INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN
ICED 05 MELBOURNE, AUGUST 15-18, 2005

METHOD USING A SELF-ORGANISING MAP FOR DRIVER CLASSIFICATION AS A PRECONDITION FOR CUSTOMER ORIENTED DESIGN

Albert Albers and Marc Albrecht

Abstract

The presented paper describes a method using a Self-Organising Map algorithm (SOM) for driver classification and to consequently assign drivers to a certain customer group according to the similarities in their style of driving. The number of customer groups can be specified by the topology of the SOM. With this method one precondition for customer oriented design and individualisation of vehicles is created.

Keywords: Driver modelling; Classification; Customer oriented design

1 Introduction

Today's cars already show a large amount of automated systems, already allowing the tuning of vehicle characteristics to a special customer group's demands (e. g. engine control, transmission control [1]). The use of adaptable software and therefore the possibility of individualisation of vehicle characteristics to meet customer demands, is expected to become more and more important in future. If models for driver sensation exist, this adaptation can already be done during virtual product development [2], [3]. One precondition for managing this is to determine which customer group a driver belongs to and consequently what he likes or dislikes. Based on this, the right vehicle characteristic which will satisfy this customer can be decided upon. For this, a tool is required that allows the assignment of the driver type (customer type) from measurable data.

2 Data Acquisition in Drive Tests

A group of 23 drivers was asked to drive a course representing an urban drive situation three times, driving in the manner which they would usually drive. A car with a 2.5l 6-cylinder combustion engine with a power of 125 kW, manual gear shift and rear wheel drive was used as the test-vehicle. To guarantee constant driving conditions and to avoid interactions with traffic, the traffic training area of the Deutsche Verkehrswacht e.V. in Karlsruhe was used as the test track. The driving task was specified as a circuit with three stops. The shape of the track is illustrated in figure 1 with the numbered arrows indicating the course. Stop 1 and stop 3 are placed on a hill, while stop 2 is placed on a flat part of the track. The length of the track is 1 mile (1.6 km) and the drivers needed between 2.5 and 4 minutes per drive, which corresponds to an average speed of 14.9 to 23.9 miles/h (24.0 to 38.4 km/h).

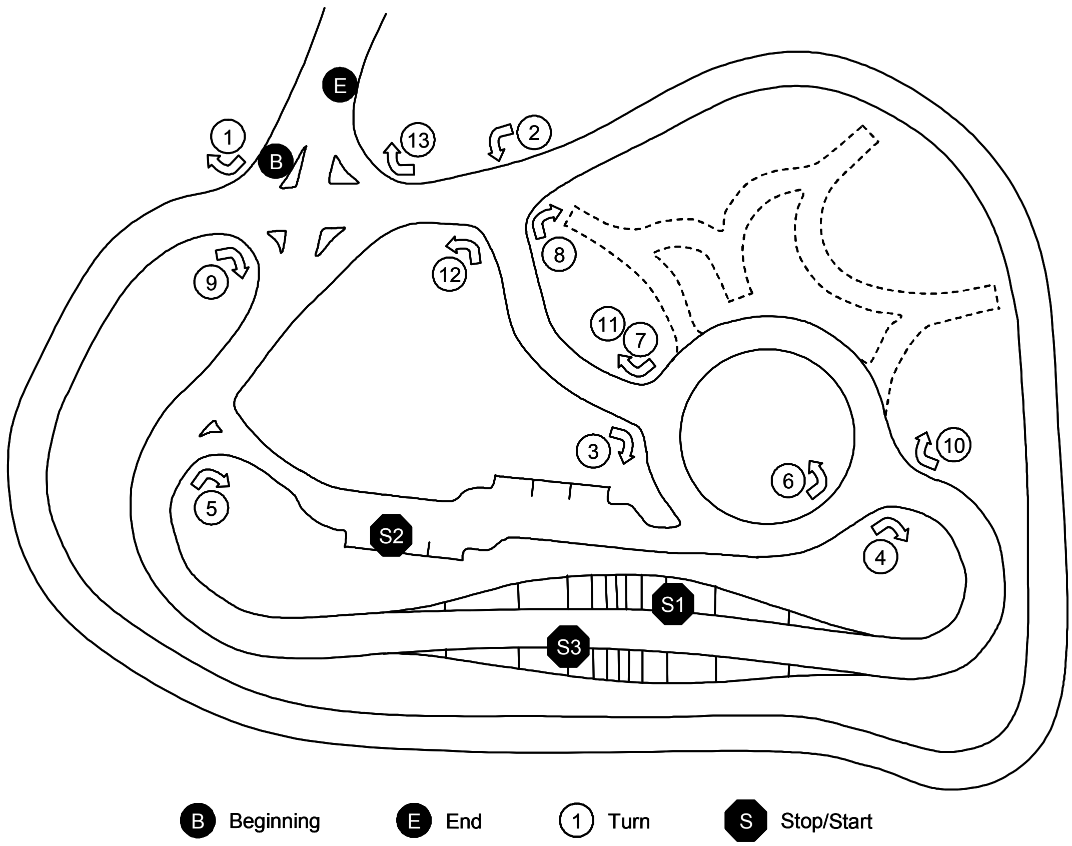

Figure 1. Track and driving task

Figure 2 exemplarily shows the time course of the rear wheel speeds during one drive.

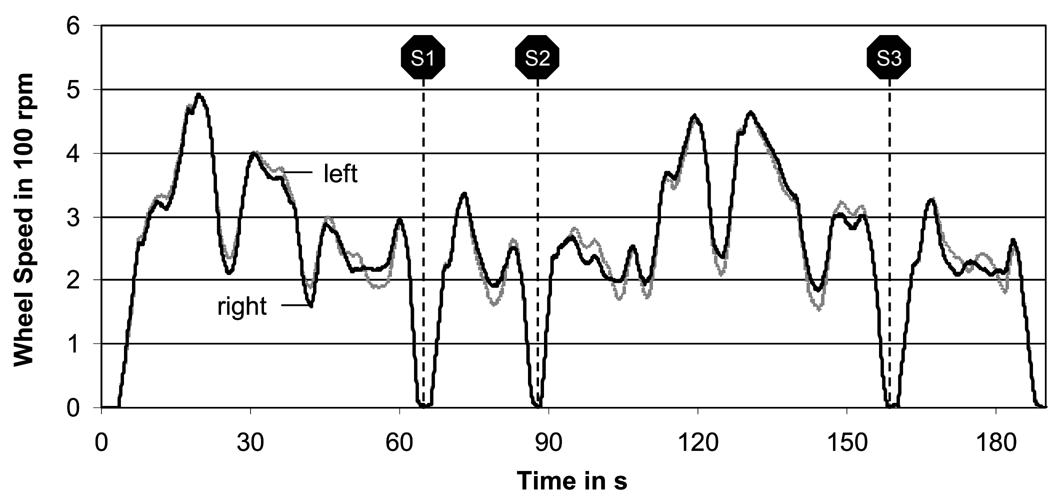

Figure 2. Time course of rear wheel speeds during one exemplary drive

The test vehicle is equipped with sensors allowing it to capture relevant vehicle data. During each drive the signals of the acceleration (measured triaxially at the driver's seat rail); rotary speeds of the engine, transmission and wheels; as well as the paths of the gas- and clutch pedal are captured. From this, the following 18 characteristic values describing the driving style are deduced:

Table 1: Characteristic values

(1)	Mean driving speed	(10)	Max. acceleration in driving direction
(2)	Max. driving speed	(11)	Max. accel. against driving direction
(3)	Mean gas pedal position	(12)	RMS value of longitudinal acceleration
(4)	Max. gas pedal position	(13)	Max. of lateral (curve) acceleration
(5)	RMS value of gas pedal gradient	(14)	RMS value of lateral (curve) acceleration
(6)	Max. gas pedal gradient at tip-in	(15)	Mean shifting time
(7)	Max. gas pedal gradient at let-off	(16)	Mean shifting engine speed
(8)	Mean engine speed	(17)	Max. shifting engine speed
(9)	Max. engine speed	(18)	No. of shifts

This way, 69 sets of data (23 drivers with 3 drives each) consisting of 18 characteristic values per drive are available. Of course some of the defined characteristic values are only valid for the chosen test track and hence for real use in industry, characteristic values which are independent of the track have to be defined. Nevertheless, the method of classification can be applied in the same manner as presented in this paper.

3 Method of Classification

3.1 Clustering Patterns by Self-Organizing Maps

Self-Organizing Map (SOM) is the most popular artificial neural network algorithm in the unsupervised learning category. It allows the display of multi-dimensional connections in a two-dimensional map according to their similarities [4]. This way, the SOM can be used for clustering patterns, where the patterns within a cluster have something in common i.e. they are judged as being similar. For example, we are given the task of grouping furniture according to use and appearance. All chair-like objects are placed in one group and all table-like objects in another. These groups are then inspected and the table-like group is split in order to separate the desks. The desk group is similar to the table-like group and so these two groups are placed close to one another away from the chair-like group. Cluster algorithms do a similar job for patterns of any data. These groups are referred to as clusters and the arrangement of clusters should reflect two properties: Patterns within a cluster should be similar in some way and clusters that are similar in some way should be close together [5]. This means, that on the map similar data lie close to each other while dissimilar data lie at a distance.

The SOM has a set of input units that correspond in number to the dimension of the training vectors (here 18) and output units that act as prototypes. The number of output units is chosen according to the number of desired groups. The input units only serve to distribute the input vector to the network's output units - the cluster units. The cluster units are arranged in a two-dimensional array in the manner shown in figure 3. During training all units can be considered as competing to be awarded the training vectors. When any training vector is presented, the distance to all cluster units is calculated and the unit that is closest to the training vector is denoted as the winning unit. The winning unit will then adapt its weights in a way that moves that cluster unit even closer to the training vector. Units within a pre-specified neighbourhood of the winning unit will update their weights as well. A unit is in the neighbourhood if it falls within a specified radius that is centred on the winning unit. The radius is reduced during training.

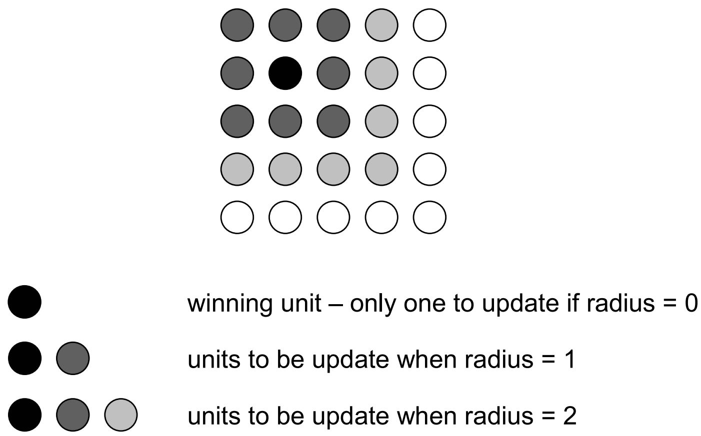

Figure 3. Cluster units

A learning rate determines the amount by which a cluster unit moves towards the training vector and like the radius, it is also decreased over time. At the end of the training the cluster units provide a summarised representation of the input pattern space. The cluster units act so as to map out the features of the input space.

3.2 Sammon’s Mapping

Sammon's mapping is also an iterative method with the objective of mapping points in multi-dimensional space usually into two dimensions [6]. The algorithm finds the locations in the target space so that as much as possible of the original structure of the measurement vectors in the multi-dimensional space is conserved. However, the numerical calculation is more time-consuming than the SOM algorithm, which can be a problem with a massive data set. On the other hand, it is able to represent the relative distances between vectors in a measurement space and is thus useful for determining the shape of clusters and the relative distances between them. It is therefore beneficial to combine these two algorithms. Sammon's mapping is thus applied to the stage where the SOM algorithm has already achieved a substantial data reduction by replacing the original data vectors with a smaller number of representative prototype vectors.

4 Results

As described in chapter 3.1, an SOM algorithm is applied to the data described in chapter 2.1. The number of neurons is first chosen to be three. Accordingly, figure 4 shows the result of classifying the 69 drives (3 cycles per 23 drivers) in three groups. For this a SOM with three neurons (cluster units) in a grid topology trained for 500 epochs is applied. In figure 4 each drive is characterised by a letter and a number. The letters A to W represents the 23 drivers, while the number 1 to 3 indicates the three drives of each driver. This way, A1 means the first drive of the first driver and W3 means the third drive of the last driver.

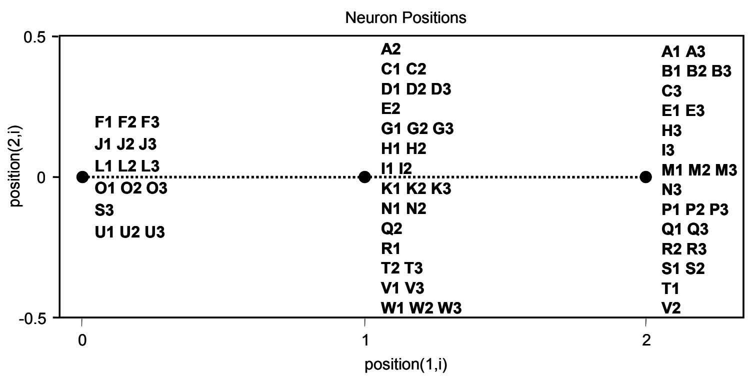

Figure 4. Trained SOM with 1 x 3 grid topology after 500 epochs

Using the map pictured in figure 4 you can detect that the three drives of driver F, J, L, O and U are assigned to the same neuron (0/0). This indicates that these drivers belong to one customer group according to their style of driving. Also all three drives of drivers B, M, and P are assigned to one other neuron (2/0). This means that those drivers also represent one customer group. The fact that the two neurons are far away from each other (as far as they can be in the used topology) shows that these drivers show a very different driving style relative to the ones mentioned before. The drivers D, G, K, and W show a driving style which is classified as being between those. Drives of some drivers, like A, C, E, H, I, N, Q, R, T, and V, are assigned to two neighboured neurons, which indicates that they did not show a constant driving style in the three drives. The drives of driver S are even assigned to neurons which lie far away from each other in the map (0/0) and (2/0), which implies that drive S3 is very different to S1 and S2, although driven by the same driver. In fact, the driving profiles shown in figure 6 prove this assumption.

To allow a more detailed classification, a network of 25 neurons in a grid arrangement of 5 x 5 neurons is used. The result is a map with 25 neurons where the 69 single drives are depicted. Figure 5 shows the classification of the single drives after 500 epochs. It can be seen that the drives assigned to the left neuron (0/0) in figure 4 are now all assigned to neurons located in the upper left area (0/2, 0/3, 0/4, 1/3, 1/4 and 2/4), while those assigned to the neuron (2/0) in figure 4, are now figure 5 assigned to neurons located in the lower right area in (1/0, 1/1, 2/0, 2/1, 3/0, 3/1, 3/2, 4/0,4/1, 4/2 and 4/3).

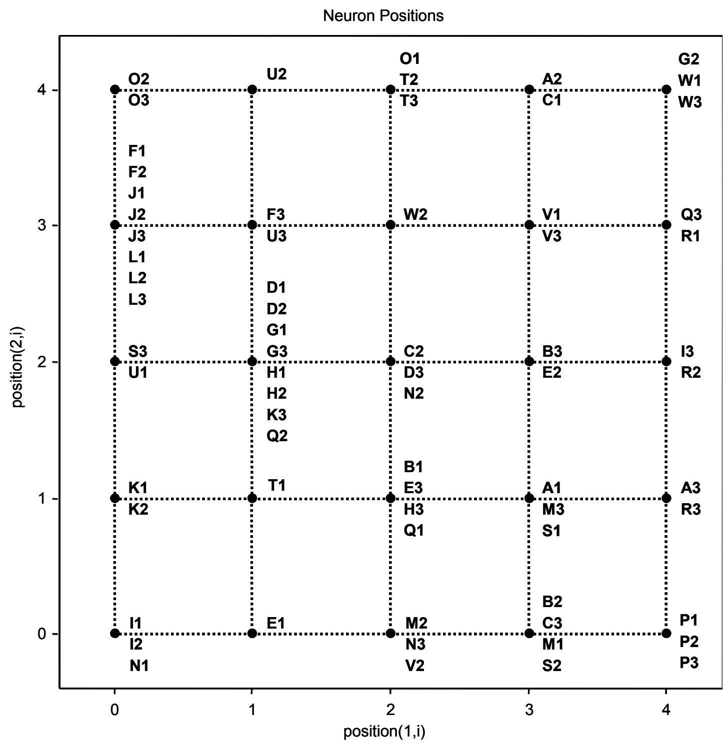

Figure 5. Trained SOM with 5 x 5 grid topology after 500 epochs

Also in this more detailed map, the three drives of some drivers can be found close to each other on the map while drives of some drivers are separated from each other. It can be found that driver P is an extreme and typical driver because all three drives are grouped at the neuron (4/0) and this neuron is at a corner position i.e. its distance to another corner neuron lying diametrically is extremely large. All three drives of drivers J and L are also grouped at one neuron (0/3). This indicates firstly, that the three drives of each driver are very similar to each other and secondly that the driving style of driver J and L is very similar. Because of the large distance between the two neurons by which these drives are classified, it can be further deduced that the driving style of driver P differs very much with respect to drivers J and L.

There are also drives of drivers which are more distributed in the map. While the first and the second drive of driver S are classified to a neuron in the neighbourhood (3/1) (3/0) of neuron (4/0) where driver P is typically classified to, it can be deduced that these drives (P1, P2, P3 and S1, S2) are all similar. However the third drive of driver S is classified to a neuron relatively far away from this area. This drive is classified to neuron (0/2), which is in the direct neighbourhood of neuron (0/3), where drivers J and L are represented. As already shown in figure 4 it seems that this third drive of driver S differs from his first and second drive and is more similar to the drives of drivers J and L. So driver S shows an inconsistent driving style.

Having information about some reference drivers enables new drivers to be classified and according to the type of drivers they are classified as, their demands can be derived. In the example discussed here, driver P is the driver showing the highest engine speeds and the highest acceleration values, indicating that this is a very sporty driver. Drivers J and L show very low engine speeds, indicating that their driving style is more comfort oriented.

As mentioned, it is striking that in the map shown in figure 5 all three drives of some individual drivers can be found at exactly the same neuron while the three drives of other drivers are distributed. By having a closer look at this data it can be found that these drives were quite different although one and the same driver drove them. The following diagrams verify this interpretation. The characteristic values here correspond to the ones named in table 1. The values are standardized to minimum and maximum occurring values of the 69 drives. Except for the shifting time (15), all values are transformed to 1 as maximum and 0 as minimum.

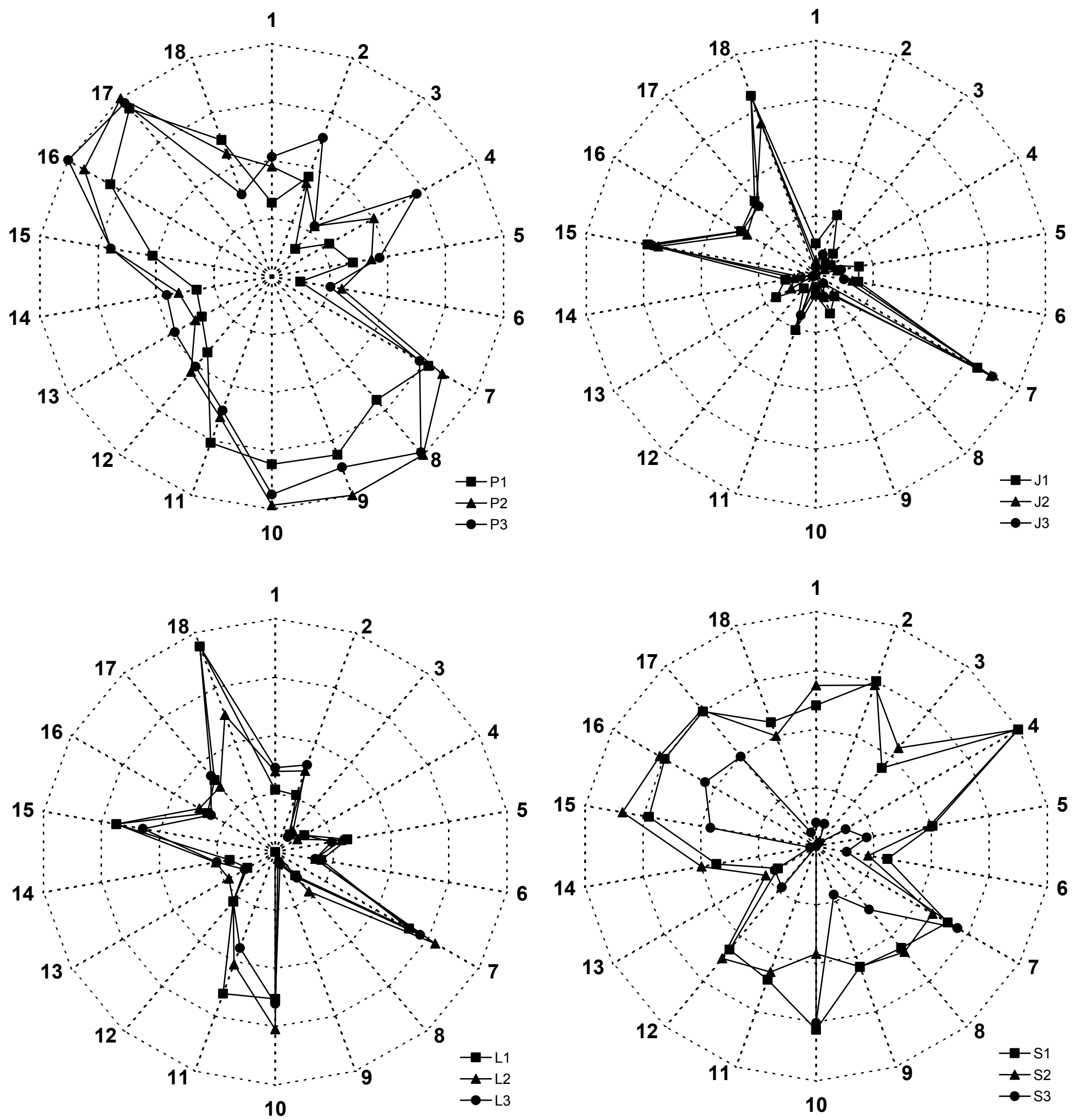

Figure 6. Driver profiles for three drives of drivers P, J, L, and S

The SOM algorithm already achieves the classification of the drives by assigning the drives to the neurons. The Sammons mapping algorithm shows the distances of the single neurons and therefore allows a further judgement.

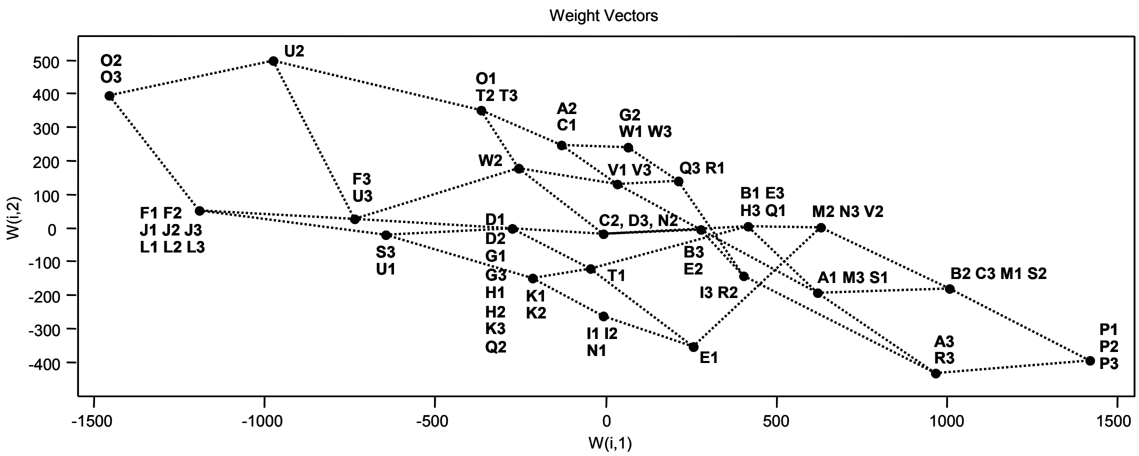

Figure 7. Sammons mapping of the SOM shown in figure 5

The drives here are assigned exactly to the same neurons as shown in figure 5. The distances of the single neurons in the two-dimensional map represent the distances of the characteristic values in the 18-dimensional space. This allows a closer consideration of the similarities. While neurons lying far away from each other represent different characteristics of the assigned drives, drives assigned to neurons lying close to each other are similar.

5 Conclusion

The results of this feasibility study show that SOM is a suitable tool to classify drivers according to their style of driving. A set of data of a driver can be found in the map according to the similarity to the other drivers so that it can be specified to which style of driving this data corresponds. Knowing the driving style and therefore the demands of some reference drivers, driving styles and consequently customer type of other drivers can be interpreted according to the distance to the reference drivers.

This method may help to determine the driving type and consequently the customer group a driver belongs to. A driver can either be classified during a test drive and therefore get the corresponding software tuning for this customer group or the vehicle may adapt its behaviour (e.g. shifting characteristics of an automated transmission [1]) accordingly during use.

References

[1] Nelles, O., “Lernfähige Fuzzy-basierte Fahrstrategie für automatische Getriebe”, Kapitel C.15 in Modellgestützte Steuerung, Regelung und Diagnose von Verbrennungsmotoren. Editor R. Isermann, Springer 2003, pp. 233-250.

[2] Albers, A.; Albrecht, M., “Driver Start-up Comfort Rating of Automated Clutch Systems – an Objectification Using Artificial Neural Networks”, FISITA 2004 World Automotive Congress, Barcelona, Spain, 23 - 27 May 2004. Book of Abstracts p. 325. Paper No. F2004U108.

[3] Schoeggl, P.; Ramschak, E.; Bogner, E.; Dank, M., “Driveability Design – Entwicklung eines kundenspezifischen Fahrzeugcharakters” ATZ Automobiltechnische Zeitschrift 103 (2001), Nr. 3, S. 186-195.

[4] Kohonen, T., “Self-Organizing Maps”, Springer Series in Information Sciences, Vol. 30, Springer, Berlin, Heidelberg, New York, 1995, 1997, 2001.

[5] Callan, R., “The Essence of Neural Networks”. Prentice Hall, 1999.

[6] Sammon Jr., J. W., "A nonlinear mapping for data structure analysis", in: IEEE Transactions on Computers, C-18(5), pp.401-409, May 1969.

o. Prof. Dr.-Ing. A. Albers
University of Karlsruhe (TH)
IPEK - Institute of Product Development
Kaiserstr. 12,
76131 Karlsruhe, Germany
Phone: +49 721 608 2371
Fax: +49 721 608 6051
E-mail: albers@ipek.uni-karlsruhe.de