International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-4, Issue-7)

Research Article

July 2015

Scene based Approach for Horse Image Categorization using Support Vector Machines

Nagarajan .A
Assistant Professor,
Dept of Computer Applications
PSNA College of Engg & Tech,
Dindigul, Tamilnadu, India

Devendran .V
Associate Professor
Dept of Computer Applications
PSNA College of Engg & Tech,
Dindigul, Tamilnadu, India

Abstract-

Scene classification, the classification of images into semantic categories is a challenging and important problem now a days. Categorization of scenes is a fundamental process of human vision that allows us to efficiently and rapidly analyzes our surroundings. This paper is classifying the images as 'images with horses' and 'images without horses' using support vector machine with radial basis kernel with p1=5. This work is double folded as to categorizing the images using support vector machine and to find better feature extraction method among the ones which have been used by the research community often i.e., wavelet features, invariant moments, singular value decomposition and co-occurrences matrix. Also, this paper shows the usages of blocking method and shows that it is having high impact on the performances of classifier. The sample images are taken from the real world dataset.

Keywords: Gray level co-occurrence matrix, Invariant Moments, Scene Categorization, Support Vector Machine and Wavelet features.

I. INTRODUCTION

In computer vision, the examples are representations of photographic images and the task of the classifier is to indicate whether or not a specific object or phenomena of interest is present in the image [13][10][17][11][24]. Simple texture analysis [26] of the image can provide useful cue towards rapid scene identification. In order to successfully accomplish this, the classifier must have sufficient prior knowledge about the appearance of the object[6]. Scene classification system is trained to recognize a type of example or differentiate between examples that fall in a separate category in changing viewpoints, photo category classification and indoor-outdoor photo classification [12][21][8][2]. Since performance of classification of scenes should be robust to real world circumstances, preprocessing techniques such as noise removal, enhancements, smoothening is not used. Feature extraction techniques are applied directly to the raw images without missing any information from the images. Various feature extraction methods are tested and compared in scene categorization. Regular moment invariant, one of the most popular and widely used contour-based shape descriptors, is a set derived by Hu [15]. In [3][19], authors present a scene description and segmentation system capable of recognizing natural objects (e.g., sky, trees, grass) under different outdoor conditions. Paper [1] investigates whether and how visual representations of individual objects are bound in memory to scene context.

This paper is trying to recognize the scenes of two different categories called 'With Horses' and 'Without Horses'. Haar Wavelet features, Singular Value Decomposition, Co-occurrence Features and Invariant Moments are the various methods used to extract features from the scenes. Radial Basis kernel function is used for classification of scenes in support vector machine.

This paper is organized as follows: Section 2 describes feature extraction methods; Section 3 explains Support Vector Machine Classifier; Section 4 gives details of database preparation; Section 5 deals with proposed work; Section 6 describes implementation and discussion; Section 7 concludes with a conclusion.

II. FEATURE EXTRACTION METHODS

2.1 Wavelet Features

There are many motivations for using features rather than the pixels directly [20]. The most common reason is that feature extraction is used to reduce the dimension of the input data and in turn helps to minimize the training time taken for the classifier. Haar wavelet [11] is widely used technique for the feature extraction, which is single-level one-dimensional wavelet decomposition and gives both an approximation and detailed coefficients. Approximation coefficients which are of size 128x1 are considered for the training of classifier.

2.2. Texture Features

We base our texture feature extraction [16][4][5] on the spatial gray-level cooccurrence-matrix (SGLCM). More specifically, a SGLCM gives a joint histogram of the quantized gray-level value pairs of two image pixels bearing a certain spatial relationship. Eight texture features are calculated i.e., energy; inertia; entropy; homogeneity; maxprob; contrast; inverse; correlation by averaging over four uniformly distributed angular directions, 0^0, 45^0, 90^0 and 135^0 degrees. Total of 32 features are computed for each image as eight features for each angle is considered.

© 2015, IJERMT All Rights Reserved

2.3 Singular Value Decomposition

SVD is a powerful linear algebra technique for solving linear equations in the least-square sense, and works even for singular matrices or matrices numerically close to being singular. The basic information needed to use SVD can be found in [22] and a rigorous mathematical treatment is given in [14]. Singular value decomposition (SVD) can be looked at from three mutually compatible points of view. On the one hand, we can see it as a method for transforming correlated variables into a set of uncorrelated ones that better explain the various relationships among the original data items. At the same time, SVD [18] is a method for identifying and ordering the dimensions along which data points exhibit the most variation. This tie in with the third way of viewing, namely, the best approximation of the original data points using fewer dimensions. Hence, SVD can be seen as a method for data reduction.

2.4. Invariant Moments Features

Invariant Moment feature descriptors were derived from the theory of algebraic invariants and are used to evaluate seven distributed parameters of an image. This technique is chosen to extract image features since the features generated are Rotation Scale Translation (RST) invariant. In any process, the images are processed to extract features that uniquely represent properties of a given category. Invariant moment was successfully applied in texture classification [23]. The set of seven invariant moments ($\phi_1 - \phi_7$) was first proposed by Hu [15] for 2D images. Two-dimensional moments of a digitally sampled M x M image that has gray function f(x,y) (x, y = 0,…,M-1) is given as,

$$m_{pq} = \sum_{x=0}^{x=M-1} \sum_{y=0}^{y=M-1} (x)^p \bullet (y)^q f(x,y)$$

where p, q = 0, 1, 2, 3… (1)

The moments f(x,y) translated by a position (a, b) are defined as,

$$\mu_{pq} = \sum_x \sum_y (x+a)^p \bullet (y+b)^q f(x,y) \qquad (2)$$

Thus the central moments μ_{pq} can be computed from (2) on substituting $a = -\bar{x}$ and $b = -\bar{y}$ where $\bar{x} = \frac{m_{10}}{m_{00}}$ and

$$\bar{y} = \frac{m_{01}}{m_{00}} \text{ as,}$$

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y) \qquad (3)$$

When a scaling normalization is applied the central moments change as,

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \text{ where } \gamma = \left[\frac{(p+q)}{2}\right] + 1 \qquad (4)$$

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned} \qquad (5)$$

In particular, Hu [15] defines seven values, computed by normalizing central moments through order three, that are invariant to object scale, position and orientation. In terms of the normalized central moments, the seven moments are given (5).

III. SUPPORT VECTOR MACHINE

Support vector machine is a relatively new pattern classifier introduced by Vapnik [25]. A SVM classifies an input vector into one of two classes, with a decision boundary developed based on the concept of structural risk minimization (of classification error) using the statistical learning theory. The SVM learning algorithm directly seeks a separating hyperplane that is optimal by being a maximal margin classifier with respect to training data. On the basis of its learning approach, the SVM is believed to have good classification rate for high-dimensional data. Consider the problem of image classification where X is an input vector with 'n' dimensions. The SVM performs the following operation involving a vector $W = (w_1, \ldots, w_n)$ and scalar b:

$$f(X) = \text{sgn}(W \bullet X + b) \qquad (6)$$

Positive sign of f(X) may be taken as 'With Horses' images and negative value of f(X) may be regarded as 'Without Horses' images. Consequently, the optimal

Fig. 1 Positive Test Samples of Inria-Horses-v103 Data Se

Fig. 2 Negative Test Samples of Inria-Horses-v103 Data Set

hyperplane is only determined by the support vectors in the training data. Burges [7] recommends that the average value of b be used in the classification. With this solution, the SVM classifier becomes

$f(X) = \mathrm{sgn}(W \bullet X + b) = \mathrm{sgn}(\sum_{\forall i, \alpha_i > 0} y_i \alpha_i (X_i \bullet X) + b)$ (7) Note that, in (6 and 7), only needs to make use of X_i, y_i and α_i of the support vectors, while X is the input vector to be classified.

Some commonly used kernel functions are: Polynomial function: $K(X_i, X_j) = (X_i . X_j + 1)^d$ (8)

Radial basis function: $K(X_i, X_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}}$ (9)

Sigmoid function $K(X_i, X_j) = \frac{1}{1 + e^{[v(X_i . X_j) - \delta]}}$ (10)

Radial basis kernel function with p1=5 is used for this non-linear classification.

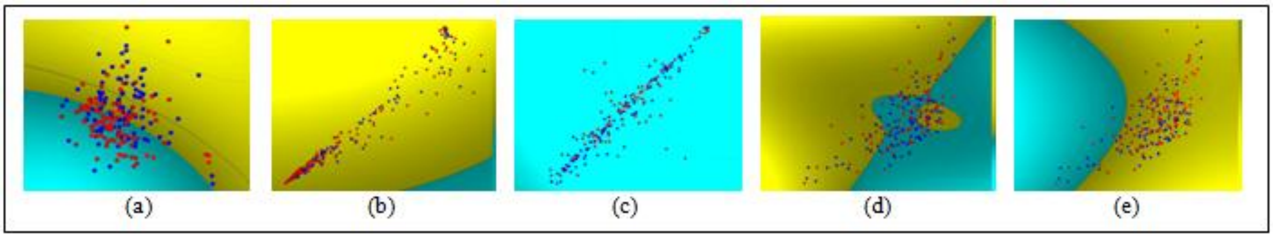

Fig. 3 Classification using a) Singular Value Decomposition b) Co-occurrence Matrix c) Haar Wavelet Features d) Invariant Moments without blocking e) Invariant Moments with blocking factor of 2.

IV. DATA PREPARATION

The sample data [9] used in this work has been taken from CALVIN - INRIA Horse Dataset, University of Edinburg, Calvin. The database consists of images positive samples - having horses and negative samples - not having horses. The sample images of the dataset ar

© 2015, IJERMT All Rights Reserved

Table 1. Detailed Experiment results using various Feature Extraction Methods

Feature Extraction	No. of features	True Pos %	True Neg %	False Pos %	False Neg %	Duration (CPU Time)	Average Classification Rate
Singular value Decomposition	256	71.42	28.57	57.14	42.86	71.18	64.28
invariant moments	7	54.29	45.71	41.42	58.57	39.09	47.855
invariant moments with blocking	28	100	0	96	04	43.56	98
Gray level co-occurrence matrix	32	65.72	34.28	65.72	34.28	180.38	65.72
wavelet feature (haar)	128	71.43	28.57	67.14	32.86	60.54	69.285

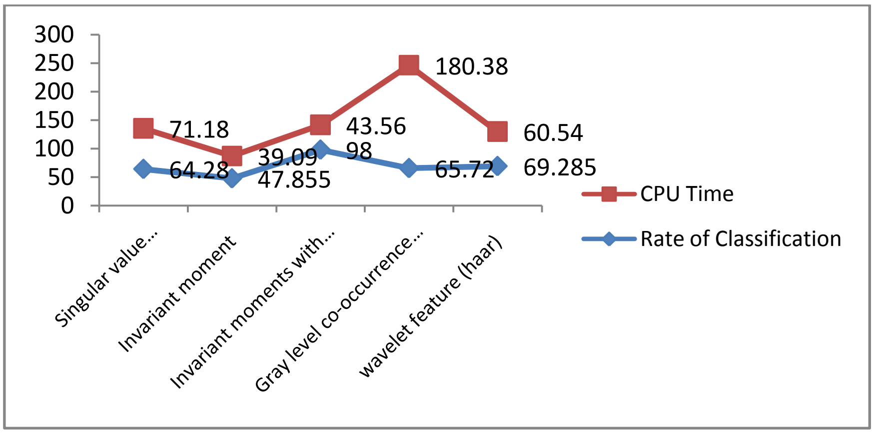

Fig. 4 Performance of a) Singular Value Decomposition b) Invariant Moments without blocking c) Invariant Moments with blocking factor of 2 d) GLCM e) Haar Wavelet Features

Given in Fig 1 and Fig 2. In each classification, 170 samples in total, which consists of 100 samples for the training phase and another set of 70 samples for the testing phase separately without overlapping, are considered for our work.

V. PROPOSED WORK

This paper aims to find the better feature extraction method which suits well for the scene classification problem. Wavelet features, Singular Value Decomposition, invariant moments and Gray level cooccurrence matrix are used to extract features from the images. Support Vector Machine is used for the classification. Support Vector Machine is trained to recognize the scene categories. The sample images of scenes are taken from the Calvin Group [9] which contains 170 samples in each category of with horses and without horses. Kernel Radial basis kernel function is used in SVM for scene classification with p1=5. Results are shown in the Table 1. Each feature extraction method is analyzed with their performances along with their CPU time taken for its execution. The details are clearly depicted in the Table 1 above. The performance too shown diagrammatically to make us understand the efficiency of feature extraction method selected for comparing with blocking factor and without blocking factor along with their reduced CPU time.

VI. IMPLEMENTATION AND DISCUSSION

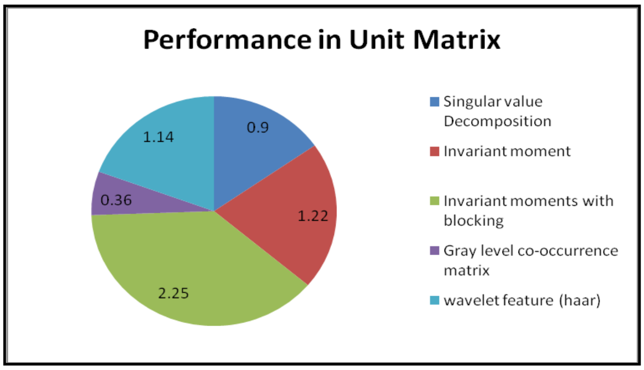

Fig. 6. Performance of FE methods using Unit Matrix

Support Vector Machine is taken as the Classifier, which is used to check the feature extraction methods. In line, SVD gives 64.28% of classification rate with the execution time of 71.18 CPU time having 256 no. of features extracted from (256x256) sized image. Invariant moments, had 7 features, without blocking, gives 47.86 % of classification rate, with

© 2015, IJERMT All Rights Reserved

39.09 CPU time. GLCM gives 65.72 % with 180.38 CPU time. Wavelet (haar) features gives 69.28 % of classification rate with 60.54 CPU time. In all the cases, True Positive and False Positive are considered to calculate the Average Classification Rate as

$$= \frac{\text{True Pos} + \text{False Pos}}{2}$$

The Fig. 4 shows the details of classification and performance. Finally, the invariant method with blocking factor 2, fetches 28 features, improves the classification rate drastically, i.e., it gives 98 % with the CPU time of 43.56. Ultimate goal is the feature selection, done with, calculating the Unit Matrix Percentage whose percentage is represented in Fig. 6.

VII. CONCLUSION

This paper concentrates on the selection of suitable feature extraction methods among the one which are often used by the research community for the scene categorization problems. The overall classification rate of SVD, invariant moments, invariant moments with blocking, GLCM and Haar Wavelet Features is 64.28, 47.85, 98.0, 65.72 and 69.28 respectively. It shows that invariant moment feature with blocking is very suitable than other features for the image categorization problems. This is clearly understood from the Fig. 6 showing the performance of feature methods whose efficiency is calculated via unit matrix considering the CPU time with the classification rate. Blocking proves better in the Moment Features with SVM for the image categorization problem. This work can be further extended to classify other categories of scenes using many other feature extraction methods. Support vector Machine classifier is used for scene categorization with Radial basis kernel function. Support vector Machine is implemented using SVM Toolbox in Matlab 6.5. Work has been done in the Windows 2000 platform with AMD Athlon 64 3000+ CPU Processor with 512 MB RAM using Matlab 6.5. Normalization is programmed using Image Processing Toolbox. Neural classification is done in Neural Network Toolbox. Support Vector Machine with various kernel functions are trained and tested using SVM toolbox for Matlab 6.5.

REFERENCES

[1] Andrew H, "Scene and Position Specificity in Visual Memory for Objects", Journal of Experimental Psychology: Learning, Memory, and Cognition, Vol. 32, No. 1, pp. 58–69, 2006.
[2] Andrew P and Sameer S, "Indoor vs. outdoor scene classification in digital photographs", Pattern Recognition, Vol. 38, pp. 1533-1545, 2005.
[3] Anna B, Xavier M and Robert M, "Which is the best way to organize/classify images by content?", Image and vision computing, Vol. 25, pp. 778-791, 2007.
[4] Arivazhagan S, Ganesan L and Padam S, "Classification using Gabor wavelets based rotation invariant features", Pattern Recognition Letters, Vol. 27, pp. 1976-1982, 2000.
[5] Arivazhagan. S, Ganesan L and Subash Kumar T G, "Texture classification using ridgelet transform", Pattern Recognition Letters, Vol. 27, pp. 1875-1883, 2006.
[6] Awcock G and Thomas R, Applied Image Processing, McGraw Hill International Editions, 1996.
[7] Burges, C.J.C., "A tutorial on support vector machines for pattern recognition", data mining and knowledge discovery, Vol. 2(2), pp. 121-167, 1998. Available from http://svm.first.gmd.de
[8] Boutell M and Luo J, "Beyond pixels: Exploiting camera metadata for photo classification", Pattern Recognition, Vol. 38, pp. 935-946, 2005.
[9] Calvin http://groups.inf.ed.ac.uk/calvin/datasets.html 08.10.2014
[10] Chella A, Frixione M and Gaglio S, "Understanding dynamic scenes", Artificial Intelligence, Vol. 123, pp. 89-132, 2000.
[11] Delorme G, Thorpe M, "Rapid processing of complex natural scenes: A role for the magnocellular visual pathways?", Neurocomputing, Vol. 26, pp. 663-670, 1999.
[12] Florica M, "Moment invariants for recognition under changing viewpoint and illumination", Computer Vision and Image understanding, Vol. 94, pp. 3-27, 2004.
[13] Guerin A and Oliva A, "Classification of scene photographs from local orientations features", Pattern Recognition Letters, Vol. 21, pp. 1135-1140, 2000.
[14] Golub G and Van L, Matrix Computations, The Johns Hopkins University Press, Baltimore, MD, 2nd Edition, 1989.
[15] Hu, M. K., "Visual pattern recognition by moments invariants", IRE Trans. Information Theory, Vol. 8, pp. 179- 187, 1962.
[16] Haralick M, Shanmugam K and Dinstein I, "Texture Features Image Classification", IEEE Transactions on Systems, Man and Cybernetics, SMC-3, Vol. 8, No. 6, pp. 610-621, 1973
[17] Ilkka A and Tapio E, "Flexible view recognition for indoor navigation based on Gabor filter and support vector machines", Pattern Recognition, Vol. 36, pp. 2769-2779, 2003.
[18] Lee K, Ou S and Fang C, "Application of SVD Noise Reduction Technique to PCA based Radar Target Recognition", Progress in Electromagnetics Research, PIER 81, pp. 447 – 459, 2009.
[19] Martin I, "Robust Learning and Segmentation for scene Understanding", PhD Thesis, Dept. of Electrical Engineering and Computer science, Massachusetts Institute of Technology, May 2005.

© 2015, IJERMT All Rights Reserved

[20] Mahi H, Isabaten F, and Serief C, "Zernike Moments and SVM for Shape Classification in Very High Resolution Satellite Images", The International Arab Journal of Information Technology, Vol. 11, No. 1, January 2014
[21] Pietikainen M, Nurmela T, Maenpaa T, Turtinen M, "View-based recognition of real-world textures", Pattern Recognition, Vol. 37, pp. 313-323, 2004.
[22] Press H, Teukolsky A, Vettering T and Flannery P, Numerical Principles in C, Cambridge University Press, Cambridge, England, 2nd Edition, 1992.
[23] Rizon M, "Object Detection using Geometric Invariant Moment", American Journal of Applied Sciences, Vol. 2(6), pp. 1876-1878, 2006.
[24] Singh S, Haddon J, Markou M, "Nearest-neighbour classiers in natural scene analysis", Pattern Recognition, Vol. 34, pp. 1601-1612, 2001.
[25] Vapnik, V.N., "The support vector method of function estimation in Generalization in Neural Network and Machine Learning", Springer-Verlag, New York, NY, pp. 239-268, 1998.
[26] Walker L, Malik J, "When is scene identification just texture recognition?", Vision Research, Vol. 44, pp. 2301–2311, 2004.
© 2015, IJERMT All Rights Reserved