Indoor Outdoor Scene Classification in Digital Images

[1]Shwetha T.M., [2]Shaila H Koppad

[1,2]Department of MCA Programme, Jnanabharathi Campus, Bangalore University, Bangalore-56, India
Email: [1]shwethamahesh678@gmail.com, [2]shaila.koppad@gmail.com

Abstract — In this paper, we present a method to classify real-world digital images into indoor and outdoor scenes. Indoor class consists of four groups: bedroom, kitchen, laboratory and library. Outdoor class consists of four groups: landscape, roads, buildings and garden. Application considers real-time system and has a dedicated data-set. Input images are pre-processed and converted into gray-scale and is re-sized to "128x128" dimensions. Pre-processed images are sent to "Gabor filters", which pre-computes filter transfer functions, which are performed on Fourier domain. The processed signal is finally sent to GIST feature extraction and the images are classified using "kNN classifier". Most of the techniques have been based on the use of texture and color space features. As of date, we have been able to achieve 80% accuracy with respect to image classification.

Keywords—Indoor, Outdoor, Image Classification, Scene.

I. INTRODUCTION

Image processing is a mechanism to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input could be an image like video frame or a photograph and output may be image or characteristics associated with that image.

Usually, image processing system includes treating images as two dimensional signals while applying already set signal processing methods to them. It is among the rapidly growing technologies in the current business world, with a wide spectrum of implementation to help resolve real-world issues. Image processing basically includes the following three steps:

i. Importing the image with optical scanner or by digital photography.
ii. Analyzing and manipulating the image which includes data compression, image enhancement and recognizing patterns that are not visible to human eyes similar to satellite photographs.
iii. Output is the last stage in which result can be altered image or report that is based on image analysis.

A. Purpose of Image Processing

The purpose of Image processing is divided into 5 groups:

i. **Visualization** – Observe the objects that are not visible
ii. **Image Sharpening and restoration** – to create a better image.
iii. **Image retrieval** – seek for the image of interest
iv. **Measurement of pattern** – Measures various objects in an image.
v. **Image Recognition** – Distinguish the objects in an image.

B. Types of Image Processing

The two types of methods used for Image processing are:

i. Analog Image Processing
ii. Digital Image Processing

Analog or visual techniques of Image processing can be used for the hardcopies like print-out and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. The Image processing is not just confined to area that has to be studied but on knowledge of analyst. Association is another important tool in image processing through visual techniques. So, it is essential for image analysts to apply a combination of personal knowledge and collateral data to image processing.

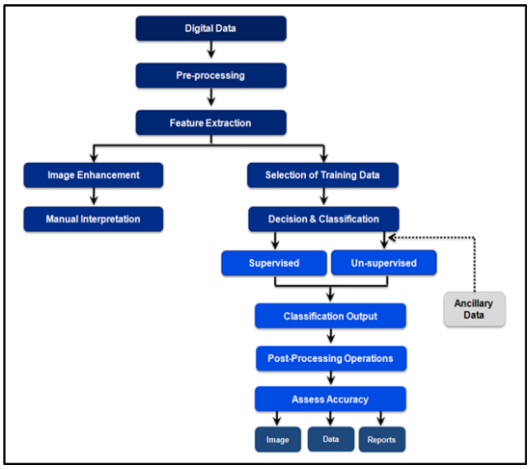


Figure 1: Block diagram for simple Image processing scenario

Digital processing techniques help in manipulation of the digital images by using computers. As raw data provided by image sensors from satellite platform contains deficiencies. To get over such flaws and to get originality of information, it has to undergo various phases of processing. The three general phases that all types of data have to undergo while using digital techniques: pre-processing, enhancement and display, information extraction.

Generally, a learning-based approach is used to resolve 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 human tendency. 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 is 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 prediction about successive inputs.

There are basically two types of approach in scene recognition. The first one is the top-down approach where we consider only the global features without segmentation and processing of the objects. The second is a bottom-up approach, where we get into detailed analysis of color, object and texture. We chose to opt for the first one as we can easily collect the information needed on a low resolution spatial configuration. In this method, we don't process the objects and classify scene by collecting its global information.

II. PROBLEM DEFINITION

Through this work, we propose techniques based on edge-analysis. Our approach towards the problem is based on finding the straightness of the edge and based on the decision we are classify the images as indoor or outdoor images. We also consider this application in real-time system. Practically we are developing our own dataset.

III. METHOD

In this paper, we classify the digital images as indoor or outdoor by using many image processing techniques such as:

- Gabor Filtering
- GIST Feature Extraction
- kNN Classification

Figure – 2 represents the flowchart which depicts the dataflow of this project. The input images are processed, in the pre-processing the image is converted into grayscale and is re-sized to 128*128 dimensions. These pre-processed images are sent to Gabor filters. The filtered digital images is sent to GIST feature extraction for further processing and is sent to kNN classifier for classification.

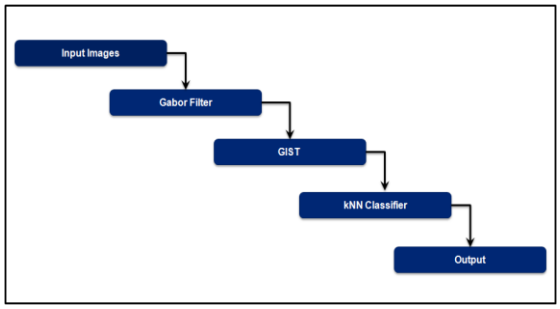


Figure 2: Data flow diagram for image classification

A. Creating a Gabor Filter

The local spectrum is established through features which are obtained by filtering with a set of 2 dimensional Gabor filter. It acts as a local band-pass filter with certain optimal joint localization properties in both the spatial domain and the spatial frequency domain. Typically, a multi-channel filtering scheme is used: an image is filtered with an a set of Gabor filters with different preferred orientations and spatial frequency, which covers appropriately the spatial frequency domain and the features which are obtained from a feature vector field which is used further. Gabor feature vectors can be used directly as input to a classification or segmentation operator or they can first be transformed into new feature vectors.

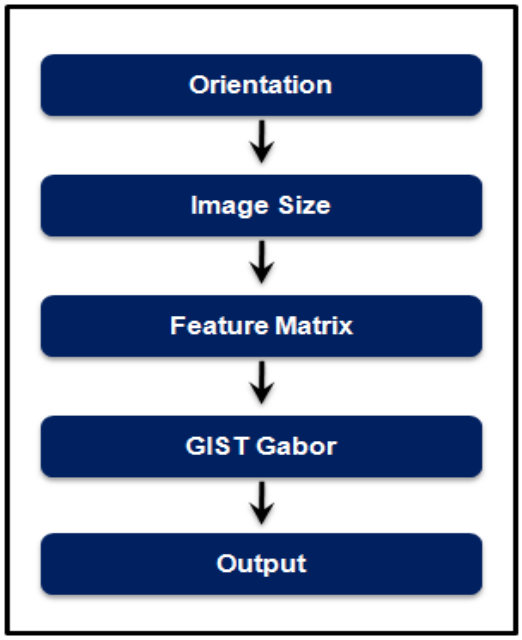


Figure 3: Data flow of the Model.

Gabor filter pre-computes filter transfer functions. The process signal is sent to GIST feature extraction.

B. GIST Descriptor

The GIST descriptor describes the spatial layout of an image using global features derived from the spatial envelope of an image. It is shown to be an efficient scene categorization mechanism. In implementation, they calculate GIST descriptors of two variants of the original image. The first variant is the re-sized version (128 x 128) and the second one is the square size of the center of the image. The reason is that both full scene and the focused scene which is usually at the center. Re-size the image for smaller computation cost. Generally, it is advanced features which are needed though we still need to further investigate on the independence of each feature. This is validated by the fact that the selective combination using only five features gives better performance than the total combination of features. The approach tends to prefer common concepts to the uncommon ones, thus, leaving some concepts totally undetected. This is because we use kNN where the algorithm assigns the most common concepts of the (k) nearest neighbors to the test image. Therefore, the selection K is important in kNN Classifier.

C. kNN Classifier

The k-nearest neighbor algorithm is a method for classifying objects based on closest training samples in the feature space. kNN is a type of instant based learning or lazy learning, where the function is only approximated locally and all computation is de-referred until classification. The k-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms; an object is classified by a majority vote of its neighbor, with the object being assigned to the class most common amongst its k-Nearest neighbor (k is a positive integer typically small). If k is 1, then the object is simply assigned to the class of its nearest neighbor. The same method can be used for regression, by simply assigning the property value for the object to be the average of the values of its k-nearest neighbors. It can be useful to weigh the contributions of the neighbors, so that the nearer neighbors contribute more to the average than more to the distant ones.

We have GIST feature matrix as an input to the kNN classifier and the number of images per class as parameter along with the suitable k-value. We tried iterating the process for different values of k and for the different member of images per class. At the end we found k=2 gave the best accuracy results. Based on the result we did critical analysis and were able to classify an image as an Indoor or outdoor image.

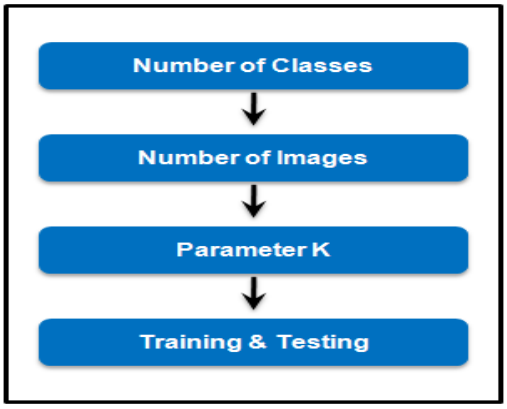


Figure 4: Data flow of kNN Classifier

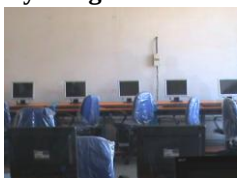
D. Dataset

This work facilitates the classification of scene as an indoor or outdoor image. We have assumed 8 classes of which 4 are Indoor classes and 4 are outdoor classes. They are as follows:

1. Indoor Classes

i. Library Images


ii. Laboratory Images

iii. Kitchen Images


iv. Bedroom Images


2. Outdoor Classes

i. Building Images

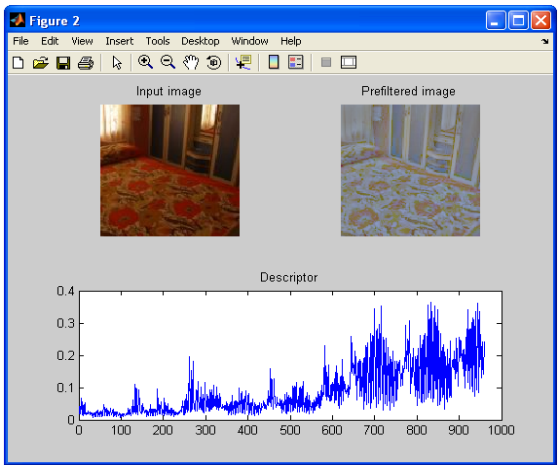

ii. Road Images

iii. Garden Images

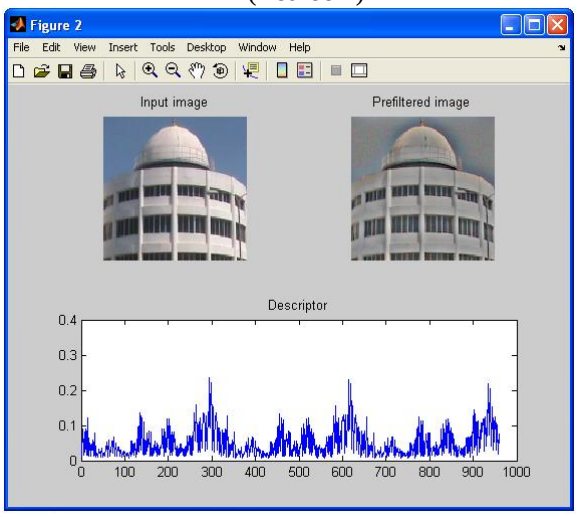
iv. Landscape Images

IV. EXPERIMENTAL RESULTS

This section provides a glimpse of the experimental results which was carried out as part of this work.



Snapshot 1: Pre-filtered image for an Indoor class (Bedroom)



Snapshot 2: Pre-filtered image for an Outdoor class (Building)

V. DISCUSSION

Image is a concept of presenting information in a best possible way. Image processing is a subject where images are analyzed based on the objects present in image. In this work, problem of scene analysis has been taken up where the Image is broadly classified into Indoor and Outdoor Images based on the feature present in them. This is achieved through a combination of features from two different types of multi resolution and multi channel filters, in general it provides superior classification of textured images. A combination of Gabor and GIST features may be sufficient for segmentation over wide variety of digital images.

To reduce the complexity of classification we considered eight (8) categories where in four (4) categories are Indoor and four (4) categories are Outdoor respectively, for the experiments on analyzing the classification of images. The efficiency of the method can be improved by combining it with filter bank approach. It was observed that accuracy responses varied for different parameter such as k value, orientation and image size. The experimental results indicate that Gabor feature and GIST feature are quite robust. Further, rotation and scale variance is important in many applications and our preliminary results on rotation invariant classification using Gabor features are very encouraging.

VI. CONCLUSION

Scene classification in itself is a complex scenario. On top of it, selection of features is much more complex. Classification becomes a very complicated problem because of the constant changes which happen in everyday life and the changes in their corresponding images. We may need to understand the human vision in depth, in order to improve computer vision on the same lines as human's capability to process abundant information in less than a fraction of a second. Till date many research studies have been carried out on image classification, but none have been able to give complete accuracy. We took many excerpts from such research works and decided to arrive at Gabor Feature vectors and find the gist of the image. We opted for a tree classification and we optimize the value 'k'. The results are satisfactory. We would like to continue this research work by applying color and texture feature vectors.

ACKNOWLEDGMENT

We are grateful to Prof. Muralidhara B.L., Co-ordinator, MCA Programme, Bangalore University, for his timely suggestions and encouragement

REFERENCES

[1] Oliva A., Torralba A., Guerin-Dugue A., and Herault J., "Global semantic classification of scenes using power spectrum templates," CIR99, Elect. work in ComputingSeries, Springer-Verlag, Newcastle. 1999, 1999.

[2] Fei-Fei L. and Perona P., "A bayesian hierarchical model for learning natural scene categories," Proceedings of the 2005 IEEE Computer Society, 2005.

[3] Lazebnik S., Schmid C., and Ponce J., "Beyond bags of features: Spatial pyramid matching for recognising natural scene categories," CVR-TR-2005-04, 2004. 22

[4] Torralba A. and Oliva A., "Semantic organisation of scenes using discriminat structural templates," Proceedings of International Conference on Computer Vision ICCV99 Korfu Greece, pp. 1253–1258, 1999.

[5] Julia Vogel , Bernt Schiele , 2004, 'A semantic typicality measure for natural scene categorization', Pattern Recognition Symposium, DAGM

[6] N. Serrano, A. Savakis and J. Luo, "Improved scene classification using efficient low-level features and semantic cues", Pattern Recognition, (2004), in press.

[7] J. Luo, A. Savakis, "Indoor vs. outdoor classification of consumer photographs using low-level and semantic features", Proceedings of the International Conference of Image Processing, Thessaloniki, Greece, (2001).

[8] M. Szummer and R. Picard, "Indoor-outdoor image classification", IEEE International Workshop on Content-Based Access of Image and Video Databases, ICCV'98, (1998).

[9] Le Saux B. and Amato G., "Image classifier for scene analysis," Computer Vision and Graphics International Conference, ICCVG 2004, Warsaw, Poland, September 2004.

[10] Torralba A., Oliva A., 1999, 'Semantic Organisation of scenes using discriminat structural templates', Proceedings of International Conference on Computer Vision, ICCV99, Korfu, Greece, 1253-1258

[11] G. Nathalie, L. B. Herve, H. Jeanny and G. D. Anne, "Towards the Introduction of Human Perception in a Natural Scene Classification System", IEEE Conference NNSP, Switzerland, (2002) September 4-6.

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

[13] 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.

[14] Oliva A. and Schyns PG, "Diagnostic colors mediate scene recognition," Cognitive Psychology, vol.41, pp. 176–210, 2000.

[15] Oded Maron, Aparna Lakshmi Ratan, 1998, 'Multiple-Instance Learning for Natural Scene Classification', Proceedings of the Fifteenth International Conference on Machine Learning, 341 – 349

[16] A. Vailaya, M. Figueiredo, A. K. Jain and H. J. Zhang, "Image classification for content-based indexing", IEEE Transactions on Image Processing, vol. 10, (2001), pp. 117-130.

[17] P. Fitzpatrick, "Indoor/outdoor scene classification project", http://www.mit.edu/ people/paulfitz/pub/indoor-outdoor.pdf.

[18] S. Srinivasan and L. Kanal, "Qualitative Landmark Recognition using Visual Cues", Pattern Recognition Lett., vol. 18, (1997), pp. 1405-1414.

[19] Y. Su and F. Jurie, "Improving Image Classification Using Semantic Attributes", International Journal of Computer Vision, vol. 100, no. 1, (2012), pp. 59-77.

[20] N. Serrano, A. Savakis and J. Luo, "A computationally efficient approach to indoor-outdoor scene classification", 16th International Conference on Pattern Recognition, vol. 4, (2002), pp. 146-149.

[21] W. Kim, J. Park and C. Kim, "A Novel Method for Efficient Indoor–Outdoor Image Classification", Signal Processing Systems, vol. 61, no. 3, (2010), pp. 1-8.

◈◈◈