Object Recognition using Multiresolution Transforms

Ahila Priyadharshini. R

Dept. of Electronics and Communication Engineering,
Mepco Schlenk Engineering College, Sivakasi – 626 005, Tamil Nadu, India

Abstract: **Recognizing the objects in images is a challenging task due to the presence of occlusion, clutter, variation in shape, scale, color, illumination, position and size of objects in an image. In this paper, the potential efficiency of mutiresolution transforms such as Ridglet transform and Log Gabor transform for the Object Recognition task is investigated. To classify objects from images,local features such as patches are extracted over the interest points detected from the original image using Wavelet based interest point detector. Then Ridgelet features and Log Gabor features are computed for each and every patch. Then these features are trained, tested and classified using SVM classifier. The experimental evaluation of proposed method is done using the Graz01 database.**

Key Words: **Object Recognition, Salient Point, Patch, Log Gabor features, Ridgelet Features.**

I. INTRODUCTION

Object recognition is of greater task in computer vision. It is the task of finding an object in image or video sequence. In case of humans, recognition task is much easier, that is they recognize millions of objects with little effort even when the objects look different in different circumstances. They could also recognize objects that are partially obstructed from the view. It is still a challenging task for computer systems to recognize objects that show different appearances in different surroundings [1].

Global features describe image as a whole and are less successful in recognition. Salient points are the points which maximize the discrimination between the objects. Salient point detection plays an important role in content based image retrieval in order to represent the local properties of the image. Since classic corner detectors cannot support natural images, detector based on wavelet transform represents global variations and local ones to detect the salient points [2, 3]. Schmid and Mohr (1997) proposed Local gray invariants for Image Retrieval, where local gray invariants are automatically extracted over the detected salient points [4]. Weber et al (2000) proposed the computation of K-means clustering algorithm at Forstner points for object recognition [5].

In order to address the scale difference of the objects, the patches have to be extracted at different scales. Moreover, the occlusions of images can easily be handled by these patches [6].Arivazhagan and Ahilapriyadharshini (2015) extracted Gabor features over the patches extracted from the images for the object recognition task [7]. The work done on object recognition using Gabor wavelet and SVM classifier is proposed by Shen and Zhen (2008) where Gabor wavelet features with 5 scales and 8 orientations are calculated and classified using SVM classifier for face recognition [8].

A lot of work has been done previously using ridgelet transform for various applications such as Iris Recognition, Face Recognition etc [9].In all the above applications, global features are extracted using ridgelet transform. Here, Ridgelet transform and Log- Gagor transform are applied to image patches for recognizing various kinds of object categories.

The paper is structured as follows. The next section deals with feature extraction using Ridgelet transform and Log- Gagor transform. Section 3 discusses the results for bject recognition task. Finally Section 4 gives the conclusion of the proposed method.

II. FEATURE EXTRACTION

The transformation of input data into set of features is called feature extraction. Thereduced information that is set of features instead of full size input is used to recognize various complex images with better accuracy. The kinds of features used here are Ridgelet features, and Log-Gabor features

A. Ridgelet Transform

The representation ofobjects with line singularities is a special form of Ridgelet Transform when compared to wavelet transforms [10].The Ridgelet transform can be represented in both continuous and digital domain. Given an integrable bivariate function $f(x)$, the Continuous Ridgelet Transform (CRT) of $f(x)$ is defined in Eqn. (1)

$$R(a, b, \theta) = \int \varphi_{a,b,\theta}(x) f(x) dx \qquad (1)$$

where $\varphi(a, b, \theta)$ is the Ridgelets and is given in Eqn.(2)

$$\varphi_{a,b,\theta}(x) = a^{-1/2} \varphi(x_1 cos\theta + x_2 sin\theta - b/a) \qquad (2)$$

Here, φ is smooth univariate function with sufficient decay and satisfying the admissibility condition.

Ridgelet function is oriented at angles θ, and is constant along the lines, i.e. $x_1 cos\theta + x_2 sin\theta = constant.$

As a sequence, wavelets are very effective in representing objects with isolated point singularities, while ridgelets are very effective in representing objects with singularities along lines(Do& Martin Vetterli 2003). In 2-D, points and lines are related via the Radon transform, thus the wavelet and ridgelet transforms are linked via the Radon transform.

The Radon transform of an object f is the collection of line integrals indexed by (θ, t) ϵ [0,2π] is given in Eqn. (3)

$R_f(\theta, t) = \int f(x_1 x_2)\, \delta(x_1 cos\theta + x_2 sin\theta - t) dx_1 x_2$(3)

where δ is the dirac distribution.

Then the Ridgelet transform is the application of the 1-D wavelet transform to the slices of the Radon transform and can be represented in Eqn.(4)

$$R(a, b, \theta) = \int R_f\,(\theta, t) \varphi_{a,b}(t) dt \qquad (4)$$

In this work, the ridgelet-based features are extracted for every patch detected overthe image. Each patch image is decomposed into number of sub-bands using digital ridgelet transform. In case of multiscale ridgelets the plane is subjected to an infinite series of partitions, based on dyadic scales, where each partition consists of squares of the given dyadic side length. The structure of ordinary Ridgelet Transform (Qiao et al 2010) using this theory is shown in Fig.1.

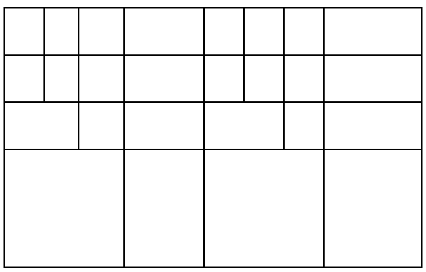


Fig. 1 Structure of Ordinary Ridgelet Transform

BLog Gabor Transform

Gabor filter is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function. It is optimally localized as per the uncertainty principle in both the spatial and frequency domain. This implies Gabor filters can be highly selective in both position and frequency, thus resulting in sharper texture boundary detection [12]. However, they have two main limitations. The maximum bandwidth of a Gabor filter is limited to approximately one octave and Gabor filters are not optimal if one is seeking broad spectral information with maximal spatial localization.

An alternative to the Gabor function is the Log-Gabor function proposed by Field [13]. Log-Gabor filters can be constructed with arbitrary bandwidth and the bandwidth can be optimized to produce a filter with minimal spatial extent. Field suggests that natural images are better coded by filters that have Gaussian transfer functions when viewed on the logarithmic frequency scale. Gabor functions have Gaussian transfer functions when viewed on the linear frequency scale. On the linear frequency scale the log-Gabor function has a transfer function as shown in Eqn.(5).

$$G(w) = e^{\left(-\log\left(\frac{w}{w_o}\right)^2\right)/\left(2\log\left(\frac{k}{w_o}\right)^2\right)} \qquad (5)$$

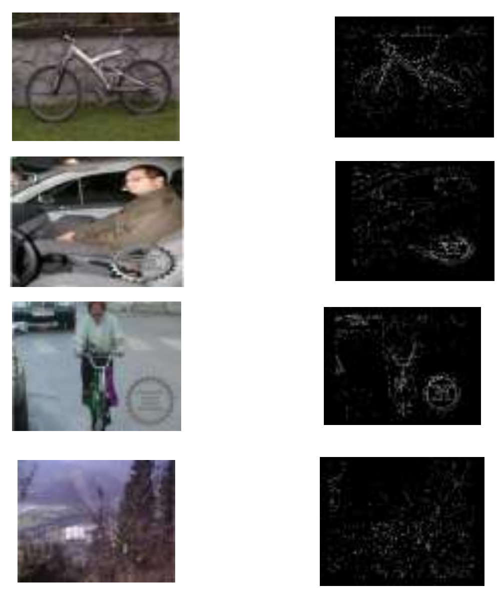
where w_o is the filter's centre frequency, $\frac{k}{w_o}$is a constant.

There are two important Characteristics of Log Gabor functions. (i) Log-Gabor functions, by definition, always have no DC component, which contributes to improve the contrast ridges and edges of images. (ii) The transfer function of the Log-Gabor function has an extended tail at the high frequency end, which enables us to obtain wide spectral information with localized spatial extent and consequently helps to preserve true ridge structures of images. In this article, Log Gabor filters with 4 scales and 6 orientations are used. Then mean and standard deviation are computed for every filtered patch.

III. RESULTS AND DISCUSSIONS

Here experiments are conducted with the images of Graz01databasesfor the two class problem. Both object images and background images are used for training and testing. The task is to determine whether an object is present in a given image or not. There are 373 bike images, 460 person images, 210 both bike and person images and 270 mixed background images in Graz01 database.In Graz01 database, the images are highly complex with high intra-class variability in scale, view point, color, location and illumination. There is much background clutter in the image.

Initially 200 most prominent points are taken using wavelet based salient point detector for every image. The patch of size 32×32 is extracted over each detected salient point. Fig.2 shows the Saliency Map for sample images of Graz01 database.



(a) (b)

Fig. 2 (a) Categories in the Graz01 Database (b) Saliency Map

Then Ridgelet transform is applied for each and every patch. Every patch is decomposed into 16 numbers of sub-bands. Mean and Standard deviation are computed for every decomposed sub bands and classified using SVM classifier. The kernel used here is RBF Kernel. In addition to the Ridgelet features, Log gabor features are also extracted and classified. To extract Log Gabor features, Log Gabor filters with 4 scales and 6 orientations are used which results in 24 filtered images. Then mean and standard deviation are computed for every filtered patch.The experiment is carried with 100 randomly selected images for training and the remaining images for testing. The performance of the object recognition system is measured in terms of the Error rate as given in Eqn (6).The performance measure is given in Table1.

$$Error\ Rate = \frac{FN+FP}{TP+FN+TN+FP} \times 100 \quad (6)$$

where

TP- Number of True Positives,

FN- Number of False Negatives

TN- Number of True Negatives

FP-Number of False Positives

Table 1: Performance Measure of Graz01 Database

Category	No of Images used for Training		No of Images used for Testing		TP		FN		TN		FP		Error Rate (%)	
	PI	NI	PI	NI	RT	LG	RT	LG	RT	LG	RT	LG	RT	LG
Bike	100	100	273	170	164	208	109	65	112	144	58	26	37.69	**20.54**
Person	100	100	360	150	251	279	109	81	158	141	12	29	22.83	**20.75**
Bike & Person	100	100	110	150	90	90	20	20	131	149	39	21	21.07	**14.64**

PI-Positive Images, NI- Negative Images, RT-Ridgelet features, LG- Log Gabor features

In Table 1, the better performance measure obtained is shown in bold. The lesser the error rate, the better the performance will be.In this experimentation, Log gabor transform outperforms the Ridgelet transform.

IV. CONCLUSION

In this work, the potential efficiency of mutiresolution transforms such as Ridglet transform and Log Gabor transform for the object recognition task is investigated. The extracted features based on Log Gabor transformperforms better in challenging Graz01databasewhere objects suffer from severe occlusions, high intra-class variability in scale, and view point reflecting real world scenes more accurately.

REFERENCES

[1] Daniel Martin Keysers (2006), Modeling of Image Variability for Recognition, RWTH Aachen University, Germany.

[2] NicuSebe, Michael S. Lew (2003),“ Comparing salient point detectors”, Pattern Recognition Letters, Leiden Institute of Advanced Computer Science, Leiden University, 24, pp.89–96.

[3] Loupias E, Sebe N, Bres S, Jolion J M (2000), “Wavelet-Based Salient Points for Image Retrieval”, In International Conference on Image Processing, vol.2, pp.518-521.

[4] Schmid C, Mohr R (1997), “Local Gray Value Invariants for Image Retrieval”, In Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 530-534.

[5] Weber M, Welling M, Perona P (2000), “Unsupervised Learning of Models for Recognition”, In Proceedings of the Sixth European Conference of Computer Vision, Springer-Verlag, London.

[6] Thomas Deselaers, Daniel Keysers, and Hermann Ney (2005), “Improving a Discriminative Approach to Object Recognition using Image Patches”, In proc. of DAGM.

[7] S. Arivazhagan, R. AhilaPriyadharshini (2015), Generic visual categorization using composite Gabor and moment features, Optik 126 (2015) 2912–2916.

[8] Shen Lin-Lin, JI Zhen (2008), “Gabor Wavelet Selection and SVM Classification for Object Recognition”, ACTA AUTOMATICA SINICA, vol.35, No.4, pp. 350-355.

[9] SatyajitKautkara, RahulkumarKocheb, TusharKeskarc, AniketPanded, MilindRanee, Gary A. Atkinsona (2010),Face Recognition Based on Ridgelet Transforms, Procedia Computer Science (ICEBT 2010), 35–43.

[10] ShadiAlZubi, Naveed Islam and MaysamAbbod (2011), Multiresolution Analysis UsingWavelet, Ridgelet,andCurvelet Transforms for Medical Image Segmentation, International Journal of Biomedical Imaging, 18pages.

[11] Yu-Long Qiao , Chun-Yan Song and Chun-Hui Zhao (2010), M-band ridgelet transform based texture classification, In Pattern Recognition Letters, 31(3) , 244-249.

[12] Ma L., Wang Y. and Tan T. (2002) ‘Iris Recognition Based on Multi-channel Gabor Filtering’, Proceedings of ACCV’2002, vol. I, pp279-283.

[13] Field D. (1987) ‘Relations between the Statistics of Natural Images and the Response Properties of Cortical Cells’, Journal of the Optical Society of America A, Vol.4, No.12, pp2379-2394