Evaluation of correlation methods applying neural networks

Abstract This paper deals with a new perception of classic correlation methods on the basis of neural nets. We present and examine neural nets that evaluate the similarity between two arbitrary vectors in a process of measurement or pattern recognition. Thus, classifying patterns by means of feature vectors is feasible just as by correlation methods. Furthermore, we show that the difference correlation procedure and the squared-distance correlation procedure can be presented directly as special cases of the neural methods. Using an example of a typical recognition problem and Gaussian-distributed measuring errors, computer simulations have yielded that neural and correlation procedures are almost identical in behaviour regarding the error rates. Consequently, the neural procedures presented can be understood as a generalisation of correlation procedures.

Keywords Classification · Correlation methods · Neural networks · Pattern recognition · Pre-processing · Structure analogy

1 Introduction

It is well-known that the object of pattern recognition techniques is to decide whether a given unclassified pattern belongs to an already known category of pattern. For this purpose, the pattern not yet classified is often compared to those patterns whose classification is known. To compare the patterns, first, from each pattern a feature vector is formed. There are numerous possibilities for this feature extraction which in general depend on the application (see e.g. [1, 2]).

This paper deals with comparing two feature vectors with the aid of neural nets. At least one of the vectors is based on measurands which are typically disturbed by systematic and stochastic errors. To solve the comparison problem, we present neural procedures and compare them with classic correlation procedures. This is reasonable because correlation procedures are known to be standard procedures for pattern recognition. On the other hand, the significance of neural networks keeps increasing for pattern recognition due to the intensive development of special hardware, e.g. [3, 4]. Numerous neural networks in the field of pattern recognition have already been successfully tested (see e.g. [5, 6, 7, 8]). The special thing about the neural networks described in literature is that the reference sets are saved in the weight values of the neural nets. It is disadvantageous that the reference patterns must be trained anew for each application. Therefore, in this paper – deviating from the applications mentioned above – neural networks are described allowing the direct comparison of two arbitrary feature vectors. Comparing the neural approaches with known classic correlation techniques allows a new interpretation of the correlators.

2 Comparison procedures

To compare two patterns $\mathbf{M}_1$ and $\mathbf{M}_2$, the two feature vectors

\[
\mathbf{M}_1 = (f_1(\mathbf{M}_1), f_2(\mathbf{M}_1), \ldots, f_N(\mathbf{M}_1)) = (m_1(1), m_1(2), \ldots, m_1(N))
\]

(1)

and

\[
\mathbf{M}_2 = (f_1(\mathbf{M}_2), f_2(\mathbf{M}_2), \ldots, f_N(\mathbf{M}_2)) = (m_2(1), m_2(2), \ldots, m_2(N))
\]

(2)

are formed first by an operation set $F = \{f_1, f_2, \ldots, f_N\}$. The choice of the functions $f_1, \ldots, f_N$ depends on the application. As a rule, there are numerous possibilities to generate sensible feature vectors. Basing on the two vectors $\mathbf{M}_1$ and $\mathbf{M}_2$, comparison procedures provide a measure for the similarity of the two patterns $\mathbf{M}_1$ and $\mathbf{M}_2$. 

Wolfgang Holzapfel · Manfried Sofsky
Institute of Measurement and Automation,
University of Kassel, D-34109 Kassel
E-mail: holzapfel@uni-kassel.de
For classifying a new pattern \( M \), normally \( L \) different classes are given. These classes can be described by \( N \)-dimensional vector spaces indicating the class boundaries or more simply by the vectors defining the class centers \( M_{1}, M_{2} \) to \( M_{1,L} \). In case a class is given by a vector space one has to prove whether the feature vector \( M \) lies inside the class boundaries. If the class centres are known, pattern \( M \) can be assigned by applying one of the correlation or neural procedures described below.

One has to check whether a pattern disturbed by systematic and stochastic measuring errors belongs to one of the \( L \) different classes. If stochastic properties of the classes are known, decision boundaries can be calculated which minimise any defined cost functions. In practical applications, these pre-conditions are often not fulfilled so that a method has to be chosen showing good performances for current disturbances. For this purpose, the similarity between the feature vector \( M \) and each vector \( M_{1}, M_{2} \) to \( M_{L} \) is determined. Then \( M \) is assigned to that class having the highest degree of similarity. In addition, a minimum similarity can be fixed which, if not reachable, generates a new class or the pattern is classified as not to be assigned.

2.1 Correlation procedures

Well-known suitable procedures for comparing two feature vectors are the difference correlation and the squared-distance correlation [9, 10, 11]. For the difference correlator the amounts of the \( N \) elements of the difference vector between \( M_{1} \) and \( M_{2} \) are summarised according to Eq. (3), standardised by the factor \( N \cdot C \) and subtracted from 1:

\[
K_{\text{Diff}} = 1 - \frac{1}{N \cdot C} \sum_{i=1}^{N} |m_{1}(i) - m_{2}(i)| .
\]  

(3)

Here, magnitude \( C \) is a completely optional standardisation factor which merely requires \( C > 0 \).

For the squared-distance correlator, the squares of the differences are summarised instead of the difference amounts \( |m_{1}(i) - m_{2}(i)| \). As this sum equals the square of the geometric distance of the two vectors, this procedure is called squared-distance correlation:

\[
K_{\text{Absq}} = 1 - \frac{1}{N \cdot C} \sum_{i=1}^{N} (m_{1}(i) - m_{2}(i))^{2} .
\]  

(4)

If the correlators of squared distance and difference are generalised the result is

\[
K = 1 - \frac{1}{N \cdot C} \sum_{i=1}^{N} |m_{1}(i) - m_{2}(i)|^{P}
\]  

(5)

where \( P = 1 \) provides the difference correlator and \( P = 2 \) the squared-distance correlator.

These correlation procedures have one thing in common: independently of the standardisation factor \( C \), they provide the correlation value \( K = 1 \) in case of completely coincident feature vectors \( M_{1}, M_{2} \). The bigger the deviation between \( M_{1} \) and \( M_{2} \), the smaller the correlation value \( K \) is. The magnitude of the correlation value, with a given feature number \( N \), merely depends on the standardisation factor \( C \). Only if calculation accuracy is poor does \( C \) influence the probability of a classification error.

2.2 Neural comparison procedures without pre-processing

To perform a pattern recognition process on the basis of neural nets, first the network shown in Fig. 1 is investigated [12]. This net directly detects (without any pre-processing) with its input neurons the components of the feature vectors \( M_{1}, M_{2} \) to be compared. The input layer is divided in two ranges with the same number of artificial neurons. On the one side, the feature vector \( M_{1} \) is put in whereas the other side is reserved for the input of \( M_{2} \). The net has one neuron in the output layer that provides the output signal \( K_{\text{neuro}} \). In principle, the number of hidden layers is arbitrary as well as their number of units.

Each unit in Fig. 1 is described by the function

\[
o_{j} = \Theta(\text{net}_{j}) \quad \text{with} \quad \text{net}_{j} = \sum_{i=1}^{n} w_{ji} \cdot u_{i} .
\]  

(6)

Here the \( n \) input signals \( u_{i} \) of the \( j \)-th neuron are weighted with the factors \( w_{ji} \) and the superposition of the weighted signals yields the net input \( \text{net}_{j} \) of the unit. The signals \( u_{i} \) can be input signals of the neural net as well as output signals from other units. By means of the