

ROCHESTER INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE

ADVANCED COMPUTER VISION

Natural Scene Classification

Final Report

Authors:

Karina DAMICO

Pauline MARTIAL

Supervisor:

Dr. Roger GABORSKI

February 14, 2013

CONTENTS

I	Introduction	1
II	Overview of the Approach	2
III	Gabor-Gist Descriptor	2
III.1	Gabor Bank Formation	2
III.2	Gabor Histogram Approach	2
III.3	Gabor "Manual-Bin" Histogram Approach	2
III.4	Gabor-GIST Visual Descriptor	3
1	Image Preprocessing	3
2	GIST Feature Extraction	3
3	Classification Results	3
IV	Centrist Descriptor	4
IV.1	Computation of CT Value	4
IV.2	Parameters	4
IV.3	Results with a Neural Network	4
V	Color Histogram	5
VI	Hybrid Descriptors	5
VI.1	Method 1: Voting Algorithm	5
VI.2	Method 2: Neural Network	6
1	Gabor + CENTRIST Hybrid Results	6
2	Gabor + Color Histogram Hybrid Results	6
3	CENTRIST + Color Histogram Hybrid Results	7
4	Gabor + CENTRIST + Color Histogram Hybrid Results	7
VII	Conclusion	7

Abstract—Scene categorization is a fundamental problem in image understanding. It is a challenging task in computer vision and widely applied in many fields, such as image retrieval, video surveillance, travel navigation or medical browser. Large image collections such as the Internet require efficient organization which is why a lot of attempts were performed to design automated systems to solve scene classification task in recent years.

The study presented in this paper aims to obtain high accuracy of scene classification through creating an advanced feature vector based on multiple feature sets extracted using Gabor filters, CENTRIST and color histograms. The goal is to produce feature vectors that describe the scene as accurately as possible. Each feature extraction method is first examined alone, then is combined with features obtained with other methods and tested on accuracy improvement. This paper shows the high performance obtained when combining these techniques.

I. INTRODUCTION

SCENE classification is different from conventional object detection. The scene is a unique composition of several objects related and organized in often unpredictable layout. The relationships between objects and their organization mostly define the scene class. It makes classification task even more difficult, as a computer is unable to "understand" such relationships. Object-based judgement is able to provide more or less satisfactory results for man-made scenes as the object functional representation may itself supply sufficient information for labelling a scene. The situation is much more complicated for classifying natural scenes, where separate object detection does not provide enough information to relate the image to a particular category. High level features stay beyond any existing technique: semantic descriptors are not recognizable by computers. Thus, images have to be described by low level descriptors that aim to mimic semantic content of the image and make it understandable for computers.

In the study presented in this paper, each feature is investigated independently before being combined with other features. Features are evaluated using statistical methods as well as individual classification results. Statistical evidence is gathered primarily from the training set as inclusive or exclusive validation runs. This further provides insight into how representative is the feature for images of the same category. Classification results are generated from an exclusive subset of the training set as well as the test set.

The following of the paper is organized as follows. In section II, our global approach chosen to solve scene recognition problem is explained. Sections III, IV and V respectively detail how Gabor, CENTRIST and color histogram feature descriptors are defined and used to describe the image. Obtained feature vectors are combined in section VI and their performance is evaluated. Section VII concludes this paper and summarizes the best results obtained.

II. OVERVIEW OF THE APPROACH

In order to have a scene classification as accurate as possible, several methods are used and combined. This choice was made with the assumption that because methods have their best efficiency for different classes, combining them should improve the overall accuracy. Results presented in section VI further revealed that this assumption was true. The chosen methods for scene classification are Gabor descriptors (described in section III), CENTRIST feature vectors (described in section IV) and color histogram feature vectors (described in section V). Those three descriptors are very different and use complementary information of the images. This is the key that enable an efficient combination of the feature vectors.

III. GABOR-GIST DESCRIPTOR

Gabor descriptors proved their efficiency as robust texture descriptors. Similarity between low level processing in biological vision and Gabor filter banks was found and proved by Daugman [1]. Thus, Gabor descriptor is the way to make computer "understand" images in the way human brain does it.

The Gabor feature extraction consists of designing a set of Gabor filters in the spatio-frequency domain, and forming a bank of filters to be applied to the images in order to transform information from the pixel space into the feature space. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. The Gabor filters are self-similar: all filters can be generated from one mother wavelet by dilation and rotation [2].

In this study, all the experiments were conducted with single 12 filter Gabor bank, but different techniques to preprocess data and reduce dimensionality were used.

III.1. Gabor Bank Formation

Gabor descriptor for an image is computed by passing the image through the Gabor filter bank. Gabor filter is a linear band-pass filter whose impulse response is defined as a Gaussian function modulated with a complex sinusoid. Mathematically, a Gabor filter of a particular orientation and frequency is a sinusoid modulated by a Gaussian:

$$h(x, y) = s(x, y)g(x, y)$$

where $s(x, y)$ is a sinusoidal plane and $g(x, y)$ is a 2-D Gaussian. The Gaussian is represented by:

$$g(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-1/2(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2})}$$

and $s(x, y)$ as:

$$s(x, y) = e^{-j2\pi(u_0x + v_0y)}$$

The filter has a real and an imaginary component representing orthogonal directions. The two components may be formed into a complex number or used individually [3].

Confusion Matrix						
Output Class	1	2	3	4	5	
	8 5.3%	3 2.0%	3 2.0%	3 2.0%	1 0.7%	44.4% 55.6%
	5 3.3%	21 14.0%	4 2.7%	2 1.3%	2 1.3%	61.8% 38.2%
	4 2.7%	1 0.7%	17 11.3%	3 2.0%	1 0.7%	65.4% 34.6%
	4 2.7%	1 0.7%	2 1.3%	12 8.0%	6 4.0%	48.0% 52.0%
	9 6.0%	4 2.7%	4 2.7%	10 6.7%	20 13.3%	42.6% 57.4%
						Target Class
						1
						2
						3
						4
						5

Fig. 1. Confusion matrix for Gabor Histogram method 1-Coast; 2-Forest; 3-Highway; 4-Mountain; 5-Tall Building

A Gabor filter in 2 dimensions can be used to approximate the frequency and spatial response of the different scene categories. Gabor filters can be viewed as specific edge detectors, which are able to percept edges of a certain orientation and strength. A filter bank consisting of filters with 3 different scales (4, 8 and 16) and 4 orientations (0, 45, 90 and 135 degrees) was considered for feature extraction, giving a total of 12 Gabor filters bank.

III.2. Gabor Histogram Approach

The idea of this method is to calculate intensity histogram for Gabor filtered images. A color image converted into grayscale was split into four horizontal pieces in order to preserve some spatial information and potentially increase the classification accuracy. Each quarter of an image in the training and testing sets was convolved with the filter bank. Intensity histogram was calculated from each piece of the image (32 bin histograms showed similar accuracy to 64 bins). Thus, the feature vector dimensionality is 1x128. Testing was performed in two ways:

- Calculating average histogram for each class from training images, and then comparing it with histograms of testing images using Euclidian distance,
- Using Multilayer Perceptron neural network.

Both classification methods performed similarly with 52% overall accuracy (see figure 1). Note that error rates differ from a category to another category. Because Gabor histogram approach performed very poorly, it was decided to continue experimenting with other approaches.

III.3. Gabor "Manual-Bin" Histogram Approach

As mentioned above, the histogram approach leads to the loss of the spatial information: it only indicates the strength of responses in certain directions and scales. The idea of "manual-bin" histogram approach is to calculate

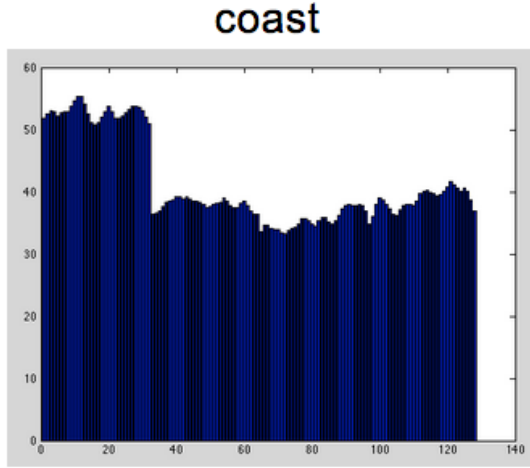


Fig. 2. Averaged histogram for Coast category with Gabor "manual-bin" histogram approach

the value for each bin separately and thus save the response spatial information and intensity.

The method was tested using MLP on 150 testing images and performed even worse than previous approach - misclassification rate of 51%. Overall histogram dimensionality reduction approaches on grayscale images showed unsatisfactory classification accuracy, hence it was decided to leave histogram idea and move to Gabor-GIST descriptor.

III.4. Gabor-GIST Visual Descriptor

4.1 Image Preprocessing

Gabor-GIST feature extraction was performed on color image in contrast to previous approaches. Gabor filter bank (12 filters) was applied to each color plane of RGB images. For better results, color images were first pre-processed before being filtered: images were whitened and contrast was adjusted.

The purpose of whitening is removing pairwise, linear correlations, and thus remove noise and "simplify" an image. Whitening can be done through designing a filter in the frequency domain that will flatten the spectrum of a natural image. Amplitude spectrum of such filter has to rise linearly with frequency, to compensate for the amplitude spectrum of natural images. This can be done through defining a grid of frequency coordinates. Because convolution in the space domain is equivalent to multiplication in the frequency domain, whitening can be performed through multiplying the spectrum of the image with the spectrum of the filter and taking the inverse Fast Fourier Transform of the filtered image (see figure 3.b). Contrast normalization is the standard operation to be done after whitening. For this project, normalization was done by dividing the output of the whitening step by the local luminance variance [3] [4].

4.2 GIST Feature Extraction

In the scope of this project, GIST dimensionality was reduced using an averaging over non-overlapping square

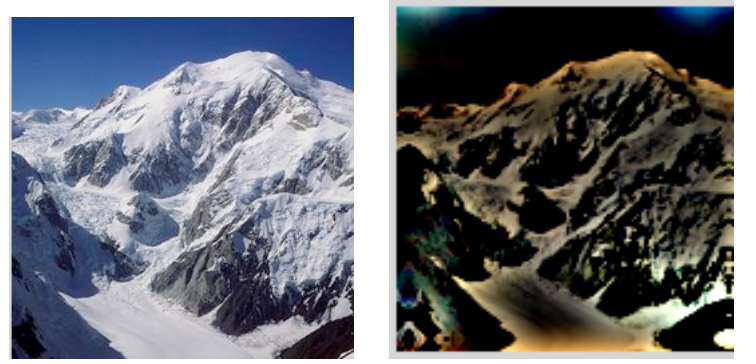


Fig. 3. a) Original Image and b) Whitened Image

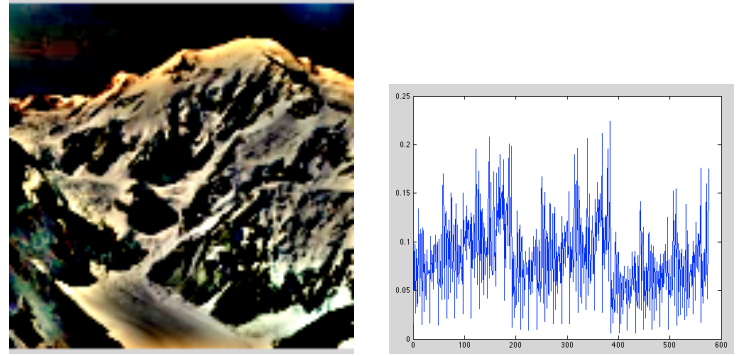


Fig. 4. a) Contrast Normalized Image and b) Gabor GIST visualization

image blocks (for this project, fixed 4x4 grid of sub-regions over the map was considered) [5]. Thus, the size of feature vector is 1x576: 4x4 grid (16 windows where average values is computed) x 12 (filters in the Gabor bank) x 3 (number of color planes for RGB image).

4.3 Classification Results

Similarly to previous approaches, MLP neural network with single hidden layer was considered as a classifier. All images from the training set were converted into 1x576x69 feature matrixes for each category (1x576 feature vector, 69 images in each category in the training set). Then training matrixes were converted into input matrix 576x345. The corresponding target matrix has 5x345 for the training set and 5x150 for the testing set. Testing matrix size is 576x150 (testing set contains 150 images). Training set was divided into three categories: training (70% of training images), validation (15%) and testing (15%). The number of input neurons is equal to the number of features in the feature vector: 576. The number of hidden nodes was found experimentally and the value of 300 was the optimal number of hidden nodes. The number of outputs is the number of categories: 5 classes of scenes in the dataset.

Overall classification accuracy with Gabor-GIST features on 150 training set reached 86%. Error rate differs from category to category, for example Forest images were classified correctly in 100% of cases, when Highway had the low-

Confusion Matrix						
Output Class	1	2	3	4	5	
	26 17.3%	0 0.0%	4 2.7%	1 0.7%	0 0.0%	83.9% 16.1%
	1 0.7%	30 20.0%	3 2.0%	0 0.0%	0 0.0%	88.2% 11.8%
	1 0.7%	0 0.0%	21 14.0%	0 0.0%	1 0.7%	91.3% 8.7%
	2 1.3%	0 0.0%	2 1.3%	23 15.3%	0 0.0%	85.2% 14.8%
	0 0.0%	0 0.0%	0 0.0%	6 4.0%	29 19.3%	82.9% 17.1%
						86.7% 13.3%
						100% 0.0%
						70.0% 30.0%
						76.7% 23.3%
						96.7% 3.3%
						86.0% 14.0%
						1
						2
						3
						4
						5
						Target Class

Fig. 5. Confusion matrix for Gabor-GIST. 1-Coast; 2-Forest; 3-Highway; 4-Mountain; 5-Tall Building

$$\begin{array}{cccccc}
 32 & 64 & 96 & 1 & 1 & 0 \\
 32 & \mathbf{64} & 96 & \Rightarrow 1 & 0 & \Rightarrow (11010110)_2 \Rightarrow CT = 214 \quad (1) \\
 32 & 32 & 96 & 1 & 1 & 0
 \end{array}$$

Fig. 6. CT value calculation

est accuracy - only 70% of correctly classified images. See confusion matrix in figure 5 for more details. Comparably to Gabor histogram approaches, Gabor-GIST over color image performed significantly better. Hence, Gabor-GIST features will be used for the hybrid feature vector creation.

IV. CENTRIST DESCRIPTOR

IV.1. Computation of CT Value

CENsus TRansform hISTogram (CENTRIST), is a visual descriptor used to identify scene categories [6]. The CENTRIST feature vector is perfectly suitable for scene classification because this descriptor encodes the structural properties of a given image without keeping the detailed textural information that would not be helpful for this problem.

The CENTRIST feature vector of a given image is calculated in 2 distinct steps:

1. First, the CT value for each pixel is computed. This is performed by a call to MATLAB function `compare.m`



Fig. 7. CT value calculation for 2 adjacent pixels. Red circles reveal dependencies between CT values: because $36 < 37$ when calculating CT value of green box center pixel, the relation will be $37 > 36$ when calculating CT value of blue box center pixel

for each pixel of the image.

2. Once the CT values of all pixels of the image have been calculated, the histogram of the CT values is computed using function `CENTRIST4lines.m`. This histogram is the CENTRIST feature vector.

The CT value is calculated as follows: for each pixel, the 8 neighbours are compared to the center pixel value and are assigned value 0 if they are greater than the pixel value, 1 if they are smaller (see figure 6). Concatenating all 8 bits give a binary number with value between 0 and 255. This is the CT value. Figure 7 shows the obtained CT values for two adjacent pixels. It reveals dependencies between CT values: the right neighbour of a pixel for a given 3x3 pixels square will be the left neighbour for the adjacent 3x3 square. Thus if the right neighbour pixel value is greater (for example in figure 7, $37 > 36$), then the left neighbour will be smaller for the next CT value computation ($36 < 37$).

IV.2. Parameters

The MATLAB function `CTfeatureVector.m` creates the feature vector CENTRIST for all training and testing images and saves data in a `.mat` file. In order to obtain the best performance, different parameters have to be tested. The following parameters have been used for the rest of the project since they produced the smallest misclassification rate:

- Downsizing the image to a ratio of 0.125. This reduces the computation time and keeps the relevant information of the image.
- Dividing the image into 4 lines seemed to be the most efficient: instead of considering the image as a whole entity, it is divided into 4 lines. The CENTRIST vector is computed for each line and the 4 CENTRIST vectors are concatenated to obtain the CENTRIST vector of the image. This enable to keep some of the spatial information in the image that can help classifying scenes.
- Histogram Intersection Method. In order to compare the CENTRIST feature vector of a tested image to the mean of the CENTRIST vectors of images belonging to a given class, the measure of difference must be calculated. This can be done using the Euclidian distance or the Histogram Intersection Method. The Histogram Intersection Method generated much better classification than when using the Euclidian distance.
- 40 bins are used for the histograms. This is the best compromise between the classification rate, the computation time and the size of the generated CENTRIST feature vector.

IV.3. Results with a Neural Network

When fed into a neural network, the CENTRIST method gives a misclassification rate of 14% (see figure 8) which is a correct value for a single classification method. Note that even if the overall misclassification rates for Gabor-GIST

Confusion Matrix						
Output Class	1	2	3	4	5	
	23 15.3%	0 0.0%	5 3.3%	0 0.0%	0 0.0%	82.1% 17.9%
	2 1.3%	28 18.7%	0 0.0%	2 1.3%	0 0.0%	87.5% 12.5%
	4 2.7%	0 0.0%	25 16.7%	4 2.7%	0 0.0%	75.8% 24.2%
	1 0.7%	1 0.7%	0 0.0%	23 15.3%	0 0.0%	92.0% 8.0%
	0 0.0%	1 0.7%	0 0.0%	1 0.7%	30 20.0%	93.8% 6.2%
						Target Class
						1 2 3 4 5

Fig. 8. Confusion matrix for CENTRIST features. 1-Coast; 2-Forest; 3-Highway; 4-Mountain; 5-Tall Building

(figure 5) and for CENTRIST (figure 8) are both 14%, the misclassification rates vary for each category. This means that the method used will affect each category's misclassification rate. For example the Tall Building category has a misclassification rate of 0% with the CENTRIST descriptor and 3.3% with Gabor-GIST. At the opposite, the Coast category is much better classified if Gabor-GIST is used (13.3% instead of 23.3% with CENTRIST).

V. COLOR HISTOGRAM

In scope of scene classification problem, color descriptors cannot be considered reliable for the reason that colors are the matter of constant change. The same scene appears in different colors depending on the time of the day or the season for example.

A lot of papers still offer color descriptors in the combination of texture descriptors and overall experiments show that such combinations improve the performance of classification.

For this project, a simple color histogram descriptor was implemented. The histogram is computed by counting the number of pixels corresponding to the following colors: red, yellow, green, cyan, blue, magenta, white and black. It is performed first through converting the image to HSV color space and evaluating each plane in order to determine the number of pixels belonging to each of the selected color categories. The number of black and white pixels are estimated through Saturation and Value planes, and the rest of color categories are determined through Hue plane. Every pixel color category is then normalized by dividing the count of the pixels by the number of pixel in the image. A 8-bin color histogram formats 1x8 feature vector (see figure 9.b).

Classification accuracy for color histogram features was evaluated using MLP Neural Network (8 input, 8 hidden, 5 output neurons) using 150 testing images from five differ-

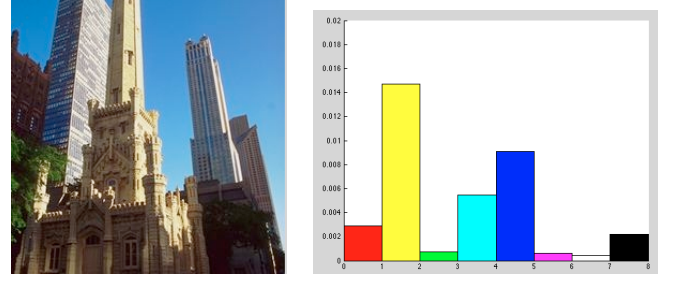


Fig. 9. a) Tall Building Image and b) Color Histogram Image Representation

Confusion Matrix						
Output Class	1	2	3	4	5	
	4 2.7%	0 0.0%	4 2.7%	6 4.0%	2 1.3%	25.0% 75.0%
	6 4.0%	25 16.7%	2 1.3%	0 0.0%	1 0.7%	73.5% 26.5%
	8 5.3%	4 2.7%	19 12.7%	3 2.0%	7 4.7%	46.3% 53.7%
	6 4.0%	0 0.0%	0 0.0%	11 7.3%	4 2.7%	52.4% 47.6%
	6 4.0%	1 0.7%	5 3.3%	10 6.7%	16 10.7%	42.1% 57.9%
						Target Class
						1 2 3 4 5

Fig. 10. Confusion matrix for Color Histogram features. 1-Coast; 2-Forest; 3-Highway; 4-Mountain; 5-Tall Building

ent categories of scenes. Classification error for color histogram descriptor is 50%, i.e. half of images were misclassified (see figure 10). Color histogram features performed the worst for Coast category images (only 13% accuracy), the best performance is seen for Forest images (83% accuracy). Experiments proved that color cannot be used as a reliable feature for scene categorization problem, but as mentioned before, our study is concerned with improving classification accuracy by adding color features to Gabor-GIST and CENTRIST features and forming the hybrid feature vector that is expected to provide the minimum misclassification rate.

VI. HYBRID DESCRIPTORS

VI.1. Method 1: Voting Algorithm

In order to combine the different feature vectors, a voting algorithm has been coded in the MATLAB function `votingAlgo`. The file `testVotingAlgoFunction.m` can be used to test the function using Gabor feature vector, Gabor-GIST descriptor and CENTRIST feature vector, and display the output. It can also run the tests several times and give an average performance. For the voting algorithm, training images are split randomly into a training

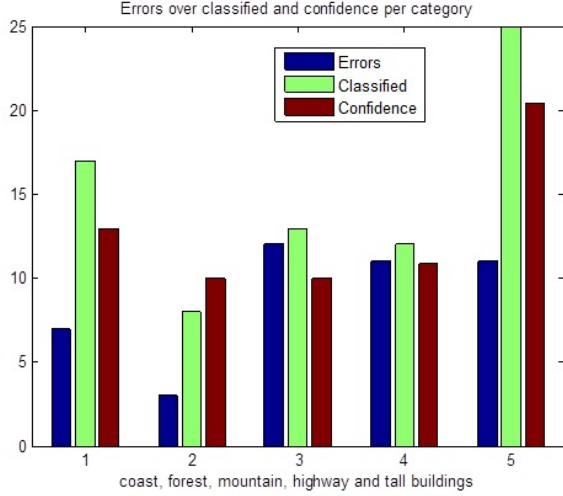


Fig. 11. Performance for the voting algorithm: results obtained for classification of 15 images of 5 different classes

and a testing set with a chosen percentage. The results for the voting algorithm were not as good as expected. Indeed, figure 11 shows that a lot of images are misclassified, knowing that there should be 15 images classified in each category.

The following of the paper details the classical classifier, the neural network, that has been trained with the feature vectors in order to solve the scene classification problem.

VI.2. Method 2: Neural Network

The first step that needs to be done in order to create the hybrid feature vector is data scaling. Features extracted with each approach have to be equally weighted. Scaling values of vector $b = [b_{min}, b_{max}]$ in the range of vector $a = [a_{min}, a_{max}]$ values is performed using equation (1).

$$y = (x - b_{min}) \frac{a_{max} - a_{min}}{b_{max} - b_{min}} + a_{min} \quad (1)$$

where x is the initial value of vector b and y the rescaled value of b in the range of a .

Figure 12 shows the effect of applying equation (1) to Gabor-GIST and CENTRIST feature vectors. This will later enable the neural network to take them both into account with same weight.

In previous stages, input matrices for each descriptor were built and targets stayed the same for the same dataset. In order to create the hybrid vector of features, all that needs to be done is appending one matrix to another ([features1; features2]) and train the MLP network on the new combined input matrix. The same procedure applies for testing input matrix. In the following sub-sections, different combinations of features are analysed and classification accuracy is evaluated.

2.1 Gabor + CENTRIST Hybrid Results

The confusion matrix produced by testing the images based on Gabor-GIST and CENTRIST descriptors is shown

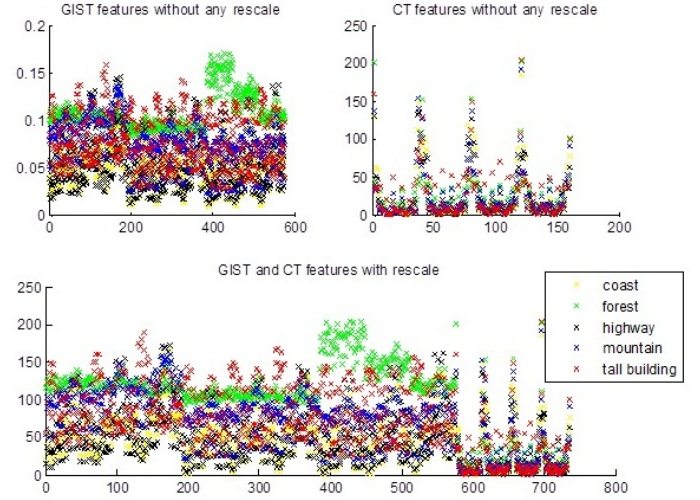


Fig. 12. a) Gabor-GIST and CENTRIST without rescale and b) Gabor-GIST and CENTRIST after rescale

Confusion Matrix						
Output Class	1	2	3	4	5	
	28 18.7%	0 0.0%	3 2.0%	1 0.7%	0 0.0%	87.5% 12.5%
	1 0.7%	30 20.0%	0 0.0%	0 0.0%	0 0.0%	96.8% 3.2%
	0 0.0%	0 0.0%	25 16.7%	1 0.7%	0 0.0%	96.2% 3.8%
	1 0.7%	0 0.0%	2 1.3%	27 18.0%	0 0.0%	90.0% 10.0%
5	0 0.0%	0 0.0%	0 0.0%	1 0.7%	30 20.0%	96.8% 3.2%
						Target Class
						1 2 3 4 5

Fig. 13. Confusion matrix for Gabor + CENTRIST hybrid features. 1-Coast; 2-Forest; 3-Highway; 4-Mountain; 5-Tall Building

in figure 13. It indicates potential improvement from combining the features together (if mutual classification compensation occurs). On practice, significant classification error disease occurred, and overall accuracy raised from 86% (figure 5) to 93.3%(figure 13). Features compensated each other deficiencies and didn't create any noticeable confusion.

2.2 Gabor + Color Histogram Hybrid Results

Our team was pretty sceptical about combining color histograms with features extracted with other methods, as color histograms performed poorly. The guess was that color histogram features would add more confusion than do any good. However, experiments demonstrated slight performance improvement from 86% accuracy (Gabor-GIST only) to 89.3% (Gabor-GIST and color histogram). Gabor-GIST and color histogram hybrid improved the correct

Confusion Matrix						
Output Class	1	2	3	4	5	
	27 18.0%	0 0.0%	1 0.7%	0 0.0%	0 0.0%	96.4% 3.6%
	1 0.7%	30 20.0%	0 0.0%	0 0.0%	0 0.0%	96.6% 3.2%
	0 0.0%	0 0.0%	29 19.3%	1 0.7%	0 0.0%	96.7% 3.3%
	2 1.3%	0 0.0%	0 0.0%	27 18.0%	0 0.0%	93.1% 6.9%
	0 0.0%	0 0.0%	0 0.0%	2 1.3%	30 20.0%	93.8% 6.2%
	1	2	3	4	5	
Target Class						
	90.0% 10.0%	100% 0.0%	96.7% 3.3%	90.0% 10.0%	100% 0.0%	95.3% 4.7%

Fig. 14. Confusion Matrix for Gabor-GIST + CENTRIST + Color histogram hybrid features. 1-Coast; 2-Forest; 3-Highway; 4-Mountain; 5-Tall Building

classification rate for Coast, Highway and Mountain categories. However, performance decreased for Forest and Tall Building scene classes.

2.3 CENTRIST + Color Histogram Hybrid Results

Adding the color histogram descriptor to CENTRIST also improved the classification accuracy. Improvement occurred for Coast, Forest and Mountain categories but a performance decrease was observed for Highway and Tall Building scene classes. Not that for both Gabor-GIST and CENTRIST, adding color histogram feature vector improves Coast and Mountain correct classification and reduces Tall building classification. However, the behaviour differs for Highway and Forest images.

2.4 Gabor + CENTRIST + Color Histogram Hybrid Results

Finally, all the features extracted with the three methods studied in this work (Gabor-GIST, CENTRIST and color histogram) were combined. As expected, this is the combination that performed the best. Compared to Gabor-GIST + CENTRIST hybrid, the classification accuracy increased from 93.3% (figure 13) to 95.3% (figure 14). Significant improvement occurred for Highway category (83.3% category accuracy to 96.6%), and slight accuracy decrease was demonstrated in Coast category (93.3% to 90% class accuracy).

VII. CONCLUSION

This paper shows that combining Gabor-GIST descriptor with CENTRIST and color histogram feature vectors is particularly efficient for scene classification. Indeed, the accuracy obtained using a Neural Network as a classifier is above 95% for the provided dataset of images. This very high rate of correctly classified images is made possible by

the complementarity of the methods used: because Gabor-GIST, CENTRIST and color histogram have their maximum efficiency for different categories, the overall accuracy is increased when they are combined.

Finally, we can say that this combination of three descriptors is very interesting and competitive when compared to other existing methods that can be found in scene classification papers.

REFERENCES

- [1] J. Daugman, "Complete discrete 2-d gabor transforms by neural networks for image analysis and compression," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 36, no. 7, pp. 1169–1179, 1988.
- [2] D. Dunn, W.E. Higgins, J. Wakeley, "Texture segmentation using 2-d gabor elementary functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 2, 1994.
- [3] P. Kruizinga, N. Petkov, S.E. Grigorescu, "Comparison of texture based features based on gabor filters," *IEEE Conf. on Image Analysis and Processing*, 1999.
- [4] A.K. Jain, F. Farrokhnia, "Unsupervised texture segmentation using gabor filters," *Pattern Recognition*, vol. 24, 1991.
- [5] C. Siagian, L. Itti, "Rapid biologically-inspired scene classification using features shared with visual attention," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, Feb 2007.
- [6] Jianxin Wu, James M. Rehg, "Centrist: A visual descriptor for scene categorization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, August 2011.