



An Exploration of Neural Networks & Transformation Techniques for Image Classification

K. Mahantesh*

Department of ECE, SJBIT
Bengaluru, India

V. N. Manjunath Aradhya

Department of MCA, SJCE
Mysuru, India

Abstract— *The problem of facilitating machine to learn & detect patterns of objects and categorizing them in the rapidly expanding image databases has become increasingly challenging for image retrieval systems. In this regard, we propose transform domain techniques to identify the line singularities on the edges and low frequency components present in an image signal to generate highly discriminative geometrical features preserving color and texture information which are analogous to human vision perceptual models. Firstly, image is segmented based on color homogeneity using k-means clustering in hybrid color space. Later, transform techniques such as Ridgelets and Coslets are derived by applying mono-dimensional wavelets on 2-D radon and DCT sub-bands respectively to analyse pixel distribution of segmented regions in multi-resolution framework. The obtained ridgelet and coslet coefficients are projected on to a lower dimension feature space using PCA to obtain compressed and de-correlated feature vectors. For classification purpose, we use five different distance measures and neural networks such as GRNN and PNN to obtain an average recognition rates. The demonstration of proposed techniques on very large scale image datasets such as Caltech-101 & Caltech-256 has resulted in producing progressive classification rates in comparison to recent biologically inspired transformation and subspace based models.*

Keywords— *Image Classification; Feature Extraction; DCT; Radon; Ridgelet; Coslet; PCA; Distance Measures (DM); GRNN; PNN.*

I. INTRODUCTION

Nowadays, searching a large database for images that match a query is becoming challenging in computer vision and pattern recognition applications. Traditionally, Textual annotations (e.g. keywords, filenames, and captions) were used to index and retrieve images leading to large human intervention. Recently, Content Based Image Retrieval (CBIR) is a technique which uses visual contents such as color, texture & shape to classify images in large scale image datasets [1]. Color being most extensively used visual content and usually represented by color histograms, color correlograms, color coherence vector and color moments under a certain color space has proved efficient in representing color distributions of images [2]. MPEG-7 uses Color Layout Descriptor (CLD) & Scalable Color descriptor (SCD) derived in HSV color space with fixed quantization for feature extraction and matching procedure [3]. Color connectivity between regions are extracted and then for each regions - energy, inertia, entropy & uniformity are calculated in color co-occurrence matrix to preserve texture & color information in order to obtain leading results compared to gray-level co-occurrence matrix and color histogram methods [4]. Block encoding technique [5] can be used to code floating point color histogram bin values and converted into integer values, XOR based bitwise similarity measure technique is applied between query and trained coded features to improve classification rate.

II. RELATED WORKS

In recent years, the trend on "similarity learning" is observed. Discriminative methods such as linear discriminant analysis (LDA) and Support Vector Machines (SVM's) aim to find a mapping or decision function that best separate the data from different classes. In CBIR, generative methods are usually used for modeling the image generation process for image content representation [7], class descriptors [8] and visual dictionary [9] under different assumptions. The use of prior knowledge within a constellation model framework was suggested by Fei Fei et al. [10] who extended the work of Fergus et al. [11] into a full Bayesian setting, thereby allowing for the introduction of prior distributions. Anna et al. explored multi-way classifier by selecting the ROI and using random forest and ferns classifier [12]. NBNN, (Naive-Bayes Nearest Neighbor) classifier employs NN distances in the space of the local image descriptors and computes direct 'Image-to-Class' distances without descriptor quantization [13]. Nearest Neighbor classifier is remodeled by picking K-nearest neighbors and computed pairwise distances between those neighbors, further distance matrix is converted into a kernel matrix and then multi-class SVM is applied to recognize objects on large scale multiclass datasets outperforming nearest neighbor and support vector machines for shape matching [14].

Wavelet and curvelet coefficients are used to extract higher order moments at multiple resolution scale which increased the discriminativity between the images. Wavelets [15] being popular in representing point singularities and isolating the discontinuity across edge, failed to represent line singularities and smoothness along the edges. Discrete

Cosine Transforms (DCT) and wavelet transforms are widely used in image and video compression and for still images it is revealed that wavelets outperforms DCT by 1DB in PSNR where as for video coding it is less noticed [16]. DCT & Wavelets with proper thresholding techniques shows significant improvement in compression ratio in terms of MSE and PSNR [18]. Lower frequency DCT coefficients are considered neglecting high frequency components from normalized image patches, and also selecting the number of DCT coefficients has shown the affect on retrieval performance [17]. Malik & Baharudin considered 1 DC with first 3 AC coefficients in DCT transformed matrix at different histogram bins and analyzed distance metrics for classification in Content Based Image Retrieval (CBIR) [19].

Being inspired from the fact of representing line singularities in higher dimensional multi-resolution feature bands and motivated by the facts that DCT efficiently preserves local affine information and robust to the variant factors such as illumination, occlusion & background clutter present in an image. We explore Coslets & Ridgelets by applying 1-D wavelets on DCT & radon sub-bands respectively to extract highly discriminative visual features for object categorization. As a part of pre-processing we first segment dataset images in hybrid color space generating perceptually homogeneous regions.

The rest of the paper is organized as: Section III briefs proposed Coslets & Ridgelets in hybrid color space along with different classifiers. In Section IV, classification results on two widely used datasets and performance analysis on several benchmarks are reported. Conclusion and some future research issues are discussed at the end.

III. PROPOSED METHODOLOGY

Addressing the issue of human perception of visual data content and deriving discriminating image features to reduce the semantic gap, we carried out image analysis in frequency domain by integrating transform & subspace based techniques for better image representation and categorization of very large image datasets. Since, Wavelet transform provides sparse representation for piecewise-linear signals. Coslets and ridgelets are derived using wavelets in DCT and radon sub-bands from perceptually segmented images to obtain corresponding transformed coefficients. Transformed coefficients representing line/curve singularities along with color-texture features are further processed in reduced feature space using PCA. Finally, Neural Networks and four distance measures are used as classifiers to obtain an average per class recognition rate. As mentioned above, in-detail explanation regarding segmentation, transform domain techniques along with different classifiers can be found in following subsections.

A. Segmentation in Complex Hybrid Color Space

As a part of pre-processing, a fast and color-clustering-based segmentation method combining color spaces and k-means is developed to generate perceptually meaningful regions in complex hybrid colorspace as investigated in computer vision [6]. The crucial steps involved in segmentation process are as given below:

- Step 1: Input RGB image is transformed into YCbCr & HSI color spaces.
- Step 2: Consider H component of HSI & CbCr components in YCbCr.
- Step 3: Augmenting three higher dimensional matrices to generate hybrid color space – HCbCr.
- Step 4: Further HCbCr is transformed into LUV color space.
- Step 5: Finally k-means ($k=3$) clustering is applied to partition image into regions (3 disjoints).

Above steps are sequentially processed in supervised context to obtain segmented image of highly coherent regions preserving low frequency components and reducing (or sometimes completely eliminating) the effect of background, which further exploited for better image representation.

B. Coslets

For most of the images, much of the signal strength lies at low frequencies and DCT significantly separates out image spectral sub-bands with respect to image's visual features. Hence DCT is very popular transform technique used in image compression [20] and face recognition [21].

1-D DCT can be defined as:

$$F(u) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \sum_{x=0}^{N-1} \Lambda(x) \cos\left[\frac{(2x+1)u\pi}{2N}\right] f(x) \quad (1)$$

Where $f(x)$, $x=0, 1, 2, \dots, N-1$, be a signal sequence of length N . Unlike 1-D signal, DCT of given image $f(x, y)$ of size $m \times n$ is defined as:

$$C(u, v) = \frac{2}{\sqrt{mn}} \alpha(u)\alpha(v) \sum_{x=1}^m \sum_{y=1}^n f(x, y) \cos\left[\frac{(2x+1)u\pi}{2m}\right] * \cos\left[\frac{(2y+1)v\pi}{2n}\right] \quad (2)$$

Where $u = 1, 2, \dots, m$ & $v = 1, 2, \dots, n$ are scaling factors.

Due to high energy compaction property exhibited by the DCT, conventional zigzag scanning technique is applied to obtain transformed coefficients. Transformed DCT coefficients preserve low frequency components to identify the local information which is invariant to illumination, occlusion, clutter and almost nullifying the effect of high frequency coefficients.

Multi-resolution representations are very effective for analyzing information content of images and since choice of wavelet bases is highly subjective in nature as it depends on data in hand, we selected 'Haar' as the best wavelet basis for image representation [22]. Let $\psi(x)$ be the orthogonal wavelet, an orthogonal basis of H_2^j is computed by scaling the wavelet $\psi(x)$ with a coefficient 2^j and translating it on a lattice with an interval 2^{-j} . Haar wavelet is given as:

$$\Psi(x) = \begin{cases} 1 & \text{if } 0 \leq x < \frac{1}{2} \\ -1 & \text{if } \frac{1}{2} \leq x < 1 \\ 0 & \text{Otherwise} \end{cases} \quad (3)$$

Let 'x' be a given 1D signal, wavelet consists of $\log_2 N$ stages at the most. At the first step, we obtain two sets of coefficients: Approximation coefficients *CAI* and Detailed coefficients *CDI*. These coefficients are obtained by convolving 'x' with low-pass filter for approximation, and with the high-pass filter for detail, followed by dyadic decimation.

To conclude this section, DCT is used to convert 2-D signal into elementary frequency components of lower frequency and wavelets representing point singularities by isolating edge discontinuities and capturing maximum energy of the given image put together to form 'Coslets'. For more information kindly refer to [37].

C. Ridgelets

Ridgelets introduced by Candes and Donoho [23] represents higher dimensional singularities, later evolved to be very effective in representing line singularities in 2-D. With an idea of projecting point singularities onto a line, ridgelets are proposed by applying 1-D wavelet to the slices of radon sub-bands.

Given an integrable bivariate function $f(x_1, x_2)$, its Radon transform (RDN) is expressed as:

$$RDN_f(\theta, t) = \int_{R^2} (f(x_1, x_2) * \delta(x_1 \cos\theta + x_2 \sin\theta - t)) dx_1 dx_2$$

The radon transform operator plots the information from spatial domain to projection domain (θ, t) , in which one of the directions of the radon transform coincides and maps onto a straight line in the spatial domain. On a contrary each point in the spatial domain turns out to be a sine wave representation in the projection domain.

The Continuous Ridgelet Transform (CRT) is merely the application of 1-D wavelet to the slices of radon.

$$\Psi_{ab}(t) = \frac{1}{a} \psi\left(\frac{t-b}{a}\right) \quad (4)$$

It is evident that ridgelet function is constant along lines when $x_1 * \cos\theta + x_2 * \sin\theta = \text{const}$. Our goal of applying ridgelets to 2-D image was simplified by Do and Vetterli [24] by exploring Finite Ridgelet Transform (FRIT). FRIT is developed based on Finite Radon Transform (FRAT) and can be defined as summation of image pixel intensity values over certain set of predefined line in Euclidian space.

The FRAT of real discrete function f on the definite grid Z_2P can be defined as:

$$FRAT_f(k, l) = \frac{1}{\sqrt{p}} \sum_{(i,j) \in L_{k,l}} f(i, j) \quad (5)$$

Where 'p' is the prime number with modulus operation in finite field Z_p & $L_{k,l}$ denotes the set of points that make up a line on the lattice Z_2P and expressed as:

$$L_{k,l} = \begin{cases} (i, j) : j = (ki + l) \pmod{p}, i \in Z_p \text{ if } 0 \leq k \leq p \\ (l, j) : j \in Z_p \text{ if } k = p \end{cases} \quad (6)$$

Excellent energy information is found to be accumulated in the low-pass band of the ridgelet decomposed image. It is also worth noting that ridgelet coefficients include the information about the smoothness in t and θ of the radon transform. Radon transform exhibits a certain degree of smoothness in particular; we can instantly see that the ridgelet coefficients decay rapidly as θ and/or b moves away from the singularities.

Large numbers of coefficients generated due to the transformation process are projected onto a reduced feature space using PCA to improve computational efficiency. More information can be found in [37].

D. Classification

In order to calculate visual similarities between a query and database image, many classification techniques have been developed for image retrieval based on empirical estimates of the distribution of features in recent years [1]. We carried out two separate classification experiments, firstly by considering similarity/distance measure techniques and secondly neural network architectures such as PNN & GRNN are applied to obtain an average classification rate. Detailed explanation can be found in the following sections.

1) Similarity Distance Measures

Different similarity distance measures will affect the recognition rate, in this regard we consider five different distance measure techniques such as Minkowski distance, Euclidian distance, Modified Squared Euclidian distance, Correlation coefficient based distance and Angle Based distance to acquire an average classification rate.

2) Generalized Regression Neural Network (GRNN)

Neural networks have been motivated right from its inception by the recognition that the human brain computes in an entirely different way from the conventional digital computer. GRNN are paradigms of the Radial Basis Function (RBF) used to functional approximation was rediscovered to perform general regressions.

An input vector 'x' (PCA projected feature matrix F_i) is formed, GRNN as classifier calculates weight vector 'W' using the following equation:

$$y(x) = \frac{\sum_{i=1}^n T_i W_i}{\sum_{i=1}^n W_i} \tag{7}$$

Where $W_i = e^{\left[\frac{-\|x-x_i\|^2}{2h^2}\right]}$ is the output $y(x)$ is weighted average of the target values T_k of training cases x_i close to a given input case x .

3) Probabilistic Neural Network (PNN)

Due to its excellent generalization performance SVM is most promising classifiers in machine learning. However, SVM's are slow and still remains to be a bottleneck for large datasets and multiclass classification [26]. Donald Specht introduced PNN, network is based on concepts used for conventional pattern recognition problems. In particular PNN models the Bayesian classifier [25] in order to minimize the risk of misclassification. Bayes' classifier is usually criticised due to lack of information about the class probability distributions and makes use of nonparametric techniques to calculate apriori probabilities π_i , but probability density functions $f_i(x)$ are generally more difficult to estimate. Parzen [27] constructs a family of estimates from kernels, general form of the estimator is given as:

$$F(x) = \frac{1}{n\lambda} \sum_{i=1}^n l\left(\frac{x-x_i}{\sigma}\right) \tag{8}$$

Where x_i are identically distributed and independent random variables with weighting function l to be bounded. Parzen extended to the multivariate distribution and weighting function l is the multivariate exponential (Gaussian) function and equation expressed in the form:

$$F(x) = \frac{1}{(2\pi)^{\frac{p}{2}} \sigma^p m} \sum_{i=1}^n e^{-\frac{(x-x_i)^T(x-x_i)}{2\sigma^2}} \tag{9}$$

The inherent advantage of PNN is the better generalization and convergence properties when compared to that of Bayesian classifier in classification problems. Proposed method is divided into two phases i) Training & ii) Testing, brief description of proposed algorithm is as given in table below:

TABLE I ALGORITHM OF PROPOSED METHOD

<p>Algorithm: Training Phase Input: A set of training sample images Output: A Knowledge Database (KDB) Method:</p> <ol style="list-style-type: none"> 1. Acquire training samples (Segmented images). 2. Apply Ridgelet/Coslet transformation models on set of segmented training samples to obtain feature vector. 3. Store the feature matrix in Knowledge Database. <p>Algorithm: Training Phase Ends</p>	<p>Algorithm: Testing Phase Input:</p> <ol style="list-style-type: none"> i. A Knowledge Database (KDB) ii. Test/Query image, I. <p>Output: Class label / Identity of I. Method:</p> <ol style="list-style-type: none"> 1. Obtain the feature matrix F of the pattern I (Segmented image). 2. Apply GRNN/PNN for successive classification. 3. Label the class/identity using step2. <p>Algorithm: Testing Phase Ends</p>
---	---

IV. EXPERIMENTAL RESULTS

In this section, we present result and analysis carried out on Caltech-101 & Caltech-256 datasets, consisting of various object categories. Caltech-101 dataset consists of 101 different object categories ranging from 31 to 800 images per category, and are subjected to high intensity variations, occlusions and affected by corner artifacts [10]. Caltech-256 generated by Griffin et al. comprises 256 different object categories which has been manually screened out of Google images exhibiting high variations in intensity, clutter, object size, location, pose, and also increase in number of category with at least 80 images per category [28].

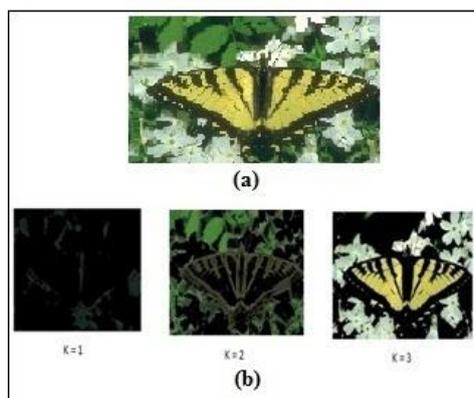


Fig. 1 (a) Original image. (b) Resultant Segmented images for different values of “k”.

The system is initialized by transforming image from RGB to hybrid color space and then k-means technique is applied to generate clusters based on color homogeneity (please refer Fig.1.). Segmented RGB image is converted to a gray scale image and further scaled down the image size to give a sensible number of features per image. Transformed feature co-efficients can be obtained by applying 1-D wavelet in radon and cosine domain respectively. These coefficients are vectorized from 2-D to 1-D to get resultant feature vector containing 10000+ co-efficients per image, PCA is applied to extract principal components of these features and to reduce the feature space. Further, classification is performed using Neural Network and Distance measure (DM) classifiers to record an individual average classification rates respectively in compressed domain.

As a standard experimental procedure mentioned in [10 & 28], proposed PCA transformation models are experimented by varying the training samples. We divided entire dataset into 15 and 30 images/category as training phase and remaining as testing to obtain an average of per class recognition in each stage for each of the classifiers. Performance analysis for well known techniques & benchmarking datasets are explained in the following paragraphs.

Table II shows the recognition rate for transform technique and subspace models considering both distance measures and neural networks classifiers for (a) Caltech-101 & (b) Caltech-256 datasets. From the table it is noticed that Ridgelet and PCA model along with distance measure (DM) as a classifier outperforms the other subspace and transformation models.

Ridgelet and coslet transforms extract the energy information of the image from all the angle with respect to x-axis. Most of the energy information are present in terms of low frequency coefficients. Application of PCA for this purpose can indeed extract the adequate information in all the possible direction. Whereas applying PCA merely extracts adequate information in one angle and hence combining transforms with PCA creates more discriminant power compare to other subspace methods.

In general, performance of transformation models can be improved for segmented images in hybrid color space, especially Ridgelet with Distance Measure (DM) classifier proved to be very efficient and progressive when compared to the other amalgamation of transformed subspace and classifier models.

The following observations can be made:

- Performance of PCA is better compared to other standard subspace methods.
- Coslets preserve low frequency coefficients to identify the local information which is invariant to illumination, occlusion, clutter and almost nullifying the effect of high frequency coefficients.
- Ridgelets preserves edge information by representing line singularities in higher dimensional feature sub-bands.
- Application of subspace & transformation models has showed increase in recognition accuracy for segmented images considering both DM and PNN classifiers.
- Our approach allows using more features and reduces corner artifacts; occlusion and/or relative scale problems in the dataset (please refer Fig.1 (a) for few sample images).
- Feature representation in segmented images proved to be very effective way of image representation.
- Both Ridgelet and Coslets proved efficient in extracting highly discriminative features present in an image which exhibits corner artifacts, rotation, and varying illumination.

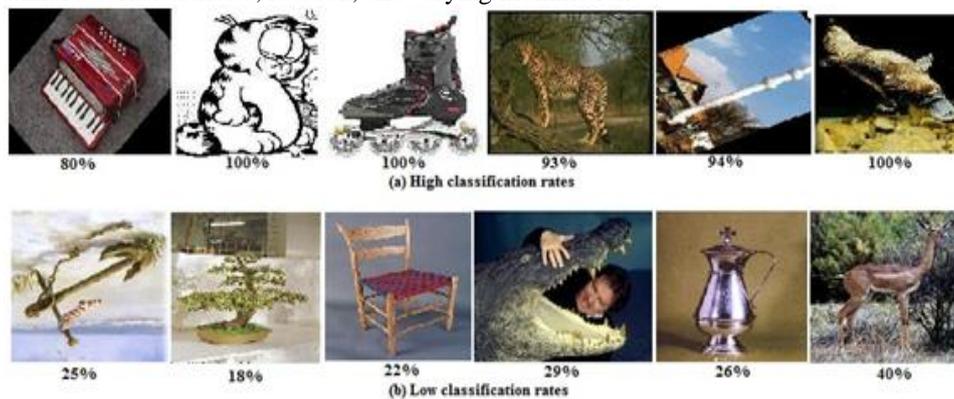


Fig. 2 Sample images of Caltech 101 dataset with (a) High & (b) Low classification rate.

TABLE II RECOGNITION ACCURACY OF PROPOSED METHODS FOR CALTECH – 101 & CALTECH – 256

Method	15 Train	30 Train	Method	15 Train	30 Train
Recognition rate (%): Caltech – 101			Recognition rate (%): Caltech – 256		
Ridgelets + DM	38.05	45.36	Ridgelet + DM	18.7	22.4
Ridgelets + segmentation + DM	48.3	56.07	Ridgelet + segmentation	27.06	31.3
Ridgelets +PNN	21	28.02	Ridgelets + PNN	12	16
Ridgelets +PNN(segmentation)	33.6	41.4	Ridgelets + PNN(segmentation)	15	20
Ridgelets +GRNN	17	23.5	Ridgelets + GRNN	11	15

Ridgelets +GRNN(segmentation)	32.2	39.6	Ridgelets + GRNN(segmentation)	14	18
Ridgelets + NN_{average}	46	52	Ridgelets + NN_{average}	25	28
Coslets + DM	37.05	43.10	Coslets + DM	16.9	21
Coslets + segmentation + DM	46	54.7	Coslets + segmentation	25.3	30.6
Coslets + PNN	18	24	Coslets + PNN	10	14.5
Coslets + PNN(segmentation)	31	39.4	Coslets + PNN(segmentation)	13	17
Coslets + GRNN	17	24.5	Coslets + GRNN	9	14
Coslets + GRNN(segmentation)	30.2	38.6	Coslets + GRNN(segmentation)	12	16
Coslets + NN_{average}	41	49	Coslets + NN_{average}	23	28

Table III gives the comparison of the proposed model across few benchmarking techniques mentioned in the literature for both caltech-101 & 256 datasets. Complex features obtained due to combined transforms and subspace techniques with segmentation outperform biologically motivated visual cortex model [29], hybrid model [30] & shape matching technique [14] considering caltech-101. Whereas for caltech-256 dataset ridgelet with segmentation along with distance measures as a classifier achieved high classification rate for lesser number of training samples in comparison with kernel codebook method [32]. Highly competitive results can be noted when compared with the techniques mentioned in [28, 31 & 33]. Figures 2 & 3 shows few sample images of Caltech-101 & 256 datasets with high and low classification rates.

TABLE III PERFORMANCE ANALYSIS OF THE PROPOSED METHODS

Method	15 Train	30 Train
Recognition rate (%): Caltech – 101		
Serre et al. [29]	35	42
Holub et al. [30]	37	43
Berg et al. [14]	45	-
Sancho et al. [31]	57.8±0.3	65.2±0.4
Ridgelets +GRNN(segmentation)	32.2	39.6
Ridgelets +PNN(segmentation)	33.6	41.4
Ridgelets + NN_{average}	46	52
Ridgelets + segmentation + DM [36]	48.3	56.07
Coslets + GRNN(segmentation)	30.2	38.6
Coslets + PNN(segmentation)	31	39.4
Coslets + NN_{average}	41	49
Coslets + segmentation + DM [37]	46	54.7
Recognition rate (%): Caltech – 256		
Van et al. [32]	-	27.17 ± 0.46
Griffin et al. [28]	28.3	34.1
Sancho et al. [31]	33.5	40.1
Jianchao et al. [33]	28.30	34.10
Ridgelets +GRNN(segmentation)	14	18
Ridgelets +PNN(segmentation)	15	20
Ridgelets + NN_{average}	25	28
Ridgelets + segmentation + DM [36]	27.06	31.3
Coslets + GRNN(segmentation)	12	16
Coslets + PNN(segmentation)	13	17
Coslets + NN_{average}	23	28
Coslets + segmentation + DM [37]	25.3	30.6

In order to draw comparison with few of the very recent techniques, we considered the parametric settings mentioned in [34 & 35] i.e. by training 15 and 30 images/category and not more than 20 & 50 images/category for testing respectively. For Caltech 101 dataset, we obtained highest classification rate of 67% & 73% respectively in accordance with 67% & 74% of DLMM [34], 66% & 72% of ScSPM [35] techniques. Similarly for Caltech-256 dataset, we obtained highest recognition rates of 29.6% & 36% outperforming 29% & 35% of [34] and 28% & 34% of [35] considering 15 & 30 images/category as labeled images respectively.

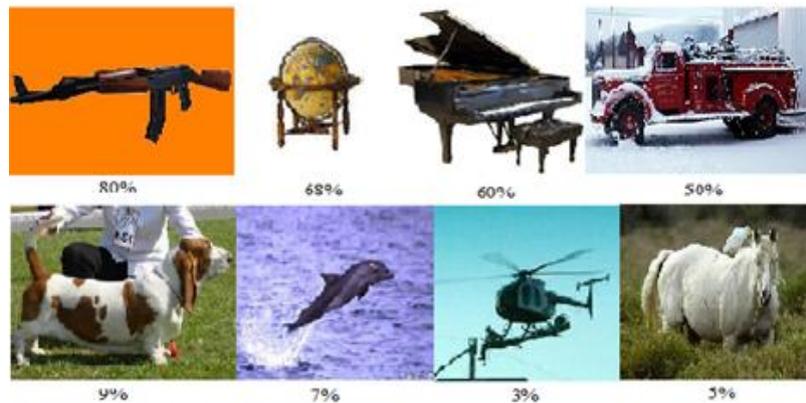


Fig.3 Sample images of Caltech 256 dataset with high (top) & low (bottom) classification rates.

V. DISCUSSIONS & CONCLUSION

Despite the earlier efforts made in the image retrieval research, we do not yet have a generally acceptable algorithmic means of distinguishing human vision. In this regard, a remarkable effort to extract discriminative patterns more distinctively in the context of interpreting images is realized by integrating transformation and subspace based techniques. Radon and DCT extracts highly discriminative localized features exhibiting linear properties and low frequency information respectively, later wavelets decomposes these signals into multi-resolution sub-bands giving inherent benefit of spectral analysis in developing robust and invariant geometrical features with structural information. Though distance measures proved to be efficient as classifier, neural networks have augmented the classification rates in comparison with individual distance measure technique. Introducing PCA not only reduces the feature space but also enhanced the classification rate. Segmentation as pre-processing has gained lot of importance by improving the classification performance.

We demonstrated our proposed methods on two widely used multiclass datasets and achieved leading classification rates in contrast with nearest neighbor classifier [14], found more flexible than feed forward visual cortex model in representing edge features [29], competitive to dictionary learning manifold method and sparse coding techniques [34, 35]. In the near future, we investigate classification performance of our approaches using SVM and various neural network architectures.

REFERENCES

- [1] Ritendra Datta, Dhiraj Joshi, Jia Li, James Z. Wang, Image retrieval: ideas, influences, and trends of the new age, *ACM Computing Surveys*, vol. 40, no. 2, pp. 1-60, 2008.
- [2] A. Mojsilovic, B. Rogowitz, Capturing image semantics with low-level descriptors, *Proceedings of the ICIP*, pp. 18-21, 2001.
- [3] B.S. Manjunath, et al., *Introduction to MPEG-7*, Wiley, New York, 2002.
- [4] WangXing-yuan , ChenZhi-feng, YunJiao-jiao. An effective method for color image retrieval based on texture. *Computer Standards & Interfaces*, vol 34, pp 31-35, 2012.
- [5] Shaila S G, Vadivel A. Block Encoding of Color Histogram for Content Based Image Retrieval Applications. *Procedia Technology*, vol 6, pp 526 - 533, 2012.
- [6] K. Mahantesh, V.N. Manjunath Aradhya, S.K. Niranjana. An impact of complex hybrid color space in image segmentation. *2nd International Symposium on Intelligent Informatics (ISI'13)*, vol 235, pp 73-82, 2013.
- [7] T. Westerveld and A. P. de Vries, "Experimental evaluation of a generative probabilistic image retrieval model on 'easy' data," in *Multimedia Information Retrieval Workshop*, Toronto, Canada, 2005.
- [8] W. Jiang, K. L. Chan, M. Li, , and H. Zhang, "Mapping low-level features to highlevel semantic concepts in region-based image retrieval," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2005.
- [9] R. Zhang and Z. Zhang, "Hidden semantic concept discovery in region based image retrieval," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 996-1001, 2004.
- [10] L. Fei-Fei, R. Fergus, and P. Perona. Learning generative visual models from few training examples. *Computer Vision and Pattern Recognition (CVPR) Workshop on GMBV*, 2004.
- [11] R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale invariant learning. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, page 264, 2003.
- [12] Anna Bosch, Andrew Zisserman, Xavier Munoz, Image Classification using Random Forests and Ferns, *IEEE 11th International Conference on Computer Vision (ICCV)*, pp 1 - 8, 2007.
- [13] Boiman O, Shechtman E, Irani, M., In Defense of Nearest-Neighbor Based Image Classification, *IEEE CVPR*, pp 1-8, 2008.
- [14] A. C. Berg. Shape Matching and Object Recognition. PhD thesis, Computer Science Division, University of California, Berkeley, 2005.
- [15] Fionn Murtagh, Jean-Luc Starck. Wavelet and curvelet moments for image classification: Application to aggregate mixture grading. *Pattern Recognition Letters*, vol 29, pp 1557-1564, 2008.

- [16] Zixiang Xiong, Kannan Ramchandran, Michael T. Orchard, and Ya-Qin Zhang. A Comparative Study of DCT-andWavelet-Based Image Coding. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 9, No. 5, pp 692-695, 1999.
- [17] Stepan Obdrzalek and Jir Matas. Image Retrieval Using Local Compact DCT-based Representation. *DAGM'03, 25th Pattern Recognition Symposium*, Springer, LNCS 2781, pp 490-497, 2003.
- [18] Prabhakar.Telagarapu, V.Jagan Naveen, A.Lakshmi..Prasanthi, G.Vijaya Santhi. Image Compression Using DCT and Wavelet Transformations. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol. 4, No. 3, pp 61-74, 2011.
- [19] Fazal Malik, Baharum Baharudin. Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the DCT domain. *Journal of King Saud University - Computer and Information Sciences*, vol 25, issue 2, pp 207-218, 2012.
- [20] Prabhakar.Telagarapu, V.Jagan Naveen, A.Lakshmi..Prasanthi, G.Vijaya Santhi. Image Compression Using DCT and Wavelet Transformations. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol. 4, No. 3, pp 61-74, 2011.
- [21] Zhu.J, Vai.M, and Mak.P. Face recognition using 2D DCT with PCA. In *The 4th Chinese Conference on Biometrics Recognition (sinbiometrics'03)*, pp 150-155, 2003.
- [22] Jen-Tzung Chien and Chia-Chen. Discriminant wavelet faces and nearest feature classifiers for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 24, no 12, pp 1644-1649, 2002.
- [23] E. J. Candes and D. L. Donoho. Ridgelets: a key to higher-dimensional intermittency?. *Phil. Trans. R. Soc. Lond. A.*, pp 2495–2509, 1999.
- [24] M. N. Do and M. Vetterli. Discrete ridgelet transforms for image representation. *Info-science, LCAV-Report*, 2004.
- [25] Specht, A general regression neural network, *IEEE Trans. Neural Networks*, vol.2, pp 568-576, 1991.
- [26] C. Burges. A tutorial on support vector machines for pattern recognition. *Knowledge Discovery and Data Mining*, vol 2, pp 1–47, 1998.
- [27] E. Parzen. On estimation of a probability density function and mode. *Analysis of Mathematical Statistics*, vol 33, pp 1065–1076, 1962.
- [28] G. Griffin, A. Holub, and P. Perona. Caltech 256 object category dataset. *Technical Report*, California Institute of Technology, 2007.
- [29] T. Serre, L. Wolf, and T. Poggio. Object recognition with features inspired by visual cortex. In *CVPR*, San Diego, June 2005.
- [30] A. Holub, M. Welling, and P. Perona. Exploiting unlabelled data for hybrid object classification. In *NIPS Workshop on Inter-Class Transfer*, Whistler, B.C., December 2005.
- [31] Sancho Mc Cann and David G. Lowe, Local Naive Bayes Nearest Neighbor for Image Classification, *IEEE CVPR*. pp 3650 - 3656 , 2012.
- [32] J. C. van Gemert, J.-M. Geusebroek, C. J. Veenman, and A. W. M. Smeulders. Kernel codebooks for scene categorization. In *ECCV*, 2008.
- [33] Jianchao , Kai Yuz, Yihong Gongz, Thomas Huang . Linear Spatial Pyramid Matching Using Sparse Coding for Image Classification, *IEEE CVPR*. pp 1794 - 1801 , 2009.
- [34] Bao-DiLiu , Yu-XiongWang, Yu-JinZhang, BinShen, Learning dictionary on manifolds for image classification. *Pattern Recognition* , vol. 46, pp. 1879–1890, Elsevier 2013..
- [35] B. Liu, Y. Wang, Y. Zhang, Y. Zheng, Discriminant sparse coding for image classification, in: *Proceedings of the 37th International Conference on Acoustics, Speech and Signal Processing*, pp. 2193–2196, IEEE 2012.
- [36] K. Mahantesh, V. N. Manjunath Aradhya, An Exploration of Ridgelet Transform to handle Higher Dimensional Intermittency for Object Categorization in Large Image Datasets, In the proceedings of *The 2nd International Conference on Applied Information and Communications Technology – ICAICT*, Procedia Technology, Elsevier, Muscat, Oman, pp 515-521, 2014.
- [37] K. Mahantesh, V. N. Manjunath Aradhya, S K Niranjan, Coslets: A Novel Approach to Explore Object Taxonomy in Compressed DCT Domain for Large Image Datasets, In the Proceedings of *Advances in Intelligent Systems and Computing*, Springer, Vol. 320, pp 39-48, 2015.