Performance Comparison Of 2-D DCT On Full/Block Spectrogram And 1-D DCT On Row Mean Of Spectrogram For Speaker Identification

Dr. H. B. Kekre hbkekre@yahoo.com
Senior Professor,
MPSTME, SVKM's NMIMS University
Mumbai, 400-056, India.

Tanuja K. Sarode tanuja_0123@yahoo.com
Ph.D. Scholar, MPSTME, SVKM's NMIMS University
Assistant Professor, TSEC, Bandra (W),
Mumbai, 400-050, India.

Shachi J. Natu shachi_natu@yahoo.com
Lecturer, TSEC, Bandra (W),
Mumbai, 400-050, India.

Prachi J. Natu prachi.natu@yahoo.com
Lecturer, TSEC, Bandra (W),
Mumbai, 400-050, India.

Abstract

The goal of this paper is to present a very simple approach to text dependent speaker identification using a combination of spectrograms and well known Discrete Cosine Transform (DCT). This approach is based on use of DCT to find similarities between spectrograms obtained from speech samples. The set of spectrograms forms the database for our experiments rather than raw speech samples. Performance of this approach is compared for different number of coefficients of DCT when DCT is applied on entire spectrogram, when DCT is applied to spectrogram divided into blocks and when DCT is applied to the Row Mean of a spectrogram. Performance comparison shows that, number of mathematical computations required for DCT on Row Mean of spectrogram method is drastically less as compared to other two methods with almost equal identification rate.

Keywords: Speaker identification, Speaker Recognition, Spectrograms, DCT, Row Mean

1. INTRODUCTION

With an extensive use of internet technology and a switch over from single user applications to multi-user applications, security has become a major issue. To provide security, it has become crucial to identify users and to grant access only to those users who are authorized. Problem of identifying users can be handled using various approaches either separately or in combination with each other. More and more sophisticated techniques are used with the increase in need of security. Uses of login and password, retinal blood vessel patterns, face recognition, fingerprint recognition are some of the widely used techniques. Login and password technique is not secure enough. This is because attackers can easily steal the password using sophisticated electronic eavesdropping techniques [1]. Techniques like face recognition, fingerprint recognition and retinal blood vessel patterns also have their own drawbacks. To identify an individual by these methods, he/she should be willing to undergo the tests and should not get upset by these procedures. Speaker identification allows non-intrusive monitoring and also achieves high accuracy rates which conform to most security requirements. Speaker recognition is the process of automatically recognizing who is speaking based on some unique characteristics present in speaker's voice [2]. For this recognition purpose, preserving the speaker specific characteristics present in the speech signal is important. Speaker recognition can be classified into two main categories, namely speaker identification and speaker verification. Speaker identification deals with distinguishing a speaker from a group of speakers. In contrast, speaker verification aims to determine if a person is the one who he/she claims to be from a speech sample. Speaker identification problem basically consists of two stages: feature extraction stage and pattern classification stage. For the given test utterance, classifier finds out which speaker has pronounced this utterance. To perform this job, models are constructed for each speaker using training data. Speaker specific information from the test utterance is then compared with these models to generate similarity measure so that test utterances can be related to each speaker. These classifiers are of various types and can be grouped into template based and stochastic based classifiers [3]. Template based classifiers are the simplest one. Examples of template based classifiers are: Dynamic Time Warping and Vector Quantization. Stochastic models provide better flexibility and more meaningful results in the form of probabilistic scores [4]. Gaussian Mixture Model, Hidden Markov Model, Neural Networks are the examples of stochastic models.

Speaker identification can be further categorized into text-dependent and text independent speaker identification based on the relevance to speech contents [2]. The text dependent speaker identification can be either a 'closed set' or an 'open set' speaker identification [2]. In closed set problem, from N known speakers, the speaker whose reference template has the maximum degree of similarity with the template of input speech sample of unknown speaker is obtained. This unknown speaker is assumed to be one of the given set of speakers. In the open set text dependent speaker identification, matching reference template for an unknown speaker's speech sample may not exist. In this paper, closed set text dependent speaker identification is considered. In the proposed method, speaker identification is carried out with spectrograms and DCT [15-18]. Thus an attempt is made to formulate a digital signal processing problem into pattern recognition of images.

The rest of the paper is organized as follows: in section 2 we present related work carried out in the field of speaker identification. In section 3 we discuss spectrograms. In section 4 we present our proposed approach. Section 5 elaborates the experiment conducted. Results are tabulated in section 6. Conclusion has been outlined in section 7.

2. RELATED WORK

Many approaches are available in literature for speaker identification process based on various approaches for feature extraction. Feature extraction is the process of extracting subset of features from the entire feature set. The basic idea behind the feature extraction is that the entire feature set is not always necessary for the identification process.

The Mel Frequency Cepstrum Coefficients (MFCC) is one of the popular techniques of feature extraction. The MFCC parameter as proposed by Davis and Mermelstein [5] describes the energy distribution of speech signal in a frequency field. Wang Yutai et. al. [6] has proposed a speaker recognition system based on dynamic MFCC parameters. This technique combines the speaker information obtained by MFCC with the pitch to dynamically construct a set of the Mel-filters. These Mel-filters are further used to extract the dynamic MFCC parameters which represent characteristics of speaker's identity.

Sleit, Serhan and Nemir [7] have proposed a histogram based speaker identification technique which uses a reduced set of features generated using MFCC method. For these features, histograms are created using predefined interval length. These histograms are generated first for all data in feature set for every speaker. In second approach, histograms are generated for each feature column in feature set of each speaker.

Another widely used method for feature extraction is use of linear Prediction Coefficients (LPC). LPCs capture the information about short time spectral envelope of speech. LPCs represent important speech characteristics such as formant speech frequency and bandwidth [8].

Vector Quantization (VQ) is yet another approach of feature extraction [19-22, 25]. In Vector Quantization based speaker recognition systems; each speaker is characterized with several prototypes known as code vectors [9]. Speaker recognition based on non-parametric vector quantization was proposed by Pati and Prasanna [10]. Speech is produced due to excitation of vocal tract. Therefore in this approach, excitation information can be captured using LP analysis of speech signal and is called as LP residual. This LP residual is further subjected to non-parametric Vector Quantization to generate codebooks of sufficiently large size. Combining nonparametric Vector Quantization on excitation information with vocal tract information obtained by MFCC was also introduced by them.

3. SPECTROGRAMS [11]

A spectrogram is an image that shows how the spectral density of a signal varies with time. Spectral density describes how the energy of a signal is distributed with frequency. If f(t) is a finite energy signal, its spectral density is the square of the magnitude of continuous Fourier transform of the signal. The most common format of showing a Spectrogram is a graph with two geometric dimensions. The horizontal axis represents time, whereas the vertical axis represents frequency. A third dimension indicating amplitude of a particular frequency is represented by the intensity or color of each point in the image.

Spectrograms can be created in one of the two ways: using a series of bandpass filters or by calculating Short Time Fourier Transform (STFT) for the signal. The first approach usually uses analog processing, while the second one is a digital process. In the approach using STFT, digitally sampled data are divided into chunks of specific size say 128, 256 etc. which usually overlap. Fourier transform is then obtained to calculate the magnitude of the frequency spectrum for each chunk. Each chunk then corresponds to a vertical line in the image, which is a measurement of magnitude versus frequency for a specific moment in time.

4. PROPOSED APPROACH

In the proposed approach, first we converted the speech samples collected from various speakers into spectrograms. This was done using the second approach of creating spectrogram as mentioned in section 3. Thus we converted the speech sample database into image database. These spectrogram images are then resized to 256 x 256 sizes. The Discrete Cosine Transform [12, 23, 24] is then applied to these images in three different ways to obtain their feature vectors. In the first one, DCT is applied to entire image. Out of total database, 60% of images were used as trainee images and 40% images were used for testing purpose. Euclidean distance between test image and trainee image is used as a measure of similarity. Euclidean distance between the points X(X_1, X_2, etc.) and point Y(Y_1, Y_2, etc.) is calculated using the formula shown in equation (1).

$$D = \sqrt{\sum_{i=1}^{n}(X_i - Y_i)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots..(1)$$

Smaller is the Euclidean distance between test image and trainee image, more accurate speaker identification is achieved. Fig. 1 shows the flowchart for the first method using DCT.

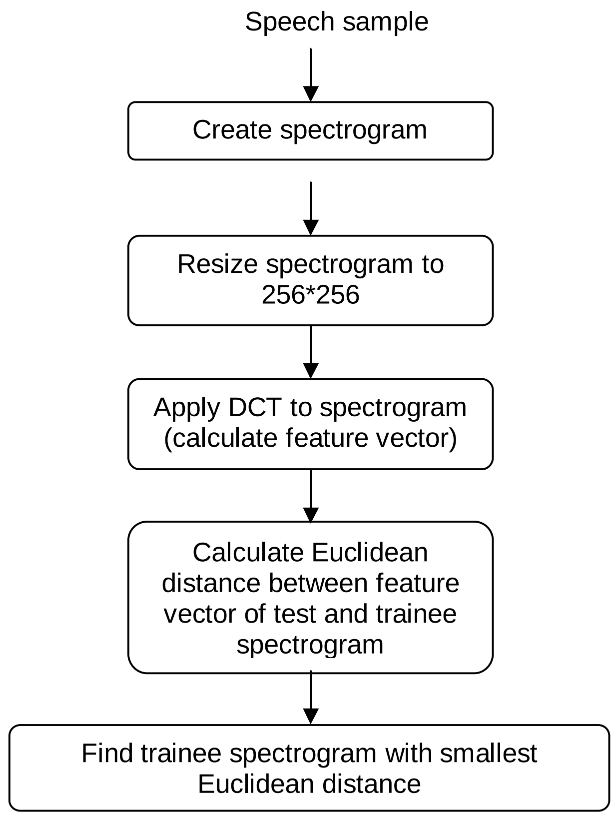

Fig.1: Flowchart for the proposed approach 1

In the second method, resized image is divided into four equal parts as shown in Fig.2 and then DCT is applied to each part. DCT for each block when appended as columns forms a feature vector for an image. Again Euclidean distance is used as a measure of similarity. Fig. 3 shows the flowchart for second method.

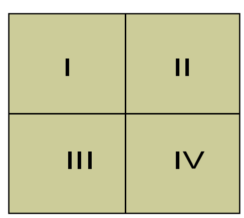

Fig.2: Image divided into four equal nonoverlapping parts

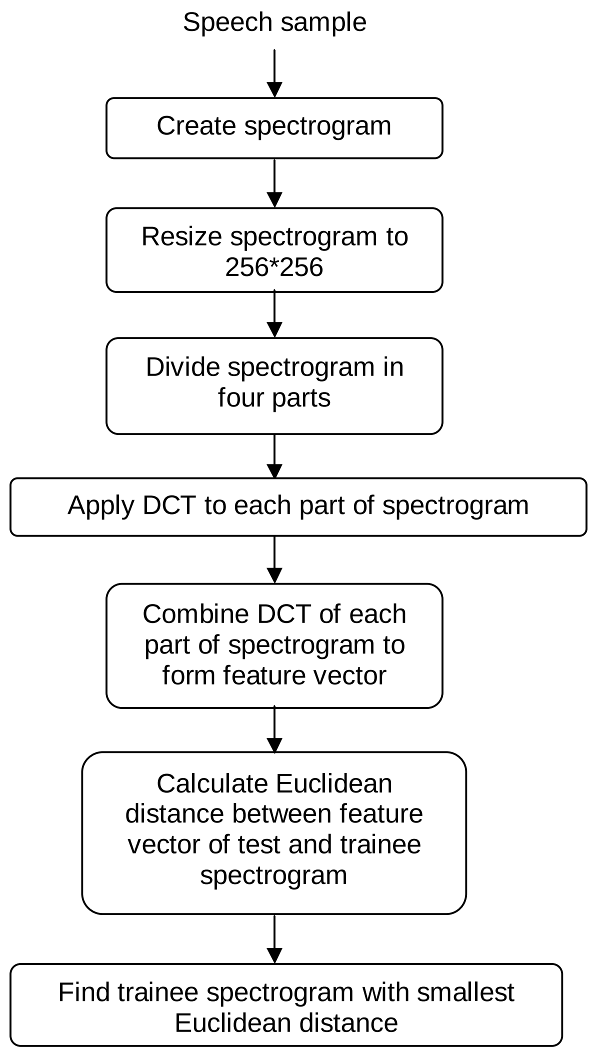

Fig.3: Flowchart for the proposed approach 2

In the third method, Row Mean of an image is calculated [26]. Row mean is nothing but an average of pixel values of an image along each row. Fig. 4 shows how the Row Mean of an image is obtained. DCT is then calculated for this Row mean of an image and Euclidean distance is used to identify speaker.

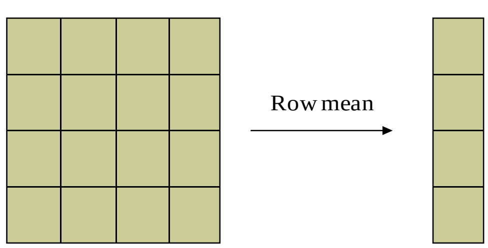

Fig.4: Row Mean of an image

5. EXPERIMENTS

To study the proposed approach we recorded six distinct sentences from 30 speakers: 11 males and 19 females. These sentences are taken from VidTIMIT database [13] and ELSDSR database [14]. For every speaker 10 occurrences of each sentence were recorded. Recording was done at varying times. This forms the closed set for our experiment. From these speech samples spectrograms were created. Before creation of spectrograms, DC offset present in speech samples was removed so that signals are vertically centered at 0. After removal of DC offset, speech samples were normalized with respect to amplitude to -3 dB and also with respect to time. Spectrograms generated from these speech samples form the image database for our experiment. In all we had 1800 spectrograms in our database.

For every speaker 6 spectrograms were used as trainee images and 4 spectrograms were used as test images per sentence, i.e. we had 1080 spectrograms for training purpose and 720 spectrograms for testing purpose. DCT was then applied to the trainee images and result was stored as feature vectors for trainee images.

Similarly, feature vectors for test images were obtained by applying DCT to test images. Euclidean distance between the test image and trainee images was calculated to determine the most probable match i.e. to identify speaker.

Being a text dependent approach, Euclidean distance for a test image of speaker say 'x' for a particular sentence say 's1' is obtained by comparing the feature vector of that test image with the feature vectors of all the trainee images corresponding to sentence 's1'. Results are calculated for set of test images corresponding to each sentence.

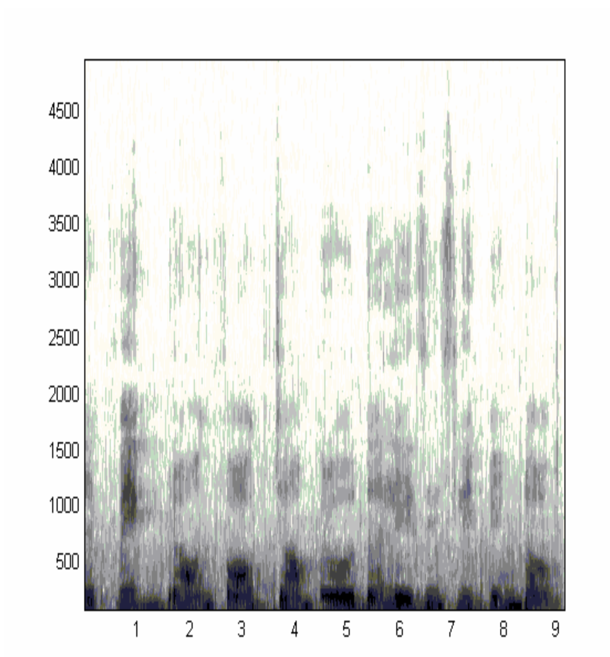

Fig.5: Spectrogram of sentence s1 for speaker 1

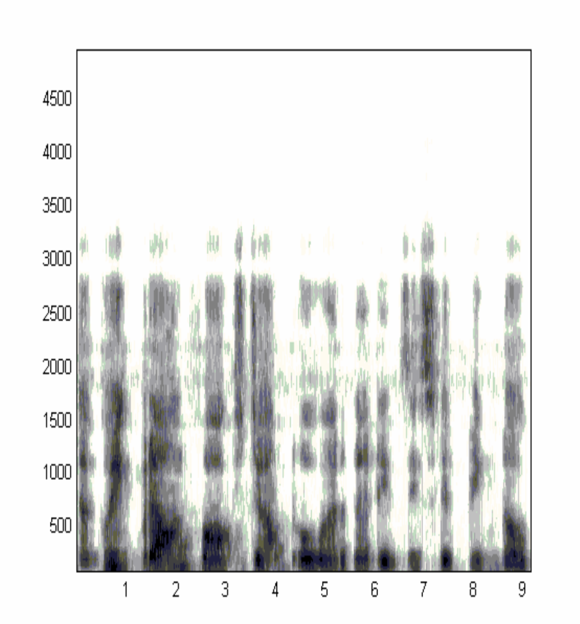

Fig.6: Spectrogram of sentence s1 for speaker 5

Fig.5 and Fig.6 show that the spectrogram for the same sentence, uttered by different speakers is different. The three approaches/methods that were carried out are described in the following subsections.

5.1. Method 1: DCT On Entire Image:

i) As shown in Fig.1, feature vector is obtained by applying DCT on full image.
ii) Euclidean distance between feature vector of test image and trainee image is calculated
iii) Trainee Image with the smallest Euclidean distance is declared as identified speaker.
iv) Steps ii) and iii) are repeated for selected portion of feature vector.

This selection of feature vector is illustrated in following Fig.7. and is based on the number of rows and columns that we selected from the feature vector of an image. For example, we had selected full feature vector (i.e. 256*256), then portion of size 192*192, 128*128, 64*64, 32*32, 25*25, 20*20, 18*18 and 16*16 was selected from the feature vector. For these different sizes, identification rate was obtained.

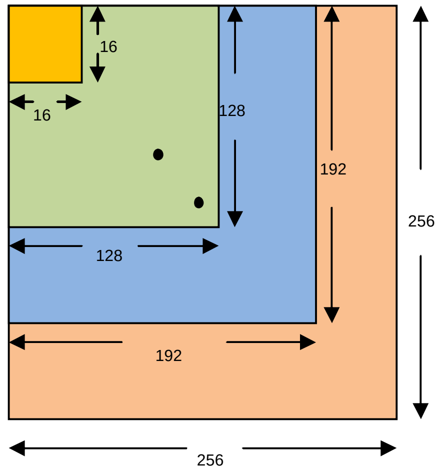

Fig.7: Selection of varying size portion from feature vector

5.2. Method 2: DCT on image block:

i) As shown in Fig.3, feature vector is obtained by taking DCT of image blocks. These image blocks are obtained by dividing image into four parts as shown in Fig.2.
ii) Euclidean distance between feature vector of test image and trainee image is calculated
iii) Trainee Image with the smallest Euclidean distance is declared as identified speaker
iv) Steps ii) and iii) are repeated for selected portion of feature vector.

Selection of feature vector is similar to the one shown in Fig.7. But in this method, size of feature vector is 128*512, 96*384, 64*256, 32*128, 16*64 and 8*32.

5.3. Method 3: DCT on Row Mean of an image:

a) Row mean of full image -

i) Row Mean of an image is obtained.
ii) DCT is applied to this Row Mean to obtain the feature vector.
iii) Euclidean distance between feature vectors of test image and trainee image is calculated.
iv) Trainee image with the smallest Euclidean distance is declared as identified speaker.

b) Row Mean of image blocks -

i) Image is divided into blocks of size 128*128.
ii) Calculate Row Mean of each block.
iii) Apply DCT on Row Mean of each block to form the feature vector of image.
iv) Euclidean distance between feature vector of test image and trainee image is calculated.
v) Trainee image with the smallest Euclidean distance is declared as identified speaker.

Steps ii) to v) in Row Mean of image blocks are repeated for the block size 64, 32, 16 and 8.

6. RESULTS AND COMPLEXITY ANALYSIS

6.1. Results

Following tables show the identification rate obtained for different number of coefficients. These different numbers of coefficients are based on selection of varying sized feature vector portion as shown in Fig. 7.

Table 1 shows the identification rate for sentences s1 to s6 when different numbers of DCT coefficients are taken to find the matching spectrogram i.e. to identify speaker using DCT on full image. Portion from feature vector selected for these coefficients is 256*256, 192*192, 128*128, 64*64, 32*32, 25*25, 20*20, 18*18 and 16*16 respectively.

Portion of feature vector selected	Sentence					
	S1	S2	S3	S4	S5	S6
256*256	54.16	59.16	56.66	56.66	68.33	62.50
192*192	58.33	65	67.5	65	73.33	69.16
128*128	65.83	64.16	71.66	67.5	74.16	72.5
64*64	70.83	70.83	71.66	72.50	77.50	75.83

32*32	75	73.33	74.16	75	80	77.5
25*25	75.83	75	75.83	73.33	80	81.66
20*20	78.33	75.33	78.33	71.66	81.66	80
18*18	75.83	75	77.5	75	80.83	78.33
16*16	72.5	76.66	74.16	74.16	76.66	79.16

Table 1: Identification rate for sentences s1 to s6 for varying portion of feature vector when DCT is applied to full image

Table 2 shows the overall identification rate considering all sentences, for various percentages of DCT coefficients i.e. for portions of different sizes from the feature vector in first approach.

Portion of feature vector selected	Number of DCT coefficients	Identification rate (%)
256*256	65536	60
192*192	36864	66.38
128*128	16384	69.30
64*64	4096	73.19
32*32	1024	75.83
25*25	625	76.94
20*20	400	**77.63**
18*18	324	77.08
16*16	256	76.66

Table 2: Overall Identification rate for varying number of DCT coefficients when DCT is applied to full image

Similarly Table 3 shows the identification rate for sentences s1 to s6 when different numbers of DCT coefficients are taken to identify speaker using DCT on image blocks, whereas, Table 4 shows the overall identification rate considering all sentences, for various number of DCT coefficients using the same approach.

Portion of feature vector selected	Sentence					
	S1	S2	S3	S4	S5	S6
128*512	54.16	59.16	57.5	57.5	68.33	63.33
96*384	60	63.33	65.33	65	73.33	68.33
64*256	65	65	70.83	66.66	74.16	71.16
32*128	70.83	70.83	70.83	71.66	76.66	75
16*64	75.83	74.16	75	75.83	81.66	77.5
8*32	69.16	76.66	75	75.83	75	75.83

Table 3: Identification rate for sentences s1 to s6 for varying portion of feature vector using DCT on image blocks

Portion of feature vector selected	Number of DCT coefficients	Identification rate (%)
128*512	65536	60
96*384	36864	65.97
64*256	16384	68.88

32*128	4096	72.63
16*64	1024	**76.66**
8*32	256	74.58

Table 4: Identification rate for varying size of feature vector portion using DCT on image blocks

Table 5 and Table 6 show the sentence wise identification rate and overall identification rate when DCT of Row Mean is taken by dividing an image into different number of non-overlapping blocks.

No. of blocks for image split	**Sentence**					
	S1	**S2**	**S3**	**S4**	**S5**	**S6**
Full image (256*256)	57.5	66.66	64.16	60.83	60.83	62.5
4 Blocks (128*128)	60.83	70.83	63.33	65.83	70	65.83
16 Blocks (64*64)	69.16	75.83	70.83	65.83	73.33	71.66
64 Blocks (32*32)	75	76.66	75.83	70	78.83	75.83
256 Blocks (16*16)	76.66	75	75.83	72.5	80	82.5
1024 Blocks (8*8)	74.16	72.5	75	72.5	80.83	78.33

Table 5: Identification rate for sentences s1 to s6 for DCT on Row mean of an image when image is divided into different number of nonoverlapping blocks

No. of blocks for image split	**Number of DCT coefficients**	**Identification rate (%)**
Full image (256*256)	256	62.08
4 Blocks (128*128)	512	66.11
16 Blocks (64*64)	1024	71.11
64 Blocks (32*32)	2048	75.27
256 Blocks (16*16)	4096	**77.08**
1024 Blocks (8*8)	8192	75.55

Table 6: Overall Identification rate for DCT on Row mean of an image when image is divided into different number of nonoverlapping blocks

6.2. Complexity Analysis

For 2-D DCT on N*N image, $2N^3$ multiplications are required and $2N^2(N-1)$ additions are required. For 2-D DCT on four blocks of size N/2*N/2, N^3 multiplications are required and $N^2(N-2)$ additions are required. For 1-D DCT on N*1 image, N^2 multiplications are needed and N(N-1) additions are needed. Further for the calculation of Euclidean distance between the feature vectors of size

M*N, number of multiplications required are M*N and number of additions required are 2MN-1. These computational details are summarized in Table 7.

	No. of Multiplications	No. of Additions
2-D DCT on N*N image	$2N^3$	$2N^2(N-1)$
2-D DCT on four blocks of size N/2*N/2 each	N^3	$N^2(N-2)$
1-D DCT on N*1 image	N^2	$N(N-1)$

Table 7: Computational details for 2-D DCT on N*N image, 2-D DCT on N/2*N/2 image and 1-DCT on N*1 image respectively

Considering the above facts, we compare the number of DCT coefficients used and number of computations in terms of multiplications and additions including DCT calculation and Euclidean distance calculation, for the highest identification rate obtained using our three methods. The comparisons are given in Table 8.

Parameter	DCT on Full image	DCT on image blocks	DCT on Row Mean of image
Number of DCT coefficients used	400	1024	4096
Number of multiplications required	33554832	16778240	69632
Number of additions required	33424159	16648191	69631
Identification Rate	77.63	76.66	77.08

Table 8: Number of DCT coefficients used, number of multiplications and number of additions for DCT on full image, DCT on image blocks and DCT on Row Mean of 256*256 image

7. CONCLUSION

In this paper we considered closed set text dependent speaker identification rate using three different ways of applying DCT on spectrograms. For each method, Identification rates obtained for various numbers of DCT coefficients are compared. It has been observed that as the number of DCT coefficients chosen is smaller up to a certain limit; better identification rate is achieved in all three methods. Further it has been observed that DCT on full image gives its maximum identification rate of 77.63% for only 20*20 portion of feature vector i.e. by using only 400 DCT coefficients. DCT on image blocks gives maximum identification rate of 76.66% when 16*64 portion of its feature vector is considered which has 1024 DCT coefficients. Finally DCT on Row Mean gives maximum identification rate of 77.08% for Row Mean of 8*8 size image blocks i.e. for 4096 DCT coefficients.

Further when these maximum identification rates in all three methods are compared, it has been observed that though number of coefficients used in Row Mean method is higher, number of multiplications and additions reduce drastically as compared to other two methods. Number of multiplications in DCT on full image method is 482 times more than the number of multiplications in Row Mean method whereas for DCT on image blocks it is 241 times more. Number of additions needed in DCT on full image and DCT on image blocks are also 480 times and 239 times more than the additions required in Row mean method respectively. Identification rate obtained by Row Mean method is very much closer to the one obtained by applying DCT on full spectrogram and with considerably lesser number of mathematical computations.

8. REFERENCES

[1] Evgeniy Gabrilovich, Alberto D. Berstin: “Speaker recognition: using a vector quantization approach for robust text-independent speaker identification”, Technical report DSPG-95-9-001’, September 1995.

[2] Tridibesh Dutta, “Text dependent speaker identification based on spectrograms”, Proceedings of Image and vision computing, pp. 238-243, New Zealand 2007,.

[3] J.P.Campbell, “Speaker recognition: a tutorial”, Proc. IEEE, vol. 85, no. 9, pp. 1437-1462, 1997.

[4] D. O’Shaughnessy, “Speech communications- Man and Machine”, New York, IEEE Press, 2nd Ed., pp. 199, pp. 437-458, 2000.

[5] S. Davis and P. Mermelstein, “Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences,” IEEE Transaction Acoustics Speech and Signal Processing, vol. 4, pp. 375-366, 1980.

[6] Wang Yutai, Li Bo, Jiang Xiaoqing, Liu Feng, Wang Lihao, “Speaker Recognition Based on Dynamic MFCC Parameters”, International Conference on Image Analysis and Signal Processing, pp. 406-409, 2009

[7] Azzam Sleit, Sami Serhan, and Loai Nemir, “A histogram based speaker identification technique”, International Conference on ICADIWT, pp. 384-388, May 2008.

[8] B. S. Atal, “Automatic Recognition of speakers from their voices”, Proc. IEEE, vol. 64, pp. 460-475, 1976.

[9] Jialong He, Li Liu, and G¨unther Palm, “A discriminative training algorithm for VQ-based speaker Identification”, IEEE Transactions on speech and audio processing, vol. 7, No. 3, pp. 353-356, May 1999.

[10] Debadatta Pati, S. R. Mahadeva Prasanna, “Non-Parametric Vector Quantization of Excitation Source Information for Speaker Recognition”, IEEE Region 10 Conference, pp. 1-4, Nov. 2008.

[11] Tridibesh Dutta and Gopal K. Basak, “Text dependent speaker identification using similar patterns in spectrograms”, PRIP'2007 Proceedings, Volume 1, pp. 87-92, Minsk, 2007.

[12] Andrew B. Watson, “Image compression using the Discrete Cosine Transform”, Mathematica journal, 4(1), pp. 81-88, 1994,.

[13] http://www.itee.uq.edu.au/~conrad/vidtimit/

[14] http://www2.imm.dtu.dk/~lf/elsdsr/

[15] H.B.Kekre, Sudeep D. Thepade, “Improving the Performance of Image Retrieval using Partial Coefficients of Transformed Image”, International Journal of Information Retrieval (IJIR), Serials Publications, Volume 2, Issue 1, pp. 72-79 (ISSN: 0974-6285), 2009.

[16] H.B.Kekre, Tanuja Sarode, Sudeep D. Thepade, “DCT Applied to Row Mean and Column Vectors in Fingerprint Identification”, In Proceedings of International Conference on Computer Networks and Security (ICCNS), 27-28 Sept. 2008, VIT, Pune.

[17] H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shah, Prathmesh Verlekar, Suraj Shirke,“Energy Compaction and Image Splitting for Image Retrieval using Kekre Transform over Row and Column Feature Vectors”, International Journal of Computer Science and Network Security (IJCSNS),Volume:10, Number 1, January 2010, (ISSN: 1738-7906) Available at www.IJCSNS.org.

[18] H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shah, Prathmesh Verlekar, Suraj Shirke, “Performance Evaluation of Image Retrieval using Energy Compaction and Image Tiling over DCT Row Mean and DCT Column Mean”, Springer-International Conference on Contours of Computing Technology (Thinkquest-2010), Babasaheb Gawde Institute of Technology, Mumbai, 13-14 March 2010, The paper will be uploaded on online Springerlink.

[19] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, Vaishali Suryavanshi, “Improved Texture Feature Based Image Retrieval using Kekre’s Fast Codebook Generation Algorithm”, Springer-International Conference on Contours of Computing Technology (Thinkquest-2010), Babasaheb Gawde Institute of Technology, Mumbai, 13-14 March 2010, The paper will be uploaded on online Springerlink.

[20] H. B. Kekre, Tanuja K. Sarode, Sudeep D. Thepade, “Image Retrieval by Kekre’s Transform Applied on Each Row of Walsh Transformed VQ Codebook”, (Invited), ACM-International Conference and Workshop on Emerging Trends in Technology (ICWET 2010),Thakur College of Engg. And Tech., Mumbai, 26-27 Feb 2010, The paper is invited at ICWET 2010. Also will be uploaded on online ACM Portal.

[21] H. B. Kekre, Tanuja Sarode, Sudeep D. Thepade, “Color-Texture Feature based Image Retrieval using DCT applied on Kekre’s Median Codebook”, International Journal on Imaging (IJI), Volume 2, Number A09, Autumn 2009,pp. 55-65. Available online at www.ceser.res.in/iji.html (ISSN: 0974-0627).

[22] H. B. Kekre, Ms. Tanuja K. Sarode, Sudeep D. Thepade, "Image Retrieval using Color-Texture Features from DCT on VQ Codevectors obtained by Kekre’s Fast Codebook Generation", ICGST-International Journal on Graphics, Vision and Image Processing (GVIP), Volume 9, Issue 5, pp.: 1-8, September 2009. Available online at http://www.icgst.com/gvip/Volume9/Issue5/P1150921752.html.

[23] H. B. Kekre, Sudeep Thepade, Akshay Maloo, “Image Retrieval using Fractional Coefficients of Transformed Image using DCT and Walsh Transform”, International Journal of Engineering Science and Technology, Vol.. 2, No. 4, 2010, 362-371

[24] H. B. Kekre, Sudeep Thepade, Akshay Maloo,”Performance Comparison of Image Retrieval Using Fractional Coefficients of Transformed Image Using DCT, Walsh, Haar and Kekre’s Transform”, CSC-International Journal of Image processing (IJIP), Vol.. 4, No.2, pp.:142-155, May 2010.

[25] H. B. Kekre, Tanuja Sarode “Two Level Vector Quantization Method for Codebook Generation using Kekre’s Proportionate Error Algorithm” , CSC-International Journal of Image Processing, Vol.4, Issue 1, pp.1-10, January-February 2010

[26] H. B. Kekre, Sudeep Thepade, Akshay Maloo, “Eigenvectors of Covariance Matrix using Row Mean and Column Mean Sequences for Face Recognition”, CSC-International Journal of Biometrics and Bioinformatics (IJBB), Volume (4): Issue (2), pp. 42-50, May 2010.