Scene Image Analysis using GLCM & Gabor Filter

Ranjita Mishra[#1]
[#]Electronics & Telecommunications Department, Silicon Institute of Technology , BPUT, Odisha

ABSTRACT

In this paper, images of real- world natural scenes and manmade structures of different depth are taken. With increase in image depth , roughness increases in case of man-made structures whereas natural scene images become smooth, thus reducing the local roughness of the picture. Such kind of specific arrangement produces a particular spatial pattern of dominant orientations and scales that can be described using Gabor filter as it gives the local estimate of frequency content in an image. Here various techniques are used i.e. grey level co-occurrence matrices (GLCM), Gabor filters, combined GLCM and Gabor filters. Here the real scene images are classified in four classes such as near natural, near manmade, far natural and far manmade .Gabor filter only classify into low energy and high energy scenes. So the combination of Gabor filter and GLCM are used for classification in to four classes. In the proposed method i.e. the combination of Gabor and GLCM, first Gabor and PNN is used for classification between two groups high energy (such as near natural & far manmade) and low energy(such as near manmade & far natural) and then the GLCM and PNN is used for classification of subgroups.

Keywords— Gabor filters, GLCM, scene classification,

I. INTRODUCTION

Images of natural scenes and manmade objects exhibit a particular behaviour with respect to varying depth [9]. Close-up views of manmade objects generally contain large flat surfaces and homogeneous surface areas are likely to be replaced by objects like ,walls ,doors, windows etc thereby increasing the global roughness of the picture. On the other hand , natural scene images show a radically different behaviour with respect to mean-depth when compare with manmade objects. Close-up views of natural scenes are usually textured with regular repetition of some specific pattern. But as the depth of such image increases small grains of close-up view are indiscernible, because there is a change in spatial frequency of the image. As a result natural structures become larger and smoother giving rise to homogeneous regions i.e. low in roughness. Thus in natural scene, more we increase the mean depth of the scene, more is the energy concentrated in the low spatial frequency components contrary to the behaviour shown by manmade structures. Therefore it may be perceived that images of distant manmade objects and close natural scenes have similar energy content. Similarly, the energy content of distant natural scenes and close manmade structures are comparable.

We consider here a fundamental problem of computer vision, i.e. enabling computers to see the way we see things. We in future wish our machines would match the capabilities of human vision. It's interesting to note that, every second we receive tremendous amount of visual data and almost unconsciously we process this information very quickly. Classifying an object as table, a ball, or a scene as mountain or river is pretty trivial for us. We can in fact process amazingly more complex information. It's a well known fact that robotic vision compares miserably with our eyes. Here, we intend to make a start towards our goal by considering a very trivial problem by the standard of human vision and that is scene classification. Given an image of the scene we wish to classify it as say a mountain, forest, city, street etc. Generally, a learning based approach is used to solve problems of this nature. A training set is initially created which would contain representative images from all categories that we need to classify. Now these images are manually labelled to the class they belong as perceived by the human. Now a learning algorithm is employed, which basically is a strategy to enable us to come up with parameters which would characterize an image for doing the classification task. Now if a random image is given as an input, on basis of parameters already identified the machine would try to classify the image. This in essence a generic way in which learning algorithms work, i.e., by learning from a huge set of data and then using this learned information to make predictions about successive inputs.

In this work we are trying to recognize the scenes of four different categories namely near-natural, near-manmade, far-natural, & far-manmade. In our work different methods are used for features extraction like Gabor filter, GLCM, and combination of Gabor filter and GLCM. Gabor filter can classify the scene images into two categories based on Gabor Energy. i.e. Low energy and High energy scene images. Near-manmade and far-natural scene images have low energy whereas, Near-natural & far-manmade scene images have high energy. So Gabor

Filter can classify these scene images into two groups but it cannot classify in two four different class.

Previously the Gray Level Co-occurrence Matrix(GLCM) features are used for classification of scene images in two categories natural and artificial . But we are trying to classify the scene images in two four different classes using GLCM for features extraction and PNN as classifier. Our proposed method is the combination of both Gabor Filter and GLCM .Then the classification accuracy is compared for the two methods.

II. GLCM

This is a popular statistical texture classification Technique ever since it is introduced by Haralick *et al.* back in 1973 [11] because it is computationally simple yet useful for many texture classification problems. The GLCM calculates the occurrence of pixel pairs within the images according to the spatial distance between the pixel pairs and orientations provided [12]. The computed GLCM can be used as a feature after it is down-sampled which we named as raw GLCM in this paper [9]. A second-order feature can be obtained from the GLCM. There are 5 commonly used textural features, i.e. contrast, correlation, energy, entropy and homogeneity [12]. In this paper, the GLCMs are generated as in [5] and [8]. In this study we use the following three features that are most commonly used:

$$\text{Energy} = \sum_{i,j} \{P(i,j)\}^2$$

$$\text{Inertia} = \sum_{i,j} (i-j)^2 * P(i,j)$$

$$\text{Entropy} = -\sum_{i,j} P(i,j) * \log(P(i,j))$$

III. GABOR FILTER

This is a signal processing method, therefore it Processes on the frequency domain rather than the spatial domain [2]. In this paper, the Gabor filters are generated by using different two radial centre frequencies and eight orientations as used in [7]. The following family of two-dimensional Gabor functions was proposed by Daugman [1] to model the spatial summation properties (of the receptive fields) of simple cells in the visual cortex:

$$g_{\lambda,\Theta,\varphi,\sigma,\gamma}(x,y) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\varphi\right) \quad (1)$$

$$x' = x\cos\Theta + y\sin\Theta$$
$$y' = -x\sin\Theta + y\cos\Theta$$

where the arguments x and y specify the position of a light impulse in the visual field and (ξ,η) has the same domain as (x,y),σ, γ,λ,θ and Φ are parameters as follows :

σ & γ : The standard deviation sigma(σ) of the Gaussian factor determines the (linear) size of the receptive field. Its ellipticity and herewith the ellipticity of the receptive field ellipse is determined by the parameter gamma(γ), called the *spatial aspect ratio*. It has been found to vary in a limited range of $0.23 < \text{gamma}(\gamma) < 0.92$. The value of the $\gamma=0.5$ is used., Sigma cannot be controlled directly in the applet. Its value is determined by the choice of the parameters lambda a and b.

λ & b : The parameter lambda(λ) is the wavelength and 1 / lambda the spatial frequency of the cosine factor in Eq. (1). The ratio *sigma / lambda* determines the spatial frequency bandwidth of simple cells and thus the number of parallel excitatory and inhibitory stripe zones which can be observed in their receptive fields. The half-response spatial frequency bandwidth **b** (in octaves) and the ratio sigma / lambda are related as follows:

$$b = \log_2 \frac{\frac{\sigma}{\lambda}\pi + \sqrt{\frac{\ln 2}{2}}}{\frac{\sigma}{\lambda}\pi - \sqrt{\frac{\ln 2}{2}}}, \qquad \frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{\ln 2}{2}} \cdot \frac{2^b+1}{2^b-1} \qquad (2)$$

Neurophysiological research has shown that the half-response spatial-frequency bandwidths of simple cells vary in the range of 0.5 to 2.5 octaves in the cat (weighted mean 1.32 octaves) and 0.4 to 2.6 octaves in the macaque monkey (median 1.4 octaves). While there is a considerable spread, the bulk of cells have bandwidths in the range 1.0-1.8 octaves. Some researchers propose that this spread is due to the gradual sharpening of the orientation and spatial frequency bandwidth at consecutive stages of the visual system and that the input to higher processing stages is provided by the more narrowly tuned simple cells with half-response spatial frequency bandwidth of approximately one octave. Since lambda and sigma are not independent when the bandwidth is fixed, only one of them, lambda, is considered as a free parameter which is used in the applet.

Θ: The angle parameter theta specifies the orientation of the normal to the parallel excitatory and inhibitory stripe zones - this normal is the axis x' in Eq. (1) - which can be observed in the receptive fields of simple cells.

Φ: Finally, the parameter phi, which is a phase offset in the argument of the cosine factor in Eq. (1), determines the symmetry of the concerned Gabor function: for phi=0 degrees and phi=180 degrees the function is symmetric, or even; for phi= -90 degrees and phi= 90 degrees, the function is antisymmetric, or odd, and all other cases are asymmetric mixtures of these two.

The values of Φ used in the simulation are Φ=0 for symmetric receptive fields and Φ= -Π/2 for antisymmetric receptive fields. Due to the complexity of the features produced, the Gabor filters are down-sampled and the singular value decomposition (SVD) is further used to reduce the dimensionality of the feature set [7].

IV. COMBINED GLCM AND GABOR FILTER

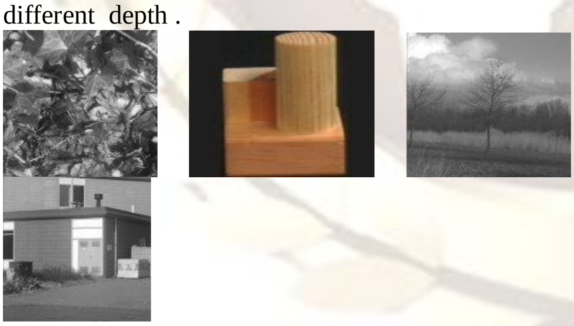
Different techniques are often combined to be used and can produce better results compared to using them individually [7, 8, 10, 13]. In this paper, we combine the Gabor filters and the GLCM feature by appending both of them into a single feature set. This combination produces a better result than either GLCM feature or Gabor filters .

V. PNN

Probabilistic Neural Network can be used for classification problems. When an input is presented , the first layer computes the distance from the input vector to the training input vectors and produces a vector whose element indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities . Finally a complete transfer function on the output of the second layer chooses the maximum of these probabilities and assigns '1' to that class and '0' otherwise. For example if there are Q input vector/target pairs and each target vector has K elements, then one of these element is 1 and the rest are 0. Thus each input vector is associated with one of the K class.

VI. EXPERIMENTAL DATASET

The dataset used in this paper is the real world dataset and the experimental tool used is MATLAB. Images of manmade objects & natural scenes download from websites[17,18] are considered for the present study. The algorithm has been verified on images having two different depth .



Sample images belonging to four classes: near-natural scenes (Row-1), near-manmade objects (Row-2), Far-natural scenes (Row-3) & Far-manmade structure (Row-4)

VII. EXPERIMENTAL RESULTS

TABLE I :GABOR ENERGY FOR DIFFERENT SCENES IMAGE

near natural	near manmade	Far natural	Far manmade
37.5	18.8	16.8	31.4
35.2	17.3	12.5	56.3
30.3	21.1	13	45.3
30.6	19.9	14	37.7
41.4	20	16.9	35

TABLE III: CLASSIFICATION ACCURACY-BY TAKING 12 IMAGES FROM EACH CATEGORY

Image Used	Classification Accuracy (%)	
	GLCM	Gabor & GLCM
Near Natural	83.3	83.3
Near manmade	66.7	78
Far_ Natural	75	91
Far manmade	66.7	83.3

TABLE IIII:**OVERALL CLASSIFICATION ACCURACY**

Different Method	No of images Test	No of images miss classify	% Correct result
GLCM	100	22	78
Gabor+GLCM	100	12	88

VIII. DISCUSSION

We have created our own database of Scenes of 4 classes each containing 100 scene images. We conducted experimentation under varying database size and we studied its effect on classification accuracy. There are three experiments conducted in this paper. They are conducted on the GLCM, Gabor filters, combined GLCM and Gabor filters, for scene image classification. Gabor Filter can classify these scene images into two groups low energy & high energy, but it cannot classify in two four different class.So we compair GLCM and the combined of Gabor & GLCM. The experimental results show that the Success rate is of 78% using GLCM & Combined rate of 88% by using both Gabor filter & GLCM.The comparison is shown in Table III.

IX. CONCLUSIONS

In this paper we have proposed a probabilistic neural network based Scene image classification method by the use of texture features. The suitable texture features such as GLCM and Gabor response is explored for the purpose of Scene image classification . We have created our own database of Scenes of four classes each containing 100 scene images. We conducted experimentation under varying database size and we studied its effect on classification accuracy.

An alternative to objects for scene representation is to directly compute space descriptors; the space descriptors (naturalness, openness, etc.) provide a scene representation at multiple levels of categorization and independently of the complexity of the scene. The descriptors of the spatial envelope are correlated with the image second-order statistics and the spatial arrangement of structures in the scene.

Result of current model: Success rate is of 78% using GLCM and Combined rate of 88% by using both Gabor filter & GLCM. Here, we discovered that the overall accuracy rate produced using feature images generated by both the Gabor filters & GLCM has the best accuracy at 88 % for the scene image classification problem. And the classification accuracy, by taking 12 images from each category is maximum (83.3%) for near natural & far natural by using GLCM, and maximum (91%) for far natural object by using the combination of both Gabor filter and GLCM.

The experimental results have shown that:

- The Gabor filter can classify the scene images into two group based on their energy but the filter gives very poor result if used for classification into four classes because the energy of the near natural and far manmade are very close to each other. Similarly the energy of the near manmade and far natural have the close values.
- The GLCM method produces a poor result as compared to combination of Gabor filter and GLCM classification techniques;

REFERENCES

IEEE Papers:

[1] Jing Yi Tou, Yong Haur Tay, Phooi Yee Lau "***A Comparative Study for Texture Classification Techniques on Wood Species Recognition Problem***" 2009 Fifth International Conference on Natural Computation, IEEE DOI 10.1109/ICNC.2009.594

[2] Devendran V , Hemalatha Thiagarajan , A. K. Santra" ***Scene Categorization using Invariant Moments and Neural Network*** "International Conference on Computational Intelligence and Multimedia Applications 2007 , IEEE DOI 10.1109/ICCIMA.2007.77

[3] R. M. Haralick, K. Shanmugam, and I.Dinstein, "***Textural Features for ImageClassification***", IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-3, No.6, November1973, pp.610-621

[4] Devendran V "***Feature Selection for Scene Categorization using Support Vector Machines***" 2008 IEEE Congress on Image and Signal Processing, DOI 10.1109/CISP.2008.579

[5] J. Garcia-Consuegra, G. Cisneros "***Integration of Gabor Functions with Co-occurrence Matrices: Application to Woody Crop Location in Remote Sensing***" 0-7803-5467-2/99/ 1999 IEEE

[6] Shuo Chen, Chengdong Wu, Dongyue Chen, Wenjun Tan "***Scene Classification Based on Gray Level-Gradient Co-Occurrence Matrix in the Neighborhood of Interest Points***" National Nature Science Foundation 978-1-4244-4738-1/09/2009 IEEE

[7] Zhihua xie , Guodong Liu , Cuiqun He, Yujie Wen "***Texture Image Retrieval Based on gray level cooccurrence matrix and singular value decomposition***" 978-1-4244-7874-3/10/2010, IEEE

[8] Peter Kruizinga and Nikolay Petkov N(1999) "***Nonlinear Operator for Oriented Texture*** " IEEE Transation,vol8,No10 pp1395-1407

[9] Oliva A,Torralba A (2002)"***Depth estimation from image structure*** ", IEEE Transaction on Pattern Analysis and Machine Intelligence , vol.24 , No.9,pp-1226-1283

[10] JOHN G. DAUGMAN "***Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression*** " IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. VOL. *36*. NO. 7. JULY 1988

Other Journal Papers:

[11] *Mari Partio, Bogdan Cramariuc, Moncef Gabbouj, and Ari Visa"* ***ROCK TEXTURE RETRIEVAL USING GRAY LEVEL CO-OCCURRENCE MATRIX "***

[12] *Jing Yi Tou1, Yong Haur Tay1, Phooi Yee Lau2 "* ***RECENT TRENDS IN TEXTURE CLASSIFICATION: A REVIEW "*** Symposium on Progress in Information & Communication Technology 2009

[13] N. Petkov, P. Kruizinga" ***Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells*** " Biol. Cybern. 76, 83–96 (1997)

[14] Brandt Tsoa* and Richard C. Olsenb "***Scene classification using combined spectral, textural and contextual information "***

[15] Julia Vogel and Bernt Schiele, "***A semantic typicality measure for natural scene categorization,***" Pattern Recognition Symposium DAGM, 2004.

[16] Oliva A. and Torralba A., "***Modelling the shape of the scene: A holistic representation of the spatial envelope,"*** International Journal of Computer Vision, vol. 42(3), pp. 145–175, 2001.

[17] http://www.cs.columbia.edu/CAVE/research/softlib/coil-100.html/

[18] http://hlab.phys.rug.nl/