Machine Learning Approach for Epileptic Seizure Detection Using Wavelet Analysis of EEG Signals

Abhishek Kumar
Department of Electrical Engineering,
Indian Institute of Technology Indore, India
abhikmr.757@gmail.com

Maheshkumar H. Kolekar
Department of Electrical Engineering,
Indian Institute of Technology Patna, India
mahesh@iitp.ac.in

Abstract—Analysis of EEG is the primary method for diagnosis of epilepsy. In this paper discrete wavelet transform is used for the time-frequency analysis of EEG signal. Using discrete wavelet transform, EEG signal is decomposed into five different frequency bands namely delta, theta, alpha, beta and gamma. Only theta, alpha and beta carry seizure information. Statistical feature like energy, variance and zero crossing rate and nonlinear feature like fractal dimension is extracted from each of the three sub bands and fed to support vector machine classifier. Support vector machine classifies the input EEG signal into seizure free and seizure signal. Experimental results show that the proposed method classifies EEG signals with excellent accuracy, sensitivity and specificity compared to the existing methods.

Keywords—EEG, fractal dimension, seizure, wavelet, Support Vector Machine, Gaussian Radial Basis Function

I. Introduction

The transient and unexpected electrical disturbances of the brain results in an acute disease called Epileptic seizures. Epilepsy is one of the most common diseases of the central nervous system. Germany has about 600,000 - 800,000 patients with epilepsy and every year;40,000 new cases are diagnosed. Altogether, approximately 10% of the population has experienced one single epileptic seizure while 3.5% of the population suffers from repeated seizures. The seizures themselves are marked with temporary impairments of motoricity, perception, speech, memory or consciousness. Seizures can vary from patient to patient, as can the degree of impairment, duration and frequency. It is probably the most prevalent brain disorder among adults and children. Over 50 million people worldwide are diagnosed with epilepsy, whose hallmark is recurrent seizures. Sometimes seizures may go unnoticed, depending on their presentation, and sometimes may be confused with other events, such as a stroke, which can also cause falls or migraines. Unfortunately, the occurrence of an epileptic seizure seems unpredictable and its course of action is very little understood. Research is needed for better understanding of the mechanisms causing epileptic disorders. Careful analysis of the electroencephalogram (EEG) signals can provide valuable insight into this widespread brain disorder [1].

EEG refers to the recording of the brain's spontaneous electrical activity over a short period of time, usually 20–40 minutes, as recorded from multiple electrodes placed on the scalp. A typical EEG signal, measured from the scalp, will have amplitude of about 10 μV to 100 μV and a frequency in the range of 1 Hz to about 100 Hz [2]. Traditionally, the EEG recordings were visually inspected by the trained neurophysiologist for detecting epileptic seizure or other abnormalities present. Since there is no definite criterion evaluated by the experts, visual analysis of EEG signals is insufficient. More ever due to human error, leads to improper diagnosis of the diseases causing fatal to human life. Therefore, some automatic computerized techniques have been used for this purpose. EEG signals are highly non-Gaussian, non-stationary and have a non-linear nature [2]. One traditional method used for EEG signal analysis is the method based on Fourier transform which is based on the assumption that EEG signal is stationary. But previous studies have shown that the frequency components of EEG signal change over time and hence EEG signal is a non-stationary process. Several time-frequency domain based methods have been developed for detection of epileptic seizure from EEG signals [2], [3], [7]. These methods include the short time Fourier transforms, the wavelet transform, and multi wavelet transform [8].

In this work an algorithm based on Discrete Wavelet Transform (DWT) and Support Vector Machine (SVM) classifier is used to detect the epileptic signal from the EEG signal. The DWT is used for time frequency analysis, giving quantitative evaluation of different frequency bands of clinical brain wave. Novelty lies in the selection of features and SVM kernels. The features extracted for classification includes both static as well as dynamic features. Static features include variance, energy and zero crossing rates (ZCR) [7] [8] while fractal dimension [9] is a dynamic feature. These features were used to train the SVM classifier to classify the seizure free and epileptic signal. This method offers the ultimate classification of the EEG segments regarding the presence or absence of seizures.

II. Dataset used for Training and Testing

The dataset includes recordings for both healthy and epileptic subjects. The data has five subsets denoted as Z, O, N, F, and S, each containing 100 single-channel EEG signals, each signal of 23.6 s in duration. The subsets Z and O have been acquired using surface EEG recordings of five healthy volunteers with eyes open and closed, respectively. The signals in two subsets have been measured in seizure-free intervals from five patients in the epileptogenic zone (subset F) and from the hippocampus formation of the opposite hemisphere of the brain (subset N). Finally, the subset S contains seizure activity

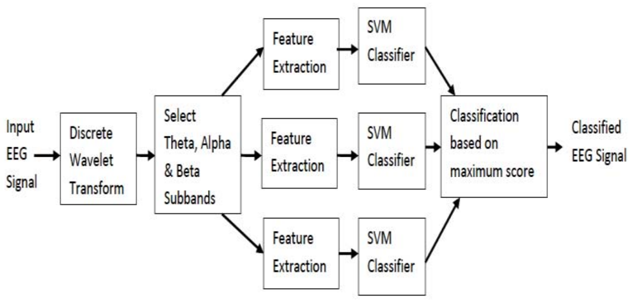

Fig-1: Block diagram of proposed scheme

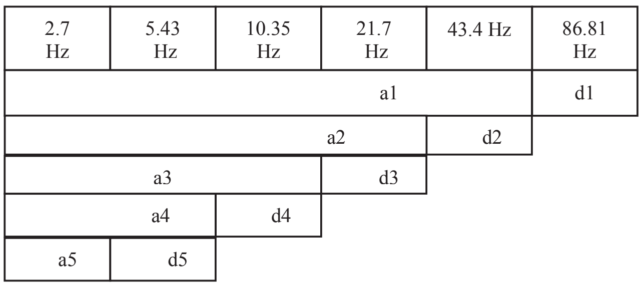

Fig-2: Wavelet decomposition of EEG

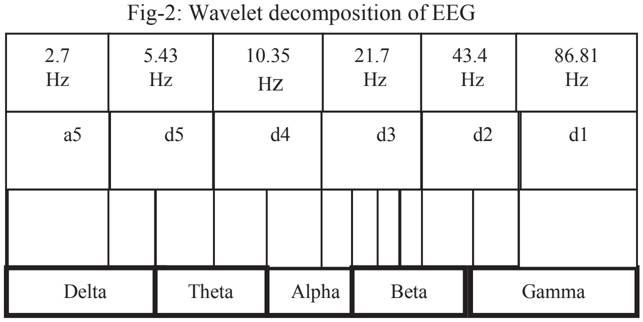

Fig-3: Generation of five EEG sub-bands by merging wavelet decomposed bands.

TABLE I: DATASET FOR TRAINING AND TESTING

	SEIZURE SAMPLES (S)	SEIZURE FREE SAMPLES (N)	TOTAL
TRAINING	50	50	100
TESTING	100	100	200

selected from all recording sites exhibiting ictal activity. The subsets N, F, and S have been recorded intra-cranially. All EEG signals were recorded with the same 128-channel amplifier system, using an average common reference as given in [1]. The data were digitized at a sampling rate of 173.61 Hz using 12-bit analog to digital (A/D) resolution. The details of the EEG samples used for training and testing are given in Table-1. The complete S and N subsets were used for experimentation. Details about the placement of electrodes to measure the EEG signals are given in the appendix section.

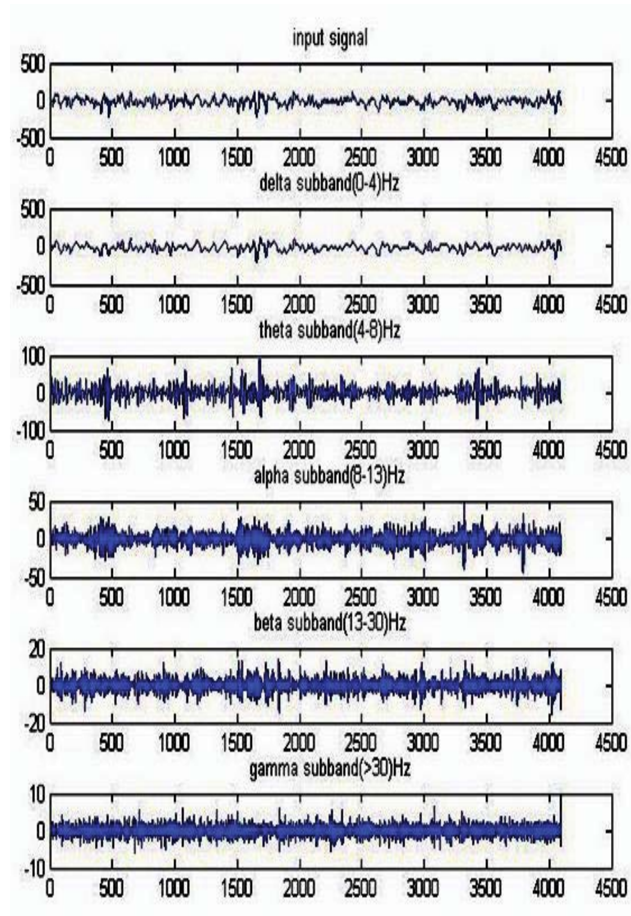

Fig-4: Decomposition of seizure free EEG signal

III. Proposed Method

The block diagram of the proposed method is shown in Fig.1. The main processing steps are Wavelet decomposition, feature extraction and applying an SVM classifier. Each EEG segment is decomposed into five EEG sub-bands using discrete wavelet transform from which three useful bands are selected. The statistical and nonlinear features are extracted from 3 significant sub-bands to form feature vectors. The SVM is trained using such feature vectors extracted from 50 EEG signals. The feature vectors from each band are fed to a separate SVM classifier which in turn makes the final decision based on maximum score.

Wavelet Decomposition

When the EEG sample is subjected to 5-level wavelet transform, it gets decomposed into different bands as shown in Fig.2. EEG spectrum contains some characteristics waveforms that fall primarily within five frequency bands: delta (<4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz) and gamma (>30 Hz) [3]. DWT is used for decomposing EEG signal into these five sub bands as shown in Fig.4 and 5. Then inverse discrete wavelet transform is used to recombine the various frequency bands to form delta, theta, alpha, beta and gamma sub-bands as shown in Fig-3. Since the seizures in recorded intracranial EEGs (IEEGs) occur between 3 and 29 Hz, the coefficients in theta, alpha and beta are chosen. The sampling frequency of EEG dataset obtained is 173.61 Hz.

According to the Nyquist sampling theorem, the maximum useful frequency is half of the sampling frequency (i.e. 86.81 Hz). We have experimented with the decomposition of EEG signal using Haar, Daubechies and Coiflets as basic functions [9], [11].

IV. Feature Extraction

A. Energy

The energy of a signal is defined as the sum of squared modulus of the sample values. The energy of the signal is expressed as

$$E = \sum_{n=0}^{N-1} |X_n|^2 \qquad (1)$$

Where X_n is the samples values and N is the total number of samples in our selected band.

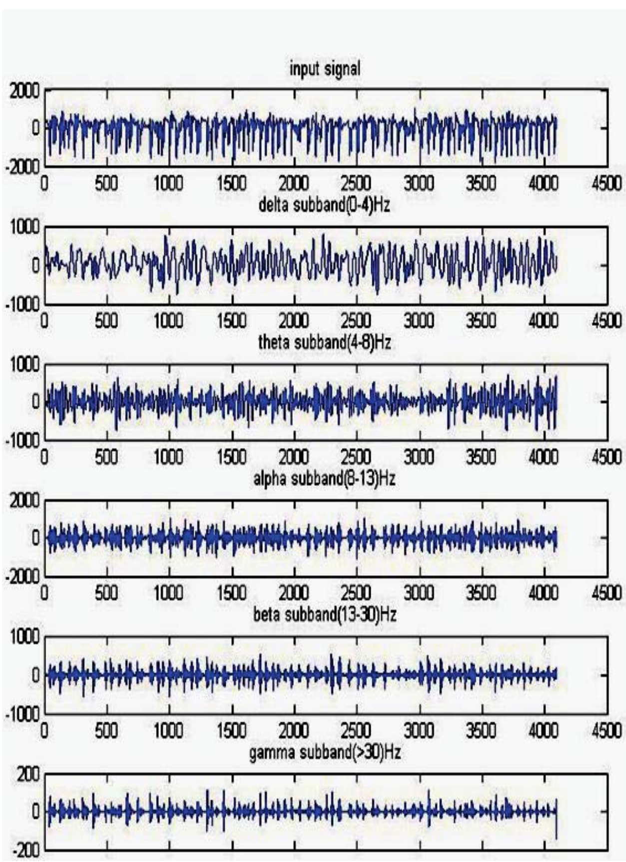

Fig-5: Decomposition of EEG signal with seizure

B. Variance

The variance is a measure of how far a set of numbers is spread out. It is one of several descriptors of a probability distribution, describing how far the numbers lie from the mean (expected value). For a set X variance can be defined as

$$V = \frac{\sum (X - \mu)^2}{N} \quad . \qquad (2)$$

Where μ is the mean value of the set X and N is the number of samples.

C. Zero Crossing Rate

The zero-crossing rate (ZCR) is the rate at which the signal changes from positive to negative or back. ZCR [7] [8] is calculated for all the five sub-bands.

D. Fractal Dimension

The Fractal dimension (FD) of a waveform is a powerful tool for transient detection. It can be used to measure the geometrical complexity of a time series. It is useful in quantifying the complexity of dynamic signals in biology and medicine. Various algorithms are available for computing the fractal dimension [9]. FD using Katz's method for EEG signal is given by:

$$D^{katz} = \frac{\log_{10}(L)}{\log_{10}(d)} \qquad (3)$$

Where L is the sum of distances between successive points, and d is the distance between the first point of the sequence and the point of the sequence that provides the farthest distance [9]. Mathematically d is expressed as follows:

$$d = \max \left(\| x(1), x(i) | \right) \qquad (4)$$

Compares the actual number of units that compose a curve with the minimum number of units required to reproduce a pattern of the same spatial extent.

V. Support Vector machines

Support vector machine is a supervised learning model used for binary classification.SVM Based on statistical learning theory and structural risk minimization is regarded as a powerful tool for pattern classification. The SVM algorithm involves training with feature vectors (X_i) of signals of two different classes i.e. seizure and seizure free classes to learn a decision boundary that separates these two classes. Once the decision boundary is learned, the SVM algorithm determines the class membership of a newly observed feature vector (X_i) based on which side of the boundary the vector falls. An example of learning of decision boundary (hyperplane) by SVM is shown in fig. 6.

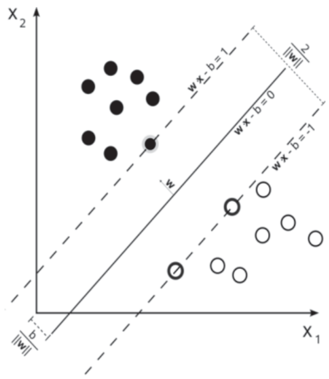

Fig 6: SVM learning

Optimal hyperplane is represented by

$$W^T X + b = 0 \quad (5)$$

Where X is the feature vector, W is normal vector to the hyperplane and b determines offset of the hyperplane with origin. Values of W and b are evaluated by solving a quadratic optimization problem. The optimization problem is formulated on the basis of maximization of margin between two parallel hyperplanes. As shown in fig. 6 margin is inversely proportional to 'W'. Hence maximization of margin is equivalent to minimization of 'W' or $\frac{1}{2}W^T W$. In this work linear soft margin formulation of SVM has been used, which finds an optimal linear separating hyperplane while performing a trade-off between training error and the model complexity. The corresponding optimization problem can be formulated as

$$\begin{aligned} &\text{minimize} && \frac{1}{2}W^T W + C\sum_{i=1}^{n}\xi_i && \\ &\text{subject to} && y_i(\mathrm{W}^T \mathrm{X}_i + \mathrm{b}) \geq 1 - \xi_i && i = 1, 2 \ldots n \quad (6) \\ & && \xi_i \geq 0 && i = 1, 2 \ldots n \end{aligned}$$

Where $\{X_i, y_i\}_{i=1}^{n}$ are N training input-output pairs, $y_i = 1$ or -1 if X_i belongs to class 1 or 2 respectively, ξ_i is called slack variable and C is an arbitrary constant.

The Lagrangian of equation 2 can be defined as

$$\begin{aligned} L(W, b, \xi, \mu, \lambda) = \frac{1}{2}W^T W + \\ C\sum_{i=1}^{n}\xi_i + \sum_{i=1}^{n}\mu_i\{1 - \xi_i - y_i(W^T \mathrm{X}_i + \mathrm{b})\} - \sum_{i=1}^{n}\lambda_i\xi_i \end{aligned} \quad (7)$$

Where μ_i and λ_i are the Lagrange multipliers. On applying Kuhn Tucker Condition to above equation and Solving the dual optimization problem of equation (3) the values of W and b (parameters of optimal hyperplane) [1] can be evaluated as

$$\begin{aligned} W &= \sum_{i=1}^{n}\mu_i y_i X_i \\ b &= y_i - X_i^T W \quad i \; s.t. \;\; 0 < \mu_i < C \end{aligned} \quad (8)$$

Once the parameters of optimal hyperplane is evaluated, SVM decides the class membership of any new feature vector X by putting all these values in decision function formulated as

$$f(X) = sign(W^T X + b) \quad (9)$$

One other popular formulation of SVM is designed with the use of nonlinear kernel function. This SVM model is well suited for the data which are not linearly separable in the feature space.

VI. Results

In this paper 100 seizure free and 100 seizures EEG signal is used to test the performance of the system. Haar, db6 and coif4 wavelets are used for the decomposition. Table-II shows the classification performance of the proposed scheme with different SVM kernels and wavelet basis functions. The classification performance of the proposed method can be determined by the computation of sensitivity, specificity and accuracy. The sensitivity (SN), specificity (SP), and accuracy (AC) are defined as follows:

1) Sensitivity: Number of true positives/the total number of seizure segments labeled by the EEG experts. True positive represents a detected seizure segment by the algorithm was also identified as seizure by the EEG experts.

$$\mathrm{SN} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100 \quad (10)$$

2) Specificity: Number of true negatives/the total number of seizure free segments labeled by the EEG experts. True negative represents a segment labeled as seizure free both by the algorithm and by the EEG experts.

$$\mathrm{SP} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \times 100 \quad (11)$$

3) Recognition accuracy: Number of correctly identified segments/total number of segments.

$$\mathrm{AC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \times 100 \quad (12)$$

Here TP and TN represent the total number of correctly detected true positive events and true negative events respectively. The FP and FN represent the total number of erroneously positive events and erroneously negative events respectively. In most of the cases, the performance indices showed by the three methods of wavelet transform were quite similar to each other. It can also be observed that the statistical parameters reach maximum while using db6 and coif4. The reason is that the basis-functions of db6 and coif4 approximate the EEG signal more accurately. Table III presents a comparison on the results with other methods. It can be observed that proposed method shows improved sensitivity (98%) compared to the existing schemes for seizure detection. The improved sensitivity is due to the features like FD, ZCR and SVM kernel functions like polynomial and GRBF.

TABLE II: CLASSIFICATION PERFORMANCE OF VARIOUS SVM KERNELS

Kernel function (parameter)	Statistical parameter	Haar wavelet	Db6 wavelet	Coif4 wavelet
Polynomial (d' = 2)	SN	97	97.5	98
	AC	95.5	96	96
	SP	95.5	96	96
Polynomial (d' = 3)	SN	97	97.5	98
	AC	96	96.5	96.5
	SP	96	96.5	94.5
Polynomial (d' = 4)	SN	97	97.5	98
	AC	96	96.5	98
	SP	95	96.5	94.5
RBF (σ=1)	SN	97	97.5	98
	AC	96	96.5	97.5
	SP	95	96	96

The usage of wavelet analysis and SVM can be justified because wavelet analysis is applied to non-stationary processes with the advantages both in the time and frequency domains.SVM has been demonstrated many unique advantages in resolving small sample, nonlinear and high dimensional pattern recognition.

TABLE-III

COMPARISON WITH DIFFERENT METHODS

Method	SN (%)	Number of seizures selected
Differential windowed variance [10]	91.525	59
A fuzzy rule based system[11]	68.9	78
Patient specific seizure detection [12]	78	63
Multistage seizure detection [13]	87.5	24
Liu et al [14]	94.46	82
Proposed method	98	100

VII. Conclusion

An expert model was developed for detection of epileptic seizures in EEG signals by using discrete wavelet transform and support vector machine. The wavelet transform is used to decompose EEG signal into five sub bands. The statistical feature like energy, variance and zero crossing rate and nonlinear feature like fractal dimension were extracted from theta, alpha and beta sub bands. Different kernel functions were used in the SVM classifier to classify seizure EEG signal and seizure free EEG signal. The performance of the proposed method is much better than other results available in the literature for detecting epileptic seizure in EEG signals. In future, we are planning to use probabilistic approaches such as Bayesian Belief Network [15], Hidden Markov Model [16] with large dataset to improve the performance of classifier.

References

[1] Andrzejak RG, Lehnertz K, Rieke C, Mormann F, David P, Elger CE "*Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state*", Physical Review E, Vol. 64, pp. 061907-1: 8, 2001.

[2] D. Puthankattil Subha, Paul K. Joseph, Rajendra Acharya U and Choo Min Lim, "*EEG Signal Analysis: A Survey*", Journal of Medical Systems, Vol.34 (2), pp 195-212,2010.

[3] B. Boashash, L. Boubchir and G. Azemi, "*Time-frequency signal and image processing of non-stationary signals with application to the classification of newborn EEG abnormalities*" IEEE Int. Symposium on Signal Processing and Information Technology, 2011.

[4] J R Panda, P S Khobragade, P D Jambhule, S N Jengthe,PRPal, T K Gandhi, "*Classification of EEG Signal Using Wavelet Transform and Support Vector Machine for Epileptic Seizure Diction*", Int. Conf. on Systems in Medicine and Biology, India 2010.

[5] Alexandros T. Tzallas, Markos G. Tsipouras, and Dimitrios I. Fotiadis, "*Epileptic Seizure Detection in EEGs Using Time–Frequency Analysis*", IEEE Transactions on Information Technology in Biomedicine, Vol 13(5), 2009.

[6] Ling Guo, Daniel Rivero, Alejandro Pazos, "*Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks*", Journal of Neuroscience Methods, pp. 156–163, 2010.

[7] M H Kolekar, S Sengupta, "*Hierarchical structure for audio-video based semantic classification of sports video sequences*", Int Conf on Visual Communications and Image Processing, 2005

[8] M H Kolekar, K Palaniappan, S Sengupta, G Seetharaman, "*Semantic concept mining based on hierarchical event detection for soccer video indexing*" , Journal of multimedia 4 (5), 298-312, 2009

[9] M. H. Kolekar, S. N. Talbar, T. R. Sontakke, "*Texture Segmentation using Fractal Signature*", IETE Journal of Research, 46(5), 2000.

[10] K. K. Majumdar and P. Vardhan, "*Automatic seizure detection in ECoG by differential operator and windowed variance*," IEEE Trans.Neural Syst. Rehabil. Eng., vol. 19, no. 4, pp. 356–365, 2011.

[11] A. Aarabi, R. Fazel-Rezai, and Y. Aghakhani, "*A fuzzy rule-based system for epileptic seizure detection in intracranial EEG*," Clin. Neurophysiol.,vol. 120, pp. 1648–1657, 2009.

[12] E. C.-P. Chua, K. Patel, M. Fitzsimons, and C. J. Bleakley, "*Improved patient specific seizure detection during pre-surgical evaluation*", Clin.Neurophysiol., vol. 122, pp. 672–679, 2011.

[13] S. Raghunathan, A. Jaitli, and P. P. Irazoqui, "*Multistage seizure detection techniques optimized for low-power hardware platforms*," Epilepsy Behav., vol. 22, pp. S61–S68, 2011.

[14] Liu, Y, Zhou, W, Yuan, Q, & Chen, S,"*Automatic Seizure Detection Using Wavelet Transform and SVM in Long-Term Intracranial EEG*", IEEE Trans. on neural Systems and rehabilitation engineering, Vol. 20(6), pp. 749-755, 2012.

[15] M H Kolekar, "*Bayesian belief network based broadcast sports video indexing*", Multimedia Tools and Applications 54 (1), pp. 27-54, 2011.

[16] M H Kolekar, S Sengupta, "*Semantic indexing of news video sequences: a multimodal hierarchical approach based on hidden Markov model*", IEEE Region Ten Conference (TENCON), pp. 1-6, 2005.