INFLUENCE OF THE NUMBER OF THE FEATURES WITH THE NEURAL NETWORK FUNCTION

Jana TUČKOVÁ, Petr BOREŠ
Dept.of Circuit Theory
Czech Technical University in Prague
Technicka 2, 166 27 Prague 6
Czech Republic

Abstract

This contribution is devoted to the evaluation of probability of success for classificatio and recognition techniques. It also depends on a proper selection of input elements-features and their number. Error probability is proportional indirectly on a quantity and quality of information provided to a classifier. It can be affected either by a learning algorithm itself and classification or by an element number on classification correcturs has been verified for Kohonen's map designed for a recognition of Czech digits.

Keywords

Speech recognition, neural network, classification

1. Introduction

A neural network approach is the one of possible technique for signal analysis. In this paper we propose an influence of a number of features of isolated Czech digit recognition.

One of reasons for difficult solution of a speechprocessing is a speech signal character. The speech is a very complicated complex signal with many technical parameters, phonetic properties of national languages and human physiological properties. Recognition problems exist because there is not sufficient knowledge how a man take in a speech, what parts of speech signal are important for a man and what parts can be put down without a decreasing of understandability.

A recognition capability of a confirmous utterances depends especially on the knowledge of a listener too. Because the noise has influenced a processed information, a man with the apriori knowledge has higher probability of success. In the automatic speech recognition we can begin partially from the technical and mathematical point of view, partially we can apply these abstract properties. To what extent, it is a question. We can assume, that a success of recognition will be greater because of increased amount of information. The increased amount of information is equal to a greater number of input elements-features. The response is not easy to express. The increased of number of features makes a great increase of a time for decision. Also the structures and the methods for the recognition are more omplicated. That is a reason why to find a compromise.

We tried to verify this influence on the neural netclassifier, the Kohonen's self-organizing feature map.

2. Neural Net Classifier

Software simulations were applied for a construction of neural net classifier. The Kohonen's map for this approach had following parameters:

- topology: square map
- dimension 12 * 12
- number of units: 144
- number of iteration: 10 000
- initial gain: 0.9 resp. 0.1
- initial neighbourhood: 5
- training set: 10 Czech digits 0, 1, ..., 9
  1967 input cepstr.vectors
  5 different mal speakers in one set
- number of features in each vector: 9, 6, 4, 2
- test set: 10 Czech digits
  5 516 input cepstr.vectors
  (1 967 from of them was trained, 3 549 was non trained)
- 13 different mal speakers

The Kohonen's map was trained by the set with 1967 cepstral vectors. KLB programm was used. Results of this experiment are:

- map of adapted synaptic weights
- maps of centroids for a separate digits
- centroids and error dependence on the order of the input vector
- average errors of the separate digits
- maps of the relative frequency of the cetroid coincidence

The results from the training procedure are copared with the test set of the particular digit by the programm REGISTER1. A principal criterion of a digit recognition was an agreement of centroid coordinates in the map. The complement criteria were average errors after the classification of the particular digits. This complement was applied when the recognition was uncertain.

3. Influence of the feature decrease

The programmes KLB (for learning) and REGISTER1 (for test) were applied to the training set and test set for 2, 4, 6, and 9 features-cepstral coefficients. We had verified these dependences.

Classifiers based on classical methods or classifers used, for example Hidden Markov Models (HMM), confirm this hypothesis. However the results of our expements allow to express the conclusion, that a situation is more complex for neural net classifiers. We cannot specify if this conclusion is true or false.

The network architectures and learning algorithms can affect considerably an answer to this question. The choice of the input features and their optimal number are important. Any methods concern to problems of decrease of number of input features are based upon discrete Karhunen-Loeve series (see [2]). The choice of the minimal numbers of features from the predeterminated feature set is the base for the other methods. It cannot specify how to choose this set. From this point of view, it cannot defined, if the features form optimal characteristic properties. It is very difficult for the speech obviously. The simple way was to use for the verification our assumptions. It is not a statistical method, but only simple decrease of numbers of cepstral coefficients, in which many informations about speech signal are contained.

4. Experimental results

Results of the correct recognition are ilustrated on figures and diagrams. The dependence of the correct recognition of the particular digits on the influence of a feature number is showed. Columns are associated with the values of the correct recognised digits (v %). Column "a" is associated with the training set, column "b" is associated to the digits prononced by the unknown speaker, column "c" is associated with the all test digits (130 test word). Almost in all events, the test results are gained for 4 and 6 cepstral coefficients. It is apparent, that the lower distorsion is obtained for the figure of 4-th resp. 6-th dimensional input space into 2-nd dimensional map than for greater differences of dimensions. On the other hand, it is possible to keep an important test information about speech.

Four cepstral coefficients characterise the frequence spectrum of the first and second formant frequencies, six cepstral coefficients the frequence spectrum of the first three formant frequencies. This information is sufficient for the vocal recognition. Two cepstral coefficients (correspond to one formant) are sufficient for the word including vocal A (the digits NULA = ZERO and JEDNA = ONE). On the contrary, the word containing vocal I must be described by six cepstral coefficients (correspond to three formant - for example TŘI = THREE and ČTYŘI = FOUR). These results are summarized on Fig.1, Tab.1 (for 130 word), Fig.2, Tab.2 (for 50 trained words), and Fig.3, Tab.3 (for 80 unknown word).

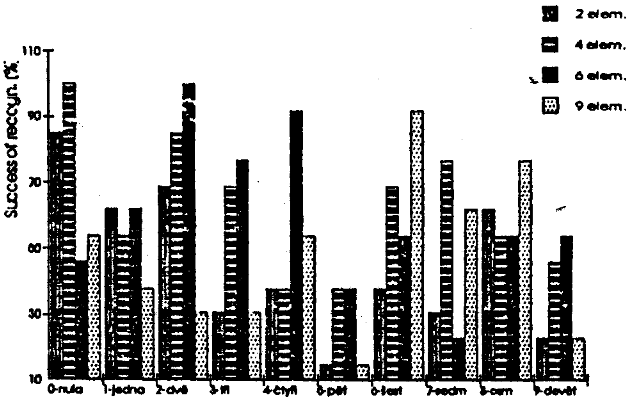

Fig.1: Influence of feature numbers for the recognition of particular digits (for the complete test set)

Digit	2 el.	4 el.	6 el.	9 el.
nula - 0	85	100	46	54
jedna - 1	62	54	62	38
dvě - 2	69	85	100	31
tři - 3	31	69	77	31
čtyři - 4	38	38	92	54
pět - 5	15	38	38	15
šest - 6	38	69	54	92
sedm - 7	31	77	23	62
osm - 8	62	54	54	77
devět - 9	23	46	54	23

Tab.1.: Influence of feature numbers for the recognition of the particular digits (in %) for the complete test set (130 word)

Table 5 summarises the results of the unambiguous recognitio of all digits (130 isolated word, 50 trained and 80 unknown words). From this point of view the succes (in %) is greater than for results obtained from particular digits. The optimal number of features is different too. From this point of view all words perform the greatest succes for 4 features, but from Tab.1 it is apparent that the the best results are for the choice of 6 features. The worst case is always for 9 features. The worst case is always for 9 features. The diagram on the Fig.4 shows this situation. Results correspond to the table presented the unambiguous and ambiguous cases (also in Tab.4).

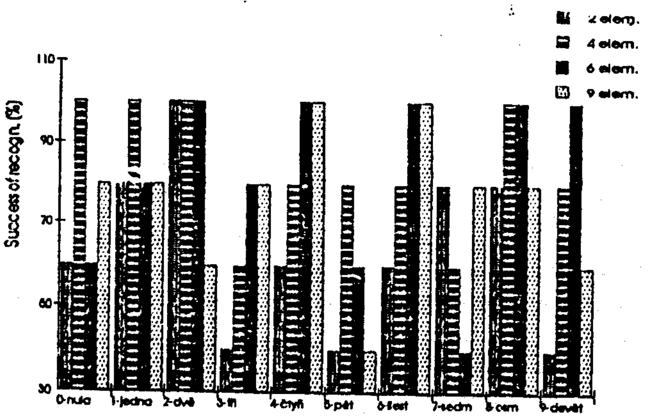

Fig.2: Influence of feature numbers for the recognition of particular digits for the training set

Digit	2 el.	4 el.	6 el.	9 el.
nula - 0	60	100	60	80
jedna - 1	80	100	80	80
dvě - 2	100	100	100	60
tři - 3	40	60	80	80
čtyři - 4	60	80	100	100
pět - 5	40	80	60	40
šest - 6	60	80	100	100
sedm - 7	80	60	40	80
osm - 8	80	100	100	80
devět - 9	40	80	100	60

Tab.2.: Influence of feature numbers for the success of the recognition of particular digits in training set (50 words from 5 speakers)

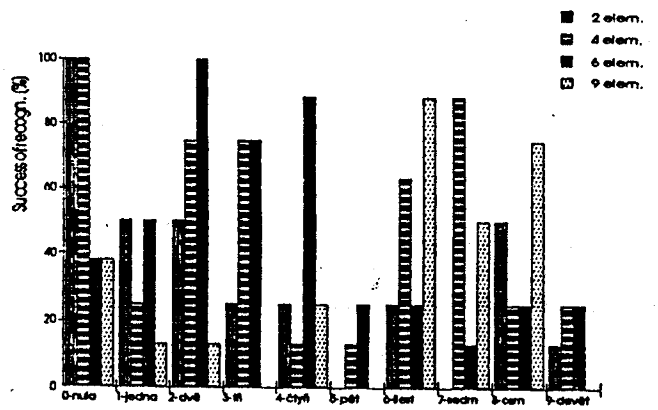

Fig.3.: Influence of feature numbers for the recognition of particular digits (for the unknown digits)

Digit	2 el.	4 el.	6 el.	9 el.
nula - 0	100	100	38	38
jedna - 1	50	25	50	13
dvě - 2	50	75	100	13
tři - 3	25	75	75	0
čtyři - 4	25	13	88	25
pět - 5	0	13	25	0
šest - 6	25	63	25	88
sedm - 7	0	88	13	50
osm - 8	50	25	25	75
devět - 9	13	25	25	0

Tab.3.: Influence of feature numbers for the success of the recognition of particular digits (in %) for the unknown digits (80 words from 8 speakers)

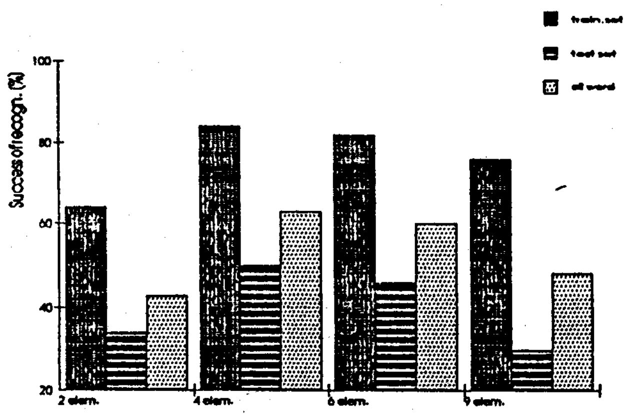

Fig.4.: Succes of the definite recognition of all digits (130 isolated word) depending on the feature numbers

Digit	2 el.	4 el.	6 el.	9 el.
training	64	84	82	76
test	34	50	46	30
all	43	63	60	48

Tab.4.: Succes of the definite recognition of all digits (130 isolated word) depending on the feature numbers

Digit	2 el.	4 el.	6 el.	9 el.
nula - 0	100	100	54	100
jedna - 1	69	54	62	54
dvě - 2	85	85	100	46
tři - 3	62	85	77	69
čtyři - 4	85	46	92	85
pět - 5	38	85	62	69
šest - 6	69	77	54	92
sedm - 7	46	92	62	92
osm - 8	85	69	77	100
devět - 9	69	62	54	46

Tab.5. Influence of feature numbers for the success of the recognition of the all digits (in %) for the the unambiguous and ambiguous case

5. Conclusion

In this paper we proposed a neural network approach for the problem of isolated Czech digit recognition. The comparison between these results and methods based on Hidden Markov Models (HMM) shows that:

- values larger than 90 labeling word or for the training based on phonems or diphones
- our results are obtained for nonlabeling isolated word

The success 89 % of the recognition of isolated word (the same database) by HMM is reached for much more cepstral coefficients. For 9 features, the success dropped to less than 80 %. Results of our experiments are close. The success 75 % can be reach by the neural net optimalisation.

The presented computer simulation results were obtained by using the original software.

References

[1] KOHONEN,T.: An Introduction to Neural Computing., Neural Networks, Vol.1.,pp.3-16, 1988

[2] KOTEK,Z.-BRŮHA,I.-CHALUPA,V.-JELÍNEK,J.: Adaptivní a učící se systémy. SNTL, Praha 1980

[3] TUČKOVÁ,J.-BOREŠ,P.-VONDRÁŠEK,F.: Neuronové sítě - moderní metoda zpracování signálu. Sborník semináře Teorie obvodů STO-5 „Moderní směry výuky elektrotechniky a elektroniky", Brno, 12-13.10.1994

About authors...

Jana TUČKOVÁ was born in Prague in 1950. She received the Ing. degree in 1974, CSc. degree in 1980, both from the Czech Technical University in Prague, Postdoctoral course Biological and Artificial Neural Networks in 1993 from the Dept.of Informatic of Swiss Federal Institute of Technology in Lausanne. Her research interests are theorie of neural networks and their applications in the signal processing, especially in speech recognition and speech synthesis.

Petr BOREŠ was born in Prague in 1949. He received the Ing. degree in 1973, CSc. degree in 1979, both from the Czech Technical University in Prague. His research interest is modelling and analysis with respect to computer oriented education of network theory.

CONTENTS OF ELECTRICAL ENGINEERING JOURNAL

(Issued by the Faculty of Electrical Engineering and Informatics, Slovak Technical University, Bratislava).

(Only papers published in English are mentioned.)

No. 5 , 1995