Indoor Object Recognition System using Combined DCT-DWT under Supervised Classifier

R. Arunkumar
Assistant Professor
Department of Computer Science and Engineering
Annamalai University

M. Balasubramanian
Assistant Professor
Department of Computer Science and Engineering
Annamalai University

S. Palanivel
Professor
Department of Computer Science and Engineering
Annamalai University

ABSTRACT
The objective of this proposed work is to recognize the real time small indoor objects from the any scene or image of our working environment for visually impaired. This will be efficiently detect and recognize the indoor objects. The objects are detected and segmented automatically by exploiting the geometrical properties of the image regions. Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT) and Combined DCT-DWT are implemented and evaluated for extracting features from the segmented object. In the Training phase, more than hundred objects were used for each category of the objects and dimension reduction of the features has been done for better result. The performance of the object recognition for visually impaired is evaluated along with the corresponding feature selection methods. The performance of the recognition system gives the recognition rate of 94.44% with the usage of Combined DCT – DWT.

Keywords
DCT, DWT, Combined DCT - DWT, Visually Impaired and SVM.

1. INTRODUCTION
Visual object recognition is a fundamental, frequently performed cognitive task. Generally, different recognition tasks [1], including categorization and identification. In order to obtain accurate recognition results, various feature extraction techniques are implemented and evaluated. Following this design, early work in computer vision is attempted to achieve recognition of objects under supervised classifier [11][12] that directly operate over low-level image features such as texture and shape[6].

In this paper, an object recognition system is established on the basis of the various extracted features. As mentioned above, an efficient method is introduced for location of object regions and the features are the main characteristic for object distinction. After extracting the entrant regions of likely target, we convert the object recognition task to region sets categorization problem, which can be solved by SVM (Support Vector Machine) classifier.

1.1 OUTLINE OF WORK
The remainder of this paper is organized as follows. Section 2 introduces the key principles and main procedure of locating object and segmentation process. Section 3 introduces the various feature extraction methods of our object recognition. Section 4 introduces the role of training data with SVM Classifier. Several experimental results and detailed explanations are provided in Section 5. Our concluding annotations and future work are enclosed in Section 6.

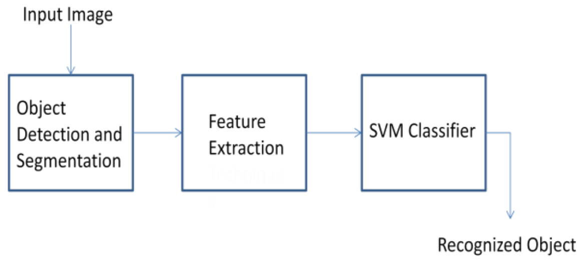


Fig.1. Recognition system of the Indoor Objects

2. OBJECT CROPPING TECHNIQUES
Cropping method using this MATLAB processing is an automatic method to spot the region of interest from the given input image easily.

2.1 CROPPING ALGORITHM
The following steps involved in cropping objects from an image

i. Read the input image with multiple objects.
ii. Convert RGB image to Gray Scale image.
iii. Convert the Gray scale image into Binary image.
iv. Morphologically open binary image (remove small objects)
v. Label the connected components in 2-D binary image
vi. Measure the properties of image regions such as the Area, Centroid, and Bounding Box to locate the objects.
vii. Crop the each objects in an input image under the bounding box.

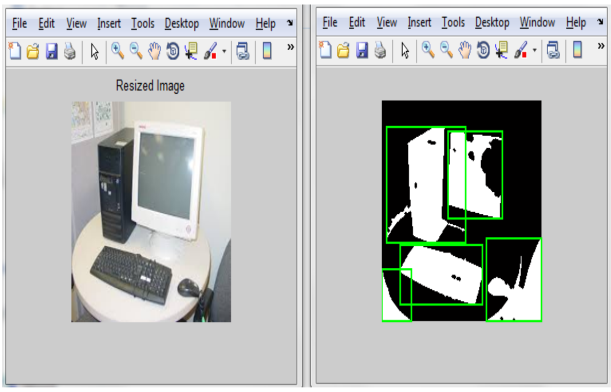


Fig. 2. Input Image and Bounding box measurements of the objects

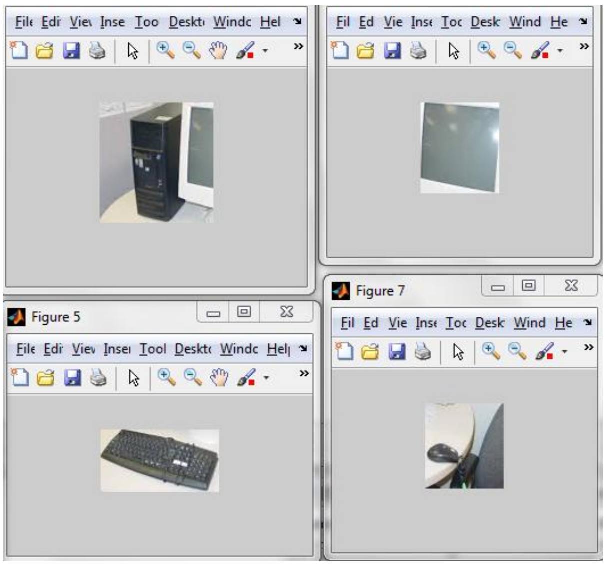


Fig 3. Cropped objects from the input image

3. INDOOR OBJECT FEATURE EXTRACTION

Feature Extraction is a dimensionality reduction method of simplifying the amount of data required to depict a large set of data. Feature extraction is a method of constructing combinations of the values of variable still describing the data with sufficient accuracy.

3.1 Discrete Wavelet Transform (DWT)

The Discrete Wavelet Transform is based on sub-band coding and it is found to yield a fast computation of Wavelet Transform. DWT is simple to implement and reduces the computation time [4][5].

The DWT is computed by successive low pass and high pass filtering of the discrete time-domain signal as shown in figure 4. and the signal is denoted by the sequence x[n], where n is an integer. The low pass filter is denoted by g[n] while the high pass filter is denoted by h[n]. At each level, the high pass filter produces detail information, while the low pass filter associated with scaling function produces coarse approximations. In the filtering or decomposition process, ↓2 represents down sampling by 2.

An image can be decomposed into a sequence of different spatial domain images using DWT. Decomposition of 2D image can be performed ensuring different frequency bands namely, LL, LH, HL and HH as shown in figure 5.

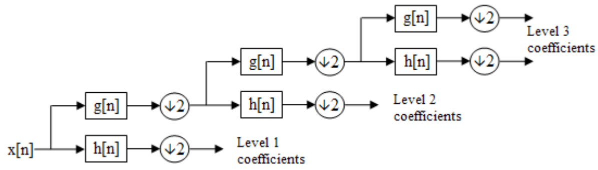


Fig. 4. DWT process

These are also known by other name, [cA, cH, cV, cD]DWT coefficients such as cA is approximation coefficients, cH is horizontal coefficients, cV is vertical coefficients and cD is diagonal coefficients.

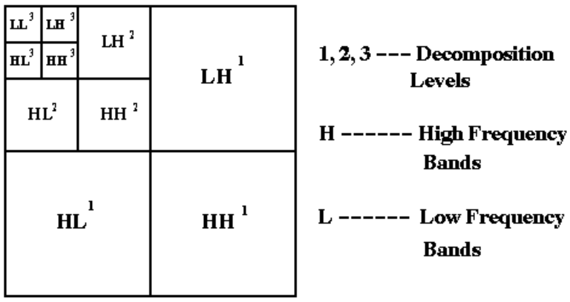


Fig. 5. 2-D DWT with 3 Level Decomposition

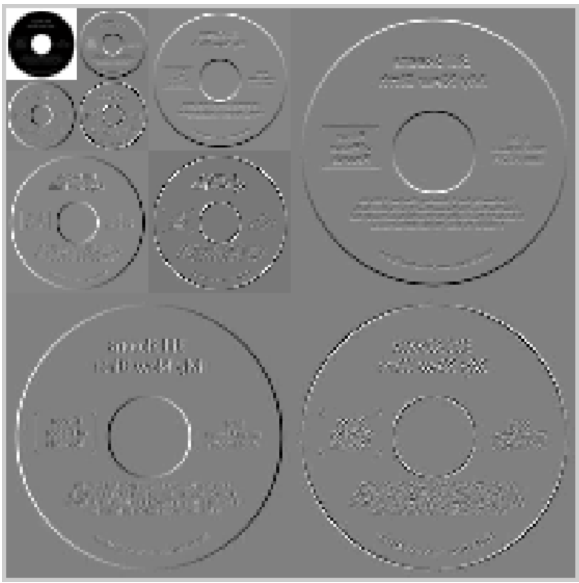


Fig. 6. 2-D DWT with 3 Level Decomposition of object 'CD'

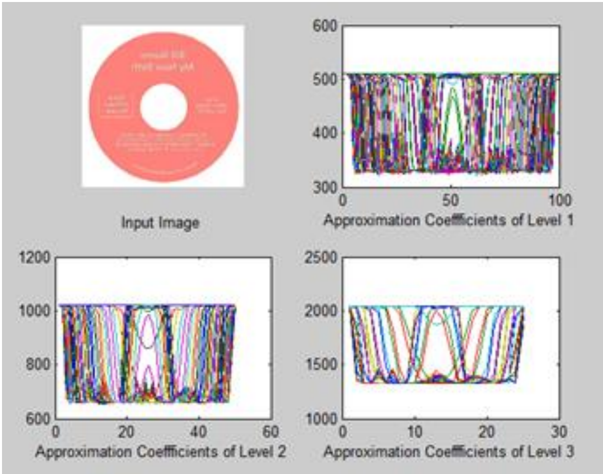


Fig. 7. 2-D DWT Approximation Coefficients of 3 Level Decomposition of Object 'CD'

We performs a Three-level two-dimensional wavelet decomposition with respect to a Daubechies wavelet and 20 features has been chosen from the final level DWT Approximation coefficients of the input image

3.2 Discrete Cosine Transform (DCT)

The discrete cosine transform is a separable linear transformation and this 2 D DCT is same as the 1-D DCT performed along a single dimension followed by a 1-D DCT in the other dimension. The description of the 2-D DCT for an input image x and output image X is expressed as

$$X(m,n) = u(m)u(n)\frac{2}{N}\sum_{i=0}^{M-1}\sum_{j=o}^{N-1} x(i,j)cos\frac{(2i+1)m\pi}{2M}cos\frac{(2i+1)n\pi}{2N},$$

$$for\ m = 0,1,2,\dots,M-1, n = 0,1,2,\dots,N-1.$$

$$\text{where}\ \ u(m), u(n) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } m, n = 0 \\ 1 & \text{otherwise} \end{cases}$$

DCT is extensively used in image compression [5], because of its outstanding energy compaction property. Plenty of frequency component of the activity acceleration data are centralized at the low-frequency. Most of the visually significant information is concentrated in just a few DCT coefficients. Therefore, the high-frequency DCT coefficients are discarded and the low-frequency DCT coefficients are selected as activity features.

In this paper, we extract the first twenty highest energy of DCT coefficients from each axis acceleration data for features. selected as activity features.

3.3 Combined features of DCT and DWT

The Combined procedure is explained below and also depicted in figure 9.

Step 1: Apply DWT to decompose the input image into four non overlapping multi-resolution sub-bands: LL1, HL1, LH1, and HH1.
Step 2: Again apply DWT again to sub-band LL1 to get four small sub-bands and choose the LL2 sub-band as shown in Fig. 8.
Step 3: Choose only the first 20 features of LL5 obtained from the five level of decomposition
Step 4: Apply DCT to the given cropped input image and choose the first 20 highest energy coefficients as a dimension reduction.
Step 5: Combine selected features of DCT and DWT together for each image.
Step 6: Finally, from each image 40 features has been selected for training phase.

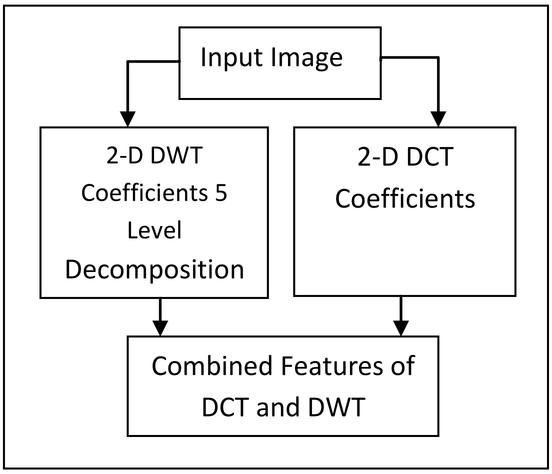
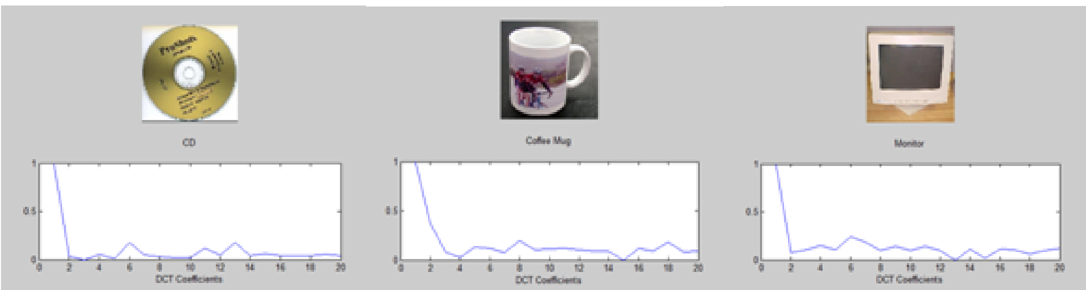


Fig. 9. Combination of DCT and DWT coefficients



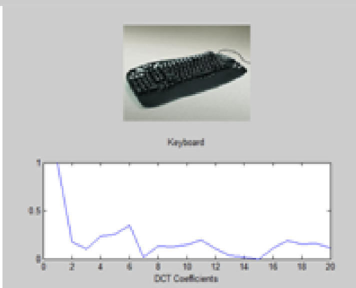


Fig. 8. DCT Coefficients of Objects such as 'CD', 'Coffee Mug', Monitor', 'Keyboard'

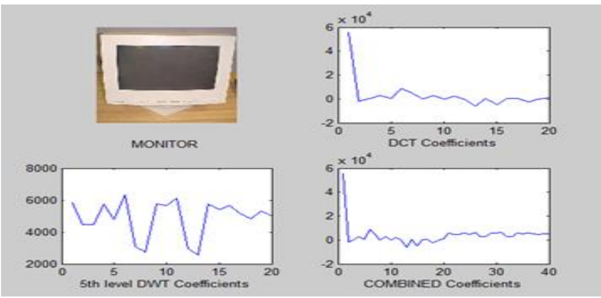


Fig. 10. Combined DCT - DWT Coefficients

4. OBJECT RECOGNITION USING SVM

Support vector machine is based on the principle of structural risk minimization, it is a classification model. Support vector are used to find hyper plane between two classes. Support vector are the training samples that define the optimal separating hyper plane. Support vectors are close to hyper plane. Like RBFNN, support vectors machine can be used for pattern classification and nonlinear regression

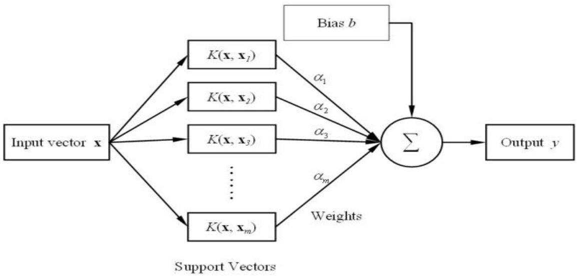


Fig. 11. SVM process

For liner separable data, SVM finds a extrication hyper plane, which split up the data with the major margin. For linearly inseparable data, t maps the data in the input space into a high dimension space $Z \rightarrow \Phi(X) \in R$ with kernel function Φ (X) to find the separating hyper plane. SVM was originally developed for two class classification problem.

$$Z=\Phi(x)=\{x_1^2, x_2^2, \sqrt{2}x_1, x_2\}$$

Support vector machine map input vectors to a higher dimensional space where a maximal separating hyper plane is conclusion. Two parallel hyper planes are construction on each side of the hyper plane, separates the data. The separating hyper is the hyper plane that maximizes the distance between the two parallel hyper planes. Each data points will represented by a p-dimensional vector. Each of these points belongs to only one or two classes. There are number of direct classifiers that might fulfill this property. The nearest distance between points is one separated hyper plane is exploited. if such a hyper plane exists. It is evidently of concern and is known as the maximum or minimum margin classifier.

$$Y = \begin{Bmatrix} n, & if \;\; d_n(X)+t>0 \\ 0, & if \;\; d_n(X)+t>0 \end{Bmatrix}$$

$$Where \;\; d_n(X) = \max\{di(X)\}_{i=1}^{n}$$

$d_n(X)$ is the distance from x to the SVM hyper plane corresponding d_i the object I, the classification threshold is t. and the class label y=0 stand for unknown. During testing, the distance from x o the SVM hyper plane is used o identify the objects. Inner product kernel maps input space to higher dimensional feature space inner product kernel $K=(x, x_i) = \Phi(x). \Phi(x_i)$, where x is the input pattern, is the support vectors.

5. EXPERIMENTAL RESULTS

The performance of the Object recognition system is evaluated using indoor objects in laboratory environment.

5.1 Object Recognition using DCT and SVM

SVM is trained to distinguish DCT features of a particular type of objects (+1) and DCT features of the other types of objects (-1) in the training set, such that, one SVM model is created for each type of objects. For training, 20 DCT features are extracted from each object and is given as input to SVM model. For Testing, 20 DCT features of a test object are given as input to SVM model and the distance between each of the feature vectors and the SVM hyper plane is obtained. The recognition of a test object is decided based on the maximum distance.

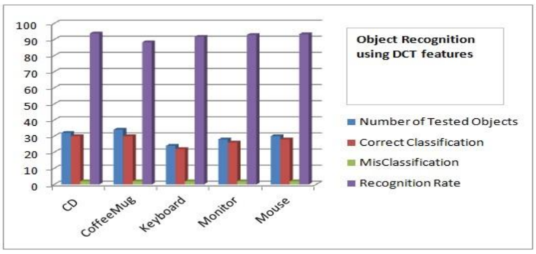


Fig. 12. Performance chart for Object Recognition using DCT features

5.2 Object Recognition using DWT and SVM

SVM is trained to distinguish DWT features of a particular type of objects (+1) and DWT features of the other types of objects (-1) in the training set, such that, one SVM model is created for each type of objects. For training, 25 DWT features are extracted from each object and is given as input to SVM model. For Testing, 25 DWT features of a test object are given as input to SVM model and the distance between each of the feature vectors and the SVM hyper plane is obtained. The recognition of a test object is decided based on the maximum distance.

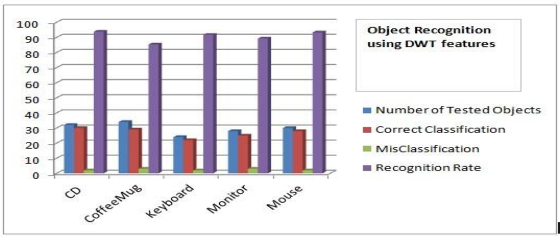


Fig. 13. Performance chart for Object Recognition using DWT features

5.3 Object Recognition using Combined DCT-DWT and SVM

SVM is trained to distinguish combined DCT & DWT features of a particular type of objects (+1) and combined DCT & DWT features of the other types of objects (-1) in the training set, such that, one SVM model is created for each type of objects. For training, 40 combined DCT & DWT features where 20 from each, are extracted from each object and is given as input to SVM model. For Testing, 40 combined DCT & DWT features of a test object are given as input to SVM model and the distance between each of the feature vectors and the SVM hyper plane is obtained. The recognition of a test object is decided based on the maximum distance.

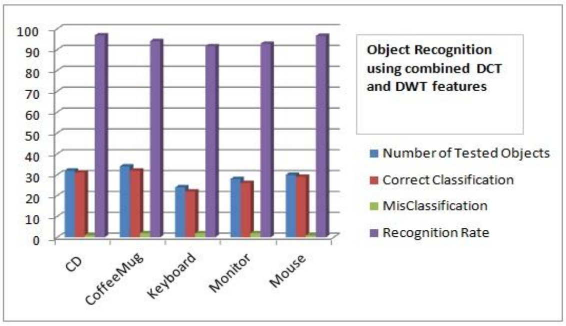


Fig. 14. Performance chart for Object Recognition using combined DCT & DWT features

An image processing method further comprising the step of displaying a result of the difference determination step and the diagnosis data, wherein in the display phase, when no region of interest is acquired from the diagnosis information, the analyzed data demonstrating that there is no region of interest is displayed.

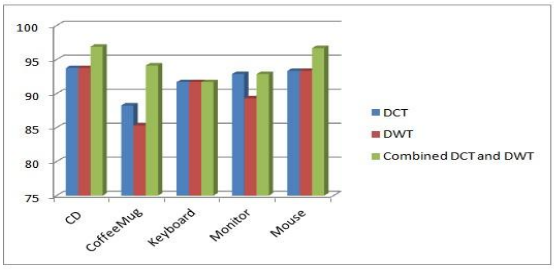


Fig. 15. Overall performance graph for Object recognition using DCT,DWT and Combined DCT&DWT.

6. CONCLUSION

In this study, the indoor object is detected by identifying various region of interest using the region properties. Different types of indoor objects are segmented from the input scene or image and DWT, DCT and Combined DCT-DWT methods are applied for best fit selection of features in order to amend the recognition performance. Using the SVM, recognition of various objects is experimented with the corresponding features and results are reported that Combination of DCT-DWT features provide better than the individuals in recognition of the above mentioned indoor objects. Our System characterized by combined DCT-DWT features report higher recognition rate of 94.44%.

7. REFERENCES

[1] Dominik Maximilián Ramík, Christophe Sabourin, Ramon Moreno, Kurosh Madani, “A machine learning based intelligent vision system for autonomous object detection and recognition” Elsevier, Applied Intelligence, August 2013

[2] Zhi-Yong Meng, Ping-Ping Yu, Guo-Qing Yu,” Copyright protection for digital image based on joint DWT-DCT transformation” Wavelet Analysis and Pattern Recognition (ICWAPR), 2012 : 15-17 July 2012 Page(s): 11 - 14

[3] Deselaers, T. , ETH Zurich, Ferrari, V. “ Global and efficient self-similarity for object classification and detection”, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE 13-18 June 2010 Pages:1633 - 1640

[4] Wahi, A, Ravi, P. ; Saranya, M.,” A neural network approach to rotated object recognition based on edge features: Recognition rate and CPU time improvement for rotated object recognition using DWT” Computing Communication and Networking Technologies (ICCCNT), 2010 , 29-31 July 2010, Pages 1 - 6.

[5] Yong Yang, “A Novel DWT Based Multi-focus Image Fusion Method”, Elsevier, Procedia Engineering, Volume 24, International Conference on Advances in Engineering 2011, Pages 177–181

[6] Y. Q. Chen, L.Q. Chen, H.J. Gu, K. Wang. “Technology for multi-focus image fusion based on wavelet transform.” In Proceedings of the Third International Workshop on Advanced Computational Intelligence; 2010, p. 405-408.

[7] B.P. Rinky, Payal Mondal, K. Manikantan, S. Ramachandran “DWT based Feature Extraction using Edge Tracked Scale Normalization for Enhanced Face Recognition” Procedia Technology, Volume 6, 2012, Pages 344–353

[8] Chih-Hsien Hsia, Jing-Ming Guo, Jen-Shiun Chiang “A fast Discrete Wavelet Transform algorithm for visual processing applications”,Elsevier,Signal Processing,Volume 92, Issue 1, January 2012, Pages 89–106

[9] Kasmani, S.A. , Naghsh-Nilchi, A. , A New Robust Digital Image Watermarking Technique Based on Joint DWT-DCT Transformation” Convergence and Hybrid Information Technology, 2008. ICCIT '08. (Volume:2) 11-13 Nov. 2008, Pages 539 – 544

[10] Huimin Qian, Yaobin Mao, Wenbo Xiang, Zhiquan Wang” Recognition of human activities using SVM multi-class classifier” ELSEVIER, Pattern Recognition Letters. Volume 31, Issue 2, 15 January 2010, Pages 100–111

[11] Zaidi, N.A., Squire, McG.D., “Local Adaptive SVM for Object Recognition”, Digital Image Computing: Techniques and Applications (DICTA), 2010 International Conference on 1-3 Dec. 2010 ,Pages: 196 – 201

[12] Maji, S., Berg, A.C. ; Malik, J.,”Efficient Classification for Additive Kernel SVMs”, Pattern Analysis and Machine Intelligence, IEEE Transactions (Volume:35 , Issue: 1) Jan. 2013, pp 66 – 77.