Video Classification using Sine, Cosine, and Walsh Transform with Bayes, Function, Lazy, Rule and Tree Data Mining Classifier

Sudeep D. Thepade, Ph.D.
Computer Engineering Department
Pimpri Chinchwad College of Engg.
Savitribai Phule Pune University, India

Madhura M. Kalbhor
Computer Engineering Department
Pimpri Chinchwad College of Engg.
Savitribai Phule Pune University, India

ABSTRACT
Video classification has become one of the important research field as hundreds of videos are generated everyday which implies the need to build the classification system. To build faster and easy classification system, the visual content of video is used. Accuracy of classification depends upon the feature extraction which is one of the most important step in video classification. This paper proposes the use of orthogonal transform to generate the feature vector and to investigate effectiveness of different transforms (Cosine, Sine, and Walsh). Experimentation is carried on different sizes of feature vectors which are formed by taking fractional coefficients of row mean of column transformed video frames. Classification algorithm from different families such as Bayes (Naive Bayes and Bayes Net), Function (RBFNetwork and Simple Logistic), Lazy (IB1 and Kstar), Rule (Decision and Part) and Tree (BFTree, J48 Random Tree and Random Forest) are used for classification. Experimental results and its analysis have shown the Simple Logistic classifier with Sine transform to be better for proposed data mining based video classification technique.

Keywords
Content based video classification; Transform, Cosine ; Sine ; Walsh ; Classifier Bayes ; Function; Lazy; Rule; Tree classifier; Fractional content.

1. INTRODUCTION
Number of videos uploaded on internet is increasing day by day. Thus, managing this video data has become one of the most important tasks. In order, to efficiently store the video data and retrieve it accurately and efficiently videos can be stored under the suitable categories. Video cataloging can be used as preprocessing step for video retrieval.

Various cataloging systems have been proposed in literature. To classify the video under a specific category, feature of video are extracted to represent the Video. Based on these features the video is classified. In this paper, transformed content of video is used to generate features. Data mining classifiers are used to correctly find the class of the given video. To improve the classification accuracy with reduced complexity, various orthogonal transforms are applied to the columns of key frames. Row mean is taken of column transformed key frames to form the feature vector. Feature vector of various sizes are formed by considering the fractional coefficients of row mean.

To get the column transformed features, orthogonal transforms such as Cosine, Sine and Walsh are used. Twelve Data mining classifiers from the families like Bayes, Function, Rule, Lazy and Tree are used for classification.

2. LITERATURE SURVEY
Features of video such as color, shape, texture, motion, audio etc. are used for classification; the process is called as content based video classification. In recent years various methodologies have been implemented to extract these features in different ways to improve the accuracy of classification.

Various Video classification methods have been proposed in literature which can be broadly classified as text based classification and content based classification. Text based classification is a process in which each image is associated with a tag and classification takes place based on the tags. Most of the research is carried out on content based video classification. Content of video can be extracted using spatial domain [1] or transform domain techniques [2]. In proposed approach transforms are used to extract the content. Various transforms used are elaborated in Section A. Once feature vector is formed using the content of video the classifier is trained to correctly classify the videos. Section B describes the various classifiers used for classification in proposed work.

2.1 Orthogonal Transforms
Orthogonal transforms are used to extract the transformed content of the video. These three transforms used here are Cosine, Sine and Walsh is elaborated.

2.1.1 Discrete Cosine Transform.
Discrete Cosine transform (DCT) is used in many applications of digital signal processing for pattern recognition, information hiding, and content based image retrieval [3]. The two dimensional DCT can be written in terms of pixel values f(x, y) for x, y= 0, 1, … , N-1 and the frequency-domain transform coefficients F(u, v) as shown in equation(1).

$$F(u,v) = \alpha(u)\,\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left[\frac{(2x+1)u\prod}{2n}\right] \cos[\frac{(2y+1)v\prod}{2n}]$$

for 0< u, v <N -1 (1)

Where, $\alpha(u) \mid \propto (v) = \frac{1}{\sqrt{N}}$ for u = 0 |

for v = 0

$\alpha(u) \mid \alpha(v) = \frac{\sqrt{2}}{N}$ for 1 <u < N-1|

for 1 <v < N-1

2.1.2 Discrete Sine Transform

The two dimensional sine transform is defined by an equation (2).In signal and image processing sine transform is widely used [5].

$$F(u,v) = \frac{2}{N+1}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\sin\left[\frac{(x+1)(u+1)\prod}{N+1}\right]\sin\left[\frac{(y+1)(v+1)\prod}{N+1}\right]$$

for 0<u, v< N-1 (2)

2.1.3 Walsh Transform

Walsh functions were established as a set of normalized orthogonal functions, Analogous to sine and cosine functions, but having uniform values ± 1 throughout their segments [6].Walsh matrix, Wj has following properties.

- Wj takes on the values +1 and -1.
- Wj[0] = 1 for all j.
- Wj x WKT =0, for j ≠ k and Wj x WKT =N, for j=k.
- Wj has exactly j zero crossings, for j = 0, 1... N-1.
- Each row Wj is even or odd with respect to its midpoint.

2.2 Data Mining Classifiers

Classification is a process which has a set of predefined classes and determines which class a new object belongs to [7]. There are large numbers of classifiers available which are used to classify the data. This paper uses bayes, lazy, function, rule, and tree family classifier.

From the Bayes family, Navie Bayes and BayesNET algorithm are used. Class for a Naive Bayes classifier use estimator classes. Numeric estimator precision values are chosen based on analysis of the training data. Bayes Network learning using various searches algorithms and quality measures. Base class is used for a Bayes Network classifier. It provides data structures (network structure, conditional probability distributions, etc.) and facilities common to Bayes Network learning algorithms like K2.

From the Function family RBFNetwork and Simple logistic algorithms are used. Function family algorithm works using the mathematical equations. RBFNetwork implements a normalized Gaussian radial basis basis function network. Simple logistic Classifier is for building linear logistic regression models.

IB1 and KStar algorithms are used from Lazy family. IB1 is Nearest-neighbour classifier. IT uses normalized Euclidean distance to find the training instance closest to the given test instance, and predicts the same class as this training instance. If multiple instances have the same (smallest) distance to the test instance, the first one found is used. KStar is an instance-based classifier, which is the class of a test instance is based upon the class of those training instances similar to it, as determined by some similarity function.

DecisionTable and Part algorithms are used from Rule family. DecisionTable is used for building and using a simple decision table majority classifier. Part is for generating a PART decision list. Uses separate-and-conquer. Builds a partial C4.5 decision tree in each iteration and makes the "best" leaf into a rule.

From the Tree family BFTree, J48, RandomTree and RandomForest algorithms are used. BFTree is Class for building a best-first decision tree classifier. This class uses binary split for both nominal and numeric attributes. J48 is Class for generating a pruned or unpruned C4.5 decision tree. RandomForest is Class for constructing a forest of random trees. RandomTree is Class for constructing a tree that considers K randomly chosen attributes at each node.

3. PROPOSED IMAGE CLASSIFICATION APPROACH

In this paper fractional content of row mean of column transformed video frames is video as feature vector. Feature vector are formed of different sizes. Content of videos are extracted using Sine, Cosine and Walsh transform. Then classifiers belonging to different families are used for classification. The video data set is divided into training and testing dataset. Five key frames from each video are used to form the feature vector.

3.1 Feature Vector Generation

The process is elaborated in Fig1. The steps for feature vector generation can be given as.

1) Extract Key frames of video (5 frames per video).
2) Extract red, blue and green planes of each key frame.
3) To get Column transformed image apply transform (Cosine, Sine and Walsh) on column of each plane.
4) Calculate row mean.
5) Form feature vector by considering fraction coefficients of each plane. For example 100 coefficients of red plane, 100 coefficients of blue plane and 100 coefficients of green plane. This will form total 300 features.
6) Feature extracted for each key frame using step five are concatenated to form final feature vector. For example here 100 coefficients of red plane, 100 coefficients of blue plane and 100 coefficients of green plane. This will form total 300 features. Final feature vector size will be 300 X 5).

3.2 Query Execution

In query execution, for Query video feature vector is formed using Feature vector generation algorithm. Feature vector of query video is given to trained classifier to correctly classify the video.

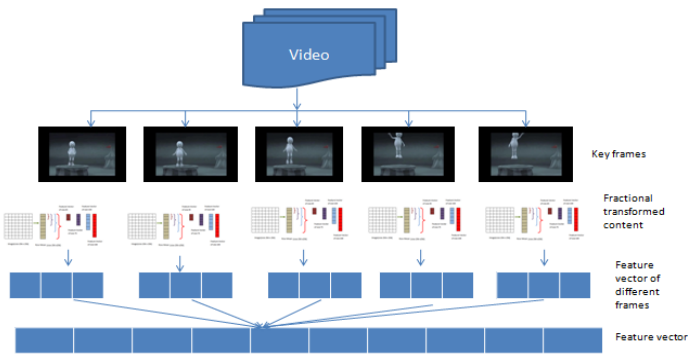

Fig1. Feature Extraction from visual video content.

4. EXPERIMENTATION ENVIRONMENT

The proposed technique is implemented using Matlab on computer with Intel core i5 processor and 4 GB RAM. Video database used which contains 300 videos of 6 different classes. Fig. 2 shows dataset sample.

Fig. 2. Testing Dataset Sample video from each of the six classes.

Classification Accuracy is used for performance evaluation to compare the techniques. The training is done using 180 videos. Total 120 queries are tested to get average accuracy. Accuracy is calculated for different feature vector size.

5. RESULTS AND DISCUSSION

For training 30 videos from each class are used which formed total 180 videos belonging to 6 different classes. 20 videos from each class are used for testing. Five key frames from each video is extracted to form the feature vector. After extracting the feature vector using different Cosine, Sine and Walsh transforms 12 different classifiers belonging to different families are used for classification. The Performance comparison of the considered transforms for respective data mining classifies is shown in Fig3 with reference to the average classification accuracy in percentage. Here Sine transform out performs the Cosine and Walsh transform. Table 1, Table 2 and Table 3 shows the percentage classification accuracy when Sine, Cosine and Walsh transforms are used respectively with different classifiers in proposed video classification technique.

From the results it can be seen that Sine transform with simple logistic Algorithm have given the best performance with the feature vector size as 25x3x5. Thus we can say using fractional content better classification is done with reduced complexity. Fig 4 gives the performance comparison of the considered data mining classifiers in video classification for respective orthogonal transforms. Here the highest video classification accuracy is observed in simple logistic classifier with sine transform.

Table 1. Average percentage classification accuracy of proposed content based image classification for various classifiers using cosine Transform.

Classifier		Feature Vector Size (Number of Cosine Coefficients Considered)						Average Classification Accuracy (%)
Family	Name	(15x3)	(25x3)	(50x3)	(75x3)	(100x3)	(256x3)	
Bayes	Navie Bayes	68.33	69.16	70.83	69.16	70.83	70.83	69.81
	Bayes Net	78.33	75.00	73.33	73.33	75.83	76.66	75.51
Fun-ction	RBFNetwork	66.66	66.66	69.16	67.05	67.05	75.83	68.88
	Simple Logistic	65.83	70.83	73.33	75.00	75.00	75.83	72.63
Lazy	IB1	70.83	70.83	70.83	70.00	70.00	67.05	69.99
	Kstar	55.83	37.05	31.66	26.66	24.16	16.66	32.08
Rules	DecisionTable	58.33	58.33	58.33	58.33	58.33	60.83	58.74
	Part	72.05	61.66	61.66	66.66	69.16	56.66	64.72
Tree	BFTree	71.66	55.00	56.66	60.00	59.16	57.05	60
	J48	70.83	60.00	58.33	62.05	62.05	62.05	62.77
	RandomForest	70.00	78.33	72.05	78.33	70.00	62.05	71.94
	RandomTree	63.33	59.16	61.66	51.66	59.16	67.05	60.41
							Total Average	63.95

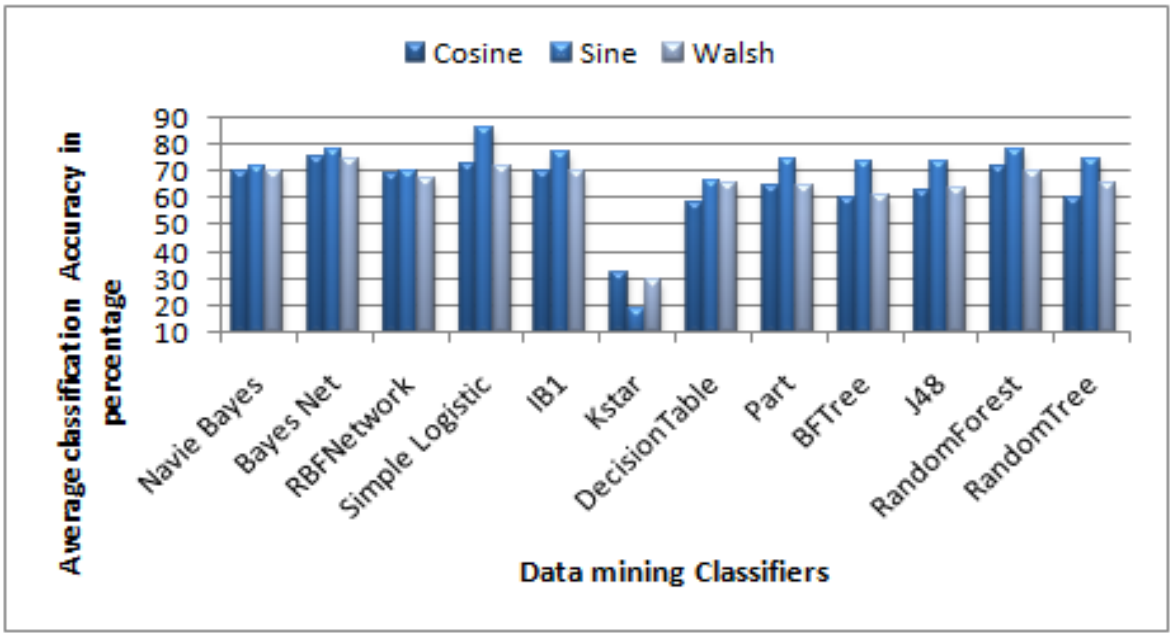

Fig 3. Performance comparison of considered transforms for respective data mining classifiers in proposed content based video classification technique

Table 2. Average percentage classification accuracy of proposed content based image classification for various classifiers using sine Transform.

Classifier		Feature Vector Size (Number of sine Coefficients Considered)						Average Classification Accuracy (%)
Family	Name	(15x3)	(25x3)	(50x3)	(75x3)	(100x3)	(256x3)	
Bayes	Navie Bayes	71.66	70.00	68.33	68.33	76.66	75.83	71.8
	Bayes Net	82.05	74.16	72.05	79.16	79.16	79.16	77.77
Fun-Ction	RBFNetwork	73.33	69.16	65.83	65.83	74.16	73.33	70.27
	Simple Logistic	83.33	90.83	90.00	88.33	80.00	85.83	86.38
Lazy	IB1	85.00	78.33	65.83	78.33	78.33	77.05	77.22
	Kstar	25.83	16.66	16.66	16.66	20.83	16.66	18.88
Rules	DecisionTable	63.33	67.05	69.16	69.16	70.83	60.00	66.66
	Part	68.33	74.16	78.33	80.83	71.66	70.83	74.27
Tree	BFTree	74.16	80.00	71.66	73.33	76.66	67.05	73.88
	J48	70.83	66.66	80.00	83.33	72.05	67.05	73.47
	RandomForest	76.66	75.83	83.33	75.00	79.16	81.66	78.61
	RandomTree	77.05	74.16	73.33	83.33	64.16	74.16	74.44
							Total Average	70.33

Table 3. Average Percentage Classification Accuracy Of Proposed Content Based Image Classification For Various Classifiers Using Walsh Transform.

Classifier		Feature Vector Size (Number of Walsh Coefficients Considered)						Average Classification Accuracy (%)
Family	Name	(15x3)	(25x3)	(50x3)	(75x3)	(100x3)	(256x3)	
Bayes	Navie Bayes	70.00	75.00	70.83	69.16	67.05	68.33	70.13
	Bayes Net	74.16	75.00	73.33	76.66	75.00	75.00	74.86
Func-Tion	RBFNetwork	65.83	66.66	65.83	67.05	67.05	70.00	67.22
	Simple Logistic	66.66	70.00	67.05	77.05	79.16	76.66	72.19
Lazy	IB1	67.05	70.83	71.66	71.66	72.05	66.66	70.13
	Kstar	52.05	32.05	28.33	25.00	21.66	16.66	29.44
Rules	DecisionTable	66.66	66.66	66.66	64.16	64.16	67.05	65.97
	Part	70.00	64.16	64.16	77.05	64.16	50.833	65.13
Tree	BFTree	66.66	70.83	67.05	53.33	51.66	60.00	61.33
	J48	70.83	72.05	65.83	65.83	58.33	50.00	63.88
	RandomForest	70.83	72.05	66.66	70.83	68.33	70.00	69.81
	RandomTree	62.05	69.16	69.16	60.83	63.33	68.33	65.55
							Total Average	64.63

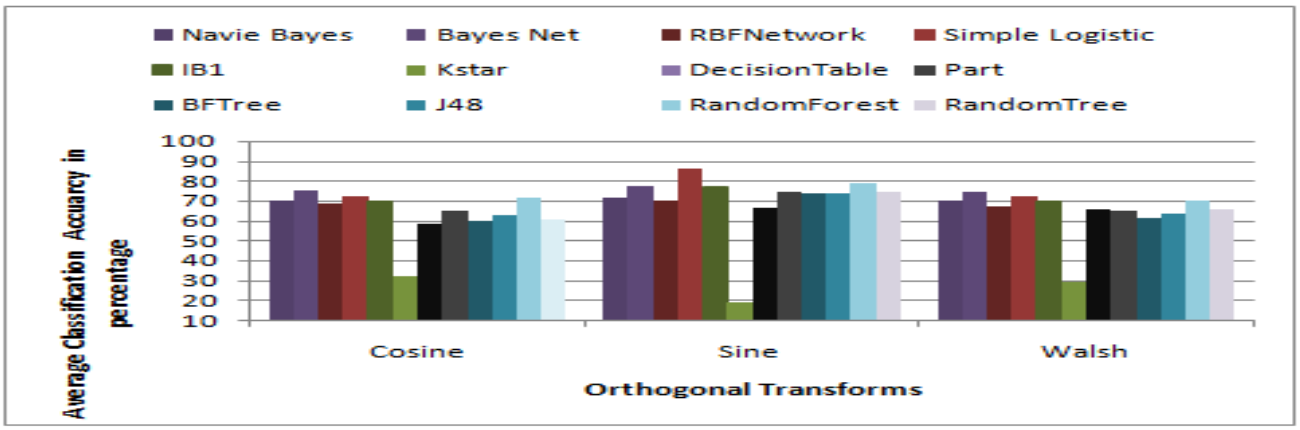

Fig 4. Comparison of data mining classifier for respective orthogonal transforms in proposed content based video classification techniques.

6. CONCLUSION

Novel data mining based approach of classifying video using data mining classifiers with orthogonal transform has been proposed here. Variations of proposed video classification techniques using twelve assorted data mining classifiers with three orthogonal transforms Cosine, Sine, Walsh and six feature vector sizes are analyzed with a video dataset having 300 videos. To reduce the complexity, transforms are applied to only columns of the key frames and row mean of such column transformed key frames are considered to generate feature vectors. The performances of proposed variations are composed with help of average percentage of video classification accuracy.

From the results is observed that use of sine transform have given better content based video classification with Simple Logistic. The fractional coefficients of row mean of column sine transformed key frames of video gives better result with less complexity, saving number of calculations and time than using whole images. Overall Simple Logistic classifier gives better classification among twelve classifiers considered here.

7. REFERENCES

[1] Dr. H.B. Kekre, Dr. Tanuja K., Sarode, Jagruti K. Save, "Image Classification in Transform Domain", (IJCSIS) International Journal of Computer Science and Information Security,Vol. 10, No. 3, March 2012

[2] Dr. Sudeep D. Thepade, Madhura M. Kalbhor, "Novel Data Mining based Image Classification with Bayes, Tree, Rule, Lazy and Function Classifiers using Fractional Row Mean of Cosine, Sine and Walsh Column Transformed Images." 2015 International Conference on Communication, Information & Computing Technology (ICCICT), Jan. 16-17, Mumbai, India

[3] H. B.Kekre, Sudeep D. Thepade, Akshay Maloo "Performance Comparison for Face Recognition using PCA, DCT &WalshTransform of Row Mean and Column Mean", ICGST International Journal on Graphics, Vision and Image Processing (GVIP), Volume 10, Issue II, pp.9-18, June 2010.

[4] http://infolab.stanford.edu/~wangz/project/imscreen/IDMS98/wang.ps.Z

[5] Dr. Sudeep D. Thepade, Rik Das," Performance Comparison of Feature Vector Extraction Techniques in RGB Color Space using Block Truncation Coding for Content Based Image Classification with Discrete classifiers" INDICON 2014.

[6] Sudeep Thepade, Rik Das, and Saurav Ghosh, A Novel Feature Extraction Technique Using Binarization of Bit Planes for Content Based Image Classification, Journal of Engineering Volume 2014 (2014).

[7] J. Han and M. Kamber, (2000) "Data Mining: Concepts and Techniques," Morgan Kaufmann.

[8] Mizianty, M. ; Kurgan, L. ; Ogiela, M.," Comparative Analysis of the Impact of Discretization on the Classification with Naïve Bayes and Semi-Naïve Bayes Classifiers". Machine Learning and Applications, 2008. ICMLA '08. Seventh International Conference.

[9] Ian H.Witten and Elbe Frank, (2005) "Datamining Practical Machine Learning Tools and Techniques," Second Edition, Morgan Kaufmann, San Fransisco.