© 2016 IJSRSET | Volume 2 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Animal Face Detection Technique Using DCT

Prasannakumar Eluru, Santhosh Kumar. B. R, Santhosh Dhanajaya, Vinodh. K, R. Gunasekari

Department of EEE, Sri Sai Ram College of Engineering, Anekal, Bengalurum Karnataka, India

ABSTRACT

The problem of facilitating machine to learn & detect patterns of animals and categorizing them in the rapidly expanding image databases has become increasingly challenging for image retrieval systems. The problem to be solved is detection of animal faces in an image. In this regard, we propose transform domain techniques to identify the low frequency components present in an image signal to generate highly discriminative geometrical features preserving color and texture information which are analogous to human vision perceptual models. A human can do this easily, but a computer needs precise instructions and constraints. Animal detection based researchers are useful for many real life applications. Animal detection methods are helpful on the research related to locomotive behavioral of targeted animal and also to prevent dangerous animal intrusion in residential area. There are a few branches of research related to animal detection. Therefore, in this paper, Face Detection is done by Discrete Cosine Transform (DCT) approach.

Keywords : Digital Image Processing; Animal Detection; DCT;

I. INTRODUCTION

The human capacity to recognize particular individuals solely by observing the human face is quite remarkable. Face perception refers to an individual's understanding & interpretation of the face, particularly the human or animal face, especially in relation to the associated information processing in the brain. During face perception, neural networks make connections with the brain to recall memories.

According to the Seminal model of face perception, there are three stages of face processing including recognition of the face, the recall of memories & information that are linked with the face of the animal & finally name recall. This capacity persists even through the passage of time, changes in appearance & partial occlusion. Because of this remarkable ability to generate near perfect positive identifications, considerable attention has been paid to methods by which effective face recognition can be replicated on electronic level [1]. Certainly, if such a complicated process as the identification of animal faces based on few methods is possible, it is very useful in counting the number of animals etc. Here we use DCT which seeks to classify various types of animals & also used to classify varieties among themselves. The DCT approach exploits the feature extraction capability of the Discrete Cosine Transform invoking low frequency part in the animal's face. This method is used in order to have a better understanding on animal behavior. Besides, these applications also can act as a warning system to human being from intrusion of dangerous wild animal for early precaution measures. The animal identification is used to identify the detected animal. It has been used in health monitoring system for domestic herds. Animal care management becomes an important issue as animals have a direct impact on human psychological and physical health. Even though there are a lot of solutions, maltreatment of animals and risks in animal health are increasing. Detection of animal helps human to monitor and manage their animals easier. The transformation of a picture using DCT is quite helpful when comparing patterns or faces of animals. Fig.1 shows the image retrieval process.

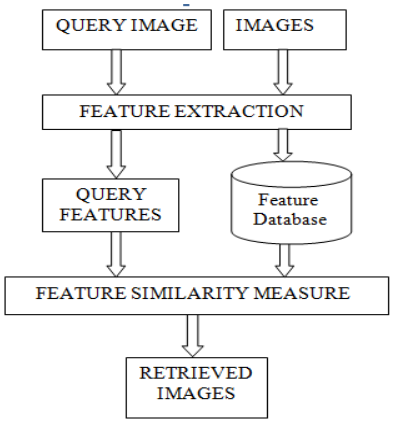


Figure 1. Image Retrieval Process

II. METHODS AND MATERIAL

1. Related Works

A. Researches on Animal Detection by Human Eyes

Early researches on animal detection are to observe how fast human eyes can detect the presence of animal in natural scene. Animal detection by human eyes has been considered as the most reliable detection method if seen from the computational point of view. This is because the image structure in natural images is complex. In [2], it is found that a human observer is able to decide whether a briefly flashed animal scene contain an animal as fast as 150ms. In [3], median reaction time results indicate a speed accuracy of 92 percent for reaction time of 390ms and increase to 97 percent of correctness for 570ms. Though human detection is effective and achieve satisfactory level, human eyes can easily get tired causing decreasing of effectiveness. Furthermore, human eyes cannot work 24 hours a day to perform animal detection. These flaws can be curbed by applying computer vision in image processing for animal detection.

B. Researches on Power Spectral

The researchers also have tried to find whether the presence of animal in the image scene will change the power spectral of the image or not. The power spectral can be defined as the amplitude of the signal in the frequency domain. This can be constructed by transforming the images from spatial domain into the frequency domain, by using Fourier transform. The main idea is to help the human observer to realize the presence of the animal in the scene by inspecting the power spectral. Work in [2] found that the human observer will not prefer to use this approach if they want to quickly detect the animal.

C. Animal Detection Using Face Detection Approach

For research regarding locomotive behavior of wild animal, method combining detection and tracking of targeted animal faces has been applied in [4] using Haar-like feature and Ad boost classifiers. The video recorders is only turn on when it is positive that targeted animal been detected to prolong battery life time and to ensure recorded video contain research value. This method especially crucial in situation whereby video man is not suitable to present at the recording scene for safety issue or video man might scare off some timid animal away. The animal faces are measured by utilizing face detection method with different local contrast configuration of luminescence channel to detect the image region of animal faces.

D. Animal Detection Based on Thresholding Segmentation Method

Target extraction from background can be performed by using threshold segmentation method. In [5], the object is found by using background subtraction method after obtaining the background image. In [6], threshold segmentation method based on the pixel values is performed. However, in this technique, researchers should carefully choose the threshold value as they also should consider the negative value obtained at certain pixel point by direct subtraction. The idea of threshold segmentation is simple, which pixel of gray that greater than threshold are set to white (i.e. intensity 255) and those less than the threshold value will be set to black (i.e. intensity 0). As stated in [6], it is difficult to select the threshold accurately as the background image periodically changes. Therefore, different appropriate threshold should be chosen for different background scene.

2. Proposed Methodology

Addressing the issue of human perception of visual data content and deriving discriminating image features to reduce the semantic gap, we carried out image analysis in frequency domain by using DCT technique for better

image representation and categorization of very large image datasets. Finally, a distance measure is used to obtain an average per class recognition rate. As mentioned above, in-detail explanation regarding transform domain techniques along with distance measure can be found in following subsections.

A. Discrete Cosine Transform (DCT)

The Discrete Cosine Transform (DCT) was introduced for the first time by Ahmed, Natarajan and Rao in 1974. Ever since, with the increase in its popularity, till now many other variants have been proposed (Rao and Yip, 1990). Wang (1984) categorized DCT into four different transformations, which are DCT-I, DCT-II, DCT-III, and DCT-IV. Out of these four classes defined by Wang, DCT-II is used as DCT and DCT-III as its inverse operation.

Every transform is a type of mathematical operation and when it is applied to a signal, being processed converts it to another different domain and it can also be converted back to the original domain by the use of the inverse transform operation. The transform gives us a set of coefficients which acts as feature vectors describing the given signal and helps us to regain the original samples of the input signal. There are some mathematical transforms which can produce de-correlated coefficients such that maximum of the signal energy is concentrated in a less number of coefficients.

The Discrete Cosine Transform (DCT) can be described as a finite sequence of data points which are in terms of summation of cosine functions oscillating at different frequencies. Like other transforms, it also attempts to de-correlate a given signal. After being de-correlated, the transform coefficient is encoded independently in such a way that there is no loss in compression efficiency. The DCT coefficients are reflection of the different frequency components which are present in it. The coefficient at the first place of the DCT refers to the DC component of the signal which is its lowest frequency and most of the time; it carries the maximum of the relevant information present in the input signal. The signal's higher frequencies are represented by the coefficients present at the end and these generally represent the finer details about the original signal. The remaining coefficients carry other levels of information of the input signal given.

DCT is a very popular transform technique used in image compression [7] and face recognition [8]. For most of the images, much of the signal strength lies at low frequencies and DCT significantly separates out image spectral sub-bands with respect to image's visual features. It transforms a signal from time domain to frequency domain.

Let f (x), x =0, 1, 2, ..., N−1, be a sequence of length N, and then 1D DCT is defined by the following equation: Unlike 1D signal, the DCT of an m x n image f (x,y) is given by:

$$F(u) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \sum_{x=0}^{N-1} \Lambda(x) \cos\left[\frac{(2x+1)u\pi}{2N}\right] f(x) \quad (1)$$

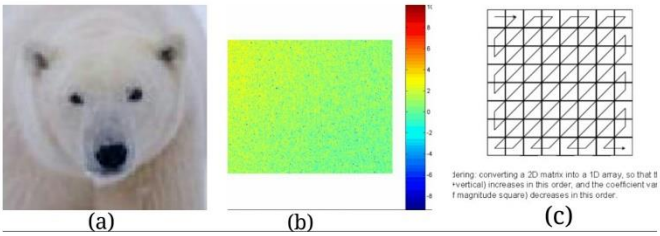
Unlike 1-D signal, DCT of given image f(x, y) of size m x n is defined as:

$$C(u,v) = \frac{2}{\sqrt{mn}} \alpha(u)\alpha(v) \sum_{x=1}^{m} \sum_{y=1}^{n} f(x,y) \cos\left[\frac{(2x+1)u\pi}{2m}\right] * \cos\left[\frac{(2y+1)v\pi}{2n}\right] \quad (2)$$

where, ∀u = 1,2, ...,m and ∀v = 1,2, ...,n are scaling factors.

Fig.1 (a) shows 256 X 256 pixels size color airplane image, Fig.1 (b) shows the color map of log magnitude of its DCT signifying high energy compaction at the origin and Fig.1(c) illustrates the scanning strategy of DCT coefficients for an image using conventional zigzag technique.

DCT coefficients preserve low frequency coefficients to identify the local information which is invariant to illumination, occlusion, clutter and almost nullifying the effect of high frequency coefficients.



(a) (b) (c)

Figure1 (a) Original image; (b) Color map of quantized magnitude of DCT; (c) Scanning strategy of DCT coefficients

DCT coefficients preserve low frequency coefficients to identify the local information which is invariant to illumination, occlusion, clutter and almost nullifying the effect of high frequency coefficients.

B. Similarity distance measure

Different similarity distance measures will affect the recognition rate, in this regard we consider Euclidean distance. There are four more different distance measure techniques such as Minkowski distance, Modified Squared Euclidean distance, Correlation coefficient based distance and Angle Based distance to acquire an average classification rate. Minkowski distanceis a metric in a normed vector space which can be considered as a generalization of both the Euclidean distance and the Manhattan distance.

Euclidean distance between two points is Euclidean space. With this distance, Euclidean space becomes a metric space.the Euclidean distance between points p and q is the length of the line segment connecting them. In cartesian coordinates, if p= (p1,p2,….,pn) and q= (q1,q2,….,qn) are two points in Euclidean n-space, the the distance (d) from p to q, or from q to p isn given by the Pythagorean formula:

$$d(p,q)= d(q,p)= \sqrt{\sum(qi-pi)^2} \quad (3)$$

In Squared Euclidean distance, the standard Euclidean distance can be squared in order to place progressively greater weight on objects that are farther apart. In this case, the equation becomes: $d^2(p,q)=(p1-q1)^2 + (p2-q2)^2 +…..+(pn-qn)^2$. Squared Euclidean distance is not a metric as it does not satisfy the triangle inequality, however, it is frequently used in optimization problems in which distances only have to be compared. Here we use either of these methods to calculate the distance between query matrix and the transformed matrix.

A correlation coefficient is a coefficient that illustrates a quantitative measure of some type of correlation and dependance, meaning statistical relationships.

Proposed method is divided into two phases i)Training & ii)Testing, brief description of proposed algorithm is as given below:

Algorithm: Training Phase
Input: A set of training sample images
Output: A Knowledge Database (KDB)
Method:

1. Acquire training samples.
2. Apply Discrete Cosine transformation models on set of segmented training samples to obtain feature vector.
3. Store the feature matrix in Knowledge Database.

Algorithm: Training Phase Ends

Algorithm: Testing Phase
Input:

i. A Knowledge Database (KDB)
ii. Test/Query image, I.

Output: Class label / Identity of I.
Method:

1. Obtain the feature matrix F of the pattern I.
2. Apply Euclidean distance.
3. Label the class/identity using step2.

Algorithm: Testing Phase Ends

III. RESULTS AND DISCUSSION

Experimental Results & Performance Analysis

In this section, we present result and analysis carried out on dataset consists of 20 different animal categories having around 100 images per category which are subjected to high intensity variations and occlusions. The database comprises 20 different animal categories which has been manually screened out of Google images exhibiting high variations in intensity, clutter, object size, location, pose, and also increase in number of category with at least 90 images per category.

In animal face detection algorithm using DCT, the system gets an image as input containing an animal's face. It is then compared with other animal faces in the training set, under the same nominal size, position, orientation and illumination conditions. It is compared on the basis of features extracted from the DCT coefficients. Basically we have to compute the DCT of

the animal faces and keep a subset of DCT coefficients as the feature vectors which describes the face.

The feature vector contains all the low-to-mid frequency DCT coefficients, having the maximum variance. For recognizing a particular animal's face which is given as input to the system, it compares the feature vector of the given animal's face to the feature vectors of the images in the database using the Euclidean distance.

Figure 2. (a) Query image. (b) Resultant relevant image.

Transformation models are experimented by varying the training samples. We divided entire dataset into 5 and 30 images/category as training phase and remaining as testing to obtain an average of per class recognition in each stage. Performance analysis for well-known techniques & benchmarking datasets are explained in the following paragraphs.

Table I shows the recognition rate for transform technique and considering the distance measure for datasets. From the table it is noticed that DCT will extract the energy information of the image from all the angle with respect to x-axis. Most of the energy information are present in terms of low frequency coefficients.

TABLE I. RECOGNITION ACCURACY OF PROPOSED METHODS

Method	05 Train	10 Train	15 Train	20 Train	25 Train
DCT	14.21	15.66	16.94	17.93	19.13

In this table, the DCT approach was put into test under a wide variety of conditions. Specifically, images with many differences between them were used in this experiment. Efficiency for all the conditions is calculated. The first two images taken for the testing was from our training set. So it can be seen that the max and min Euclidean distance is within the specified range.

IV. CONCLUSION

Despite the earlier efforts made in the image retrieval research, we do not yet have a generally acceptable algorithmic means of distinguishing human vision. In this regard, a remarkable effort to extract discriminative patterns more distinctively in the context of interpreting images is realized by using DCT. The results of DCT approach are quite satisfactory and DCT extracts low frequency information later we make use of Euclidean distance to retrieve the image similar to query image.

V. REFERENCES

[1] Guruprasad, "Face Recognition using Eigenfaces". Monday, june 11, 2012.

[2] M. F. Thorpe, A. Delorme, and S. T. C. Marlot, "A limit to the speed processing in ultra-rapid visual categorization of novel natural scene,"Cognitive Neuroscience, pp. 171-180, 2003.

[3] F. A. Wichmann, J. Drewes, P. Rosas, and K. R. Gegenfurtner, "Animal detection in natural scenes: Critical review revisited," Journal of Vision, vol. 10, no. 4, pp. 1-27, 2010

[4] T. Burghardt and J. Calic, "Real-time face detection and tracking of Animals," 8th seminar on Neural Network Applications in Electrical Engineering. Pp 27-32, 2006.

[5] C. Peijiang, "Moving object detection based on background extraction," Computer Networks and Multimedia Technology (CNMT), 2009.

[6] J .C. Nascimento and J. S. Marques, "Performance evaluation of object detection algorithms for video surveillance," IEEE Transactions on Multimedia, vol.8, pp. 761-774,2006.

[7] Anna Bosch, Andrew Zisserman, Xavier Munoz, Image Classification using Random Forests and Ferns, IEEE 11th International Conference on Computer Vision (ICCV), pp 1 - 8, 2007.

[8] Boiman O, Shechtman E, Irani, M., In Defense of Nearest-Neighbor Based Image Classification, IEEE CVPR, pp 1-8, 2008.

[9] Ritendra Datta, Dhiraj Joshi, Jia Li, James Z. Wang, Image retrieval: ideas, influences, and trends of the new age, ACM Computing Surveys, vol. 40, no. 2, pp. 1-60, 2008.