fill in for the incomplete data (the "E-Step"), compute the maximum

likelihood parameter estimates using this data (the “M-Step”), and repeat until a suitable stopping criterion is reached. Based on essence of EM algorithm we can segment each image to different parts. Simulation and Fig.1 show, 3 Gaussian function will produce the best result to extract the main object of image. Choosing different number of Gaussian function will show worst segmentation results. After extracting object we apply a novel textural feature extraction method based on spectral function (SF). SF is fully studied in [15,3,5,6].

For feature extraction method, two, one dimensional signals are obtained from image by ordering of pixels row by row and column by column. Then the SCF of each signal is calculated by FAM algorithm [3]. 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 textural information of an image. Resulting feature vectors from N number of regions from each partition as follows.

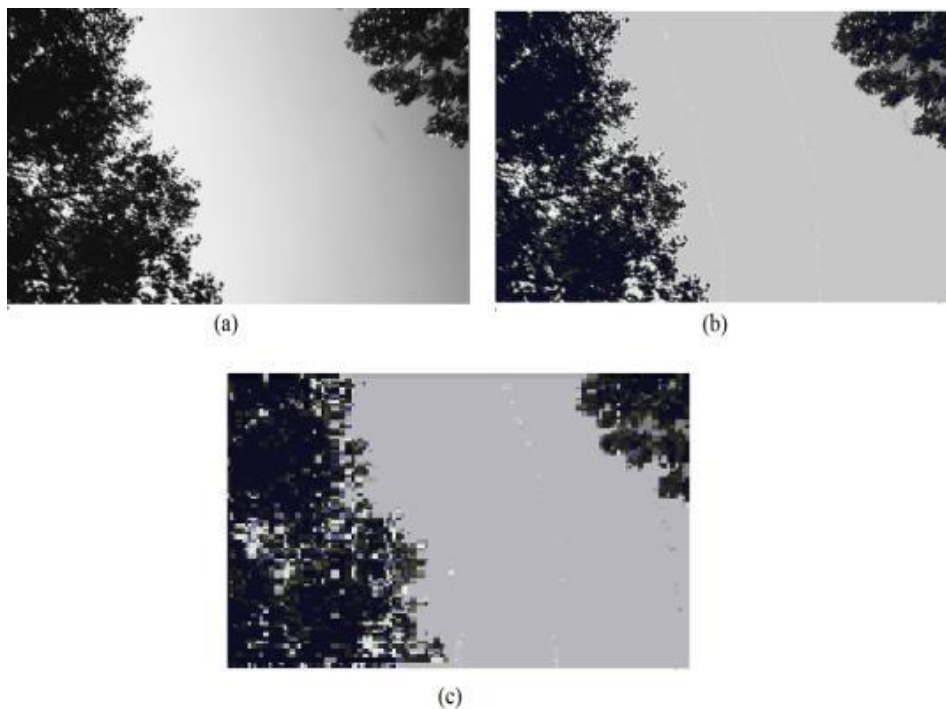


Figure 1: Result of Object extraction using EM Algorithm for Gaussian function

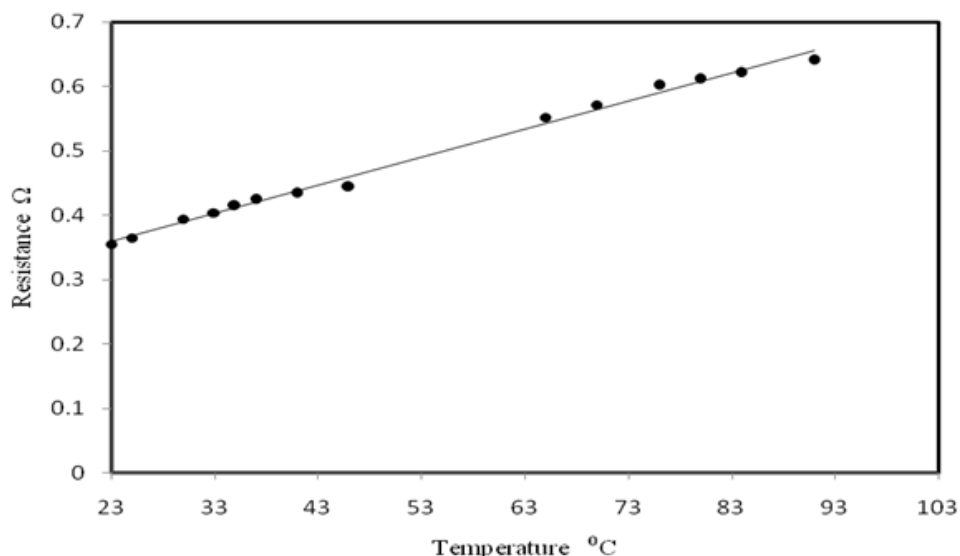


Figure 2: Average precision of different methods on Image set.

$$\begin{aligned}\bar{f}_j &= [E_{11}, \dots, E_{1N}; E_{21}, \dots, E_{2N}; \sigma_{11}, \dots, \sigma_{1N}; \sigma_{21}, \dots, \sigma_{2N}] \\ &= [f_1, f_2, \dots, f_{4N}]^T;\end{aligned}\quad (1)$$

Where E_{in} , $i\{1,2\}$, $n = 1,2,\dots,N$ and σ_{in} , $i\{1, 2\}$, $n = 1,2,\dots,N$, N is the number of diamond regions at FAM algorithm. Respectively, E_{in} and σ_{in} , are the Norm-1 energy and standard deviation of the SCF amplitude that are calculated for the i th signal at the n th region namely D :

$$E_{norml} = 1/Q \sum_{\alpha \& f \in D} |S_x^\alpha(f)| \quad (2)$$

$$\sigma = (1/Q \sum (S_x^\alpha(f) - \bar{S}_x^\alpha(f))^2)^{0.5} \quad (3)$$

Where Q and $\bar{S}_x^\alpha(f)$ are the number of estimate points and the mean of spectral function at corresponding D , respectively. Advantages of using cyclostationary analysis are the simplicity of generating features and feature vector dimension. Feature vector dimension is the one of the most important factors that determined the amount of storage space for the vector, the retrieval accuracy and retrieval time (or computational complexity). It will be shown that how simply we extract features from spectral function. The spectral function (SCF), which is the cross spectral of a signal and a frequency shifted version of itself, provides a second-order statistical description in the frequency domain of such signals. The theory and utility of cyclostationary signal models and the SCF, also called the cyclic spectrum, are discussed in references [8]. The basic time-smoothing and frequency-smoothing methods of spectral analysis were introduced in [8] and proof of their equivalence was given in [9,10]. Methods which more fully exploit the computational efficiency of the FFT, namely, the FFT accumulation method (FAM) and the strip spectral analyzer (SSCA) were introduced in [11] and discussed in.

SPECTRAL FUNCTION

The SCF for a discrete-time real-valued signal $x(n)$ is defined as the Fourier-series transform of the cyclic function $R_x^\alpha(k)$

$$S_x^\alpha(f) = \sum_{-\infty}^{\infty} R_x^\alpha(k) e^{-j2\pi f k} \quad (2)$$

Where

$$R_x^\alpha(k) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{-N}^N [x(n+k) e^{-j\pi(n+k)\alpha}] \cdot [x(n) e^{j\pi n \alpha}] \quad (3)$$

thus $S_x^\alpha(f)$ is the cross spectrum of the pair of complex valued frequency-shifted signals $x(n) e^{-j\pi n \alpha}$ and $x(n) e^{j\pi n \alpha}$ where f is the cross spectrum frequency variable and the parameter α , called the cyclic frequency, is the relative frequency shift [3][5][6]. If the signal has finite average power then there are at most a countable number of values of α for which $S_x^\alpha(f) \neq 0$. For $\alpha = 0$ the SCF reduces to the ordinary power spectrum:

$$S_x(f) = S_x^0 = \sum_k R_x^0(k) e^{-j2\pi f k} \quad (4)$$

The symmetry relationships and $S_x^\alpha(-f) = S_x^\alpha(f)$ and periodicity associated with discrete time for any integer m and n , $S_x^{\alpha+n}(f + m + n/2) = S_x^\alpha(f)$ which are easily verified from the defining expressions, imply that the entire function is determined by $S_x^\alpha(f)$ for $\{0 \leq f \leq 1/2, 0 \leq \alpha \leq 1 - 2f\}$ where f is the normalized frequency.

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. The precise discussion can be studied at [13]. All time smoothing algorithms are based on the time smoothed cyclic cross periodogram [10].

$$S_{x_T}^\alpha(n, f) = \frac{1}{T} \langle X_T(n, f + \frac{\alpha}{2}) X_T^*(n, f - \frac{\alpha}{2}) \rangle_{\Delta t} \quad (5)$$

In (8), the operation $\langle \cdot \rangle$ is the time averaging operation defined as follows:

$$(6) \langle \cdot \rangle \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} (\cdot) dt$$

and, the operation $*$ is conjugate operator.

The physical interpretation of time smoothed cyclic periodogram is the spectral components of $x(n)$ over a time span of Δt seconds. The spectral components $X_T(n, f + \frac{\alpha}{2})$ and $X_T^*(n, f - \frac{\alpha}{2})$ are the complex envelopes of narrow-band, bandpass components of a signal and called complex demodulates. For the calculating of $S_{x_T}^\alpha(n, f)_{\Delta t}$, a data tapering window of length T slides over the data for a time span of Δt 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. is done by time averaging of conjugate products over an interval of Δt seconds. To estimate the spectrum at the point (f_0, α_0) , demodulates separated in frequency by an amount α_0 and centered about a midpoint of f_0 are correlated. The quantity Δt and Δf are referred to as the time and frequency resolutions of the point estimate. Mathematically, computation of the complex demodulates is expressed as:

$$X_T(n, f) = \sum_{n=-\frac{N'}{2}}^{\frac{N'}{2}} a(r) x(n-r) e^{-j2\pi f(n-r)} \quad (7)$$

Where $a(r)$ is a data tapering window of length $T = NT_s$. Then, the complex demodulates are correlated over a time span of Δt seconds: $S_{x_T}^{\alpha_0}(n, f_0)_{\Delta t} = \sum_r X_T(r, f_1) X_T^*(r, f_2) g(n-r)$ (8)

Where $g(n)$ is a data tapering window of width $\Delta t = NT_s$; $f_1 = f_0 + \frac{\alpha}{2}$ and $f_2 = f_0 - \frac{\alpha}{2}$. It is shown in [16] that the time smoothed cyclic cross periodogram converges to the cyclic cross spectrum in the limit, as $\Delta t \rightarrow \infty$ followed by $\Delta f \rightarrow 0$, if the time windows $a(n)$ and $g(n)$ are properly normalized. Therefore, if $\sum_n a^2(n) = \sum g(n) = 1$ Then,

$$\lim_{\Delta f \rightarrow 0} \lim_{\Delta t \rightarrow \infty} S_{x_T}^{\alpha_0}(n, f_0)_{\Delta t} = S_X^{\alpha_0}(f_0) \quad (9)$$

In the FAM algorithm, time smoothing is done by Fourier transform. If frequency is shifted from α_0 to $\alpha_0 + \epsilon$, the output of the system is given by:

$$S_{x_T}^{\alpha_i + \epsilon}(n, f_0)_{\Delta t} = \sum_r X_T(r, f_1) X_T^*(r, f_2) g(n-r) e^{-j2\pi r \epsilon T_s} \quad (10)$$

If several values of ϵ are desired, evaluation of the sum can be simplified by discretizing the values of ϵ to be $\epsilon = q\Delta\alpha$. In this case the output of the algorithm is expressed as:

$$S_{x_T}^{\alpha_i + q\Delta\alpha}(n, f_j)_{\Delta t} = \sum_r X_T(r, f_1) X_T^*(r, f_2) g(n-r) e^{-\frac{j2\pi r q}{N}} \quad (11)$$

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 [7]. In this approach the frequencies of the filter bank are discretized to: $f_k = k \left(\frac{f_s}{N'}\right)$, $k = -\frac{N'}{2} \dots \left(\frac{N'}{2}\right) - 1$ (12)

The locations of SCF estimation associated with the pairs of complex demodulates are (f_j, α_i) where the frequency coordinates are: $f_j = \frac{f_k + f_l}{2} = \frac{k+1}{2} \left(\frac{f_s}{N'}\right)$ (13)

And the cycle frequency coordinates are:

$$\alpha_i = f_k - f_l = (k - 1) \left(\frac{f_s}{N'} \right) \quad (14)$$

Tilling 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 regions (one quadrant of the bifrequency plane) [2]. The more details of the FAM algorithm can be studied at [13].

ANALYSIS

To demonstrate the efficiency of system, sub set of Corel date base containing 600 image (grouped into six classes each consisting of 100 images) was used for query images in a simulated analysis. For experimentation, the images have been categorized into six classes. For each class, the performance is evaluated, as a percentage of the correct number of returned images and is presented in Table I. Average precision in present method significantly is better than simplicity and edge based methods in all categories. Comparison between present method and Horng Lin method demonstrate to some extent same results and average precision at four categories from six increased. Main advantage of present system is computation complexity which is the least one among different methods. Table I demonstrate Comparison of accuracy (ACC%) of retrieved image. Fig.1 show result of Object extraction using EM Algorithm for different Gaussian functions. Upper line of Fig.1, present three functions and lower two functions. Meanwhile, Fig.2 display Average precision of different methods on Image set.

TABLE I: Comparison of average precision (%) with proposed method and other standard retrieval systems [15][4][7] when 100 image returned.

Semantic name of Class	Simplicity [15]	Edge Based [4]	Horng Lin Method[7]	Present Method
Building	35%	35%	36%	42%
Bus	36%	60%	69%	67%
Dinosaur	95%	95%	96%	99%
Elephant	38%	25%	55%	64%
Flower	42%	65%	89%	69%
Horses	72%	65%	70%	75%
Mean	53%	57.7%	69.16%	69.33%

It is also interesting to notice that the proposed method has a high speed performance, which is faster than presented method in [7]. Moreover, unlike the proposed method in [7] the proposed method has a simple structure.

TABLE II: Comparison of accuracy (ACC%) of retrieved image on Image set.

Returned Image	1	3	4	5	10	30	50	100
Simplicity [15]	60.3	64.7	73.0	78.1	79.5	85.4	88.0	90.7
Edge Based [4]	62.4	74.8	76.6	79.0	84.0	90.2	92.3	94.6
Lin[7]	85.5	92.3	93.2	93.6	95.1	97.7	98.9	99.2
Present Method	88.3	93.9	94.4	94.6	95.7	98.1	99	99.3

CONCLUSION

This paper presents a full-automatic tool for object-based content retrieval through the structure of EM algorithm and spectral analysis. Novel features extraction based on spectral function has been introduced. Subsequently, statistical information from spectral function descriptors are retrieved and matched accordingly. The use of defined objects rather than the entire image leads to increase the flexibility of the system, it make simple to search for an object in an already prepared different kind of objects. Object-based image retrieval is not limited by the averaging properties associated with analyzing the entire image and can use local properties.

REFERENCES

1. Brune O. Synthesis of a finite twoterminal network whose driving-point impedance is a prescribed function of frequency. *J Math Physics*. 1931;10:191-236.
2. Craig JJ. 2005. *Introduction to Robotics Mechanics and Control*. 3rd ed., Pearson Prentice Hall.
3. Eppinger S, Seering W. "Three Dynamic Problems in Robot Force Control. *IEEE Trans Robotics & Auto*. 1992;8(6):751-758.
4. Fairlie-Clarke AC. Force as a Flow Variable. *Proc Instn Mech Engrs*. 1999;213:77-81.
5. Foster RM. A reactance theorem. *Bell System Tech J*. 1924;3:259-267.
6. Hamill DC. Lumped Equivalent Circuits of Magnetic Components: The Gyrator-Capacitor Approach. *IEEE Transactions on Pow Electronics*. 1993;8:97.
7. Hayward V, Choksi J, Lanvin G, Ramstein C. 1994. Design and Multi-Objective Optimization of a Linkage for a Haptic Interface". *Proc. Of ARK „94, 4th Int. Workshop on Advances in Robot Kinematics (Ljubliana, Slovenia)*, pp. 352-359.
8. Paynter HM. 1961. *Analysis and Design of Engineering Systems*. MIT Press.
9. Sass L, McPhee J, Schmitke C, Fiset P, Grenier D. "A Comparison of Different Methods for Modelling Electromechanical Mult Systems. *Multibody System Dynamics*. 2004;12:209-250.
10. *ligent Robots & Systems*, (Raleigh, NC), pp. 1103,-1110.
11. Stocco L, Yedlin M. 2006. Closing the Loop on the Electro-Mechanical System Analogy". Submitted to: *IEEE J. Circuits & Systems*.
12. Tilmans HAC. Equivalent circuit representation of electromechanical transducers: I. Lumpedparameter systems. *J Micromech Microeng*, 1996;6:157-176.