# AN ENHANCED FEATURE EXTRACTION DIMENSIONAL REDUCTION ECHNIQUE FOR CONTENT BASED IMAGE

[1]**Mr.Srinivasa Babu Kasturi,** [2] **Dr.A.Ramamohan Reddy**

[1]**Research Scholar,** [2]**Professor**
[1,2]**Department of Computer Science and Engineering**
[1,2] **SV University Tirupathi, A.P, India**

---

*Abstract*: Retrieval of information's form large dataset has been observed from long time. Content Based Image Retrieval (CBIR) retrieves a multimedia data based on the content of given multimedia query, i.e., given a motion query, relevant images are to be retrieved. Here, an approach of imagemining and localization in a image dataset and an improved histogram based approach for Correlative Image Object  detection named as Correlative Histogram Based CodingCR-HBC) is proposed. The proposed approach enhances the objective of processing noise and system overhead with respect to the representation and retrieval performance. In the representation of descriptive features a large feature count is observed, which results in the overhead in processing. To reduce these descriptive features, this paper also presents a Principal Component Multi-Linear Analysis (PCMLA) for dimension reduction approach to feature reduction based on feature relations. The experiments have been conducted on the benchmark datasets. The experimental results show that the processing resource overhead and retrieval efficiency are improved.

*Keywords*: CBIR, Correlative Histogram features, PCA, LDA, Accuracy, Precision and Computation Time.

---

## I.    INTRODUCTION

Multimedia Retrieval is one of the most important and fastest growing research areas in the field of multimedia technology. Large collections of scientific, artistic and commercial data comprising image, text, audio and video abound in the present information based society. There must be an effective and precise method of assisting users to search, browse and interact with these collections. Nowadays people not only use pure images for dailypurposes, video is also a popular media for recordingTV, diaries etc. As a consequence, effective methodsfor searching with large database are needed. Thissuggests the need for using Content Based ImageRetrieval (CBIR) [1] System for images or videos which allowusers to search images or particular image framesaccording to their preferences.Where in today's multimedia retrieval system are more developed toward image retrieval model, less focus is made on video content retrieval. With higher storage resources, and new capturing units, more video information's are been processed in real time applications, such as surveillance, film making, home monitoring, CCTVmonitoring etc. In such applications, where video information could result in more informative than their corresponding images, retrievals of information over such system are limited. Current multimedia databases such as YouTube use a text based searching mechanism to search video content. A video annotation based retrieval [2, 3], and a video summarization based approach [4] to obtain keyword based action retrieval is proposed. However the retrieval performance is purely dependent on the tagging factor for applications. It is very difficult to extract video based on the content, such as action in a video samples.A motion based coding following energy and a history image[5] was also used as action detection model. A local representation of the extracted spatio-temporal interest points (STIPs) [6, 7] from an action is presented. The local representation is more robust to statistical representation of action dataset [8, 9]. In the action detection model the Harris detector [10] and 3-D SIFT [11, 12], are used for action detection process. The SIFT operator performs a max/min searching over the difference of Gaussian (DoG) function. A combined format of histogram oriented gradient and histogram optical flow (HoG-HoF) [13] was presented. The histogram of gradient is observed to be an effectively applied approach for action detection model. A 3D-HoG [14] is outlined, with a set of descriptive approach to obtain effective action model. Towards optimization of action detection model, a histogram based coding [15] for video retrieval is developed. The system uses the localized temporal histogram features to detect an action model from a video dataset. However, as observed the Noise factor impacting the histogram representation is been considered over two successive frames only. The global distribution of noise distribution over the video frames are not been analyzed in representing the Histogram feature. This factor, effects in the retrieval accuracy as well increase the resource overhead in multimedia retrieval approach.In the approach of multimedia retrieval, feature sets are more important. Xin Geng et al. [16][17] proposed the Representation pattern Subspace method for multimedia retrieval based on pattern recognition. The idea of it is to model the representation pattern, which is defined as a sequence of personal representation pattern videos, by learning a representative sub-space from EM-like (expectation maximization) iterative learning Principle Component Analysis (PCA).In video processing and pattern recognition, frequency domain analysis is the most popular method for extracting video features.GuodongGuo et al. [20]investigated the biologically inspired features (BIF) for pattern recognition from given videos. Unlike the previous works in [18][19], Guo simulated the human visual process based on bio-inspired models [21] by applying Gabor filters. A Gabor filter is a

linear filter used in video processing for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and have been found to be particularly appropriate for textural representation and discrimination. Though PCA based coding is observed to be applied over multiple application, the feature representation, is higher and the approach of selective operation minimizes the feature selection accuracy. To overcome the representation of Histogram based coding considering noise factor, in this paper, an approach for inter frame computation of histogram features and selection is proposed which extracts sufficient features for retrieval. Simultaneously, this paper also proposed a dimensionality reduction technique which reduces the computational overhead of retrieval system. To present the stated work, this paper is outlined into 5sections. Where in section 1 present the objective and past development towards information retrieval in content based multimedia retrial system. Section 2 defines the system model section 3 defines the approach of histogram for inter frame localization and representation. Section 4 outlines the dimensionality reduction technique. Section 5 outlines the obtained experimental results for the developed approach. The conclusion made for the presented work is outlined in section 6.

## II.      CBIR SYSTEM MODEL

The general CBIR system consists of three stages, training, testing and classification. The training stage trains the various images and their feature to the database. Testing stage extracts the features of a query sample and gives it to classifier. The classifier classifies the given query sample by comparing it with the trained samples. The developed system model for the proposed CBIR is given below:
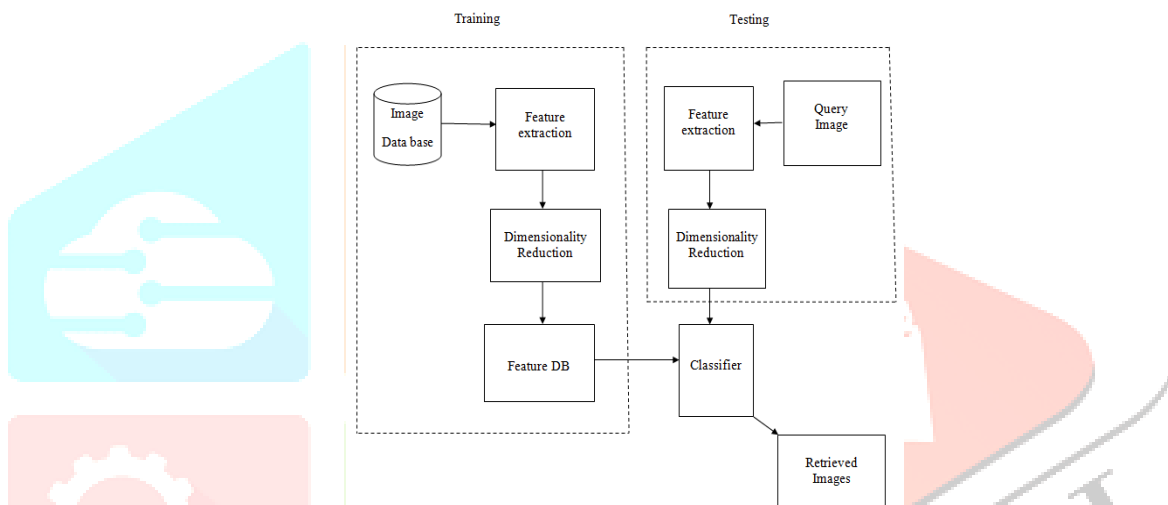


Fig.1. CBIR system model of proposed approach

The developed approach is processed into two phase of execution, of training and testing process. Wherein training process set of image sample of different types is taken, and are processed for feature selection through the proposed Correlative Histogram Based Coding (CR-HBC) approach. These features are then processed for Principal Component Multi-Linear Analysis (PCMLA) technique. Then the reduced feature set is processed for classification using a SVM classifier.

## III.      CORRELATIVE HISTOGRAM BASED CODING (CR-HBC)

In the approach of action based information retrieval, the approach of Histogram coding was presented in [8].The approach defines a temporal and spatial localization of action model based on Histogram mapping.However the noise effect leads to minimization of retrieval performance. To eliminate the noise Impact a inter frame correlation error for a set of time frames are considered. Considering a set of Histogram for k class ($H_k$), for a given video dataset of, i=1 to M,

$$H_i(k) = [H_i(kN), H_i(kN-1), ..., H_i(kN-M+1)] \ (1)$$

Where, $H_i$ is the Histogram for a video frame, N is number of frame and M are the dataset samples. To evaluate the noise effect in the temporal frames, a frame error is computed defined by,

$$e_{i,H}(k) = H_{i,t}(k) - H_{i,t+1}(k) \ (2)$$

This error defines the difference in the two frame component, and the histogram errors with lower values $\min(e_{i,H}(k))$ are considered as feature element. However this error when observed over a period of frame observation deviates a large and could be effective due to noise effect. Hence in such coding the intersection histogram would be mode concentric with noise parameter. To eliminate this problem, and to improve the feature selection more accurately, a Histogram bin selection computed over a time series is proposed. In this suggested approach, rather to taking the whole histogram from single frame information, a selection of the histogram bins is made. To derive the bin selection, the histogram bins are initially normalized using a random weight factor.

$$H_i(k) = H_i(k)\, w(k) \qquad (3)$$

Where, $w(k) = [w_0(k),\ w_1(k), \dots \dots, w_{M-1}(k)]^T$ are the allocated weight factor for each frame. The estimated error is then defined as;

$$e_{i,H}(k) = H_{i,t}(k) - H_i(k)w(k) \qquad (4)$$

The error is recursively been computed over the total frames (i=1….N), and the initial error is recorded as $e_{i,H,init}$. A weight factor is then updated as,

$$w(k+1) = w(k) + \mu \sum_{i=0}^{N-1} \frac{H_i^T(k)}{\|H_i(k)\|^2} e_{i,H,init}(k) \qquad (5)$$

Where μ is the updation step size, with an error updation factor. The objective of this computation is to select the bins satisfying the $\min\left(e_{i,H}(k)\right)$ condition. To optimize the recursion overhead, a joint adjacent weight difference is computed defined by,

$$\widetilde{w}(k) = w^o - w(k) \qquad (6)$$

Where, $w^o$ is the initial weight issued. The weight updation is then defined as,

$$\widetilde{w}(k+1) = \widetilde{w}(k) - \mu \sum_{i=0}^{N-1} \frac{H_i^T(k)}{\|uH_i(k)\|^2} e_{i,H}(k) \qquad (7)$$

The deviation in the bin variation of the histogram is then integrated over a period of 0 to N defined by,

$$E(H_{i,N}) = \int_0^N \mu \sum_{i=0}^{N-1} \left( 2E\left[ \frac{H_{i,n}(k)\widetilde{w}(k)\, e_{i,NH}(k)}{\|H_{i,N}(k)\|^2} \right] - \mu E\left[ \frac{e_{i,NH}^2(k)}{\|H_{i,N}(k)\|^2} \right] \right) \qquad (8)$$

Wherein integrating the estimate, over 'n' observation period accumulates the estimation for 'n' inter frame errors. For each frame with minimum estimate error is then selected as the selected histogram bin and an intersection bin is then derived as,

$$s\left(H_q, H_t\right) = \sum_{i=1}^{k} \left( \frac{\min(H_q - H_t)}{H_n} \right) \qquad (9)$$

Where $H_n$ is the normalized histogram for the whole dataset.
The obtained multi instance histogram features are processed for dimensionality reduction. The dimensionality reduction method reduces the dimensions of obtained feature set such that the computation overhead is minimized.

## IV.       PRINCIPAL COMPONENET MULTI-LINEAR ANALYSIS (PCMLA)

The main aim of any dimensionality reduction approach is to minimize the number of features to be processed. Less the number of features less will be the computational overhead and less computation time. So the obtained multi instance histogram features from the CR-HBC are processed for dimensionality reduction through PCMLA presented. PCMLA stands for Principal Component Multi-Linear Analysis. It is a multi linear subspace learning method that extracts features directly from multi-dimensional objects. The PCMLA is the extension to the conventional PCA [16], which operates linearly whereas PCMLA operates multi-linearly. The PCA need to reshape the multidimensional object in to the vector, whereas PCMLA operates directly on multidimensional object through two-mode processing. In this paper the histogram features obtained through CR-HBC are given as input to PCMLA. The operation for the PCA is defined by, for a given data set of N-by-N dataset sample I(x, y) as a vector of dimension $N^2$, so that the sample can be thought of as a point in $N^2$-dimensional space. A database of M samples can therefore map to a collection of points in this high dimensional "dataset space" as $\Gamma_1, \Gamma_2, \Gamma_3, \dots \Gamma_M$. With the average dataset of the sample set defined as

$$\Psi = \frac{1}{M} \sum_{n-1}^{M} \Gamma_n \qquad (10)$$

Each dataset can be mean normalized and be represented as deviations from the average dataset by $\Phi_t = \Gamma_t - \Psi$. The covariance matrix, defined as the expected value of $\Phi\Phi^T$ can be calculated by the equation

$$C = \frac{1}{M} \sum_{n-1}^{M} \Phi_{n-1} \Phi_n^T \qquad (11)$$

Given the covariance matrix C, we can now proceed with determining the eigenvectors u and eigenvalues λ of C in order to obtain the optimal set of principal components, a set of Eigen datasets that characterize the variations between dataset samples. Consider an eigenvector $u_i$ of C satisfying the equation

$$C_{u_i} = \lambda_i u_i \qquad (12)$$

$$u_i^T C_{u_i} = \lambda_i u_i^T u_j \qquad (13)$$

The eigenvectors are orthogonal and normalized hence

$$u_i^T u_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \qquad (14)$$

Combining Eq. (11) and (14), Eq. (13) thus become

$$\lambda_i = \frac{1}{M} \sum_{n-1}^{M} var(u_i \Gamma_n^T) \qquad (15)$$

Eq. (15) shows that the eigenvalue corresponding to the $i_{th}$ eigenvector represents the variance of the representative dataset sample. By selecting the eigenvectors with the largest corresponding eigenvalues as the basis vector, the set of dominant vectors that express the greatest variance are being selected.The PCA algorithm applied for dimensionality reduction reduces the dimensions of obtained feature set only by considering internal variations in that particular sample only. However the inter class variations i.e. the features of other frame in a sequence is not considered. The PCA operates in the one dimensional mode, whereas PCMLA operates along multi-dimensional mode. For a given feature space of a single class, PCA evaluates the principal components individually, whereas PCMLA evaluates by considering the feature space of remaining classes also.

The pseudo code for PCMLA is described as follows:

Step 1: take the whole feature set having MxN dimensional space

Step 2: compute the mean along each 'n' dimensions

Step 3: obtain a new matrix by subtracting the mean from all values of dataset.

Step 4: evaluate a covariance matrix

Step 5: compute Histogram vectors and the corresponding Histogram values

Step 6: sort the Histogram vectors by decreasing the Histogram values and choose k Histogram vectors with largest Histogram values from n x k dimensional matrix $W_1$.

Step 7: perform the same operation of step7 for each class of feature set and find out some more Histogram values those having effect on the retrieval accuracy.

Step 8: finally form a new projection matrix by considering inters class Histogram values and also intra class Histogram values.

Step 9: form a dimensionally reduced subspace by multiplying the projection matrix with original values.

The feature vectors are then compared with the database using SVM classifier and classify the multimedia image as to which class it match.

## V.    EXPERIMENTAL RESULTS

The experimental image database consists of 500 images from Caltech database having various categories like 'Cougar', 'Kangaroo', 'Camera', 'Cellphone', 'Pizza', 'Pigeon', 'Dragonfly', 'Lotus', 'Wheelchair', 'Watelily'.The proposed system is developed over Matlab tool, and tested over the images collected from Caltech database.To analyze the performance of this developed system an experimental analysis is carried out as presented below. The training dataset is formulated with nine classes of objects, with each class having 10 different samples. Figure 7 illustrates few content of the considered dataset samples.
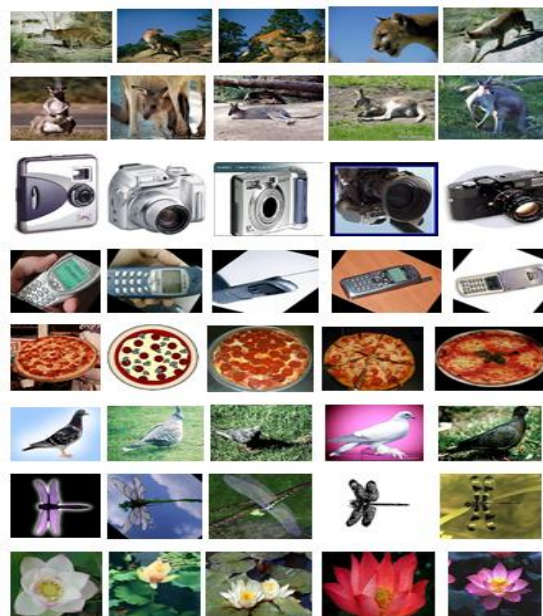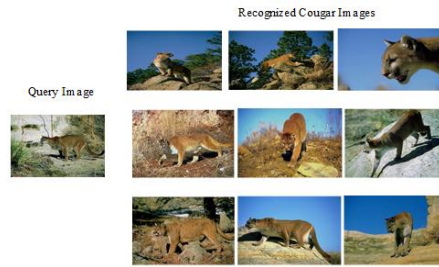


Fig.2. an example database

Fig.3. recognized cougar images for a given flower cougar image as a query

Fig.3 illustrates the results of the proposed approach, the obtained cougar image set for a given cougar image as a query. In the obtained image set, all the images are cougars only.
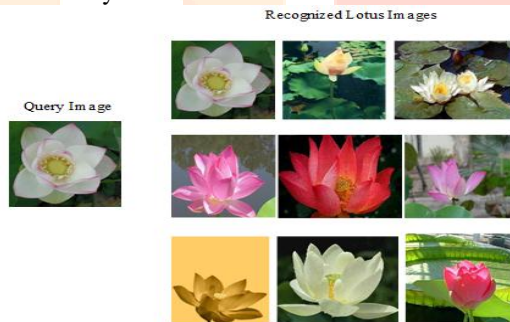


Fig.4. recognized camera images for a given flower camera image as a query

Fig.4 illustrates the results of the proposed approach, the obtained camera image set for a given camera image as a query. In the obtained image set, all the images are cameras only.



Fig.5. recognized Lotus images for a given flower camera Lotus as a query

Fig.5 illustrates the results of the proposed approach, the obtained Lotus image set for a given Lotus image as a query. In the obtained image set, all the images are Lotus only.

The main aim of CR-HBC approach is to extract efficient features those are able to increase retrieval accuracy. The CR-HBC proposed in this paper considered the correlation between the histograms of inter class image objects thus, the feature count is reduced. Then the dimensionality reduction techniques was proposed and applied on the obtained feature set of the CR-HBC, thus the feature count was reduced further. The obtained features of the 3HG, HBC, CR-HIST, PCA, LDA and the proposed PCMLA is represented in table.1

Table.1. Obtained Feature count for various approaches

| Object | Feature Extraction | | | Dimensionality Reduction | | |
|---|---|---|---|---|---|---|
| | 3HG[14] | HBC[15] | CR-HBC | PCA[16] | LDA[19] | PCMLA |
| Cougar | 59524 | 45265 | 32586 | 15862 | 14896 | 7452 |
| Kangaroo | 58236 | 44275 | 31248 | 14253 | 13968 | 6521 |
| Camera | 59964 | 43852 | 33541 | 13865 | 13230 | 6241 |
| Cellphone | 58285 | 47589 | 30124 | 12587 | 12012 | 6230 |
| Pizza | 56784 | 44125 | 32475 | 11568 | 11410 | 5894 |
| Pigeon | 56984 | 41256 | 33549 | 15236 | 14230 | 7451 |
| Dragonfly | 55230 | 44862 | 32146 | 16389 | 15351 | 6875 |
| Lotus | 56312 | 43567 | 30213 | 14857 | 13784 | 6324 |

Table.1 illustrates the obtained feature count for both feature extraction methods and dimensionality reduction methods. The feature extraction methods extract only the features. After feature extraction, the obtained entire features were trained as well as tested through the classifier. Whereas in dimensionality reduction methods, the dimensions of obtained feature set will be reduced and then only they were processed for training as well as testing. Hence the feature count of dimensionality reduction methods will be less. In table.1, the obtained feature for PCA, LDA and for PCMLA is less compared with 3HG, HBC and CR-HBC. The obtained feature count of various approaches at feature extraction stage and at dimensionality reduction sage is shown in fig.6.
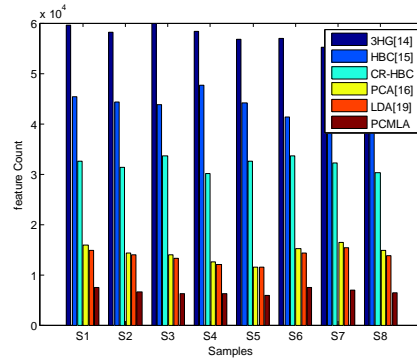


Fig.6. Feature count for four actions

The complete system is simulated in two phases, framing phase and classification phase. The training phase establishes a database with a set of features using feature extraction techniques. The classification phase performs the classification for a given query input. The time taken for training is termed as training time and the time taken for classification is termed as classification time. These two timings will vary from approach to approach. The complete analysis of the time consumption for various approaches is represented in fig.7
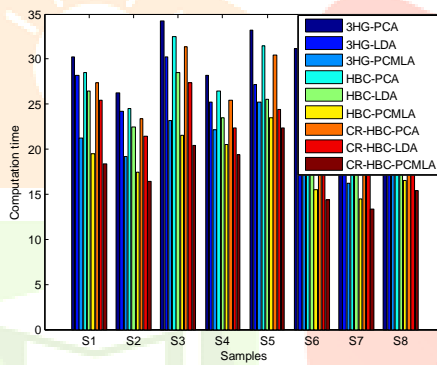


Fig.7. Computation time

Fig.7 illustrates the details of computation time in seconds. The training phase have large number of samples to process, hence the computation time for training is more when compared to classification. In fig.5, the computation time details are represented for the combination of feature extraction methods with dimensionality reduction methods. From fig.7, it is observed that, for each class, the combination of CR-HBC with PCMLA has less computation time compared remaining combinations. To evaluate the performance of the developed approach following parameters are used.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (16)$$

Where,
TP =True Positive (correctly identified)
FP= False Positive (incorrectly identified)
TN=True Negative (correctly rejected)
FN=False negative (incorrectly rejected)
For the given simulation model, totally four classes each class having five subjects were processed to training. In the testing a query sample with running action is selected and given for SVM classifier. The SVM classifier compares the given query sample features with database features. From the obtained classified results, the confusion matrix is created and it will be as,

Table.2. Confusion Matrix

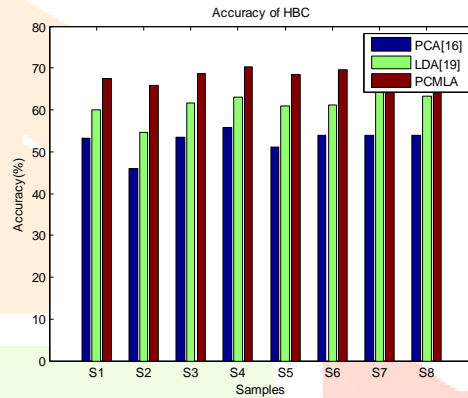| True Positive (TP) | False Positive (FP) |
|---|---|
| False Negative (FN) | True Negative (TN) |

Along with accuracy, to show the enhancement of propose approach and also to compare the proposed approach with earlier approaches, precision is evaluated as,
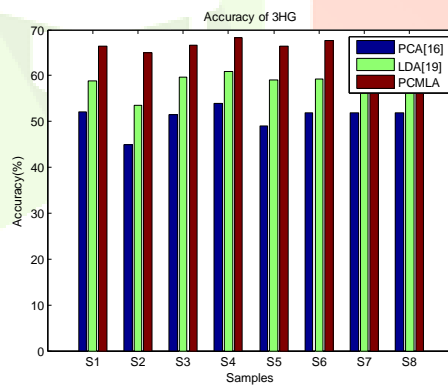
$$Precision = \frac{TP}{TP+FP} \qquad (17)$$

The obtained accuracy and precision details are illustrated in fig.8 andfig.9 respectively.
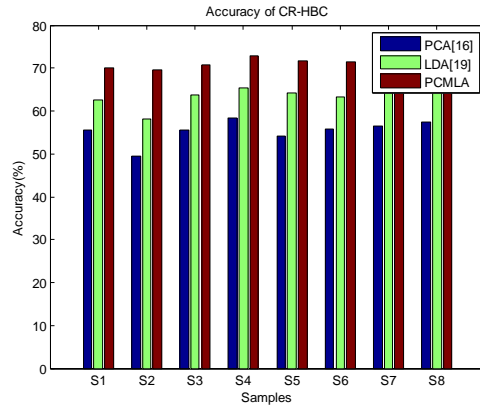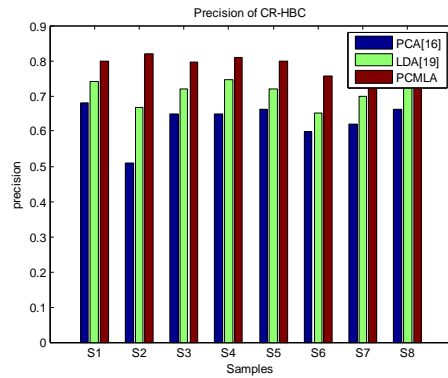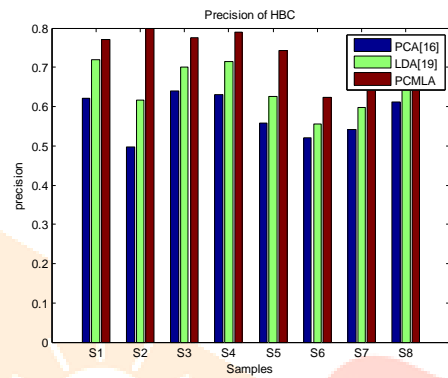


(a)



(b)



(c)

Fig.8 Accuracy of four action samples with dimensionality reduction methods, PCA, LDA and PCMLA (a)CR-HBC, (b)HBC [15], (c)3HG[14]
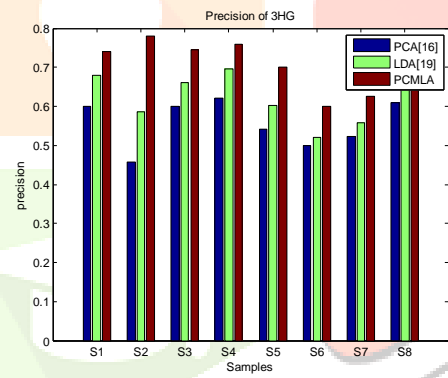
Fig.8 illustrates the accuracy details of the proposed work. The proposed work combines the feature extraction with dimensionality reduction approaches. Compared with alone, the combination will have more accuracy. The proposed work combined the 3HG, HBC and CR-HBC with PCA, LDA and PCMLA. The individual details of CR-HIST, HIST and 3HG are shown in fig.8 (a), (b) and (c) respectively. From fig.6 it is observed that for the combination of CR-HBC with PCMLA having more accuracy.

(a)



(b)



(c)

Fig.9 Precision of four action samples with dimensionality reduction methods, PCA, LDA and PCMLA (a)CR-HBC, (b)HBC [15], (c)3HG[14]

Fig.9 illustrates the precision details of the proposed work. The proposed work combines the feature extraction with dimensionality reduction approaches. Compared with alone, the combination will have more precision. From fig.9 it is observed that for the combination of CR-HBC with PCMLA having more precision.

## VI. CONCLUSION

A new coding approach for histogram based image mining is proposed. A process of inter class image histogram coding for feature selection and its representation is developed.When the Histogram was applied on image, the respective dominant features are obtained. After obtaining all possible variations from the image information, Histogram features were extracted. In conventional approach, the Histogram features are directly extracted from multimedia image thus finding the dominating features will become complex. The proposed work applied dimensionality reduction method in multi nature, reduces the dimensions of feature set by considering the intra class feature and also inter class features, whereas, the conventional approach reduces the dimensions by considering only intra class features.The retrieval performance is improved with the optimal selection of histogram features by the application of feature extraction and dimensionality reduction method. The extensive simulation had shown the effectiveness of proposed work.

## VII. REFERENCES

[1] Del Bimbo, A. (1998). "A Perspective View on Visual Information Retrieval Systems", *Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries*, Santa Barbara, California.

[2] M. Wang, X.-S. Hua, R. Hong, J. Tang, G.-J. Qi, and Y. Song, "Unified video annotation via multi graph learning", *IEEE Transaction on Circuits and System for Video Technology*, Vol. 19, No. 5, May 2009, pp. 733–746.

[3] M. Wang, X.-S. Hua, J. Tang, and R. Hong, "Beyond distance measurement: Constructing neighborhood similarity for video annotation," *IEEE Transaction on Multimedia*, vol. 11, No. 3, Apr. 2009, pp. 465–476.

[4] M. Wang, R. Hong, G. Li, Z.-J. Zha, S. Yan, and T.-S. Chua, "Event driven web video summarization by tag localization and key-shot identification", *IEEE Transaction on Multimedia,* vol. 14, No. 4, Aug. 2012, pp. 975–985.

[5] J. W. Davis and A. F. Bobick, "The representation and recognition of human movement using temporal templates,"*IEEE Conference on Computer Vision and Pattern Recognition*, pp. 928, 1997.

[6] I. Laptev, "On space-time interest points," *Journal of Computer Vision*, Vol. 64, No.2, 2005 pp. 107–123.

[7] P. Dollar, V. Rabaud, G. Cottrell, and S. Belongie, "Behavior recognition via sparse spatio-temporal features," *joint international Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, IEEE,* 2005 pp. 65–72.

[8] T. H. Thi, J. Zhang, L. Cheng, L. Wang, and S. Satoh, "Human action recognition and localization in video using structured learning of local space time features," *IEEE International Conference on Advance Video Signal Based Surveillance,* 2010, pp. 204–211.

[9] M. Ryoo and J. Aggarwal, "Spatio-temporal relationship match: Video structure comparison for recognition of complex human activities," *IEEE International Conference Computer Vision*, 2009, pp. 1593–1600.

[10] K. Mikolajczyk and C. Schmid, "An affine invariant interest point detector," *European Conference Computer Vision,* 2002, pp. 128–142.

[11] D. G. Lowe, "Distinctive image features from scale-invariant key points," *International Journal Computer Vision,* Vol. 60, No. 2, Nov. 2004, pp. 91–110.

[12] P. Scovanner, S. Ali, and M. Shah, "A 3-dimensional SIFT descriptor and its application to action recognition," *ACM Multimedia,* 2007, pp. 357–360.

[13] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld, "Learning realistic human actions from movies," *IEEEConference Computer Vision Pattern Recognition,*2008, pp.1–8.

[14] A. Kl¨aser,M. Marszałek, and C. Schmid, "A spatio-temporal descriptor based on 3-D-gradients," *Brit. Mach. Vision Conf.*, 2008, pp. 995–1004.

[15] Ling Shao, Simon Jones, and Xuelong Li," Efficient Search and Localization of Human Actions in Video Databases", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 24, No. 3, March 2014, pp. 504-512.

[16] Geng X, Zhou Z-H, Zhang Y, Li G, Dai H. "Learning from multimedia representation patterns for automatic age recognition," In ACM Conf. on Multimedia, pages 307–316, 2006

[17] GuodongGuo, Guowang Mu, Yun Fu, Charles Dyer, "A Study on Automatic Age Estimation using a Large Database", Computer Vision, 12th International Conference, pp. 1986 – 1991, IEEE, 2009.

[18] Guo G, Fu Y, Dyer, C.R., Huang, T.S. "Video-Based Human Pattern Recognition by Manifold Learning and Locally Adjusted Robust Regression," IEEE Trans.on Video Processing, Vol. 17, pp.1178-1188, 2008.

[19] Guo G, Fu Y, Huang T.S. and Dyer, C.R. "Locally Adjusted Robust Regression for Human Age Recognition," IEEE Workshop on Applications of Computer Vision, pages 1-6, 2008.

[20]GuodongGuo, Guowang Mu, Yun Fu, Thomas S. Huang, "Human Age Estimation Using Bio-inspired Features," conference on Computer Vision and Pattern Recognition, pp. 112 – 119, 2009.

[21]Serre T, Wolf L, Bileschi S, Riesenhuber M andPoggio T. "Robust Object Recognition with Cortex-Like Mechanisms," IEEE Trans. on PAMI, 29(3): 411–426, 2007.