Dynamic Feature Selection for Detecting Parkinson's Disease through Voice Signal

Meilin Su*, Keh-Shih Chuang †

* Department of Electrical and Computer Engineering
Oriental Institute of Technology, New Taipei City, 22061 Taiwan, Republic of China
Email: meilin@mail.oit.edu.tw
† Department of Biomedical Engineering & Environmental Sciences
Email: kschuang@mx.nthu.edu.tw
National Tsing-Hua University, Hsinchu City, 22061 Taiwan, Republic of China

***Abstract* — Parkinson's disease (PD) is a disorder of the central nervous system and about 89% of the people with PD suffering from speech and voice disorders. In this paper, we adopted a dynamic feature selection based on fuzzy entropy measures for speech pattern classification of Parkinson's diseases. To investigate the effect of feature selection, Linear Discriminant Analysis (LDA) was applied to distinguish voice samples between PD patients and health people. The data set of this research is composed of voice signals from 40 people, 20 with Parkinson's disease and 20 health people. The results show that various voice samples need different feature selection. We applied dynamic feature selection can get higher rate of classification accuracy than all features selected.**

***Index Terms* — Parkinson's disease, linear discriminant analysis, similarity measure, fuzzy entropy**

I. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder after Alzheimer's disease. The main symptoms of PD include tremor, muscle rigidity, and voice dysphonia. Ho A.K. et al. [1] has indicated that up to 90% of PD people display vocal impairment and a lot of research has focused on voice analysis for detecting Parkinson's disease. M.A. Little [2] extracted nonlinear speech features out of the subjects for voice disorder detection. B.E. Sakar et al. [3] designed a computed-aid data collection system and applied K-nearest neighbor (KNN) and support vector machine (SVM) classification to distinguish health subjects from PWP.

In classification procedure how to select features is an important research. Removing insignificant features from the dataset enhances the classification accuracy and reduces the computational cost. M. Shahbakhi [4] used genetic algorithm and Pasi Luukka [5] applied fuzzy entropy measures to reduce the feature dimension into classification model. In this paper, we used the fuzzy entropy based on similarity measures to remove insignificant features and adopted a dynamic feature selection for different voice samples. We evaluated the success of feature selection in discriminating the health subjects from subjects with PD.

II. Methodology

A. . Parkinson speech dataset

The dataset was built by B.E. Sakar et al. [3], it consists of 40 people at which 20 subjects suffer from PD (6 female, 14 male) and the rest are healthy (10 female, 10 male). From all subjects, multiple types of sound recordings (Table Ⅰ) were taken. A group of 26 linear and nonlinear features (Table Ⅱ) were extracted from each voice sample. UPDRs (Unified Parkinson's Disease Rating Scale) score of each patient was determined by expert physician is available in the dataset.

TABLE Ⅰ
PD DATASET (26 VOICE SAMPLES)

Sample No.	Description
1	sustained vowel (aaa)
2	sustained vowel (ooo)
3	sustained vowel (uuu)
4-13	numbers from 1 to 10
14-17	short sentences
18-26	words

B. Feature selection

Fuzzy entropy measure is considered to be a fuzziness measure and it evaluates global deviation from the type of ordinary sets. In this work, we used the fuzzy entropy based on similarity measure to remove insignificant features from the dataset. According to Shannon probabilistic entropy the measure of fuzzy entropy is equation (1)

$$\mathrm{H}(\mathrm{A}:\omega) = \sum_{j=1}^{n}(\mu_A(x_j) log\, \mu_A(x_j) + \left(1-\mu_A(x_j)\right)\log(1-\mu_A(x_j)) \quad (1)$$

We calculated the similarity value $\mathrm{S}\langle\mathbf{x},\mathbf{v}\rangle$ based on Yu's norm as the entropy value $\mu_A(x_j)$. The similarity value is,

$$\mathrm{S}\langle\mathbf{x},\mathbf{v}\rangle = T\left(Sn(\bar{\mathbf{x}},\mathbf{v}), Sn(\mathbf{x},\bar{\mathbf{v}})\right) \quad (2)$$

$$\mathrm{T}\langle\mathbf{x},\mathbf{v}\rangle = max[0,(1+p)(\mathbf{x}+\mathbf{v}-1)-\mathrm{p}\mathbf{x}\mathbf{v}] \quad (3)$$

$$\mathrm{Sn}\langle\mathbf{x},\mathbf{v}\rangle = min[1,\mathbf{x}+\mathbf{v}+\mathrm{p}\mathbf{x}\mathbf{v}] \quad (4)$$

where $\bar{x} = 1 - x$ and p is a parameter in range (-1,∞).

C. Classification

Linear discriminant analysis (LDA) is usually used to classify unknown cases to different groups. In this paper, we used LDA classifier to evaluate the performance of feature selection.

TABLE II
PD DATASET(26 EXTRACTED FEATURES)

Feature No.	Feature Description
features 1-5	Jitter (local), Jitter (local, absolute), Jitter (rap), Jitter (ppq5), Jitter (ddp)
features 6-11	Shimmer (local), Shimmer (local, dB), Shimmer (apq3), Shimmer (apq5), Shimmer (apq11), Shimmer (dda)
features 12-14	AC, NTH, HTN
features 15-19	Median pitch, Mean pitch, Standard deviation, Minimum pitch, Maximum pitch
features 20-23	Number of pulses, Number of periods, Mean period, Standard deviation of period
features 24-26	Fraction of locally unvoiced frames, Number of voice breaks, Degree of voice breaks

III. RESULTS

Fig.1 is the results of the feature selection for various voice samples. It shows the increasing in the accuracy rate while more features are considered, but it can be noticed that many cases the best accuracy does not exist on full features selected.

Table III indicates that removed insignificant features are not identical, it depends on different voice samples.

IV. CONCLUSION

In this study, the fuzzy entropy based on similarity measure was performed to discard redundant and irrelevant features in PD dataset. The results show that using fuzzy entropy we can effectively remove insignificant features.

The future work from this study is to analyze which voice samples have higher discrimination for PD patients and apply different classifiers (e.g. support vector machine, k-nearest neighbors) to verify the feature selection algorithm.

REFERENCES

[1] Ho, A.K. , Iansek, R., Marigliani, C., Bradshaw, J.L., Gates, S., "Speech impairment in a large sample of patients with parkinson's disease", *Behavioural Neurology*, Vol.11, pp. 131–137, 1998

[2] Max A. Little, Patrick E. McSharry, Eric J. Hunter, Lorraine O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease", *IEEE Transactions on Biomedical Engineering,* 56, pp. 1015-1022, 2008

[3] B.E. Sakar, M.E. Isenkul, C. Okan Sakar, F. Gurgen, S. Delil, H. Apaydin, O. Kursun, "Collection and Analysis of a Parkinson Speech Dataset with Multiple Types of Sound Recordings," *IEEE Journal of Biomedical and Health Informatics,* Vol. 17, no. 4, pp. 823-834.,2013

[4] Mohammad Shahbakhi, Danial Taheri Far, Ehsan Tahami, " Speech Analysis for Diagnosis of Parkinson's Disease Using Genetic Algorithm and Support Vector Machine", *Journal of Biomedical Science and Engineering.* Vol. 7, pp. 147-156, 2014

[5] Cesar Iyakaremye, Pasi Luukka, David Koloseni, "Feature selection using Yu's similarity measure and fuzzy entropy measures," *IEEE World Congress on Computational Intelligence.* June, 10-15,2012-Brisbane, Australia.

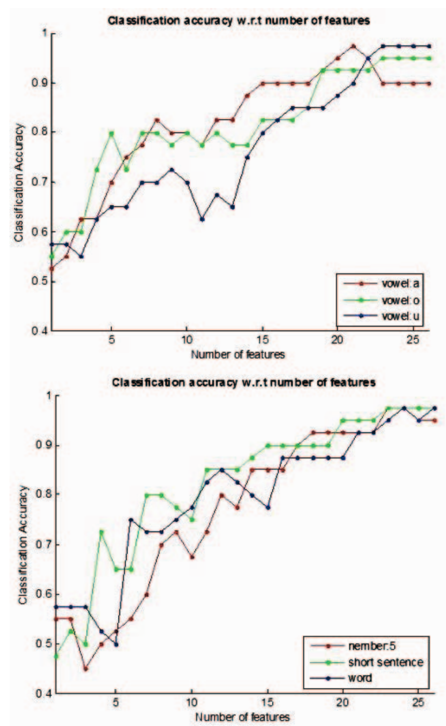

Fig.1 Classification accuracy curve vs. number of features

TABLE III
FEATURE SELECTION AND CLASSIFICATION ACCURACY

	Remove Feature(No.)	Selected features	Accuracy %
Vowel /a/	2,14,10,20,21	21	97.5
Vowel /o/	21,20,14	23	95.0
Vowel /u/	2,10,7	23	97.5
Number (#5)	24,2,10	23	97.5
sentence(#15)	19,10,24	23	97.5
Word(#23)	25,24	24	97.5

Note: #5 corresponding 5th voice samples