Voice Clustering Gender Using Fuzzy Possibilistic C-Means Standard

Arif Setiawan[1], Pratomo Setiaji[2], Putri Kurnia Handayani[3], Noor Latifah[4]
{arif.setiawan@umk.ac.id[1], pratomo.setiaji@umk.ac.id[2], putri.kurnia@umk.ac.id[3]}

Faculty of Engineering, Universitas Muria Kudus, Gondang Manis Bae Kudus, Kudus, 59324, Indonesia[1234]

Abstract. To recognize a voice pattern the computer requires a standard and logical mechanism. The main problem is how process acquisition data by generating a number of numerical data are representative and consistent.Voice recognition system use feature extraction techniques based on domain time with the two methods are short-time energy and zero crossing rate. Steps being taken is prepared ten audios, process feature extraction method based on domaintime, using a Short Time Energy, Zero Crossing Rate, and clustering using Fuzzy Possibilistic C-Means Standard. Step method voice recognition are input transducer for analyzing input of electronic signal, prepossessing to add signal condition including signal amplification, spectrum analysis and digital conversion, feature extraction to comparing template matching, response selector for selecting input pattern in software using the technique of searching, sorting, least squares analysis, and output system to show result application process.

Keywords: voice clustering, time domain,fuzzy possibilistic c-means standard, short time energy, zero crossing rate.

1 Introduction

Pattern recognitionis various theories in Artificial Intelligence (AI). It is an important component to process imitation human senses, especially seeing and hearing. For example, to replicate the human sense of hearing, computer must has standard mechanism in recognizing patternsound. It ismotivationto try a simple concept recognize patterns of sound that can be identified with either computer. The main problem when going to recognize a certain pattern is how the process of data acquisition to produce a number of numerical representative data and consistent with a sample data. This papertell a simple method to identify and classify sounds based on gender, so it can be properly identified by computer using the various theories feature extraction audio data. The main objective of this research is to analyze and prove a simple method to extract the voice data sample in ten people, consisting eight men and two women.The voice implemented either in accordance with its objectives so the computer can identify voice consistent.Interests and needs in speech-recognition community have provided a major motivation for the work on speaker clustering [1]–[3].

Voice recognition processes automatically recognizing who is speaking on the basis of individual information included sound wave, sounds recognizable by their features. These features are used to distinguish a voice with another voice. Good features are distinguishing

ICCSET 2018, October 25-26, Kudus, Indonesia

DOI 10.4108/eai.24-10-2018.2280623

characteristics, so the group sounds based on characteristics that can be owned by a high accuracy. Voice recognitioncan be classified into identification and verification1. Identify the voice is recognition process based on sound, while the voice verification is the process of acceptance or rejection votes cast.Step method voice recognition are input transducer for analyzing input of electronic signal, prepossessing to add signal condition including signal amplification, spectrum analysis and digital conversion, feature extraction to comparing template matching, response selector for selecting input pattern in software using the technique of searching, sorting, least squares analysis, and output system to show result application process (figure 1)[4]. There was still a dearth of study devoted to this problem. More recently, speaker-clustering research has enjoyed a renaissance[5]–[13], spurred by research activities in spoken document indexing for managing burgeoning collections of available speech data. It is desired that by clustering speech data from the same speakers, the human efforts required for documentation can be dramatically reduced or replaced.

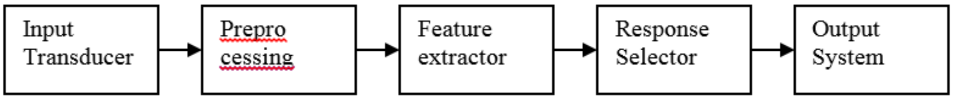

Fig. 1. Pattern Recognition Structure.

Main techniques contained in the pattern recognition system is a cluster analysis, namely identification of substructure in the data set labeled [1]. The clustering techniques is C-means, this technique using dissimilarity measure to classifying objects. Dissimilarities implemented into concept of distance, two objects be similar if it is close. The higher value of the distance, the higher its value to no resemblance. Steps method of C-means algorithm [2] are chose K initial center points randomly, cluster data using Euclidean distance (or other distance metric), calculate new center points for each cluster using points within the cluster, clusters all data using the new center points, this step could cause data points to be placed in a different cluster, repeat steps three and four until the center points has moved.

The purpose of estimating the time domain is to get the value of the auto correlation of audio signals. Auto correlation value of an audio signal will show how the shape of the wave form of a correlation to itself as a function of changes over time. The forms of the same or similar at any given time delay indicates repetition of the audio signal pattern or periodicity. Thereby we will be able to do the estimated value of the fundamental frequency. The discussion revolves around the process of look at Short Time Fourier Analisys of an audio signal, or which is also known as Power Spectral Density (PSD) on the duration an audio signal or a specific frame. By knowing the form of PSD audio signals we will be able to perform feature extraction of the audio signal.

Fuzzy C-Means (FCM) is a clustering technique where the existence of data that each data in a cluster is determined by the value of membership. The basic concept of the FCM, the first time is to determine the cluster center which will mark the location of the average for each cluster. In the initial condition, the cluster center is still not accurate. Each of the data has a degree of membership to each cluster. By improving the cluster center and membership value of each data repeatedly, it will be seen that the cluster center will move towards the right location. This iteration is based on minimizing an objective function. The objective function used in theFCM[14]:

$$J_w(U,V,X) = \sum_{i=1}^{n}\sum_{k=1}^{C}(\mu_{ik})^w (d_{ik})^2 \tag{1}$$

$w \in [1,\infty]$,

$$d_{ik} = d(x_i - v_k) = \left[\sum_{j=1}^{m}(x_{ij} - v_{kj})^2\right]^{1/2} \tag{2}$$

X is data to be in the cluster:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \tag{3}$$

V is matrix center cluster :

$$V = \begin{bmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{c1} & \cdots & v_{nm} \end{bmatrix} \tag{4}$$

Value J_w smallest is the best, so that:

$$J_w^*(U^*,V^*;X) = \min_{Mk} J(U,V,X) \tag{5}$$

theory 1 [15] :

if $d_{ik} > 0, \forall i,k; w > 1,$ has at least c elements, then $(U,V) \in M_k \text{ x } \Re^{cp}$ can minimize J_w only if :

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{w-1}}} \qquad 1 \le k \le C;\ 1 \le i \le n. \tag{6}$$

$$V_{kj} = \frac{\sum_{i=1}^{n}\left((\mu_{ik})^w * X_{ij}\right)}{\sum_{i=1}^{n}(\mu_{ik})^w} \qquad 1 \le k \le C;\ 1 \le j \le m. \tag{7}$$

FCM Algorithm is given as follows [16]:

1. Initialize Matrix X, n x m, with n = the amount of data that will be in the cluster, and m = number of variables (8)
2. Initialize :
   a. The number of clusters = C (≥ 2).
   b. weight = w (> 1)

c. Maximum iteration (MaxItr)
d. Stop criteria = ξ
e. Beginning Iteration, t = 1 (9)

3. Initialize matrix partition U

$$U = \begin{bmatrix} U_{11}(x_{11}) & U_{12}(x_{12}) & \cdots & U_{1k}(x_{1k}) \\ U_{21}(x_{21}) & U_{22}(x_{22}) & \cdots & U_{2k}(x_{2k}) \\ \vdots & & & \vdots \\ U_{i1}(x_{i1}) & U_{i2}(x_{i2}) & \cdots & U_{ik}(x_{ik}) \end{bmatrix} \tag{10}$$

Calculate the amount of each column (attribute) for normalization matrix U:

$$Qk = \sum_{k=1}^{c} \mu_{ik} \tag{11}$$

Calculate the normalized matrix U

$$\mu ik = \frac{\mu ik}{Qk} \tag{12}$$

4. Inisialize fuzzy cluster center V_{kj}, with k = 1, 2,...c and j = 1, 2,...m

$$V_{kj} = \frac{\sum_{i=1}^{n} \left((\mu_{ik})^{w} * X_{ij} \right)}{\sum_{i=1}^{n} (\mu_{ik})^{w}} \tag{13}$$

5. Calculate objective function :

$$Pt = \sum_{i=1}^{n} \sum_{k=1}^{c} \left(\left[\sum_{j=1}^{m} \left(X_{ij} - V_{kj} \right)^{2} \right] (\mu_{ik})^{w} \right) \tag{14}$$

6. Calculate new matrix partition U :

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m} \left(X_{ij} - V_{kj} \right)^{2} \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c} \left[\sum_{j=1}^{m} \left(X_{ij} - V_{kj} \right)^{2} \right]^{\frac{-1}{w-1}}} \tag{15}$$

i = 1, 2,…,n; k = 1, 2,…,c

7. Iteration process until fulfilled maximum criteria iteration :
If $\left(\left| Pt - Pt - 1 \right| < \xi \right)$ or (t > MaxIter) then stop. (16)
If not : t = t + 1, back to step 4.

Fuzzy Possibilistic C-Means Standar (FPCM) is a clustering algorithm in which each point belongs to a cluster that is based on the possibility. In this algorithm requires calculating the likelihood for a character to be used as a vector belonging to a cluster. In FCM algorithm, value μ_{ik} influenced by x_k and all cluster centers. But not so with PCM. In PCM algorithm, t_{ik} value only affected by x_k, v_i (cluster center i) and γ_i only. From here can be said that, μ_{ik} a relative distinctiveness, while t_{ik} is absolute distinctiveness of the cluster -i. Possibilistic Fuzzy C-Means (FPCM) is also based on the minimization of the objective function, as follows [15]:

$$J_{w,\eta}(T,V,X) = \sum_{i=1}^{n}\sum_{k=1}^{C}(\mu_{ik}^{w} + t_{ik}^{\eta})(d_{ik})^2 \tag{17}$$

$$w \in (1,\infty], \eta \in [1,\infty), 1 \le i \le c,$$

$$d_{ik} = d(x_i - v_k) = \left[\sum_{j=1}^{m}\left(x_{ij} - v_{kj}\right)^2\right]^{1/2} \tag{18}$$

$$\sum_{k=1}^{C}\mu_{ik} = 1;\ \forall k;$$

$$\sum_{i=1}^{n}t_{ik} = 1;\ \forall k;$$

T matrix distinctiveness :

$$T = \begin{bmatrix} t_{11} & \cdots & t_{1c} \\ \vdots & & \vdots \\ t_{n1} & \cdots & t_{nc} \end{bmatrix} \tag{19}$$

X is the data to be in clusters :

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \tag{20}$$

and V is the matrix of cluster centers :

$$V = \begin{bmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{c1} & \cdots & v_{cm} \end{bmatrix} \tag{21}$$

$J_{w,h}$ values, the smallest is the best, so:

$$J^*_{w,\eta}(U^*, T^*, V^*; X) = \min_{M_{pc}} J(U, T, V; X) \tag{22}$$

Teorema 3 [15]:

If $d_{ik} > 0, \forall i, k; w, \eta > 1,$ and X has at least c elements, the $(U, T, V) \in M_{fpc}$ x M_{fpc} x $\Re^p$ can minimize $J_{w,h}$ only if :

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{w-1}}} \qquad 1 \le k \le C;\ 1 \le i \le n \tag{23}$$

$$t_{ik} = \frac{\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{\eta-1}}}{\sum_{i=1}^{n}\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{\eta-1}}} \qquad 1 \le k \le C;\ 1 \le i \le n \tag{24}$$

$$V_{kj} = \frac{\sum_{i=1}^{n}(\mu_{ik}^{w} + t_{ik}^{\eta})x_{ij}}{\sum_{i=1}^{n}(\mu_{ik}^{w} + t_{ik}^{\eta})} \quad 1 \le k \le C;\ 1 \le j \le m. \tag{25}$$

FPCM algorithm is given as follows :

1. Determine the size of the n x m matrix X, with n = number of data to be cluster, and m = number of variables (criteria). (26)
2. Initialize :
   a. The number of clusters = C ($\ge$ 2).
   b. Weight = w (> 1) and η (>1)
   c. Maximum iteration (MaxItr)
   d. Stop criteria = ξ
   e. Beginning Iteration, t = 1 (27)
3. Take the final results of FCM algorithm, that is the partition matrix U and V cluster center, used to calculate the absolute distinctiveness matrix T:

$$t_{ik} = \frac{\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{\eta-1}}}{\sum_{i=1}^{n}\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{\eta-1}}} \qquad i = 1,2,...n;\ k = 1,2,...c;\ j = 1,2,...m \tag{28}$$

4. Fix V_{kj} cluster center, as follows:

$$V_{kj} = \frac{\sum_{i=1}^{n}(\mu_{ik}^{w} + t_{ik}^{\eta})x_{ij}}{\sum_{i=1}^{n}(\mu_{ik}^{w} + t_{ik}^{\eta})} \qquad i = 1,2,...n;\ k = 1,2,...c;\ j = 1,2,...m \tag{29}$$

5. Fix the relative distinctiveness matrix U, as follows:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{w-1}}} \tag{30}$$

6. Fix distinctiveness absolute matrix T, as follows:

$$t_{ik} = \frac{\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{\eta-1}}}{\sum_{i=1}^{n}\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{\eta-1}}} \tag{31}$$

7. Calculate the objective function at iteration - t, Pt:

$$Pt = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right]\left(\mu_{ik}^{w} + t_{ik}^{\eta}\right)\right) \tag{32}$$

8. Iteration process until the maximum iteration criterion,
   If $\left(\left|Pt - Pt - 1\right| < \xi\right)$ or (t > MaxIter) then stop.

If not : t = t + 1, back to step 4 (33)

Several similarity measures such as cross likelihood ratio[8], generalized likelihood ratio [10], and Bayesian information criterion[9]–[11] have beenexamined and compared in many literatures. The concept of the above clustering method is basically the same as a prior study reported in[17]. A similar idea has also been presented recently from the viewpoint of the socalled triangulation18, in which each utterance is modeled as a single Gaussian distribution[18], [19].

To obtain accurate and consistent data from each sample, used a feature extraction method with a sound signal in the time domain [20]. Exposure time domain is the basic technique of audio signal, where the signal amplitude is described as the unit of time, the signal may be positive or negative depending on the sound pressure. In this paper the authors use two methods, the sort time energy and zero crossing rate. The method used is as follows:

(a) Sort Time Energy

The energy associated with speech is time varying in nature. Hence the interest for any automatic processing of speech is to know how the energy is varying with time and to be more specific, energy associated with short term region of speech. By the nature of production, the speech signal consist of voiced, unvoiced and silence regions. Further the energy associated with voiced region is large compared to unvoiced region and silence region will not have least or negligible energy. Thus short term energy can be used for voiced, unvoiced and silence classification of speech. The relation for finding the short term energy can be derived from the total energy relation defined in signal processing.The total energy of an energy signal is given by[3]:

$$STE = \frac{1}{N}\sum_{n=1}^{N} X\,(n)^2 \tag{34}$$

Description:
STE = Sort time energy
N = Number of samples
X(n) = Value of the signal from the sample

(b) Zero Crossing Rate

The zero crossing rate is the rate of sign changes along a signal, the rate at which the signal changes from positive to negative or back[21]. This feature has been used heavily in both speech recognition and music information retrieval, being a key feature to classify percussive sounds. Sample sequence on a digital signal have different signs, the size of a noise signal on domain feature

$$ZC = \frac{\sum_{n=1}^{N} |\, sgn\, x(n) - sgn\, x(n-1)|}{2N} \tag{35}$$

Description:
ZC = Zero Crossing Rate
sgn x(n) = Value from x(n) , value is 1 if x (n) is positive, -1 if x (n) negative
N = Value of samples

Each method of feature taken his average, using the standard deviation, the following is a table (table 1) of average usage for each method of feature.In some cases only the "positive going" or "negative going" crossings are counted, rather than all the crossings since, logically, between a pair of adjacent positive zero-crossings there must be one and only one negative zero-crossing.

Table 1. Average usage for each method of feature.

Feature	Statistic

Sort Time Energy	Standart Deviation (std)
Zero Crossing Rate	Standart Deviation (std)

2 Method

In this study, the methodology of research is as follows (figure 2) :
(1) Reading of audio data
(2) Feature extraction process using Time Domain
(3) Feature extraction is done by two methods, namely Short Time Energy and Zero Crossing Rate
(4) Obtained matrix is stored in a file pola.txt
(5) From matrix, carried out using the C-means clustering
(6) Obtained results clustering

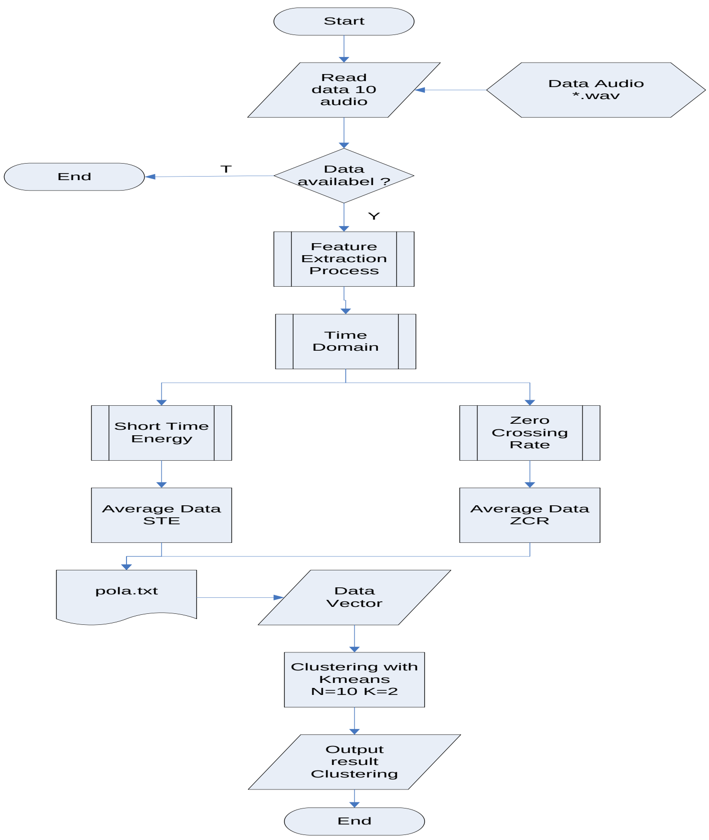

Fig. 2. Flowchart research.

3 Implementation

Before doing the voice recording must be made arrangement as in the figure 3.

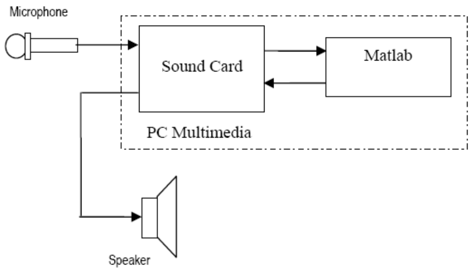

Fig. 3. Design recording device audio signal.

Computer must be equipped with multimedia tools such as sound card, active speakers and microphone. For microphone and active speakers can also be replaced with a full head set. Data on calculation of audio signal feature extraction from 10 samples obtained and recorded sound file "pola.txt", the stored data is an average of data from each of the feature extraction method.

Table 2. Data on average of feature extraction methods.

No	Name of Student	File Name	Short Time Energy	Zero Crossing Rate	Figure
1	Lia	lia.wav	0.0538388	0.02373	Fig 4
2	Rizki	rizki.wav	0.0730546	0.0305071	Fig 5
3	Edo	edo.wav	0.0662328	0.023269	Fig 6
4	Malik	malik.wav	0.137326	0.0199186	Fig 7
5	Budi	budi.wav	0.0923871	0.0449648	Fig 8
6	Rudi	rudi.wav	0.112719	0.0181066	Fig 9
7	Hendy	hendy.wav	0.115847	0.0283285	Fig 10
8	Susapto	susapto.wav	0.129537	0.0441993	Fig 11
9	Arif	arif.wav	0.0664739	0.0190522	Fig 12
10	Retno	retno.wav	0.0866736	0.0283818	Fig 13

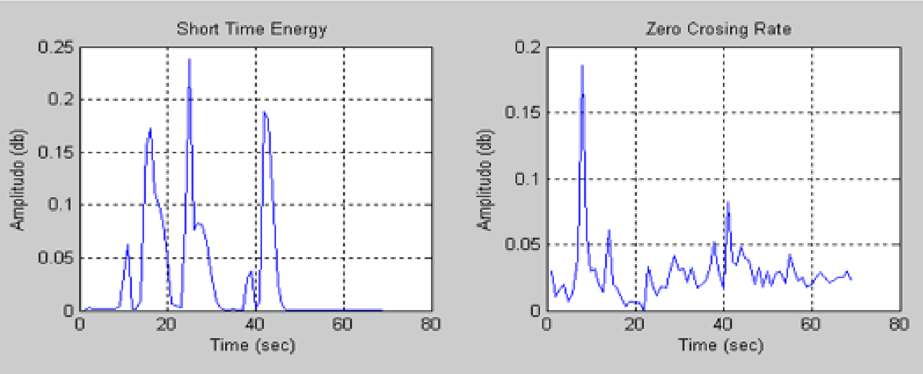

Fig. 4. Graph Lia voice with file lia.wav.

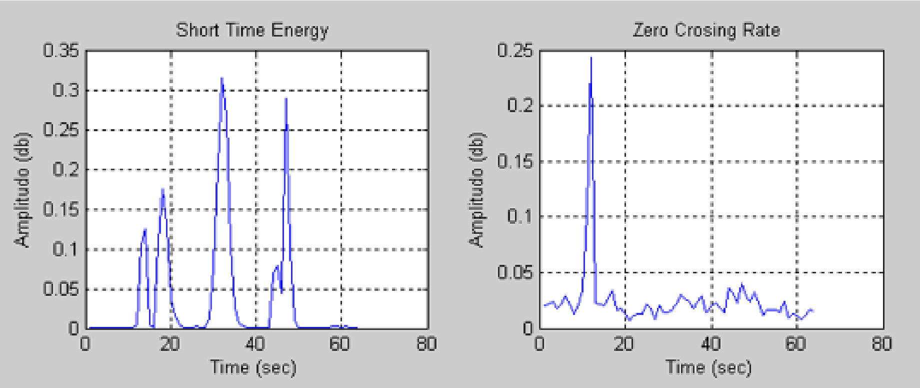

Fig. 5. Graph Rizki voice with file rizki.wav.

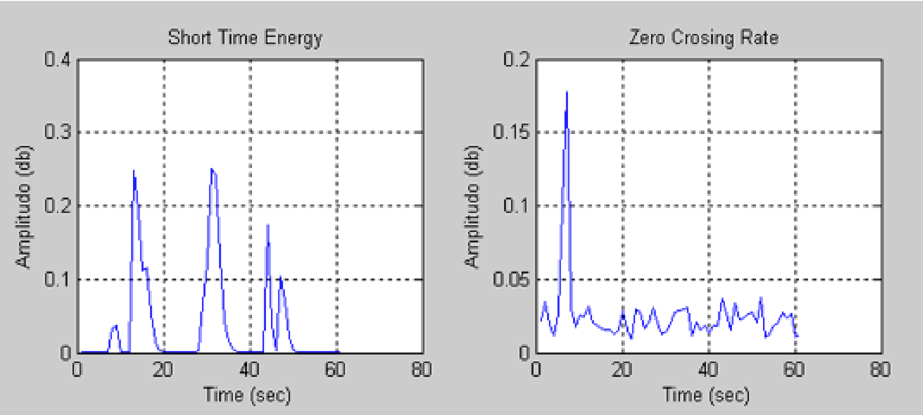

Fig. 6. Graph Edo voice with file edo.wav.

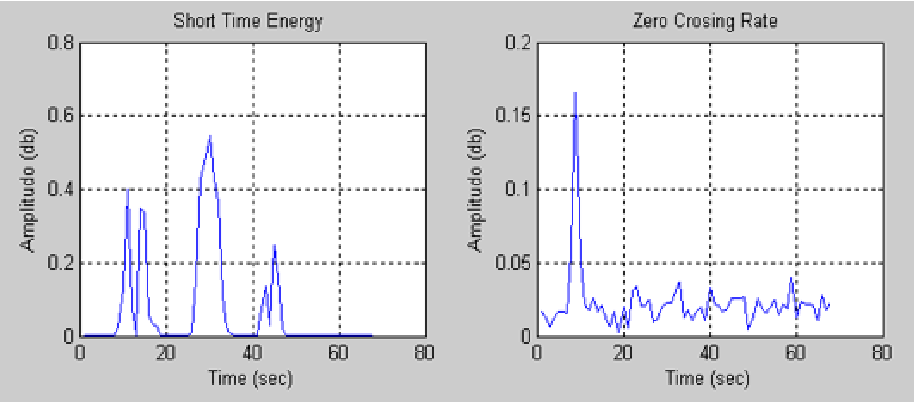

Fig. 7. Graph Malik voice with file malik.wav.

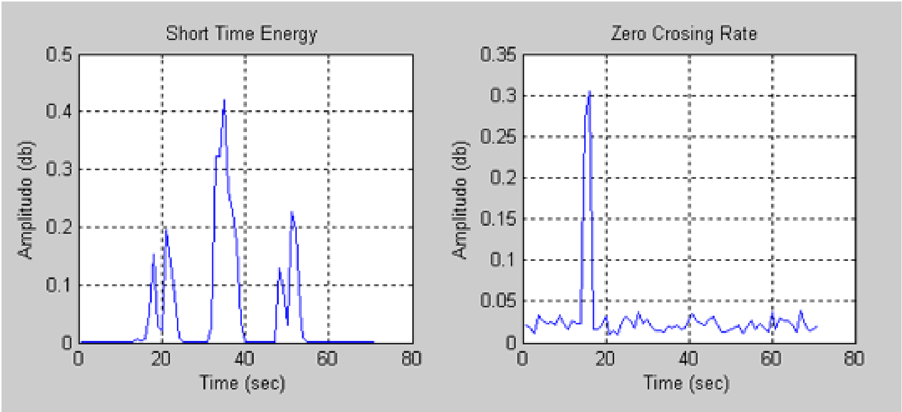

Fig. 8. Graph Budi voice with file budi.wav.

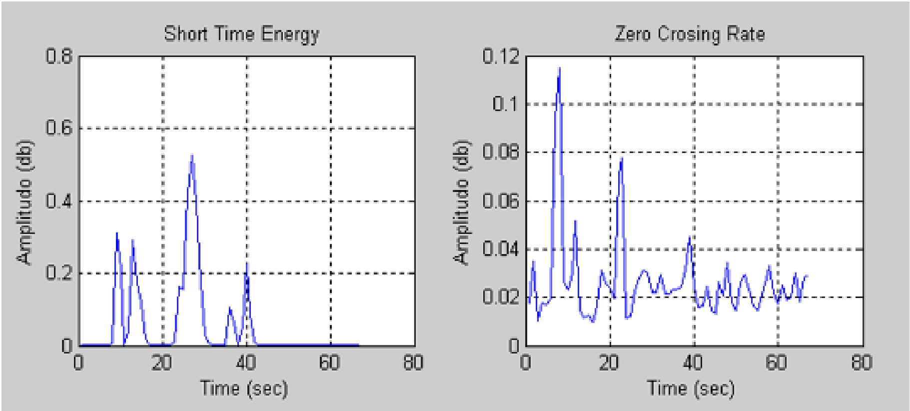

Fig. 9. Graph Rudi voice with file rudi.wav.

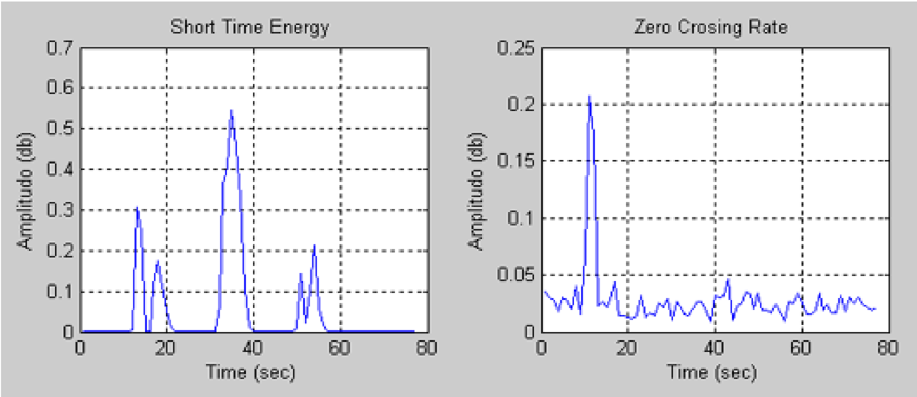

Fig. 10. Graph Hendy voice with file hendy.wav.

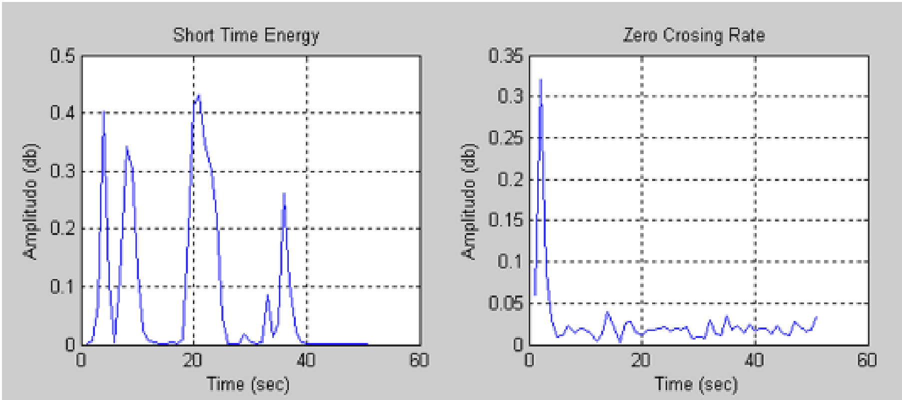

Fig. 11. Graph Susapto voice with file Susapto.wav.

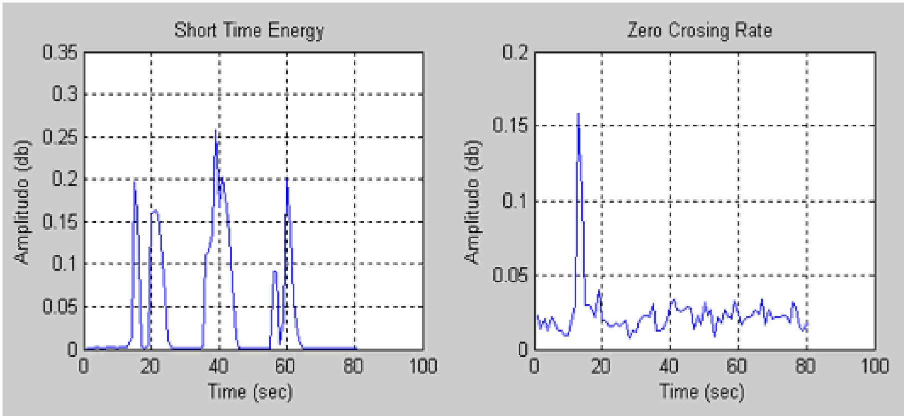

Fig. 12. Graph Arif voice with file Arif.wav.

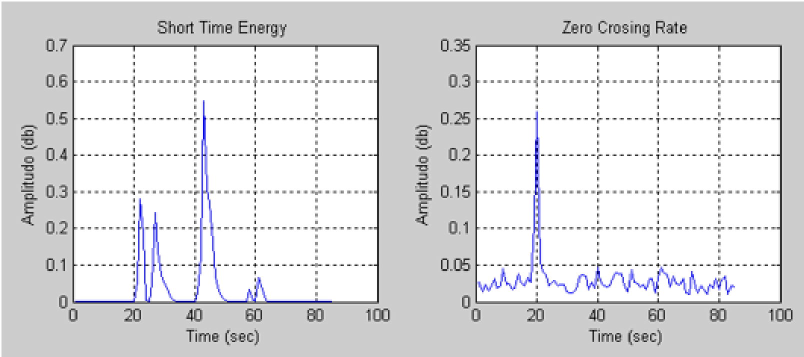

Fig. 13. Graph Retno voice with file Retno.wav.

Calling the data from ”pola.txt” changed into a matrix. Matrix M is obtained:

M=

0.0538388	0.02373	0.0893531	0.13844
0.0730546	0.0305071	0.123814	0.473905
0.0662328	0.023269	0	0.511881
0.137326	0.0199186	0	0.732067
0.0923871	0.0449648	0.124977	0.483487
0.112719	0.0181066	0.181973	0.586998
0.115847	0.0283285	0.116744	0.540119
0.129537	0.0441993	0	0.655706
0.0664739	0.0190522	0.173884	0.409927
0.0866736	0.0283818	0.126805	0.313418

From the results of the Clustering using the C-means clustering function, the obtained results:

maxRow = 10
maxCol = 4
c =

0.0538	0.0237	0.0894	0.1384
0.0731	0.0305	0.1238	0.4739

X =

0.0538	0.0237	0.0894	0.1384	1.0000
0.0731	0.0305	0.1238	0.4739	2.0000
0.0662	0.0233	0	0.5119	2.0000
0.1373	0.0199	0	0.7321	2.0000
0.0924	0.0450	0.1250	0.4835	2.0000
0.1127	0.0181	0.1820	0.5870	2.0000
0.1158	0.0283	0.1167	0.5401	2.0000
0.1295	0.0442	0	0.6557	2.0000
0.0665	0.0191	0.1739	0.4099	2.0000
0.0867	0.0284	0.1268	0.3134	1.0000

X = 1 2 2 2 2 2 2 2 2 1

Description :

1 = Women, 2 = Men

4 Conclusion

The results can be concluded that the voice clustering based on gender can be done with a sound signal feature extraction method based on time domain and frequency domain. Feature extraction that used is the Sort Time Energy, Zero Crossing Rate. The average value of each characteristic was calculated by the standard deviation, to obtain his average, and then processed for classification.

References

[1] E. Darren Ellis, “Design of a Speaker Recognition Code using MATLAB.” .

[2] Olivier Lartillot and Petri Toiviainen, “A MATLAB TOOLBOX FOR MUSICAL FEATURE EXTRACTION FROM AUDIO,” in in Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07), Bordeaux, France, 2007.

[3] S. Fu rui, “Unsupervised speaker adaptation method based on hierarchical spectral clustering,” in Proc. ICASSP’89.

[4] Richard J Hathaway and James C. Bezdek, “Recent Convergence Results for the Fuzzy c-Means,” J. Classif., vol. 5, pp. 237–247, 1988.

[5] Kosaka T and Sagayama S, “Tree-structured speaker clustering for fast speaker adaptation,” in Proc. ICASSP’94.

[6] Siegler M. A, Jain U, Raj B, and Stern R. M, “Automatic segmentation, classification and clustering of broadcast news audio,” in Proc. DARPA Speech Recognition Workshop, 1997.

[7] Jin H, Kubala F, and Schwartz R, “Automatic speaker clustering,” in Proc. DARPA Speech Recognition Workshop, 1997.

[8] Solomonoff A, Mielke A, Schmidt M, and Gish H, “Clustering speakers by their voices,” in Proc. ICASSP’98, pp. 757–760.

[9] Reynolds D. A, Singer E, Carson B. A, O’Leary G. C, McLaughlin J. J, and Zissman M. A., “Blind clustering of speech utterances based on speaker and language characteristics,” in Proc. ICSLP’98.

[10] S. Chen S and Gopalakrishnan P.S., “Clustering via the Bayesian information criterion with applications in speech recognition,” in Proc. ICASSP’98.

[11] Johnson S. E, “Who spoke when? Automatic segmentation and clustering for determining speaker turns,” in Proc. Eurospeech’99.

[12] Zhou B and Hansen J. H. L, “Unsupervised audio stream segmentation and clustering via the Bayesian information criterion,” in Proc. ICSLP’00.

[13] Lapidot I, G. H, and Cohen A, “Unsupervised speaker recognition based on competition between selforganizing maps”, IEEE Trans. Neural Networks,” IEEE Trans. Neural Networks.

[14] Ross. T. J, “Fuzzy Logic With Engineering Applications, John Wiley & Sons Inc.” United Kingdom, 2005.

[15] Pal. N. R, Keller. J.M, and Bezdek. J C, “A Possibilistic C-Means Clustering Algorithm,” IEEE Trans. Fuzzy Syst., vol. 13.

[16] Zimmermann. H. J, “Fuzzy Set Theory and Its Applications, Kluwer Academic.” Massachusetts, 1991.

[17] Moh Y, Nguyen P, and Junqua J.-C, “Towards domain independent speaker clustering,” in Proc. ICASSP’03.

[18] Tsai W. H, Cheng S.S, and Wang H.M, “Speaker Clustering of Speech Utterances Using A Voice Characteristic References Space.” .

[19] “http://iitg.vlab.co.in/?sub=59&brch=164&sim=857&cnt=1, anonym.” .
[20] P. http://www.nist.gov/speech/tests/index.htm [16] Hubert, L., Arabie, “Comparing Partitions,” J. Classif.
[21] Minh N. Do, “Digital Signal Processing Mini-Project “An Automatic Speaker Recognition System,” 2001.