INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

BLIND MAN VOTING USING SVM CLASSIFIER

[1]Panimalar S., [2]Fathima Syed Abdul Azeez, [3]Yuvashree G. V, [4]Priyanga A.,

[1]Assistant Professor, Department of CSE, Panimalar Institute of Technology.
[2,3,4] IV year-CSE, Panimalar Institute of Technology.

Abstract: - Voting in any election is important. In projected method, first voice is given as the input (The voice will be the adhar card number) and converted into text in order to compare with the record. If the adhar card number matches, then the face recognition will be taken and compared. If both adhar card number and face are valid, then the already recorded tone of voice will be played. The recorded tone of voice will have the details about the political parties, and according to the recorded tone of voice the input is given. And then the voting will be done accordingly. In this paper, we propose an HMM (Hidden Markov Model), GMM (Gaussian Markov Model) and SOM (self organized mapping) segmentation and analyze the system in CFCC and MFCC with SVM trained samples of speech. Then Voila Jones Algorithm is used for face detection. At last, once more the SVM classifier is used to recognize the face. Efficient outcome is obtained when compared to existing HMM model for speech.

Index Terms: - Hidden Markov Model, Gaussian Markov Model, Self Organized Mapping, Voila Jones Algorithm, SVM Classifier, CFCC and MFCC.

I. INTRODUCTION

Voting in any election is important as it provide blind and visually impaired citizens with the opportunity to voice their opinion regarding elected leaders and policies and to aid form the future by electing candidates who share their views.

Elected official create decision and overtake legislation that found the maximum income that a blind person can earn while still getting social security disability profit, the smallest amount wage rate paid to blind workers in sheltered workshops, and the health profit blind citizens accept below Medicare and Medic-aid.

To ensure that accessible voting machines task correctly on Election Day, Counties should guide sample workers on appropriate setup and use of the machines. The available features of these machines must be weathered to ensure their effectiveness prior to opening the machines to the municipal on Election Day. Sample workers should also have reliable access to technical support services to handle concerns as they arise on Election Day.

II. RELATED WORK

Veronique Stouten et al (2008) presented a method to habitually discover the (word-sized) phone patterns that exists in speech sounds. These patterns are educated from a set of phone lattices produced from the utterances. Just like kids obtaining language, their system does not have previous data on what the eloquent patterns are. By applying the non-negative matrix factorization algorithm to a fixed measurement lengthwise tall-dimensional vector representation of the speech utterances, decomposition in terms of additive units is attained. They illustrate that these units parallel to words in case of a small vocabulary task. Their result also raise question taking place whether open segmentation and clustering are desirable in an unsupervised learning context.

Okko Rasanen et al (2015) says that segmenting speech into syllable-like units, the system is capable to border potential word onsets and offsets to a limited number of candidate localities. These syllable tokens are then labeled using a set of features and clustered into a finite number of syllable classes. Finally, frequent syllable sequences or individual classes are preserved as word candidates. Feasibility of the approach is examined on impulsive American English and Tsonga language samples with likely results. But pattern matching stages is reliant on the reliability of the initial segment.

Alex S. Park et al (2008) says that pattern discovery can be used to inevitably attain lexical entities directly from an untranscribed audio stream .Their approach is to unsupervised word acquisition utilizes a segmental variant of a widely used dynamic programming method , which permits them to find equivalent acoustic patterns across audio streams .They establish how to group similar acoustic sequences collected to form clusters equivalent to lexical entities such as words and little multiword phrases.

C. Paramanand et al (2013) report the problem of approximating the hidden image of an unmoving bilayer sight (consisting of a foreground and a background at diverse depths) from motion blurred observations taken with a handheld camera. The camera motion is measured to be composed of in-plane rotation and translation. Since the blur at an image spot be contingent both on camera motion and depth, deblurring develops a difficult mission. They originally proposed a method to approximation the transformation spread function (TSF) conforming to one of the depth layers .The assessed TSF (which reveals the camera motion during exposure) is used to section the scene into foreground and background layers and regulate the relative depth value. The deblurred copy of the scene is finally assessed within a regularization framework by secretarial for blur variations owing to camera motion in addition to depth.

Georgios Tzimiropoulo et al (2012) present the notion of subspace knowledge from image gradient orientations for arrival-based object recognition. As image data is characteristically loud and noise is considerably dissimilar from Gaussian, traditional subspace learning from pixel intensities flops very frequently to estimation reliably the low-dimensional subspace of a given statistics people. They show that substituting pixel intensities with Gradient Orientations and the _2 norm with a cosine-based distance measure suggestions, to some extend a therapy to this difficult .Inside this framework, which they transformation IGO (Image Gradient Orientations) subspace learning, they first express and study the belongings of Principal Component Analysis of Image Gradient Orientations (IGO-PCA).Then they show its linking to before planned robust PCA method both hypothetically and experimentally .Finally they originate a amount of additional prevalent subspace learning techniques, specifically Linear Discriminant Analysis (LDA), Locally Linear Embedding (LLE) and Laplacian Eigen maps (LE). New consequences display that their algorithms outdo meaningfully general approaches such as Gabor features and attain state-of-the-art presentation for tough difficulties such as illumination- and occlusion- vigorous face recognition. Furthermore to this, the proposed IGO-methods need the Eigen-decomposition of simple covariance matrices and are as computationally well-organized as their consistent _2 norm intensity-based counterparts.

Xiangxin Zhu et al (2012) present a combined model for face detection , pose estimation, and landmark estimation in real-word , cluttered images .This model is founded on a combinations of trees with a collective pool of parts; they perfect each facemask milestone as a share and practice worldwide combinations to imprisonment topological variations due to viewpoint, They show that tree structured models are amazingly actual at capturing global elastic deformation, while being cool to enhance different dense graph structures .They present extensive consequences on standard face benchmarks, as well as a original “in the wild” annotated dataset , that proposes their system developments the state-of-the art , occasionally meaningfully, for all three tasks .Though our model is diffidently trained with hundreds of faces , it associates favorably to profitable systems trained with billions of examples (such like Google Picasa and face.com).

III. PROPOSED METHOD

3.1 SPEECH TO TEXT:

In settings where only unlabelled speech data is presented, speech technology needs to be urbanized without transcription, articulation dictionary, or language modeling text.

The trained speech is converted to text by comparing the speech within the past trained data. If the speech is matched than its word will be displayed as text.

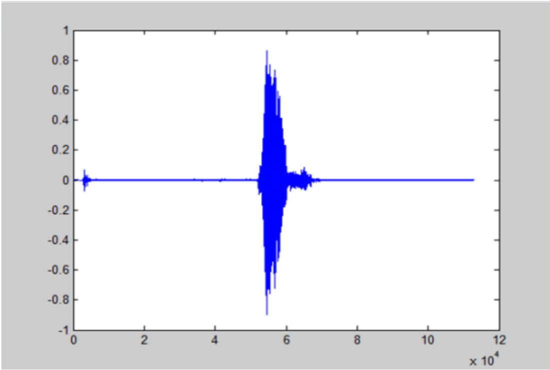


Fig 3.1

3.2 SEGMENTATION:

Speech segmentation is the process of identify the confines among words, syllables, otherwise phonemes in speaks expected languages. The name applies both to the mental processes used by humans, and to fake processes of accepted speech processing.

HMMs (Hidden Markov Model) lie at the heart of nearly all modern speech recognition system and although the necessary skeleton has not altered significantly in the last decade otherwise more, the detailed modeling technique developed within this skeleton include evolve to a state of considerable sophistication.

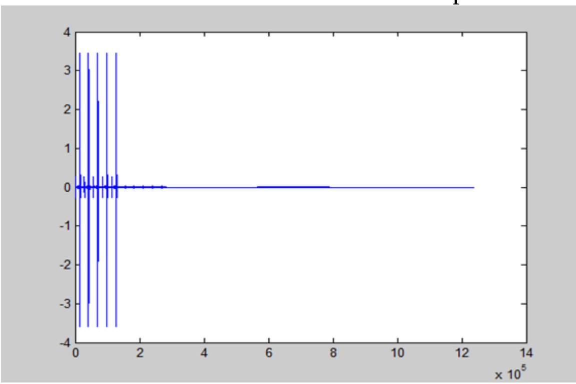


Fig 3.2.1

A range of forms of GMM (Gaussian Mixture Model) feature extraction are outlined, together with method to implement sequential smoothing and a method to incorporate a previous allocation to hamper the exact parameters. These parameters are used to take out the noises.

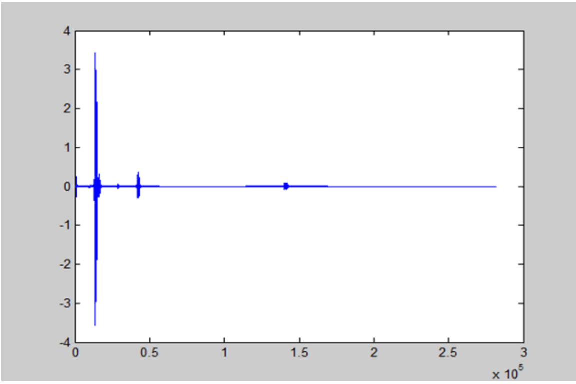


Fig 3.2.2

The feature extractor along with the SOM (Self Organized Mapping) behaved like a transducer; transform a cycle of speech samples into a series of labels.

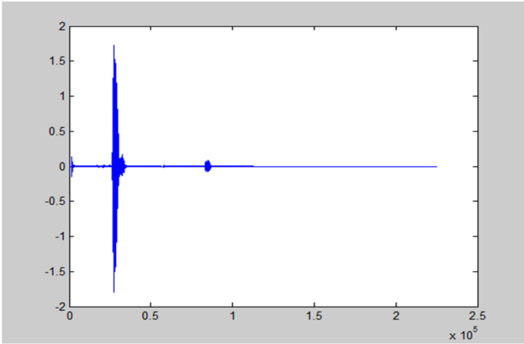


Fig 3.2.3

3.3 IMAGE FILTER:

Images are frequently contaminated by random variations in intensity, illumination, otherwise have unfortunate contrast and can't be used openly.

Filtering: transform pixel strength value to reveal positive image characteristics

Enhancement: improve contrast.

Smoothing: take away noise.

Template matching: detects known patterns.

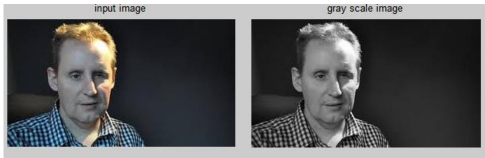
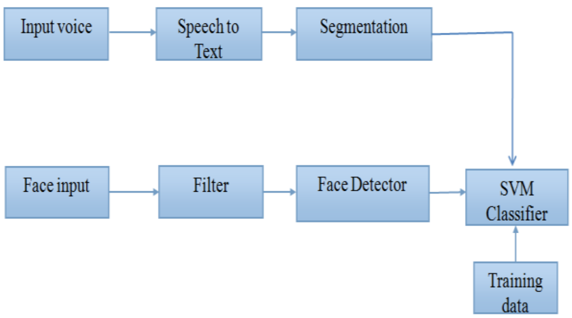


Fig 3.3



Block Diagram

3.4 THE VIOLA- JONES FACE DETECTOR:

Viola-Jones Face Detector has three well-known key aid Integral Image, variant AdaBoost learning algorithm and Cascade structure to complete elevated dispensation speed and detection rates. The detection rates of Viola-Jones face detector are as good as to the finest earlier systems. So it has been used in real-time applications.

Fig 3.4

Vol.4 Issue 3, Pg.: 43-47
March 2016

3.5 SVM CLASSIFIER:

Support Vector Machines (SVM) has lately publicized their ability in example detection and arrangement. Specified a position of points which belong to any of two classes, a linear SVM find the hyperplane exit the prime achievable division of points of the same class on the same side, while maximizing the detachment of any class from the hyperplane. According to this hyperplane minimizes the risk of misclassified examples of the test set.

IV. CONCLUSION

An efficient application has been proposed to perform voting by blind people. A HMM (Hidden Markov Model), GMM (Gaussian Markov Model) and SOM (self organized mapping) segmentation are used to analyze the system in CFCC and MFCC with SVM trained samples of speech. Viola Jones algorithm is used for face detection and SVM classifier is used to recognize the face. Efficient outcome is obtained when compared to existing HMM model for speech.

REFERENCES

[1] F. Metze, X. Anguera, E. Barnard, M. Davel and G. Gravier, ”The spoken web search task at MediaEval 2012,” in Proc.ICASSP,2013.
[2] K. Levin, A. Jansen, and B. Van Durme,”Segmental acoustic indexing for zero resource keyword search,” in Proc.ICASSP, 2015.
[3] A. Jansen and B. Van Durme,”Efficient spoken term Discovery using randomized algorithms,” in Proc.ASRU, 2011.
[4] V. Stouten ,K. DemuYnck, and H. Van hamme, ”Discovering phone patterns in spoken utterances by non-negative matrix factorization,” IEEE Signal Proc.Let,vol.15.pp.131-134,2008.
[5] O. J.Rasanen, G.Doyle, and M. C. Frank, “Unsupervised Word discovery from speech using automatic segmentation into syllable-like units,” in Proc.Interspeech. 2015.
[6] M. Athineos and D. P. W. Ellis, “Frequency-domain linear prediction for temporal features,” in Proc.ASRU, 2003.
[7] G. Zweig and P. Nguyen,”SCARF: a segmental conditional random field toolkit for speech recognition,” in Interspeech, 2010.
[8] A. S. Park and J. R. Glass, “Unsupervised pattern discovery in speech” IEEE Trans. Audio, Speech, Language process. vol.16, no.1, pp.186-197, 2008.
[9] Y. Zhang and J. R. Glass,”Unsupervised spoken keyword spotting via segmental DTW on Gaussian Posteriorgrams,”in Proc.ASRU, 2009.
[10] O. Waller, T. Korthals, R. Haeb-Umbach, and B. Raj, ”A hierarchical system for word discovery exploiting DTW-based initialization,” in Proc. ASRU,2013.
[11] Y. Zhan, R. Salakhutdinov, H.-A. Chang, and J. R. Glass, “Resource configurable spoken query detection using deep Boltzmann machines.” in Proc.ICASSP, 2012.
[12] C. Paramanand and A. N. Rajagopalan,"Non-uniform motion deblurring for bilayer scenes”, Proc. IEEE Conf. Comput. Vis. Pattern Recognit. pp.1115-1122.
[13] T. Ahonen, E. Rahtu, V. Ojansivu and J. Heikkila, "Recognition of blurred faces using local phase quantization”, Proc. 19th Int. Conf. Pattern Recognit. , pp.1-4.
[14] G.Tzimiropoloulos, S. Zafeiriou, and M. Pantic, “Subspace learning from image gradient orientations,” IEEE Trans. Pattern Anal. Mach. Intell., vol.34, no. 12, pp. 2454-2466, Dec. 2012.
[15] X. Zhu and D. Ramanan, Face detection, pose estimation, and landmark localization in the wild,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 2879-2886.