International Journal of Engineering

Journal Homepage: www.ije.ir

Phoneme Classification Using Temporal Tracking of Speech Clusters in Spectro-temporal Domain

N. Esfandian*

Department of Electrical Engineering, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran

P A P E R I N F O

Paper history:
Received 28 August 2019
Received in revised form 22 September 2019
Accepted 08 November 2019

Keywords:
Phoneme Classification
Spectro-temporal Features
Primary Auditory Cortex
Speech Feature Extraction
Weighted Gaussian Mixture Models

A B S T R A C T

This article presents a new feature extraction technique based on the temporal tracking of clusters in spectro-temporal features space. In the proposed method, auditory cortical outputs were clustered. The attributes of speech clusters were extracted as secondary features. However, the shape and position of speech clusters change during the time. The clusters temporally tracked and temporal tracking parameters were considered in secondary features. The new architecture was proposed for phoneme classification by a combining classifier using both tracked and energy-based features. Clustered based spectro-temporal features vectors were used for the classification of several subsets of TIMIT database phonemes. The results show that the phoneme classification rate was improved Using tracked spectro-temporal features. The results were improved to 78.9% on voiced plosives classification which was relatively 3.3% higher than the results of non-tracked spectro-temporal feature vectors. The results on other subsets of phonemes showed good improvement in classification rate too.

***doi**: 10.5829/ije.2020.33.01a.12*

1. INTRODUCTION

The main goal of speech features extraction methods is the extraction of valuable discriminative information in the extracted features while reducing the amount of data to a minimum. Mel scaled frequency cepstral coefficients and spectro-temporal features [1-3] are the most frequently used representations of the speech signal that are both inspired by the human auditory model. The auditory model is inspired based on psycho-acoustical and neurophysiological findings in the human auditory system. In a few years, this auditory model has been successfully employed in various applications of speech processing [4-10]. The first stage of the auditory model is inspired by the internal ear. An auditory spectrogram is obtained in the early stage. Then, in the cortical stage of this model, the spectral and temporal features are extracted using 2-D spectro-temporal receptive field (STRFs) filters which are the scaled versions of a two dimensional Gabor shaped impulse response [2, 3]. The high dimensionality of the cortical output of the auditory model makes the system impractical in this domain and affects the parameter estimation accuracy in the training phase of the phoneme classifier. In this study, the proposed method follows our previous research in which clustering methods had been used to cluster spectro-temporal feature space in order to extract secondary features vectors [11, 12]. Therefore, the phoneme was presented using the attributes of speech clusters in this feature space. One of the open issues in the previous study was the order of speech clusters in the second feature vector. This sorting order determines the consistency in the value of each element of the feature vector and dramatically affects the phoneme classification rate if arranged inappropriately. In the previous study, it was assumed that the energies of speech clusters in each frame are the intrinsic characteristics of the phonemes and can be considered as the measure of clusters sorting in spectro-temporal features vector [13]. Although this assumption is often true in central parts of phonemes, especially in long-duration phonemes, however, this presumption cannot be assumed in the gradual interchange of co-articulated phonemes which are frequently occurred in an uttered speech sentence.

*Corresponding Author Email: *nafis.esfandian@gmail.com* (N. Esfandian)

Please cite this article as: N. Esfandian, Phoneme Classification using Temporal Tracking of Speech Clusters in spectro- temporal domain, International Journal of Engineering (IJE), IJE TRANSACTIONS A: Basics Vol. 33, No. 1, (January 2020) 105-111

Variations in clusters locations are tracked during the time and the order of clusters is registered in the whole sequence based on a reference vector before sending them to the phoneme classifier. The spatiotemporal features of clusters of two frames are matched based on their weighted Euclidean distance. The clusters are sorted using two strategies in secondary feature vectors. In the first strategy, the clusters are sorted based on energy measured in each frame and in the other strategy; speech clusters are re-sorted using temporal tracking results of the secondary feature vectors sequence. Combining mechanisms of two features sorting strategies is applied to improve the phoneme classification rate.

In section 2, two stages of the auditory model and clustering-based spectro-temporal features extraction method are described. Section 3 presents an overview of the proposed feature extraction and clusters tracking algorithms in the spectro-temporal domain. Experimental results and performance evaluation of the proposed features vectors on standard datasets for phoneme classification task are provided in section 4. Section 5 concludes the paper.

2. AUDITORY MODEL

In the early stage of this model, the speech signal is transformed into the auditory spectrogram. In the cortical stage, the auditory spectrogram is analyzed using a bank of 2-D filters to obtain the spectro-temporal features. The output of the cortical stage of this model has a four-dimensional scale (Ω in cycles/octave), rate (ω in Hz), frequency, and time. The auditory cortical stage is modeled by spectro-temporal filter banks. Each filter is tuned to a range of spectral–temporal modulations. At this stage, a two-dimensional wavelet transform of the auditory spectrogram is calculated using the two-dimensional wavelet transform function (such as Gabor function). The dimensions of the spectro-temporal feature space are very large. Therefore, the reduction of features space dimensions is a crucial task to train the parameters of artificial speech classifiers efficiently. Figure 1 shows the rate-scale representation of phoneme /b/. It can be observed, the energy was concentrated in the middle section of the rate axis. These clusters are moved and reshaped along with time. The secondary feature vectors are the clusters parameters.

2. 1. Clustering-Based Features Extraction Method In Spectro-Temporal Domain In the first stage, the auditory spectrogram was calculated. In the next stage, spectro-temporal features were estimated using the spectro-temporal receptive field (STRFs) filters. Scale, rate, and frequency are spatial information of each point in the spectro-temporal domain which should be considered as the primary feature vectors. In

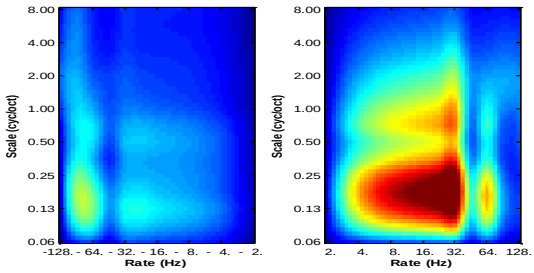

Figure 1. The rate-scale representation of phoneme /b/

the cluster-based feature extraction method, weighted Gaussian mixture model (WGMM) and weighted K-means (WKM) clustering is used to extract the features with informative discriminative attributes. It means that the primary features space was segmented into clusters clustering algorithms. As a result, the main clusters were determined and new feature vectors were extracted with reduced dimensions. The secondary feature extraction mechanisms using the WKM clustering method was shown in Figure 2. Each point in the input space was defined as a three-dimensional vector $v_i = (r_i, s_i, f_i)$. In this vector, r denotes the rate, s is the scale, f is the frequency of downward STRFs at each point of the spectro-temporal space. The magnitude components of points $w_i = A_i$ were considered as the weighting factor of input vectors. These primary feature vectors v_i were clustered using WGMM and WKM algorithm and the elements of mean vector and covariance matrix of speech clusters were considered in the secondary attributes as $V = (\mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3)$. In each frame, three clusters were assumed. μ_i and σ_i are mean and variance vectors of the i[th] cluster.

3. TWO STRATEGIES FOR CLUSTERS SORTING IN THE SECONDARY FEATURE VECTORS

In this study, two strategies were used for clusters sorting in spectro-temporal features vectors. In the first strategy, it was assumed that the center of cluster with larger magnitude has more information. Therefore, the features were sorted according to the cluster's amplitude. Time variations of clusters were not considered in this method and the centers of the clusters sorting are performed based on energy measure. It

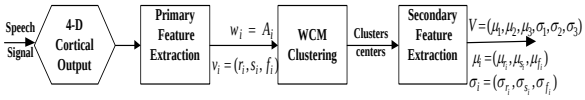

Figure 2. Spectro-temporal features extraction using WKM clustering

means that the magnitude component of clusters centers is sorted descending. Energy-based cluster sorting strategy may cause the system sensitive to noisy conditions; because the cluster locations are changed in the scale, rate and frequency axes during the time. Therefore, the results of temporal tracking of the clusters over time were used for features sorting in the second strategy. In this method, the centers of the clusters were matched to a reference vector using a distance measure. The result of this matching determines the order of clusters. Euclidian distance measure was employed to track the clusters. Two mechanisms in various conditions were used for temporal clusters tracking.

3. 1. Clusters Matching Over Consecutive Frames

In the first mechanism, the cluster centers are matched over the consecutive frames. Therefore, the distances between each cluster center of the current frame and all cluster centers of the previous frame are computed. Then, the best match of the current frame cluster centers with the cluster centers of the previous frame is determined by minimizing the Euclidean distance between their corresponding features. Three clusters were assumed for each frame. Therefore, a 3×3 distance matrix is obtained for each frame. Each element of the distance matrix $dis(i, j)$ is defined as:

$$dis(i, j) = \sum_{k=1}^{n} \left(C_{ic}(k) - C_{jp}(k)\right)^2 \quad (1)$$

where C_{ic} and C_{jp} are the i[th] cluster center of the current cluster and the j[th] cluster center of the previous frame respectively and n are the numbers of features in each cluster center. Each cluster center vector, $C_i = (\mu_i, \sigma_i)$, have six components. Each mean vector μ_i consists of three components as $\mu_i = (\mu_{r_i}, \mu_{s_i}, \mu_{f_i})$ and the variance vector σ_i consists of three components $\sigma_i = (\sigma_{r_i}, \sigma_{s_i}, \sigma_{f_i})$. Thus, the secondary feature vector had 18 elements. In this matching strategy, the centers of the clusters of the first frame of each phoneme are sorted in descending according to their amplitudes. Then, the cluster centers of the next frames are rearranged by the matching results using the status matrix that defines possible permutations of the clusters in features vectors. Status matrix is defined as:

$$S = \begin{bmatrix} 1 & 1 & 2 & 2 & 3 & 3 \\ 2 & 3 & 1 & 3 & 2 & 1 \\ 3 & 2 & 3 & 1 & 1 & 2 \end{bmatrix} \quad (2)$$

Each column of the status matrix shows the position of clusters centers in each speech frame after cluster matching over time. In fact, the column number of the status matrix shows the matching status number. Although in some frames, the positions of clusters centers remain unchanged, (e.g. the first column of the status matrix); however, the positions of most clusters centers change over time. To determine the clusters matching the result of each frame, the cost vector was defined using the status matrix.is the cost $Cost(k)$ function of the k[th] matching status according to the k[th] columns of the status matrix. The cost function was calculated for all columns of the status matrix in each frame.

$$Cost(k) = \sum_{i=1}^{3} w_i.dis(i, S(i,k)) \quad (3)$$

$w_i \in \{0,1\}$ is the death/birth factor of the i[th] cluster. This factor is zero if the minimum distance between a cluster in the current frame and the clusters of the previous frame is more than the empirical threshold value T.

$$w_i = \begin{cases} 0 & if \quad \min(dis(i, j)) \geq T \\ 1 & Otherwise \end{cases} \quad (4)$$

If this factor is zero, it means that a cluster is dead and a new cluster is born in the new frame. In this case, the i[th] cluster with $w_i = 0$ is not considered in the cluster matching procedure. Finally, the best matching is determined regarding the minimum value of the cost functions of the status matrix columns as:

$$BM = Arg\min_k(Cost) \quad (5)$$

$$Best \quad Matching(i) = S(i, BM) \quad , i = 1,2,3 \quad (6)$$

The clusters locations of the current frame (except the first frame) are permuted according to the best matching that is obtained from Equation (6).

3. 2. Clusters Matcing Using Reference Vecor

In another matching mechanism, the clusters of the primary feature vectors which are sorted using the energy measure were matched with the clusters of a reference vector. Three reference vectors are assessed in this clusters matching strategy. In local matching (LM) mechanism, the local reference vectors are determined for each phoneme separately. The local reference vector of each phoneme is calculated by averaging its feature vectors in the utterance sample. Then, all frames of each phoneme are matched according to this reference vector and the distances between all clusters centers of each frame and all cluster centers of the reference vector are computed for each phoneme. Each cluster in the current frame is matched to the cluster of the reference vector that minimizes the Euclidean distance between their corresponding features. Finally, the cluster centers are rearranged using matching results that are obtained between each frame and the reference vector according to Equation (6). In class-based matching (CBM) mechanism, the reference vectors are determined by averaging between all feature vectors of each class in the training phase. In this mechanism, the number of

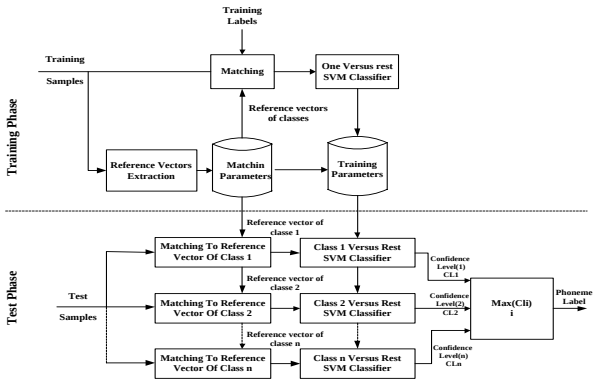

Figure 3. Block diagram CBM mechanism

reference vectors is depended on the numbers of the phonemes classes that should be classified. Assuming a class for each unknown utterance is a prerequisite for this reference vector selection mechanism, which contradicts with the main goal of a classification application. Therefore, the classifier architecture should be adapted to this tracking mechanism. The block diagram of this architecture is proposed in Figure 3. In this architecture, the distances between each cluster center in the current frame and the cluster centers of the reference vector are calculated. In the training phase, the features vectors of each class are matched to the reference vector of the same class. In contrast, in the test phase, the features vectors of an unknown phoneme are matched to the reference vectors of all classes. The feature vectors that are matched with the reference vector of the ith class are classified using the ith class versus the rest binary SVM classifier. The class of an unknown phoneme is determined with respect to the maximum value of decision levels of the classifiers' outputs. In other words, each frame of an unknown phoneme is matched with the best reference vector of the classes by using the confidence levels of the classifiers' outputs. In a global matching (GM) strategy, the frames of all phonemes are matched according to a global reference frame. This global reference vector is calculated by averaging all of the training feature vectors of all phonemes. The calculated global reference vector is also used for clusters matching in the test phase of phoneme classification. After a subtle error analysis of different proposed classification methods, it was observed that there is a considerable mismatch between error samples sets in tracked and non-tracked classifiers. This led us to design a combining classifier mechanism to tune the result. In this mechanism, parallel classifiers are trained using two types of tracked and non-tracked secondary features vectors. To have consistent architectures for classifiers, the proposed one-versus-the-rest SVM classifier architecture was used for the phoneme classification in each branch. Finally, phoneme classification is performed using the combination of confidence values of the classifiers that are obtained in each branch. The employed block diagram of fusion strategy is shown in Figure 4. $CL_1 = (CL_{11}, CL_{21}, ..., CL_{n1})$ and $CL_2 = (CL_{12}, CL_{22}, ..., CL_{n2})$ are the confidence levels vectors for each unknown phoneme that were estimated using the outputs of n classifiers in first and second branches respectively. In this paper, the overall confidence of the ith class CL_i is empirically evaluated by a few fusing rules. Finally, the maximum confidence value will indicate the winner class.

4. RESULTS AND DISCUSSION

Therefore, in this study, most of the experiments are conducted on /b, /d/, /g/ phonemes to evaluate and tune the tracking performance of proposed cluster tracking strategies. The evaluation of the proposed feature extraction method is performed on clean speech and the phonemes are selected from TIMIT acoustic-phonetic continuous speech corpus [14]. The new features vectors were classified using the proposed classifier architecture. Radial basis function (RBF) was used as the SVM kernel.

4. 1. Descriminative Analysis Discriminative capability of individual extracted features is one of the important assessments which determine the preference of features for classification. The discriminative capability of each featurewas calculated using discrimination $D^{(i)}$ measure [15]:

$$D^{(i)} = \frac{(\mu_1^{(i)} - \mu_2^{(i)})^2}{\sigma_1^{(i)^2} + \sigma_2^{(i)^2}} \quad (7)$$

where $D^{(i)}$ is the discrimination measure of the ith attribute of two different phonemes.are $\mu_2^{(i)}$ and $\mu_1^{(i)}$ the average of the ith components of first and second phonemes frames. In addition, $\sigma_1^{(i)}$ and $\sigma_2^{(i)}$ are the

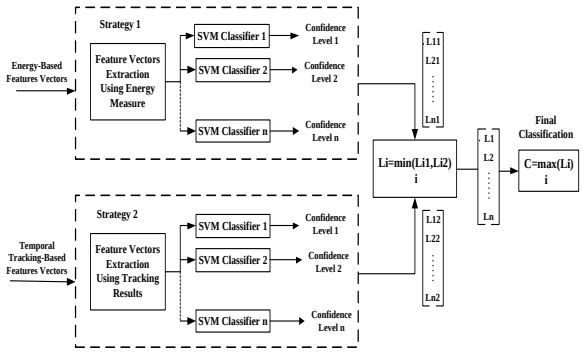

Figure 4. The block diagram of proposed classification and fusion architecture

variances of the ith components of all frames of first and second phonemes respectively. The attribute with greater values of discrimination measure is more discriminative in phoneme classification. The overall discrimination measure of the energy-based features vectors and temporal tracking-based features vectors for /b/, /d/ and /g/ phonemes are tabulated in Table 1. The overall discrimination results show that both secondary features vectors have good discrimination properties.

4. 2. Phoneme Classification Results of One-Versus-the-rest SVM The results are obtained using the best death/birth threshold value in equation (4). The best results of temporal tracking are obtained usingr these strategies respectively fo $T=1$ and $T=1.7$ and $T=1.5$ were the optimum threshold for strategies 3 (GM) and 4 (CBM). The death/birth rates for (/b/, /d/, /g/) phonemes in various matching strategy are shown in Figure 5. Death/birth rate was defined as the number of frames with $w_i=0$ respect to the numbers of all frames.

4. 3. Combining Classification Results The results of /b/, /d/, /g/ phonemes classification using WGMM clustering methods and combining mechanism for the best death/birth threshold values were tabulated in Table 2. The results show that the temporally tracked features gave better results in comparison to energy-ordered features. The best results were obtained by fusing two classifiers. The classification rates on different categories of phonemes using energy-ordered and temporally tracked features using WGMM clustering and combining mechanisms were evaluated and the results were tabulated in Table 3. As it can be observed, the classification results using temporally tracked features was improved in comparison to energy-ordered features in all categories of phonemes. In addition, phoneme classification results were fine-tuned using the combining mechanism. In nasals, significant improvements were 16.9%.

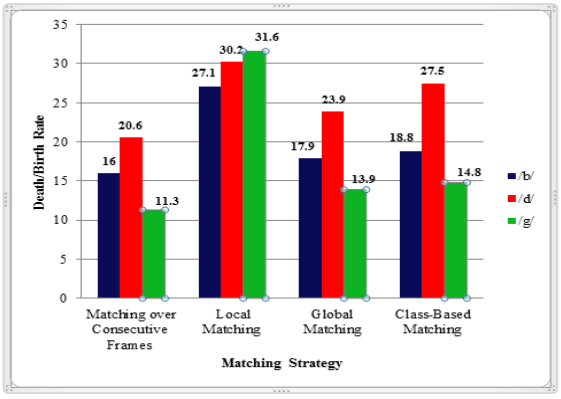

Figure 5. Death/birh rate of (/b/, /d/, /g/) phonemes for various matching strategies

TABLE 1. Discriminative analysis of /b/,/d/,/g/ phoneme

Phonmes	Energy-Based	Temporally Tracked Features using WKM	TemporallyTracked Features using WGMM
b,d	0.79	0.82	0.91
b,g	0.71	0.84	0.87
g,d	0.86	1.14	1.29

TABLE 2. Phoneme classification rates using proposed features using WGMM clustering

Matching Strategy	Energy-Based Features	Tracked Features	Combining Classification Rate	Fusion Rue
Strategy1	76.1	74.8	76.5	Summation
Strategy 2 (LM)	76.1	75.2	77.3	Maximum
Strategy 3 (GM)	76.1	77.5	78.6	Maximum
Strategy 3 (CBM)	76.1	78.9	80.4	Summation

TABLE 3. Phoneme classification rate using proposed features using WGMM clustering

Phonemes	Energy-Based Features	Tracked Features	Combining Classification Rate
Voiced Plosives (/b/,/d/,/g/)	76.1	78.9	80.4
Unvoiced Plosives (/p/,/t/,/k/)	70.6	72.5	73.4
Voiced Fricatives (/v/,/dh/,/z/)	83.6	85.3	85.9
unvoiced Fricatives (/t/,/s/,/sh/)	89.6	92.4	93.1
Nasals (/m/,/n/,/ng/)	50.3	.664	67.2
Front Vowels (/ih/,/ey/,/eh/,/ae/)	64.4	67.1	68.1
Back Vowels (/uw/,/uh/,/ow/,/aa/)	.974	76.3	77.5

5. CONCLUSION

In this paper, cluster tracking-based methods were proposed to extract the discriminative features in the spectro-temporal domain. Secondary spectro-temporal features vectors were extracted using clustering methods. In the first strategy, the clusters were sorted to consider their energy in the spectro-temporal domain. In the second strategy, the cluster centers were sorted in

feature vectors based on temporal tracking results. The various matching strategy was used for temporal tracking of the clusters. Overall, GM and CBM strategies were successful in comparison to other matching strategies. In addition, results show that temporally tracked feature vectors give better results in comparison to energy-based features vectors. Finally, fusing the classifiers showed good performance to cover the errors in tracked and non-tracked approaches.

6. REFERENCES

1. Ruiz-Muñoz, J.F., You, Z., Raich, R. and Fern, X.Z., "Dictionary learning for bioacoustics monitoring with applications to species classification", ***Journal of Signal Processing Systems***, Vol. 90, No. 2, (2018), 233-247.
2. Chi, T., Ru, P. and Shamma, S.A., "Multiresolution spectrotemporal analysis of complex sounds", ***The Journal of the Acoustical Society of America***, Vol. 118, No. 2, (2005), 887-906.
3. Mesgarani, N., David, S.V., Fritz, J.B. and Shamma, S.A., "Mechanisms of noise robust representation of speech in primary auditory cortex", ***Proceedings of the National Academy of Sciences***, Vol. 111, No. 18, (2014), 6792-6797.
4. Lu, K., Liu, W., Zan, P., David, S.V., Fritz, J.B. and Shamma, S.A., "Implicit memory for complex sounds in higher auditory cortex of the ferret", ***Journal of Neuroscience***, Vol. 38, No. 46, (2018), 9955-9966.
5. Yin, P., Shamma, S.A. and Fritz, J.B., "Relative salience of spectral and temporal features in auditory long-term memory", ***The Journal of the Acoustical Society of America***, Vol. 140, No. 6, (2016), 4046-4060.
6. Ruggles, D.R., Tausend, A.N., Shamma, S.A. and Oxenham, A.J., "Cortical markers of auditory stream segregation revealed for streaming based on tonotopy but not pitch", ***The Journal of the Acoustical Society of America***, Vol. 144, No. 4, (2018), 2424-2433.
7. Francis, N.A., Elgueda, D., Englitz, B., Fritz, J.B. and Shamma, S.A., "Laminar profile of task-related plasticity in ferret primary auditory cortex", ***Scientific Reports***, Vol. 8, No. 1, (2018), 16375. doi: 10.1038/s41598-018-34739-3
8. Winkowski, D.E., Nagode, D.A., Donaldson, K.J., Yin, P., Shamma, S.A., Fritz, J.B. and Kanold, P.O., "Orbitofrontal cortex neurons respond to sound and activate primary auditory cortex neurons", ***Cerebral Cortex***, Vol. 28, No. 3, (2017), 868-879.
9. Shamma, S. and Dutta, K.J.T.J.o.t.A.S.o.A., "Spectro-temporal templates unify the pitch percepts of resolved and unresolved harmonics", Vol. 145, No. 2, (2019), 615-629.
10. Elgueda, D., Duque, D., Radtke-Schuller, S., Yin, P., David, S.V., Shamma, S.A. and Fritz, J.B.J.N.N., "State-dependent encoding of sound and behavioral meaning in a tertiary region of the ferret auditory cortex", Vol. 22, No., (2019), 447-459.
11. Esfandian, N., Razzazi, F., Behrad, A. and Valipour, S., "A feature selection method in spectro-temporal domain based on gaussian mixture models", in IEEE 10th International Conference On Signal Processing Proceedings, IEEE., (2010), 522-525.
12. Esfandian, N., Razzazi, F. and Behrad, A., "A clustering based feature selection method in spectro-temporal domain for speech recognition", ***Engineering Applications of Artificial Intelligence***, Vol. 25, No. 6, (2012), 1194-1202.
13. Esfandian, N., Razzazi, F. and Behrad, A., "A feature extraction method for speech recognition based on temporal tracking of clusters in spectro-temporal domain", in The 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012), IEEE., (2012), 12-17.
14. Fisher, W.M., "Ther darpa speech recognition research database: Specifications and status", in Proc. DARPA Workshop on Speech Recognition, (1986), 93-99.
15. Guyon, I., Gunn, S., Nikravesh, M. and Zadeh, L.A., "Feature extraction: Foundations and applications, Springer, Vol. 207, (2008).

Phoneme Classification Using Temporal Tracking of Speech Clusters in Spectro-temporal Domain

N. Esfandian

Department of Electrical Engineering, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran

P A P E R I N F O

Paper history:
Received 28 August 2019
Received in revised form 22 September 2019
Accepted 08 November 2019

Keywords:
Phoneme Classification
Spectro-temporal Features
Primary Auditory Cortex
Speech Feature Extraction
Weighted Gaussian Mixture Models

چکیده

این مقاله، یک روش جدید استخراج ویژگی مبتنی بر ردیابی خوشه ها در فضای ویژگی های را طیفی- زمانی را معرفی می کند. در روش پیشنهادی، خروجی کورتیکال شنیداری خوشه بندی شده است. ویژگی های خوشه های گفتار به عنوان ویژگی های ثانویه استخراج شد. از آنجایی که شکل و مکان خوشه های گفتار در طول زمان تغییر می کند، خوشه ها، در طول زمان ردیابی شده است و پارامترهای ردیابی زمانی به عنوان ویژگی های ثانویه در نظر گرفته شد. یک ساختار جدید برای طبقه بندی واج ها با استفاده از طبقه بندی کننده ترکیبی و با استفاده از ویژگی های مبتنی بر ردیابی زمانی و ویژگی های مبتنی بر انرژی پیشنهاد شده است. ویژگی های طیفی- زمانی مبتنی بر خوشه بندی برای طبقه بندی دسته های مختلف واج ها از بانک اطلاعاتی TIMIT استفاده شده است. نتایج نشان داده است که نرخ طبقه بندی واج ها با استفاده از ویژگی های طیفی- زمانی مبتنی بر ردیابی زمانی بهبود یافته است. نتایج تا ۷۸/۹٪ در طبقه بندی واج های انفجاری صداداربهبود یافته است که ۳/۳ ٪ بالاتر از نتایج حاصل از بردارهای ویژگی های طیفی- زمانی ردیابی نشده بوده است. نتایج حاصل از دیگر مجموعه واج ها نیز بهبود خوبی در نرخ طبقه بندی نشان داده است.

***doi**: 10.5829/ije.2020.33.01a.12*