Efficient ECG Data Compression and Transmission Algorithm for Telemedicine

Chandan Kumar Jha
Department of Electrical Engineering
Indian Institute of Technology
Patna, India
ckjha.bit@gmail.com

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

***Abstract*—This paper presents an efficient electrocardiogram (ECG) data compression and transmission algorithm based on discrete wavelet transform and run length encoding. The proposed algorithm provides comparatively high compression ratio and low percent root-mean-square difference values. 48 records of ECG signals are taken from MIT-BIH arrhythmia database for performance evaluation of the proposed algorithm. Each record of ECG signals are of duration one minute and sampled at sampling frequency of 360 Hz over 11-bit resolution. Discrete wavelet transform has been used by means of linear orthogonal transformation of original signal. Using discrete wavelet transform, signal can be analyzed in time and frequency domain both. It also preserves the local features of the signal very well. After thresholding and quantization of wavelet transform coefficients, signals are encoded using run length encoding which improves compression significantly. The proposed algorithm offers average values of compression ratio, percentage root mean square difference, normalized percentage root mean square difference, quality score and signal to noise ratio of 44.0, 0.36, 5.87, 143, 3.53 and 59.52 respectively over 48 records of ECG data.**

Keywords—discrete wavelet transform; run length encoding; telemedicine; compression ratio;

I. INTRODUCTION

India is a developing country where to provide medical facilities to all over the country is a great challenge. Rural areas are intact with the good medical facilities and expert doctors. In this regard, telemedicine [1] can be proved as a boon for people of rural areas. Telemedicine is the application telecommunication and information technology for health care purposes at a distant rural area. Using telemedicine medical data can be sent from remote areas to expert doctor's premises. After reception and analysis of the medical data by expert doctor, important medical suggestions can be sent to remote areas for proper health care of patient. Thus using telemedicine, people of rural areas can be benefited by the service of expert doctors. At present, heart diseases are the number one cause of death globally and it increasing exponentially. In this regard, to provide medical facilities to heart patients of remote areas is a prime concern. Generally people of rural areas are suffered with the problem of unavailability of expert cardiologists. This problem can be solved by the application of telemedicine. Using telemedicine, ECG data can be sent from rural areas to expert cardiologist's premises. For proper health monitoring of the heart patient, expert cardiologists can send required medical suggestions to remote areas after analysis of the received ECG data. ECG is recording of electrophysiological signal generated due to cardiac functionality of the human heart [2]. Space required for storage of ECG data can be minimized using ECG data compression techniques. Compressed ECG data can also be transmitted efficiently using telemedicine.

During past several decades, various methods of ECG data compression and transmission have been developed. These schemes can be roughly classified into direct, transform and parametric methods. Direct method extracts significant data samples of original signals directly for compression. The commonly used direct methods include amplitude zone time epoch coding (AZTEC) [3], turning point (TP) [4], scan along polygonal approximation (SAPA) [5], co-ordinate reduction time encoding system (CORTES) [6], delta coding and entropy coding [7]. Transform methods use decomposition of original signal by means of orthogonal linear transformation. Further compression performance is improved by proper encoding of expansion coefficients. Various linear

978-1-4673-9622-6/16/$31.00 ©2016 IEEE

transformation methods include Hermite transform, discrete cosine transform (DCT) [8], Fourier transform, Walsh Transform, Karhunen-Loeve transform (KLT) and wavelet transform [9]-[12]. Parametric methods are the combination of direct and transform methods. Examples of parametric methods are beat codebook [13], artificial NN, peak picking and vector quantization [14].

There are also other techniques, which uses ECG data for arrhythmia detection and prediction based on Bayesian network [15], [16] , hidden Markov model classifier [17], [18] and support vector machine classifier [19]. After reception and analysis of the compressed ECG data, expert cardiologists can detect and predict arrhythmia using these techniques and can send required medical suggestions to the patient using telemedicine.

This paper reports an efficient ECG signal compression and transmission algorithm which uses linear transform by means of discrete wavelet transform and run length encoding for further compression improvement. The proposed algorithm offers comparatively high compression ratio and low percent root mean square difference. This paper includes four sections. Section II focuses on methodology used by the proposed algorithm. Section III elaborates results and discussion part after the implementation of the proposed algorithm over 48 records of ECG signals which are taken from MIT-BIH arrhythmia database [20]. Conclusion and future scope have been included in section IV.

II. MATERIALS AND METHODS

A. *Overview*

The proposed algorithm includes compression and decompression of ECG signal which are summarized in Fig.1 and Fig. 2 using block diagrams. Fig. 1 elaborates the compression procedure which is used at transmitter section of telemedicine system whereas decompression procedure is used at receiver section which is shown in Fig. 2. For implementation of compression procedure, 48 records of ECG signal have been pre-processed before applying the backward signal difference. The pre-processing of ECG signal involves filtering with a 3-degree Savitzky-Golay filter (SGF) using 21 points constant window. SGF provides smoothing of a noisy signal whose frequency span is large. SGF is also called digital smoothing polynomial filter which is more efficient to preserve high frequency component of signal compared to finite impulse response (FIR) filters. Polynomial degree and dimension of window of the SGF is empirically adopted after testing these parameters. Fig. 3 shows original signal and Savitzky –Golay filtered signal waveforms for record no 117. The pre-processing of ECG signal also involves notch filtering which reduces the power line interference of 50 Hz. After pre-processing, backward signal difference has been applied to the ECG signal which compresses it by 50% with lossless [21]. Threshold and quantization of transform coefficients of ECG signal have been implemented after applying discrete wavelet transform. Using run length encoding, further compression has been achieved at the transmitter. To reconstruct the signal at receiver, inverse run length encoding, inverse discrete wavelet transform and inverse backward difference have been used.

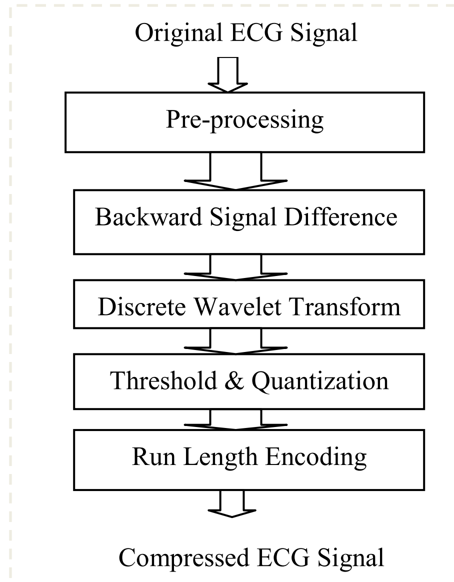


Fig. 1 Compression Procedure

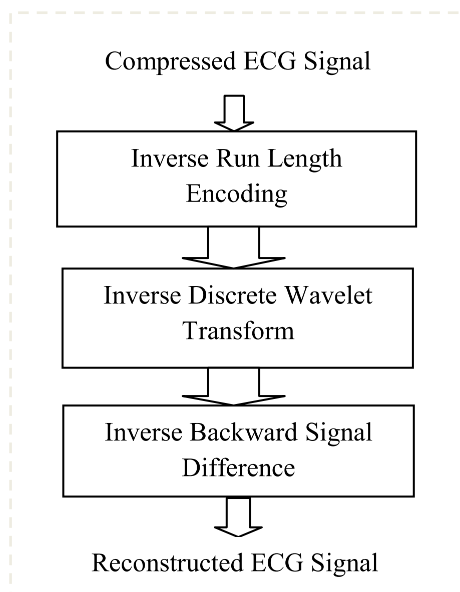


Fig. 2 Decompression Procedure

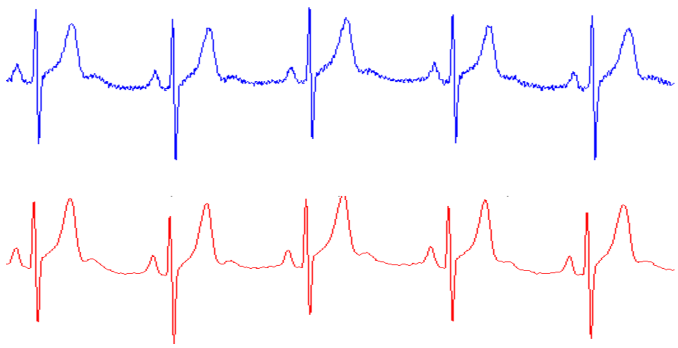


Fig. 3: For record no. 117 (a) Original Signal (b) Savitzky-Golay Filtered Signal

B. *Wavelet Transform*

Using wavelet transform [22], [23], [24] analysis of a signal can be performed in time and frequency domain both. It also preserves the local features of the signal very well. Hence it is well suited for the compression of biomedical signal, where to keep safe the features of the signal is more concerned.

A mother wavelet is a function ψ (t) with zero average in time domain [25]

$$\text{i.e,} \quad \int_{-\infty}^{+\infty} \psi(t)\, dt = 0 \qquad (1)$$

A daughter wavelet function with scale parameter p and translation parameter q and can be represented by

$$\psi^{p,q}(t) = \frac{1}{\sqrt{p}} \psi\left(\frac{t-q}{p}\right) \qquad (2)$$

The wavelet transform of a function x (t) can be computed by the cross correlation of x (t) with daughter wavelet function $\psi^{p,q}(t)$ and can be represented as

$$X_W(p, q) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\, \psi^*\left(\frac{t-q}{p}\right) dt \qquad (4)$$

Discrete wavelet transform is the digital implementation of (4).

C. *Run Length Encoding*

Run length encoding has been used to encode transform coefficients of ECG signal which is repetitive in nature. DWT coefficients contains long run of data and can be stored as a single data value. Run length encoding uses repetition of data and it provides improved compression ratio of ECG signal.

D. *Performance Parameters*

For evaluation of ECG data compression and transmission

Table 1 Compression Results of the Proposed Algorithm

Record	CR	PRD	PRDN	QS	RMS	SNR
100	39.9	0.23	7.34	167	2.16	53.8
101	48.7	0.18	4.48	267	1.73	59.4
102	27.9	0.5	15.5	55.5	5.32	69.5
103	33.9	0.33	5.59	100	3.19	58.9
104	75	0.34	7.02	217	3.71	55.2
105	59	0.34	5.62	169	3.4	55.8
106	65.2	0.36	5.73	180	3.34	58
107	39.6	0.66	4.71	59.6	7.63	59.5
108	75.2	0.3	10	249	3.47	61.2
109	38.6	0.38	5.06	99.7	4.17	60.8
111	46.1	0.24	6.1	188	2.54	61.4
112	53.9	0.22	5.12	244	1.9	60.7
113	48.3	0.32	4.06	149	3.18	62.6
114	42.9	0.21	9.51	201	2.04	64.3
115	46.8	0.28	4.3	164	2.59	61.8
116	38.7	0.5	3.75	76.9	4.15	62.7
117	47	0.31	6.4	148	2.78	61.4
118	42.6	0.67	7.1	63.6	5.71	51.2
119	38.5	0.47	3.42	81.8	3.95	63.4
121	45	0.18	2.73	248	1.54	72
122	20.2	0.29	3.7	68.9	2.52	65.3
123	45.7	0.35	6.25	128	3.11	59.2
124	43.8	0.29	3.58	148	2.5	64.1
200	50	0.37	4.83	136	3.73	55.5
201	33	0.18	4.75	175	1.8	59.2
202	54.6	0.29	6.09	186	2.92	54.3
203	43.8	0.5	4.8	85.5	4.88	55.1
205	43.6	0.17	5.04	249	1.57	59.9
207	64.3	0.34	3.57	187	2.56	67.1
208	52.8	0.47	5.33	111	5.14	58.8
209	36.2	0.24	6.49	148	2.52	56.4
210	46.6	0.18	3.82	253	1.83	64.2
212	26.9	0.4	6.59	67.2	3.96	60.7
213	48.4	1.19	11.19	40.6	11.7	45.4
214	34.7	0.43	4.96	79	4.48	61.4
215	50.7	0.51	10.4	98.9	5.24	45.7
217	29.5	0.36	3.23	82	3.93	66.5
219	30.7	0.43	4.25	70	3.61	60.6
220	46.2	0.32	5.67	144	3.02	56.2
221	46.9	0.26	4.75	174	2.62	60.7
222	41.9	0.16	6.63	247	1.72	58.6
223	33.6	0.31	3.72	107	2.84	64
228	42.1	0.44	6.64	94	4.48	58.6
230	40.8	0.48	7.91	83.9	4.8	51.5
231	40.9	0.28	5.19	142	2.85	59.3
232	53.2	0.2	8.86	265	1.97	60.4
233	29.2	0.63	5.9	46.2	6.5	55.3
234	33.6	0.23	4.16	141	2.4	60.9
Avg.	44.0	0.36	5.87	143	3.53	59.5

algorithms, various performance parameters are used [26]. Some of these performance parameters are as follows:

1) Compression Ratio (CR) can be defined as

$$CR = \frac{\text{size of original signal in bytes}}{\text{Size of reconstructed signal in bytes}} \quad (5)$$

A high value of CR indicates high compression performance.

2) Percentage Root Mean Square Difference(PRD%) can be represented by

$$PRD = \sqrt{\frac{\sum_0^{N-1}(X(n)-Y(n))^2}{\sum_0^{N-1}(X(n))^2}} \text{ x } 100\ \% \quad (6)$$

Where X(n) is the original signal and Y(n) is the reconstructed signal. For a good compression performance PRD % values should be low.

3) Normalized Percentage Root Mean Square Difference (PRDN) can be defined as

$$PRDN = \sqrt{\frac{\sum_0^{N-1}(X(n)-Y(n))^2}{\sum_0^{N-1}(X(n)-\text{mean }(X(n)))^2}} \text{ x } 100\ \% \quad (7)$$

4) Quality Score (QS) can be represented as

$$QS = \frac{CR}{PRD} \quad (8)$$

A high value of quality score indicates a high quality of compression.

5) Root Mean Square Error (RMS) can be defined as

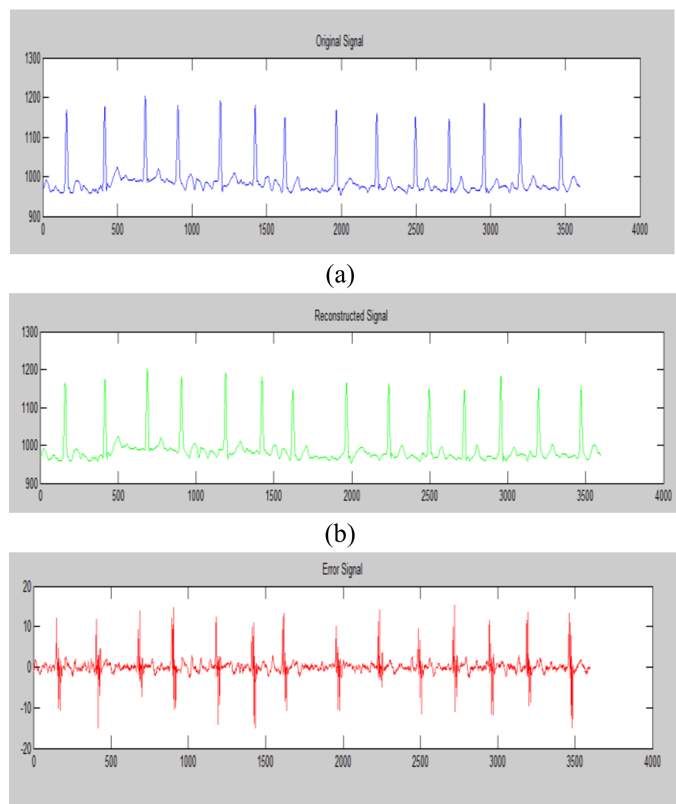
$$RMS = \sqrt{\frac{\sum_0^{N-1}(X(n)-Y(n))^2}{N-1}} \quad (9)$$

6) Signal to noise ratio (SNR) can be represented as

$$SNR = 10 \times \log\left(\frac{\sum_0^{N-1}(X(n)-\text{mean}(X(n)))^2}{\sum_0^{N-1}(X(n)-Y(n))^2}\right) \quad (10)$$

III. RESULTS AND DISCUSSION

This section describes the performance evaluation of the proposed algorithm over 48 records of ECG data taken from MIT-BIH database. Compression results are tabulated in Table 1. Performance comparison of different algorithms with the proposed algorithm for record no. 100, 117 and 119 has been shown in Table 2 and from this table it is clear that the proposed algorithm offers best compression performance. Original signal, reconstructed signal and error signal for record no. 201 are shown in Fig. 6 and it is clear that the reconstructed signal preserves characteristic features of the original signal very well. Fig. 7 shows the CR and PRD values plots for 12 records of ECG signal. From these plots, it can be seen that there is not significant variation in CR and PRD (%) values. It implies that the proposed algorithm is suitable for compression and transmission of various morphologies of ECG data.

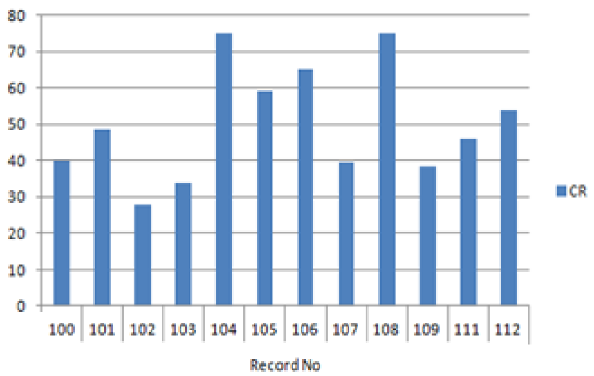


(a) (b) (c)

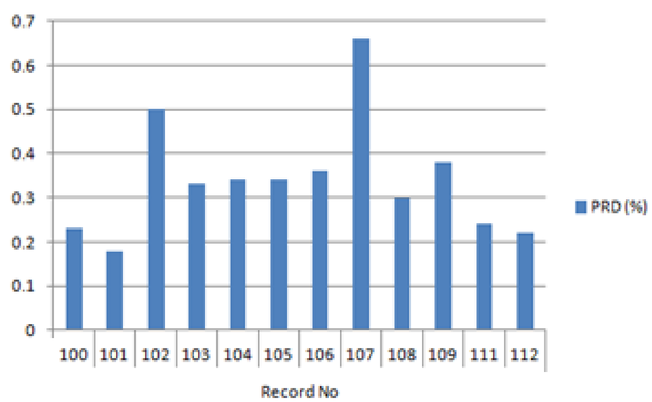
Fig. 6: (a) Original Signal (b) Reconstructed Signal (c) Error Signal of MIT-BIH record no. 201

Table 2 Comparison of Compression Results

Algorithm	CR			PRD		
	Record No.			Record No.		
	100	117	119	100	117	119
Proposed	**39.9**	**39.6**	**38.5**	**0.23**	**0.66**	**0.47**
Lee et. al [26]	24.4	18.6	3.93	1.17	3.93	2.05
Chou et. al [27]	24.0	13.0	20.9	4.06	1.18	1.81
Fira et. al [28]	20.0	12.7	18.7	0.44	0.61	1.03
Sandryhaila [29]	23.5	24.9	23.5	0.51	0.60	1.09
Kumar et.al [30]	39.8	27.1	21.7	5.39	3.91	2.92



(a)



(b)

Fig. 7: (a) CR and (b) PRD for 12 records

IV. CONCLUSION AND FUTURE SCOPE

In this paper, an efficient compression and transmission algorithm has been proposed which uses discrete wavelet transform and run length encoding. Compression results of the proposed algorithm have been compared with several algorithms and it is found that it gives the better performance. The proposed algorithm offers very less PRD (%), hence the characteristic features of the original signal are preserved very well by the reconstructed signal. In future, before implementation of compression and transmission algorithm of ECG data, we will use probabilistic approach such as hidden Markov model [31], [32] and Bayesian network [33], [34] for arrhythmia detection and prediction.

REFERENCES

[1] Tanima Dutta, " *Medical Data Compression and Transmission in Wireless Ad Hoc Networks*," IEEE sensor Journal, Vol. 15, No. 2, pp. 778-786, 2015

[2] Chandan Kumar Jha, Maheshkumar H. Kolekar "*Performance Analysis of ECG Data Compression using Wavelet Based Hybrid Transform Method*," International Confrence on Microwave, Optical and Communication Engineering, 2015

[3] J. R. Cox, F. M. Nolle, H. A. Fozzard, and C. G. Oliver, "*AZTEC, a preprocessing program for real time ECG rhythm analysis*," IEEE Trans. Biomed. Eng., vol. BME-15, no. 4, pp. 128–129, 1968

[4] W. C. Mueller, "*Arrhythmia detection program for an ambulatory ECG monitor,*" Biomed. Sci. Instrument., no. 14, pp. 81–85, 1978

[5] M. Ishijima, S. B. Shin, G. H. Hostetter, and J. Sklansky, "*Scan along polygon approximation for data compression of electrocardiograms*," IEEE Trans. Biomed. Eng., vol. BME-30, no. 11, pp. 723–729, 1983.

[6] J. P. Abenstein and W. J. Tompkins, "*A new data reduction algorithm for real time ECG analysis*," IEEE Trans. Biomed. Eng., vol. BME-29, No. 1, pp. 43–48, 1982

[7] D. A. Huffman, "*A method for the construction of minimum redundancy coders,*" Proc. IRE, vol. 40, No. 9, pp. 1098–1101, 1952.

[8] Maheskumar H. Kolekar, S. Sengupta, " *Hidden Markov Model Based Video Indexig with Discrete Cosine Transform as a Likelihood Function*", IEEE INDICON Conference, IIT Kharagpur, India, 2004

[9] A. Bilgin, M. W. Marcellin, and M. I. Altbach, "*Compression of electrocardiogram signals using JPEG2000*," IEEE Trans. Consumer Electron.,vol. 49, no. 4, pp. 833–840, 2003

[10] A. Al-Shrouf, M. Abo-Zahhad, and S. M. Ahmed, "*A novel compression algorithm for electrocardiogram signal based on the linear prediction of the wavelet coefficients*," Digit. Signal Process., vol. 13, pp. 604–622, 2003.

[11] Z. Lu, D. Y. Kim, and W. A. Pearlman, "*Wavelet compression of ECG signals by the set partitioning in hierarchical trees (SPIHT) algorithm*," IEEE Trans. Biomed. Eng., vol. 47, no. 7, pp. 849–856, 2000.

[12] M. L. Hilton, "*Wavelet and wavelet packet compression of electrocardiograms,*" IEEE Trans. Biomed. Eng., vol. 44, no. 5, pp. 394–402,1997.

[13] P. S. Hamilton, "*Adaptive compression of the ambulatory electrocardiogram,*" Biomed. Inst. Technol., vol. 27, no. 1, pp. 56–63, 1993.

[14] A. Cohen, P. M. Poluta, and R. Scott-Millar, "*Compression of ECG signals using vector quantization,*" in Proc. IEEE-90 S. A. Symp. Commun. Signal Process., 1990, pp. 45–54.

[15] Lorena S.C. de Oliveira, Rodgrigo V. Andreao, Mario Sarcinelli-Filho " *Detection of remature Vetricular Beats in ECG records usin Bayesian Networks Involving the P-wave and Fusion of Results*," IEEE EMBS Buenos Aires, Argentina, pp.1131-1134, 2010

[16] Maheshkumar H. Kolekar, "*Bayesian Belief Network Based Broadcast Sports Video Indexing,"* Int. Journal on Multimedia Tools and Applications, Springer publication, pp. 27-54, 2011

[17] W. T. Chen , K.L. Chan " *Classification of Electrocardiogram using Hidden Markov Models,*" IEEE Engineering in Medicine and Biology Society, Vol. 20, No. 1, pp. 143-146, 1998

[18] Mahesh Kumar H. Kolekar, *" Hidden Markov Model Based Highlight Generation of Cricket Video Sequence using Video and Audio Cues," CSI Communications*, vol. 28(7), pp. 25-26, 2005

[19] Maheshkumar H. Kolekar, Deba Prasad Dash "*A Nonlinear Feature based Epileptic Seizure Detection using Least Square Support Vector Machine Classifier*" IEEE TENCON, 2015

[20] MIT-BIH Arrhythmia Database. [Online]. Available: www.physionet.org/physiobank/database/mitdb/.

[21] S.J. Lee, M. Lee " *A Real-Time ECG Data Compression Algorithm for a Digital Holter System*," IEEE Engg. in Medical and Biological Socitey Conference, pp. 4736-4739, 2008

[22] Schen- Chuan Tai, Chai –Chun sun, Wen-Chien Yan " *A 2-D ECG Compression Method Based on Wavelet Transform and Modified SPIHT*," IEEE Transactions on Biomedical Engineering, Vol. 52, No. 6, pp. 999-1008, 2005

[23] Abhishek Kumar, Maheshkumar H. Kolekar, "*Machine Learning Approach for Epileptic Seizure Detection Using Wavelet Analysis of EEG Signals,*" Int. Conf. on Medical Imaging, M-health and Emerging Communication Systems, pp. 412-416, 2014

[24] Maheshkumar H. Kolekar, G. Raja, S. Sengupta, "*An introduction to wavelet based image processing and its applications*", book chapter in the book on Recent Advances in Computer Vision and Image Processing: Methodologies and Applications, IGI Global publishers, USA, 2013.

[25] M. Sharafat Hossain, Tareq Aziz, M,A. Haque *" ECG Compression using Multilevel Thresholding of Wavelet Coefficients*," IEEE Intelligent in Sensor, Sensor Network and Information Processing, pp.321-326, 2008

[26] S. Lee, J. Kim, and J.-H. Lee, " *A Real-Time ECG Data Compression and Transmission Algorithm for an e-Health Device*," IEEE Transactions on Biomedical Engineering, vol. 58, no. 9, pp. 2448–2455, 2011

[27] Hsiao-Hsuan Chou, Ying-Jui Chen, Yu-Chien Shiau, Te-Son Kuo *" An Effective and Efficient Compression Algorithm for ECG Signals with Irregular Periods,"* IEEE Transactions on Biomedical Engineering, Vol. 53, No. 6, pp. 1198-1205, 2006

[28] Catalina Minica Fira, Liviu Goras " *An ECG Signals Compression Method and Its Validation using NNs*," IEEE Transactions on Biomedical Engnieering, Vol. 55, No. 4, pp. 1319-1326, 2008

[29] A. Sandryhaila, Samir Saba, Markus Puschel, Jelena Kovacevic " *Efficient Compression of QRS Complexes Using Hermite Expansion*," IEEE Transactions on Signal Processing, Vol. 60, No. 2, 2012

[30] Ranjeet Kumar, Anil Kumar "*Dual Tree DWT Analysis based Electrocardiogram Signal Compression using Zero Coding Technique*," ICECS 2015, pp. 624-629, 2015

[31] Maheshkumar H. Kolekar, S. Sengupta, "*Semantic Indexing of News Video Sequences: A Multimodal Hierarchical Approach Based on Hidden Markov Model*", IEEE Int. Region 10 Conference (TENCON), Melbourne, Australia, Nov 2005

[32] Maheskumar H. Kolekar, S. Sengupta, "*Hidden Markov Model Based Structuring of Cricket Video Sequences using Motion and Color Features,*" Indian Conf. on Vision, Graphics and Image Processing (ICVGIP), 2004

[33] Maheshkumar H Kolekar, S Sengupta, "*Bayesian Network-Based Customized Highlight Generation for Broadcast Soccer Videos*", IEEE Transaction on Broadcasting, Vol. 61, No. 2, pp 195-209, 2015

[34] Maheshkumar H. Kolekar, K. Palaniappan, S. Sengupta and G. Seetharaman, "*Video Event Detection and Semantic Identification Using Bayesian Belief Networks*", workshop of IEEE Int. Conf. on Computer Vision (ICCV), Kyoto, Japan, 2009