INTERNATIONAL DESIGN CONFERENCE - DESIGN 2010
Dubrovnik - Croatia, May 17 - 20, 2010.

A DATA MINING METHOD FOR SELECTING THE SUITABLE EXISTING PRODUCT VARIANT AS A DEVELOPMENT BASE FOR A NEW ORDER

J. Feldhusen, A. Nagarajah and S. Schubert

Keywords: adaptable products, automotive supply industry, self organising maps

1. Indroduction and statement of problem

Today in the automotive supply industry a high percentage of products is developed through adaptive design [Tovey 2001]. The aims are to reduce the development time, to fullfill the high safety standards and to reduce the costs. The experience in the automotive supply industry shows, that these criteria are not having the same importance. Although it is expected from the suppliers that they achieve the industry's quality standards and that they have also to keep the development time. These two criteria represent only the necessary conditions for the placing of orders by OEMs (Original Equipment Manufacturer). The most decisive criterion is the demand to reduce product costs. The companies solve this problem by using adaptive design to realize the new customer's enquiry. Hereby they try to keep the necessary adaptations of the existing product as few as possible. Product changes mean for the company that the subsequent processes, for example the production process and assembly process also have to be changed and therefore the costs will increase. An optimal product design is only given, if the product, which is used as development base, fulfills the demand of the minimum changing costs. Many projects in the supply industry have shown that the developers do not take the suitable product as a development basis for a new order, but often select the last developed product.
For the solution of this problem the question arises whether the selection of a suitable product with the currently available tools can be realised.

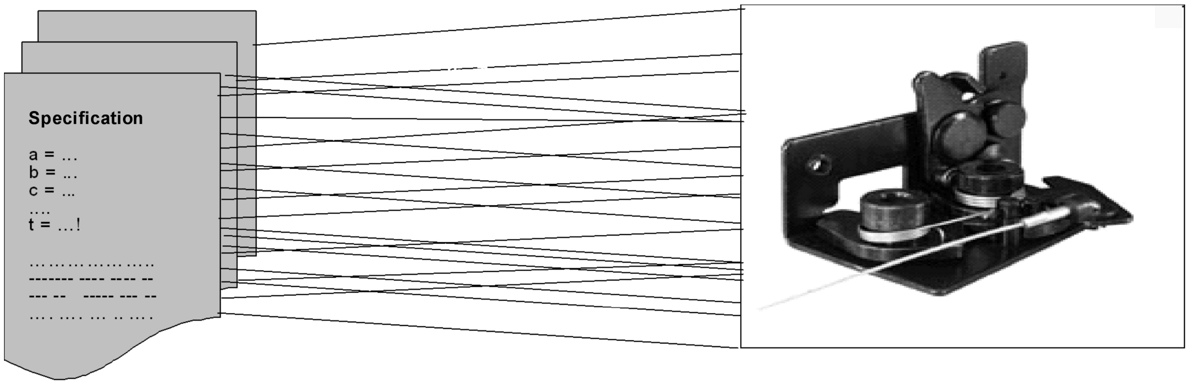

Figure 1. Allocation of requirements to the shape parameters

In order to answer this question many tools offer the opportunity to link the requirements with the product components (Fig. 1). In this approach the product components should be highlighted to the user, which is affected through the changed requirements. The disadvantage of this approach is the

increased time required for carrying out such assignments. Moreover the implicit requirements which are also realized through the design can not be identified and therefore not assigned to it.
A computer based solution, which automaticly select the suitable products would be promising for practical use.
In this research project the aim is not to link the requirements with the product components. The product selection is based on the comparison of specifications (Fig. 2).

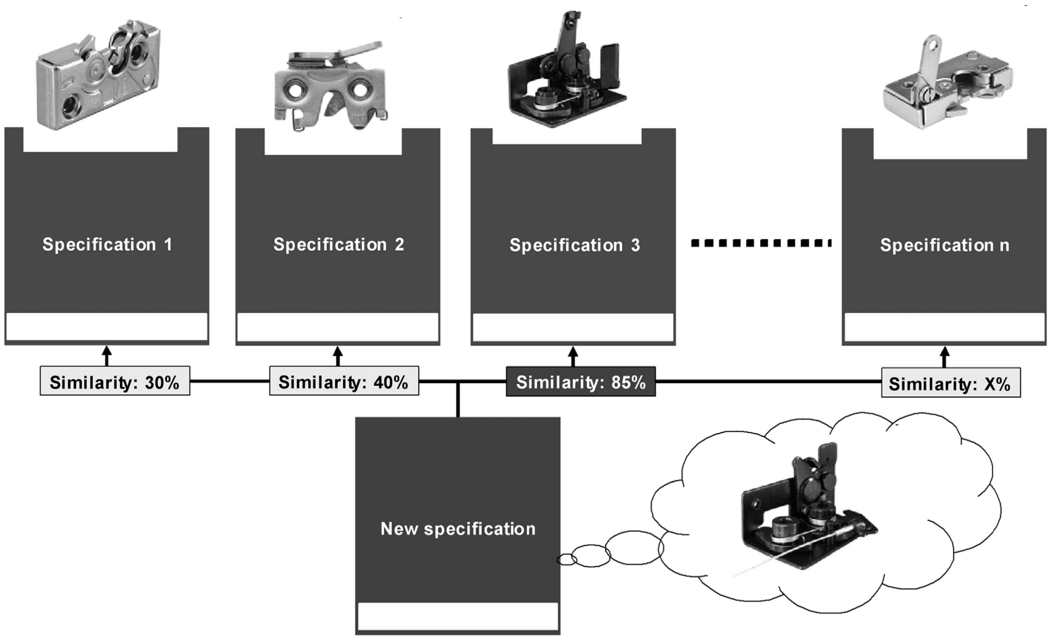

Figure 2. Approach

For a computer based comparison of specifications the following assumptions are made.
The change of requirements from the specification is decisive to the product changes. The demands of other stakeholders, such as lawgivers or standardisation institutes, have implications for all product variants and therefore are not adaequate for identifiying the changes.
The next assumption is, that the requirements discription and the number of requirements are kept constant and only the expression of the requirements changes. It is typical for reused design.
The assumptions describe on the one hand the constraints in the automotive supply industry accurately and on the other hand the prerequisite for an automatic comparison of requirement list using SOM is given. The requirements which are described by the specifications of the OEMs must be processed in the first step for the application of the SOM.

2. The approach of the research project

For a successful procedure of this research, it is necessary to divide the main goal into the following sub goals:

1. The first aim is to define the requirements from the specifications of the original equipment manufacturers (OEM) in textual elementary components and to quantify them in order to prepare them applicable for SOM.
2. The second aim is the development of a SOM that supports the developer in a meaningful way with the accomplishment of the adaptable product development for a new order, which should be suitable for the intended requirements.

2.1 Procedure for the creation of basic requirement components

In this project, specifications of different OEMs for bonnet locking systems [Fig. 3] have been analyzed. First all requirements have been extracted in original phrase and listed in a table. Specifications of bonnet locking systems contain a high number of requirements. In order to determine missing requirements a structuring would be reasonable. Therefore classes are found. The number of requirements per class is smaller and therefore missing requirements are easier to detect. Further the classification supports the engineer like a checklist not to miss a whole domain of requirements. The handling of the whole requirement list is also simplified

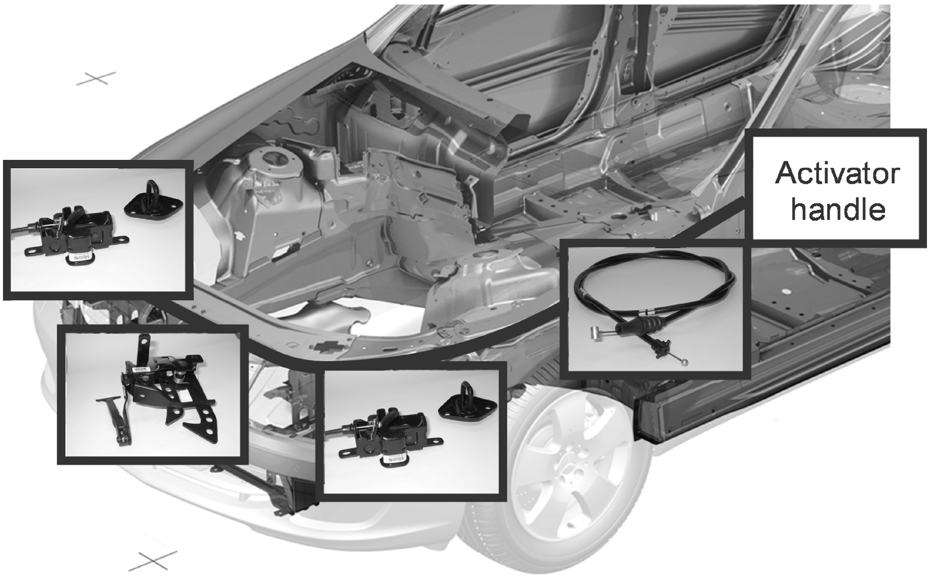

Figure 3. Bonnet locking system

A further step in the preparation process is the linguistic standardization. The German Automotive Association (VDA) provides a guideline which aims for a precise communication between OEM and supplier. Included in this guideline, recommendations for the setting up of the requirement list is included. According to this recommendation the requirements are transferred into unambiguous and clear textual components [Verband der Automobilindustrie 2006]. The user uses defined components to describe the requirements. In addition there is a “weak word-list” containing words not to be used. The requirement in direct quotation of the specification is e.g. “In case of a front impact, the bonnet must not be opened unintentionally.” Transformed according to the VDA standard the requirement will be “For a front impact (condition) the locking system (subject) have to (demand-word) keep the bonnet (object) closed (keep ... closed = action).

2.2 Procedure to quantify requirements

Basically requirements can be distinguished into qualitative and quantitative requirements. For quantification and therewith for preparation for SOM, it is reasonable not just to distinguish between qualitative and quantitative requirements, but to refine these groups [Jörg 2005] [Fig. 4]. Generally qualitative requirements are not processable for SOM, but they can be prepared to be processable. The need for a preparation of quantitative requirements is depending on the type of the quantitative requirement.

Quantitative requirements can either contain a numerical value (type 1) or a word (type 2), which defines the attribute precisely. An accepted tolerance or minimum force is normally expressed by a numerical value; a material is normally expressed by a word. The quantitative requirements described with a numerical value, can be directly processed by SOM. To process the verbally described quantitative requirements, the attribute has to be transferred into a numerical value, e.g. aluminium is set to “1”, while stainless steel is “12”. The associated number has to have no meaningful connection to the material. A definition table is set up and the assigned number is transferred into the requirement list [Fig. 5]. Thus the SOM is able to determine whether the materials of two requirement lists are identical.

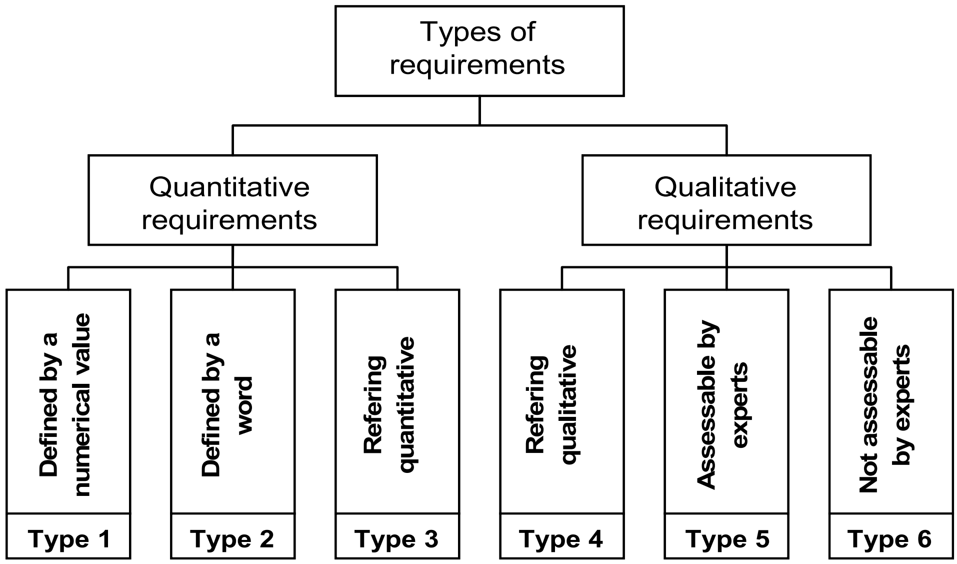

Figure 4. Types of requirements

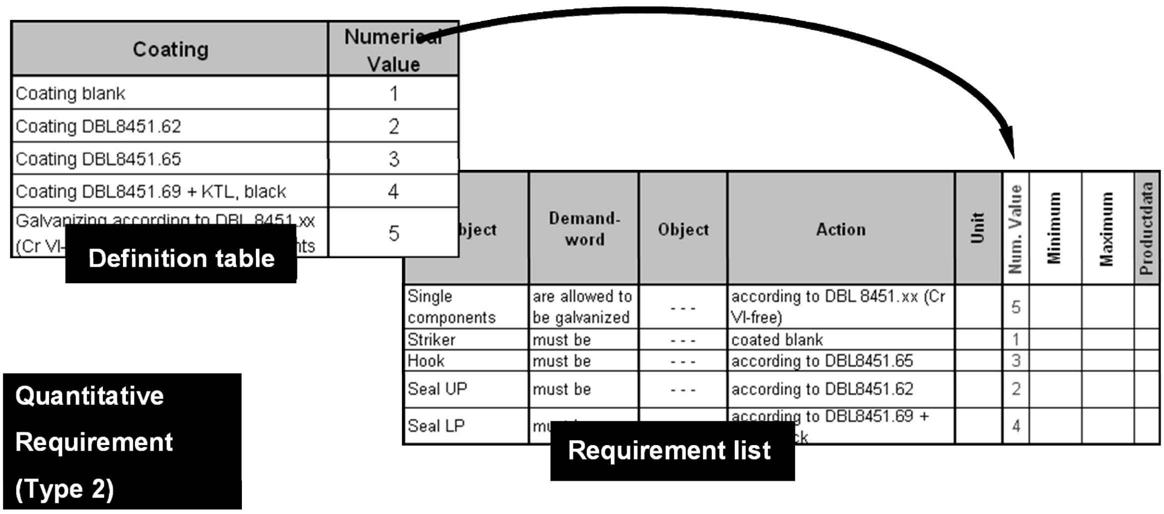

Figure 5. Assignment of numerical values to requirements (type 2)

There is existing one further group of quantitative requirements (type 3), containing requirements which refer to general documents like standards, guidelines or laws, e.g. “The bonnet locking system has to meet conditions according to a certain standard A”. For SOM it is just important that product related to a certain requirement list fulfills this standard A, while the exact conditions of the standard A can be neglected. For a requirement like this, it is sufficient just to store a Boolean value. “1” if the requirement, in this case the fulfillment of the Standard A, is considered, “0” if the requirement is not mentioned in the requirement list. There are also three different types of qualitative requirements. The first group is referring to the companies internal knowledge like a previous variant of the product (type 4), e.g. “The weight has to be reduced compared to the previous variant”. To interpret this requirement domain-specific knowledge is necessary. No exact numerical value can be determined for this type of requirements. Therefore a workshop with experts is executed. Experts are able to determine a range which is reasonable for a certain requirement. The thresholds of the range provide numerical values, which can be processed by SOM, but compared to the thresholds of a quantitative requirement they are not fixed [Fig. 6].

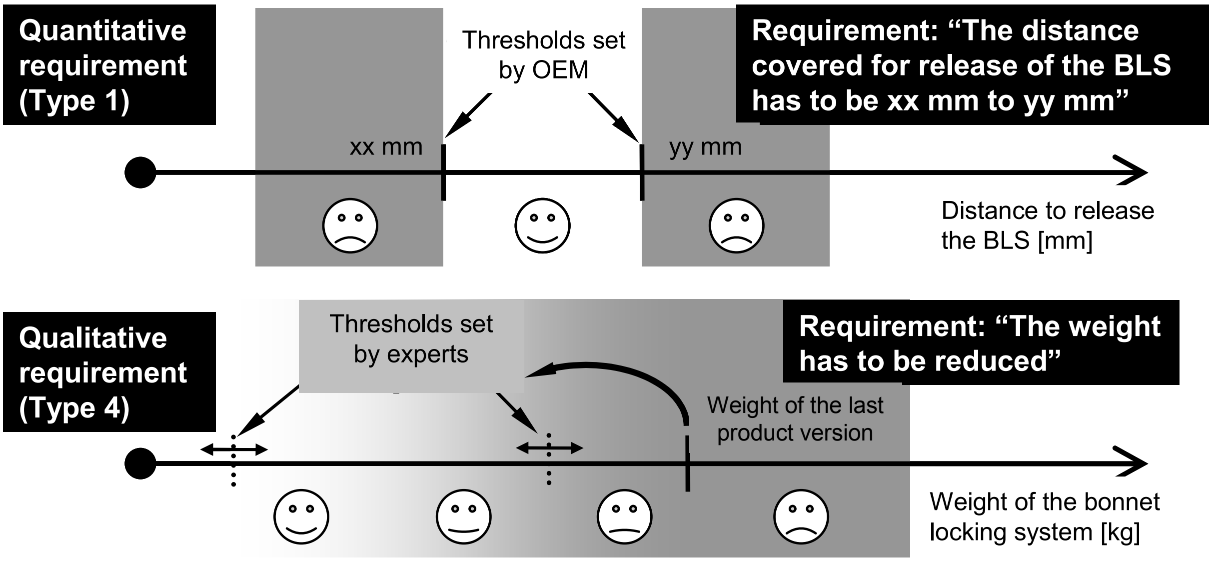

Figure 6. Assignment of numerical values to qualitative requirements (type 4)

The last two subgroups of qualitative requirements are also domain-specific, but not referring to any object. In both subgroups the requirements are phrased generally. The requirements belonging to the first subgroup (type 5) can be assessed by experts, e.g. "The sound level of the bonnet locking system during the closure has to be low." These requirements are also transferred into numerical values during the mentioned workshop.

For the remaining requirements, which constitute the last subgroup, an assessment by experts is not reasonable. A requirement demanding fail-safe can not be concretized by a numerical value (type 6). So the information, if a requirement of this subgroup is contained or disregarded in a requirement list is stored by a Boolean value, "1" if it is considered, otherwise "0".

3. SOM basics

SOM is a special data-mining-method, which is used to identify and furthermore to visualize complex numerical coherences.

The generic term 'data-mining' includes a great number of so called 'knowledge-generating' methods. These methods are mainly used to detect complex coherences in data, mainly in those cases where common statistical tools will not generate satisfying solutions [Otte 2004].

Based on historical data of a product, for example, real reasons for quality-defects of that product can be detected, even though these causes are hidden in the data [Otte 2004].

Several powerful methods and algorithms were developed, each with specific capabilities (random-forest, support-vector-machines, neural networks, etc.) [Zell 1997]. The correct use of these and other methods will finally result in an expert-system which can be used for classification, approximation and prediction [Otte 2004].

The SOM is a special architecture of artificial neural networks. Neural networks are information processing structures, which have been built up with the knowledge of the function of the human brain. They consist of a large number of highly interconnected elements, called neurons. Between the neurons, there are connection weights, which are established through the learning process. Through the adjustment of connection weights, such a network is able to in to recognize structures and to generalize them. Basically there are two kind of learning methods: supervised and non-supervised. In a supervised learning method a training record including the desired results is given. The network has to learn the relationship. For a non-supervised learning there is no desired output. Networks with this learning method are used for clustering. SOM belongs to the non-supervised learning systems. This makes them suitable for the mapping of the requirements.

3.1 Realisation of SOM for adaptable product development

An essential feature of a SOM is to map the similarity structure of a high-dimensional data space on a two dimensional chart [Kaski 1997]. This is an appropriate measure to reduce the evaluation effort by using a simple type of visualization, which can easily be analyzed.
The first step of the mapping process is to extract data vectors from complex requirement lists as shown in figure 7, in which the standardized requirements are listed. In order to get a consistent dataset, all requirement lists have to contain the same set of requirements, so the requirement vectors contain the same number and kind of elements. Thus, each variant of the already existing requirement lists is transformed into a standardized requirement vector.

								Req. List 1				Req. List 2			
Requirement Identificationnumber	Condition	Subject	Demand-word	Object	Action	Weighting	Unit	Num. Value	Minimum	Maximum	Productdata	Num. Value	Minimum	Maximum	Productdata
44	In the unlocked condition	Striker	must	opening height	have		mm		aa	bb			cc	dd	
45	- - -	Seal UP	must	according to	be coated			2				3			
46	- - -	The BLS	must	Standard No.123 (OEM)	meet the conditions according			0				1			
47	- - -	The BLS	must	the previous product version	be lighter than		kg		ee	ff			gg	hh	
48	- - -	The BLS	ought	- - -	to be disassembled quickly		s		ii	jj			kk	ll	
67	- - -	The acitvator handle	has	- - -	to be user-friendly			1				1			

Figure 7. Standardized requirement lists

3.2 Training

Next, a non-supervised learning procedure is carried out: A batch of the requirement vectors is applied to SOM.
In this case the "batch algorithm" is used, as a significantly faster computation of the SOM is performed in comparison to the incremental-learning SOM algorithm [Kohonen 2001].
The algorithm works in the following way: All Units of the grid are initialized with a model vector. Randomly a variant vector is picked out of the dataset and the distance between each model vector of the grid and the variant vector is calculated. The unit with the shortest distance to the variant vector is called best matching unit (bmu) and this variant vector will be placed on the map at this location. This operation is repeated while all variant vectors are picked out of the data set and are placed on the map. Thereafter the model vectors are adapted. Due to the inclusion of all already placed vectors at every mapping-iteration for a new variant vector, the similarity structure will be trained on a higher degree [Fig. 8].
To achieve a high reliability of the SOM, it has to be trained with a large number of requirement lists whereas a sufficient number of requirement lists is higher than the number of requirements. In case the number of available requirement lists is not adequate, the spectrum of requirements also has to be decreased. Currently, research is carried out to clarify if the application of modules and therefore a lower number of required training sets on SOM is sufficient. At this point of time a main problem is the non-consideration of the interactions between these modules. A measure to compensate this limitation has to be elaborated.

After the training process the SOM grid has a topology, consisting of nodes. In each node a requirement vector is stored, similar ones close to each other, different ones are located having a distance in between.

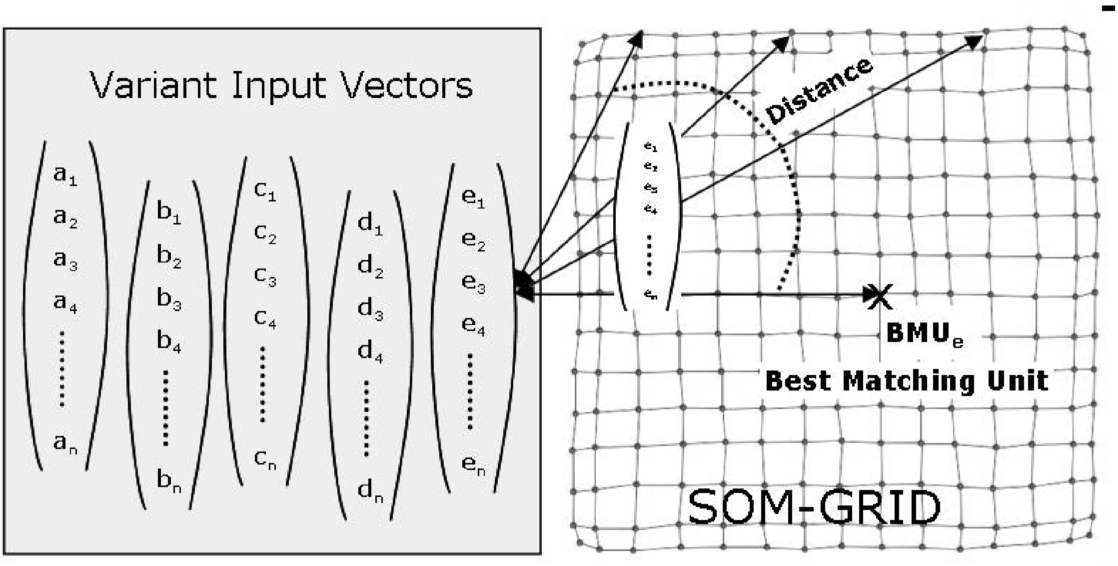

Figure 8. Training process

3.3 Interpretation of the Maps and selecting the suitable product variant

For the selection of the product variant, the visualisation of the frequency map is used. Since none of the existing product variant is identical to each other and since in addition the data set which has to be compared for the new application does not seem to be identical respectively very similar, it can be expected that there are no neurones representing two or more input vectors and therefore product variants.

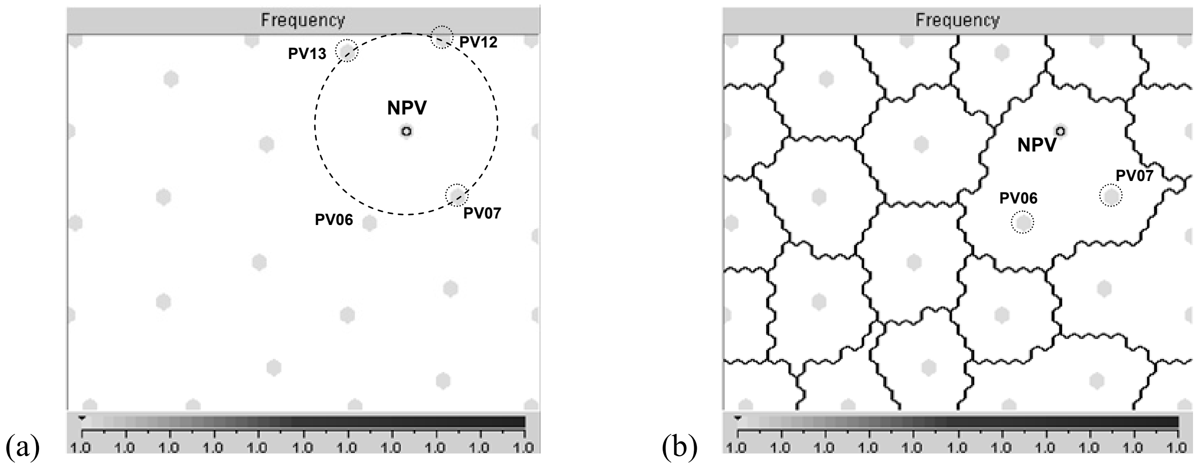

Figure 9. (a) frequency distribution, (b) frequency distribution with cluster

The colour bar in figure 9 (a) shows that no neurone represents more than one input vector. This leads to the conclusion that the illustration of the map at least does not deliver completely useless results.
For the further investigation the expert estimation for a chosen selection process is compared with the interpretation of the SOM. The outcome was very satisfying which guaranteed that the results of SOM could be used very well for the selected product variant.
For the product variant selection at first the next neighbours are identified (figure 9). For this only the distance to the neighbour solution is looked into. The analysis shows that the product variants (PV) 07, 12 and 13 have the shortest distance to the new product variant (NPV). As a next step the map was

clustered in order to check which product variant could be related to the same cluster of the new application. For the outcome see figure 1. In a workshop the results were evaluated and a ranking for the checking of the product variants was established. It was decided to weight the cluster result higher than the linear distances on the map. For a final validation of the result the requirement attributes, which showed to be the most important factors, were visualised, in order to gain conclusions for the final selection of the product variants. This procedure is to be explained with the help of the three attributes fracture strength, absolute weight and overall width. For the visualisation of these attributes see figure 10.

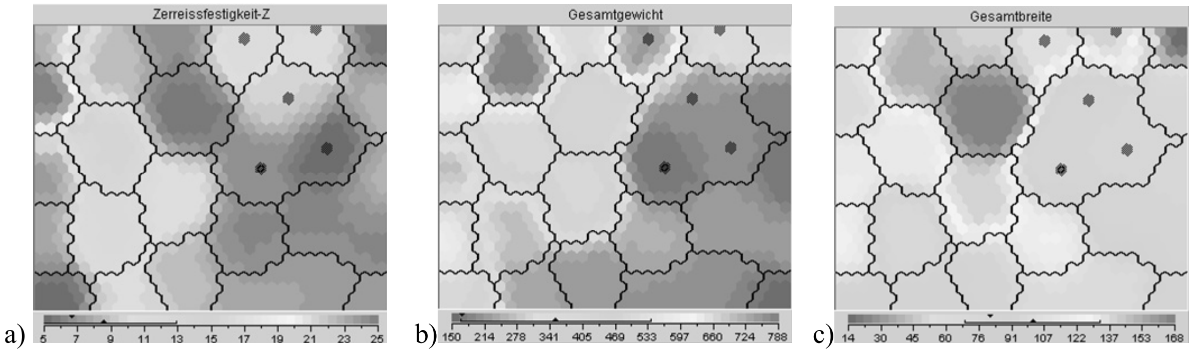

Figure 10. (a) fracture strength in z-direction, (b) absolute weight, (c) overall width

The visualisation of the attribute "fracture strength-z" shows that the product variants 12 and 13 result in very similar values as the ones required for the new application. Therefore they must have preference. The comparison with the attribute "absolute weight" shows however, that the weight of variant PV13 is higher than the one of the new product variant. The variants PV06 and PV07 do have a similar respectively lower weight as required.
With regard to the overall width, the product variants PV06 and PV07 also get preference. If it is not possible to build in a bigger bonnet locking system because of the cross-section and if the requested higher fracture strength is a given factor and if no other factors are regarded as equally important, none of the firstly identified product variants can be used. In this case the product variants would be selected which would meet those requirements. In this case the product variants in the upper-left area would be looked into.
After an analysis of the most important attributes, in the current project PV06 was regarded as the most successful variant followed by the variants PV07, PV12 and PV13. For the ranking additionally the test values of the product variant were looked into. The result was that in most cases an overachievement of the requirement by the product variant was accomplished. Using this knowledge a final decision was made.
The procedure shows that it is not possible to get an unambiguous and automatic decision. Because of this a human being as a decision-maker has to evaluate all available pieces of information and based on that make the decision.
The validation if the selected product variant definitively causes the smallest modification effort cannot be made absolutely. This implies that in reality the development of all the existing product variants was done to get a clear comparison. Because of this, the question arises if the product variants identified with SOM causes less work for the execution of the necessary changes as the last product variant. The experts could answer that question unanimously with a yes.

4. Summary

The aim of this research project is to apply a SOM based system to shorten the time for conceptual design of an adaptable product development. Therefore an SOM-based system can be applied on requirement lists of existing products to identify solutions that are suitable as a development base for a new product order. The engineer is able to select one of these proposed solutions as a development base. This can be adapted with a minimal effort for the new product order. This increases the customer satisfaction, because the new product is based on an approved product.

To apply specifications provided by OEMs to SOM, the requirements have to be extracted and edited. This process has several steps. In this paper an emphasis is put on the quantification of the requirements. First the requirements are distinguished into six different types, which are all either quantitative or qualitative. In this context a description by an exact word is also considered as quantitative. Examples are the assignment of a material or a standard, which has to be fulfilled. In case of a description by word the quantitative requirements need a preparation. Furthermore all qualitative requirements need to be prepared. Examples are shown for each type.
Currently the quantification of different types of requirements is made in workshops by experts. This is time-consuming and expensive. In future the aim is to use fuzzy-sets. Therefore the requirements have to be prepared further. The requirement demanding "the bonnet locking system to be user-friendly" is easily assessable for experts. For a fuzzy-set this requirement will be split into elementary requirements, e.g. defining a maximum hand-force needed and optimal shape of the grip.

Acknowledgements

The authors like to thank the German Federal Ministry of Education and Research for aiding this research project.

References

Ahrens, G., „Das Erfassen und Handhaben von Produktanforderungen", dissertation, TU Berlin, 2000.

Jörg, M., „Ein Beitrag zur ganzheitlichen Erfassung und Integration von Produktanforderungen mit Hilfe linguistischer Methoden", dissertation, Karlsruhe Germany, 2005.

Kaski S., "Data exploration using self-organizing maps", Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering, No. 82., 1997, pp 1—57.

Kohonen T., "Self-organizing Maps", Springer Berlin Germany, 2001.

Otte, R., Otte, V., Kaiser, V. „Data Mining für die industrielle Praxis", Hanser Fachbuch München Germany, 2004.

Tovey, M., "Design of Lean Implementation Based on Cost Analysis, Journal of Engineering Design, 2001, Verband der Automobilindustrie, „Automotive VDA-Standardvorlage Komponentenlastenheft", Henrich Druck + Medien Frankfurt Germany, 2006.

Zell A., „Simulation neuronaler Netze", Oldenbourg Verlag München Germany, 1997.

Arun Nagarajah, M.Sc.
Research Assistant
Chair and Institute for Engineering Desing (ikt)/ RWTH Aachen University
Steinbachstrasse 54B, 52074 Aachen, Germany
Telephone: +49 (0)241 80-27346
Telefax: +49 (0)241 80-22286
Email: nagarajah@ikt.rwth-aachen.de