INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED'07
28 - 31 AUGUST 2007, CITE DES SCIENCES ET DE L'INDUSTRIE, PARIS, FRANCE

UNSUPERVISED BAYESIAN MODELS TO CARRY OUT PERCEPTUAL EVALUATION IN A DESIGN PROCESS

Ben Ahmed Walid[1] and Bernard Yannou[2]
[1] Power Train Division, Renault, France
[2] Ecole Centrale Paris, Laboratoire Génie Industriel, France

ABSTRACT

This paper aims at developing unsupervised bayesian kansei models in order to help designers specifying new technical solutions as well as assessing their impact on customer's perception. In other words, these models provide designers with the three following use scenarios. The first one, denoted hereafter direct/analysis scenario, allows assessing the impact of a given decision related to the technical and/or functional design parameters on the final customer perception. The second one, denoted hereafter the inverse/synthesis scenario, allows determining the optimal levels of design technical attributes that can meet a given customer's expected perception. The third one, called henceforth compound scenario, is a combination of the two previous scenarios. To develop these kansei models, we use unsupervised bayesian learning of probabilistic relations between perceptual attributes and technical characteristics of the product under development. As a case study, we propose to predict the customer perception of car dashboards by assuming their technical characteristics. We also propose to help designers defining optimal dashboard technical characteristics to satisfy given customer expectations.

Keywords: Decision-making, kansei engineering, unsupervised bayesian learning, design analysis, design synthesis, perceptual evaluation, emotional design

1. INTRODUCTION

A design process can be seen as an iterative and complex process guided by a final and ultimate objective, which is to make the developed product fitting the customer aspirations.
Hence, predicting customers' satisfaction level when we develop a new product is fundamental.
Designers need tools to help them understanding customers' needs and thereby predicting their appreciation level of a new product. During a design process, designers are facing two main issues that we explain through the two following questions. The first one is a question of simulation of performances (deductive reasoning), denoted hereafter direct/analysis: *what is the impact of a given decision related to the design parameters (i.e. technical and/or functional parameters) on the final customer perception?* The second is more narrowly related to the design goal itself, starting from the need (expected performances or perceived impacts) and inducing a satisfactory design solution, denoted hereafter the inverse/synthesis scenario: *what are the optimal levels of design technical attributes that can meet a given expected perception by a customer?*
Certainly, the two questions are connected and can be handled if the correlation between the design attribute levels and the customer perceived impact levels is known. This correlation can be learnt through experience. In other words, experienced engineers may be capable to convert a subjective expected perception expressed by customers in concrete technical design attribute levels. However, their suggestions remain subjective and generally difficult to explain. The second way to rationalize this correlation is to construct an analytic predictive model. In this paper, we propose to construct such a model using data mining learning techniques and more specifically Bayesian Networks (BNs) learning.
As a case study, we propose to use customers' evaluations of perceptual attribute levels of existing car dashboards in order to predict the customer perception of new car dashboards by assuming their

technical attributes or main design parameters (*design analysis situation*). Conversely, we also propose to advise designers on values of dashboard physical attributes so as to better ensure meeting given customer perceptions on these dashboards (*design synthesis*). Henceforth, we call *technical* (or *physical*) attribute any parameter that deals with the design characteristics (shape type, number of colors, subsystem location, average curvature radius of shapes, etc.). In other words, technical attributes are parameters that a designer can tune to change and ultimately optimize his design. We call *perceptual* (or *subjective*) attributes all variables that deal with the customer evaluation of the design (comfort, novelty, cultural values, etc.). In the following, perceptual attributes are assessed by a set of customer in an non-hedonistic way, i.e. perceptual attributes are assessed on an absolute scale independently of personal preferences (as much as possible). This scale is automatically built through the use of pairwise comparison techniques which are widely presented in [28][34][35] and which are out of the scope of this paper.

Various scientific approaches relating physical attributes and perceptual attributes have been gathered by Japanese researchers under the name *Kansei* Engineering. Ours may be considered as such. This research aims at exploring the structure of emotions by building a database on consumer feelings. From the consumer's point of view, a *forward mapping process* from perceptual words to design elements is established, and from the designer's point of view, a *backward process* from drawings to perceptual words is proposed [24][25]. These two processes exactly correspond to, respectively, the aforementioned synthesis and analysis use scenarios. Some methods of category classification methods based on the Semantic Differential Method (SDM) have been used for the design of car interiors [15]. More sophisticated methods based on genetic algorithms, neural networks or fuzzy logic have been applied to ensure mappings between perceptual words and design elements. However, these systems are often opaque for designers and consumers. A semantic transformation method for automotive form design is proposed in [12], allowing an automatic regulation of the shape with respect to the required image required. The aim of this paper is to study the use of a still rarely used technique in design, namely the Bayesian Network. Its main advantages compared to other techniques is the fact that it is not opaque to the designer and that it can be easily used in compound analysis/synthesis scenarios, complying with actual design problems.

After presenting Bayesian Networks in section 2, the collect protocol for perceptual data is briefly evoked in section 3. Section 4 describes the model building and section 5 presents the model use scenarios.

2. THEORETICAL BACKGROUND: UNSUPERVISED LEARNING OF BAYESIAN NETWORKS

2.1. Bayesian networks

Consider "A"= "use comfort"={weak, good} as one of the perceptual attributes and "B"="control button shape"={circular, square) as one of the physical attributes characterizing the dashboards. The double functionality of our model (*direct* and *inverse*) is possible thanks to Bayesian Networks (BNs), which are based on the Bayes theorem: $\mathrm{P(A/B)} = \dfrac{\mathrm{P(B/A).P(A)}}{\mathrm{P(B)}}$.

BNs are directed acyclic graphs used to represent uncertain knowledge in Artificial Intelligence [14]. A BN is defined as a couple: *G=(S, P)*, where:

- *S=(N, A)* represents the structure (i.e. the graph);
  - *"N"* is a set of *nodes*. Each node represents a discrete variable X having a finite number of mutually exclusive states (modalities). In our case study, X may be a perceptual attribute as well as a technical attribute;
  - *"A"* is a set of *edges*; the relation "N_1 is a parent of N_2" is represented by an edge linking N_1 to N_2. In our case study, an edge may be interpreted as a causal relation.
- *P* represents a set of *probability distributions* that are associated to each node. When a node is a *root node* (i.e. it does not have a parent), *P* corresponds to the probability distribution over the node states. When a node is not a root node, i.e. when it has some parent nodes, *P* corresponds to a *conditional probability distribution* that quantifies the probabilistic dependency between that node and its parents. It is represented by a *Conditional Probability Tables* (*CPT*).

Because a Bayesian network is a complete model for the attributes and their relationships, it can be used to answer probabilistic queries about them. For example, the network can be used to find out updated knowledge of the state of a subset of attributes when other attributes (the evidence attributes) are observed. This process of computing of the posterior distribution of attributes given evidence is called *probabilistic inference*. Various inference algorithms can be used to compute marginal probabilities for each unobserved node given information on the states of a set of observed nodes. The most classical one relies on the use of a junction tree (see [14], pp. 76). Inference in BN [13] allows then taking any state attribute observation (an event) into account so as to update the probabilities of the other attributes. Without any event observation, the computation is based on a priori probabilities. When observations are given, this knowledge is integrated into the network and all the probabilities are updated accordingly.
The use of Bayesian networks in industry is continually growing up especially in risk management fields, in marketing, and generally in domains where there is uncertainty and thereby a need to predict a complex behavior such as in a decision making process.
Many commercial tools provide user with HMI allowing graphical representation of Bayesian networks to model expert knowledge. They are also providing user with a panel of supervised and unsupervised learning algorithms to extract automatically knowledge from databases.

2.2. Unsupervised learning of Bayesian networks

An unsupervised learning of a bayesian network consists in an unsupervised learning of the whole set of probabilistic relationships existing between attributes within the data. In other words, there is no target attribute to guide the learning task.
There are several techniques for learning Bayesian networks from data (see [31][27][4][16][26] for an overview). Two main techniques may be distinguished. The first one is a constraint-based method and the other is a score-based method. The constraint-based method employs statistical tests on the data set for deciding the existence of probabilistic relation between attributes (i.e. edges in the Bayesian network). The accuracy of the constraint-based method strongly depends on the size of the data set. A huge data set can provide a more accurate statistical independence test. In this paper, we use a small data set (cf. section 3), so we adopt the score-based approach.
Methods to learn Bayesian networks from data often consist of two components. The first component is a score which is used to evaluate how well the learned model fits the data. The second component is a learning algorithm (i.e. a structure search strategy) which is used to identify one or more network structures with high scores by searching through the space of possible network structures.

1. **The score function**: there have been some scores proposed for learning Bayesian networks. These includes the AIC score [1], the BIC score [32], the K2 score [7], the BDe score [11], the GU score [22] and the MDL (Minimum description length) score [19][29][3][33]. Remco R. Bouckaert [3] indicates that the performance of learning Bayesian network structure using MDL score is slightly better than the performance of the other scores. In this paper, we use the MDL, which is an information-theoretic criterion that favors models which provide the shortest description of the training data. This description includes both the description of the model and the description of the data given the model. Formally, given a Bayesian network $BN = (S;P)$, and a training data set D, the MDL score of BN is defined as $Score_{MDL}(BN;D) = MDL(BN) + MDL(D\backslash BN)$. Without going into details of MDL derivation, we just note here that the first term of the MDL score is the description length of a Bayesian network, i.e. the number of bits required to encode the network parameters, while the second term is the negative log likelihood of the model BN given data D, which gives the number of bits needed to describe D when using BN.
2. **Structure searching strategy**: the number of possible network structures (*NS*) grows exponentially with the number of nodes (*n*) (cf. formula 2 [30][9]):

$$NS(n) = \sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} 2^{i(n-1)} NS(n-i) \text{ for } n>1 \tag{2}$$

For example NS(5) = 29281 and NS(10) = 4.2×10^{18} [30]. Therefore, most algorithms for learning Bayesian networks are heuristic search algorithms. Some examples are the K2 algorithm [7], the Structure EM algorithm [8] and the Greedy Equivalent Search (GES) algorithm [21][5][6].

3. DATA PREPARATION

The data colleting protocol has been described in [35] and already experimented on another case study in [28][34]. 10 dashboards (AUDI A2, CITROEN C2, FIAT Idea, LANCIA Ypsilon, NISSAN Micra, PEUGEOT 206, RENAULT Clio, RENAULT Modus, TOYOTA Yaris, VW Polo} are evaluated by 11 customers. (cf. Figure 1).

Figure 1. The 10 dashboards evaluated by customers

We defined a set of 8 technical attributes characterizing the dashboards with corresponding modalities (two at least but the number may increase): the "*Speedometer dial position"={behind steering wheel, at the center of the dashboard}*, "*Display lay-out"={Analogue, Digital}*, "*Air conditioner control"={Button, Other}*, "*Air vent shape"={Rounded, Square}*, "*Dashboard color"={Single color, Two colors}*, "*Aerator shape"={Rounded, Square}*, "*Arrangement space"={Many, Few}* and "*Style lay-out"={Curved lines, Straight lines}*. The characterization of the 10 dashboards according to the technical attributes is objective and do not depend on the preference of customers. It is presented in table 1.

Table 1. The technical characterization of the 10 dashboards

Dashboards	Speedometer Dial position	Display lay-out	Air conditioner control	Air vent shape	Dashboard color	Aerator Shape	Arrangement space	Style lay-out
AUDI A2	behind steering wheel	analogue	button	square	signle colour	square	many	Straight lines
CITROEN C2	behind steering wheel	digital	other	rounded	signle colour	rounded	few	curved lines
FIAT Idea	at the center	analogue	other	square	two colours	square	many	Straight lines
LANCIA Ypsilon	at the center	analogue	other	square	two colours	square	many	curved lines
NISSAN Micra	behind steering wheel	analogue	button	rounded	signle colour	rounded	few	Straight lines
fewGEOT 206	behind steering wheel	analogue	other	rounded	two colours	rounded	few	curved lines
RENAULT Clio	behind steering wheel	analogue	other	square	signle colour	square	few	Straight lines
RENAULT Modus	at the center	digital	button	rounded	two colours	rounded	many	curved lines
TOYOTA Yaris	at the center	digital	other	rounded	signle colour	rounded	many	curved lines
VW Polo	behind steering wheel	analogue	other	square	signle colour	square	few	Straight lines

We also defined a set of 11 perceptual attributes, which describe the customer assessing of the *"Space organization"*, "*Control button comprehensibility"*, "*Aerator lay-out"*, "*Arrangement space"*, "*Comfort"*, "*Simplicity"*, "*Sportive lay-out"*, "*Masculinity lay-out"*, "*Quality*", "*Novelty"* and "*Harmony"* (see [10] for details on attributes). The customer evaluations of the dashboard perceptual attribute levels is made in qualitatively pairwise comparing the 10 dashboards under each of the 11 perceptual attributes (see [20] for mathematical details). It leads to 11 normalized score vectors. The advantage of this method is that the value scale is automatically built thanks to the pairwise comparison mechanism without the need to define a specific metrics (for instance, a score of 0.1 for the "*Masculinity lay-out"* means much more feminine than a score of 0.3). Next, each normalized score vector (the scores sum is 1) is transformed to fit into a standard scale of [0, 20]. Finally, continuous attribute levels are projected into discrete categories: *[0, 5]=Very low*, *[6, 10]=Low*, *[10, 14]=Medium*, *[15, 17]=High*, *[18, 20]=Very high*.

As 11 customers have participated to this study, a *110 x 19 matrix* is then constructed: *rows=10 dashboards x 11 customers*, *columns=8 technical attributes & 11 perceptual attributes*.

4. UNSUPERVIZED KANSEI MODEL CONSTRUCTION

4.1. Model learning

An unsupervised Kansei model construction consists in an unsupervised learning of the whole set of probabilistic relationships existing within the data and especially between perceptual attributes and technical attributes. We used the SopLEQ technique [17][18][23], which is a quick search based on the whole dataset. It uses a cost function based on the MDL score and a structure search method based on equivalent model classes. The description of SopLEQ as well as its advantages compared to the other search techniques is pointed out in [17][18][23]. Figure 2 represents the constructed BN without any user modifications.

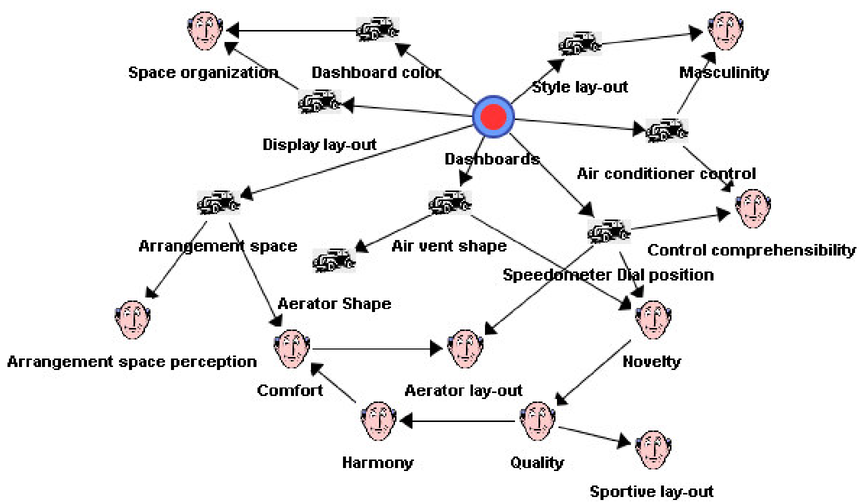

Figure 2 *Unsupervised learning to identify probabilistic relationships within the data (i.e. between dashboard physical - car icon - and perceptual - face icon - attributes)*

Edges in this bayesian network can be interpreted as causal relationships. For instance, according to Figure 2, the subjective attribute "*Novelty*" depends on the two physical attributes "*Air vent Shape*" and "*Speedometer position*". Each relation (i.e. edge) is expressed through a *conditional probability table*, which is automatically computed. For example, the relation between "*Novelty*", "*Air vent Shape*" and "*Speedometer position*" is represented through Table 2.

Table 2 *Conditional probabilities representing the causal relation between "Air vent Shape", "Speedometer position" and "Novelty". According to this table : P(novelty= very low/Speedometer dial position=at the center & Air vent shape=rounded)=13,6%*

Speedometer dial position	Aerator shape	Novelty				
		Very low	Low	Medium	High	Very High
At the centre	Rounded	13.6	36.4	31.8	9.1	9.1
	Square	27.3	36.4	27.3	0.0	9.1
Behind steering wheel	Rounded	24.2	60.6	9.1	6.1	0.0
	Square	75.8	24.2	0.0	0.0	0.0

We notice here that the constructed model (Figure 2) allows identifying three types of relationships:

- **Relationships within technical attributes**. For example, "*Air vent shape*" has a direct impact on the "*Aerator shape*".
- **Relationships within perceptual attributes**. For example, "*harmony*" perception has a direct impact on "*comfort*" perception.
- **Relationships between technical and perceptual attributes**. For example, the two physical attributes *"Air vent Shape*" and "*Speedometer position*" have an impact on the *"Novelty"* perception.

4.2. Model evaluation

The numerical accuracy of an unsupervised model is written as: *P(D|BN)* meaning the probability of the data given the Bayesian Network (BN). It is more common to use the log likelihood of BN given D provided by formula (3):

$$\log_2(P(D|BN)) = \sum_{i=1}^{n}\sum_{j=1}^{Qi}\sum_{k=1}^{Ri} -(N_{ijk}/N)*\log(N_{ijk}/N_{ij}) \qquad (3)$$

Where:

- U is the set of attributes $\{X_1, X_2, \ldots X_n\}$, $n>=1$,
- n is the total number of attributes,
- X_i is an attribute which takes values from $\{X_{i1}, X_{i2} \ldots\}$,
- R_i is the total number of values of X_i,
- D is the data set over U,
- BN is a Bayesian network structure over U,
- N is a number of instances in D,
- P_i is the set of parents of X_i in Bs
- W_{ij} The jth instantiation of P_i
- Q_i is the total number of value combinations of P_i in BN,
- N_{ijk} is the number of cases in D in which $X_i=X_{ik}$ and $P_i=W_{ij}$
- $N_{ij} = \sum_{k=1}^{Ri} Nijk$

The log likelihood measures how many bits are needed to describe D based on the probability distribution P. It also has a statistical interpretation: the higher the log likelihood, the closer BN is to model the probability distribution in the data D.

To assess our unsupervised Kansei model, we split the data into two subgroups, a training set ($D_{training} = 80\%$ of the data set) and a testing set ($D_{test} = 20\%$ of the data set). We selected the testing set to be the most representative as possible with respect to the original data. In a second step, we learned a BN using the training set. Then, we compared the log likelihood of BN given $D_{training}$ and the log likelihood of BN given D_{test}. The model is acceptable if the two values are close, i.e. $log(P(D_{training}/BN)) \infty log(P(D_{test}/BN))$. This means that the model is able to represent unseen data (i.e. D_{test}).
$Log(P(D_{training}/BN))=-21.75$ and $log(P(D_{test}/BN))=-24.38$. The model fits to the testing set as well as to the training set.

5. MODEL USE SCENARIOS

As we pointed out in the introduction, three main use scenarios of our BN model are possible. We present each of them with explaining examples in the following sub-sections.

5.1. Direct/analysis scenario

A design process can be perceived as a decision process during which the designer tunes a set of technical attributes in order to satisfy a required predefined performance. Before carrying out any decision (i.e. technical choice), the designer analyzes the impact of his choice on the others attributes (perceptual as well as physical). The **direct/analysis scenario** allows answering the question "*what is the probable impact of the choice related to physical attributes on the other design attributes and especially on the perceptual attributes*"? In a sense, this scenario consists in observing a technical attribute and analyzing its impact on the other attributes. It may typically be used to compare different technical solutions. In a direct use scenario, the input is a technical (or physical) attribute. Let us take the *speedometer dial position* as an example.

5.1.1. The impact of a technical attribute (e.g. "the speedometer dial position") on the perceptual attributes

An interactive simulation of the model we presented in Figure 2 allows us having an idea about the global impact of the choice related to the speedometer dial position. According to Figure 3, placing the speedometer dial at the dashboard center improves the novelty perception, the sportive layout, the harmony perception, the quality perception, the arrangement space perception. However, at the same time this choice may deteriorate the comfort perception as well as the control comprehensibility.

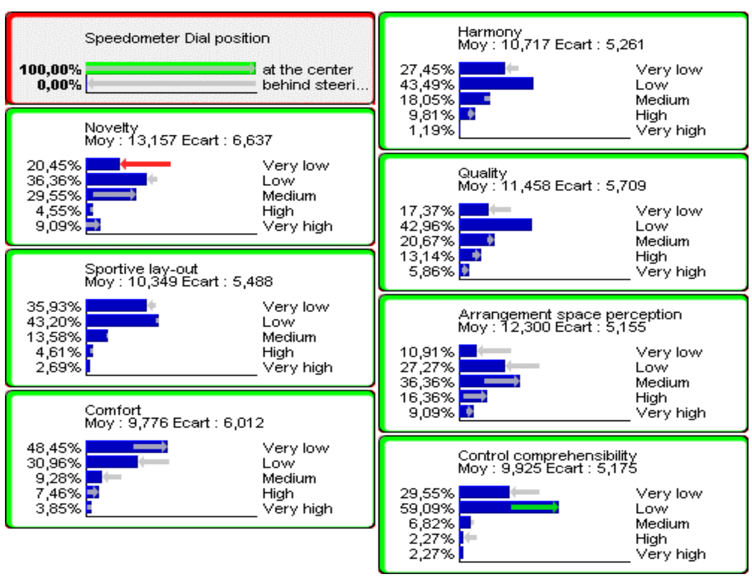

Figure 3. The influence of the speedometer dial position on the perceptual attributes characterizing a dashboard.

5.1.2. The interactions of a technical attribute (e.g. the "speedometer dial position") with the other technical attributes

A technical choice may be incompatible with other technical choices or induce hard constraints toward other technical choices (e.g. cost, technological incompatibility, functional incompatibility, etc.). Figure 4 shows that the "*speedometer dial position*" has significant probabilistic correlations with the "*dashboard arrangement space*", the "*dashboard color*" and the "*dashboard style layout*". Indeed, a change of the speedometer dial position may affect the *arrangement space* attribute and the choice of the *dashboard color* and/or the choice of the *style layout*. For instance, the model sates that positioning the speedometer at the center of the dashboard may increase the arrangement spaces (cf. Figure 4). It also states that, generally, dashboard whose speedometer dial is positioned at the center have curved lines style layout and two colors (cf. Figure 4)

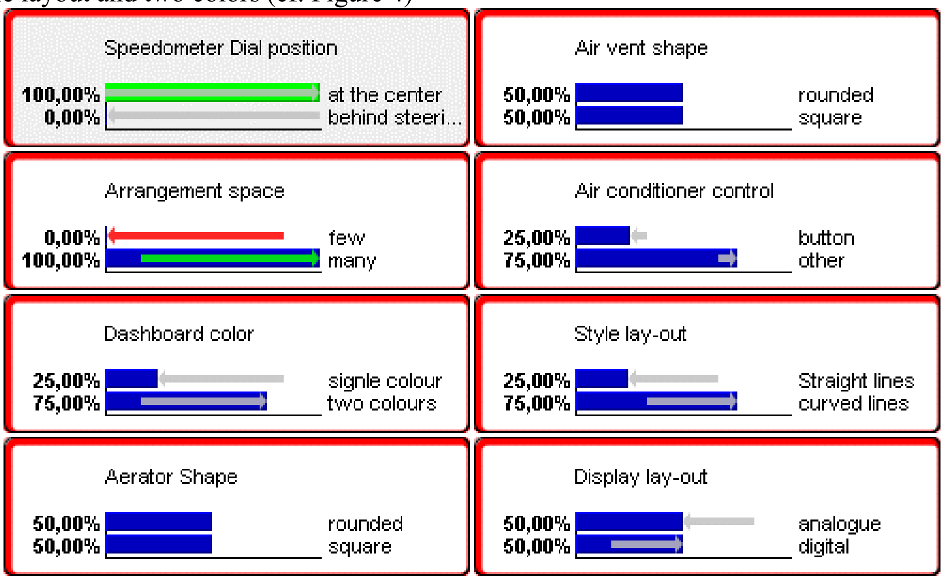

Figure 4. The interaction of the speedometer dial position with the technical attributes characterizing a dashboard. According to this figure, dashboards whose speedometer dial is located at the center have generally many arrangement spaces, two colors, curved lines style layout and digital display layout.

5.2. Inverse scenario

The **inverse scenario** allows answering the question "*what is the best choices (related to technical attributes) designer has to perform in order to maximize a perceptual attribute*"? This is a typical question designer asks when he carries out a **design synthesis task**. In the pervious section, we showed how a bayesian network allows designer to simulate the impact of a technical choice on perceptual and technical attributes: *input=design choice, output=impact on design attributes and performances*. In this section, we show how the same model allows designer identifying all possible design choices that allow him optimizing a given perceptual attribute (or performance): *input=perceptual attribute to be optimized, output=possible design choices*. In other words, this scenario consists in observing a perceptual attribute and analyzing how it interacts with the other attributes.

As an example, we take the "*dashboard novelty perception*" as target attribute to optimize and show how a BN Kansei model allows identifying the best technical choices designers can perform to improve that attribute. The same model we presented in Figure 2 can be used to carry out this optimization: it allows the two following points:

- Identifying all technical choices a designer can perform to improve the perceptual target attribute. Thereby, the designer can choose to tune the technical attributes, which are at the same time the most relevant (in term of their impact on the target attribute) and which engender the least constraints (in term of cost, time, technologies, etc.) (cf. Figure 5)

- Analyzing the impact of the optimization and avoid to deteriorate other perceptual performances of the design. For example, Figure 6 shows that improving the *novelty* perception is in coherence with the improvement of other perceptual attributes such as the "*quality*", the "*sportive layout*", the "*harmony*", the "*comfort*". However, it may at the same time worsen the "*control comprehensibility*", the "*space organization*" and the "*masculinity lay out*" of a dashboard. Providing a designer with all these information is crucial to help him optimizing his choices.

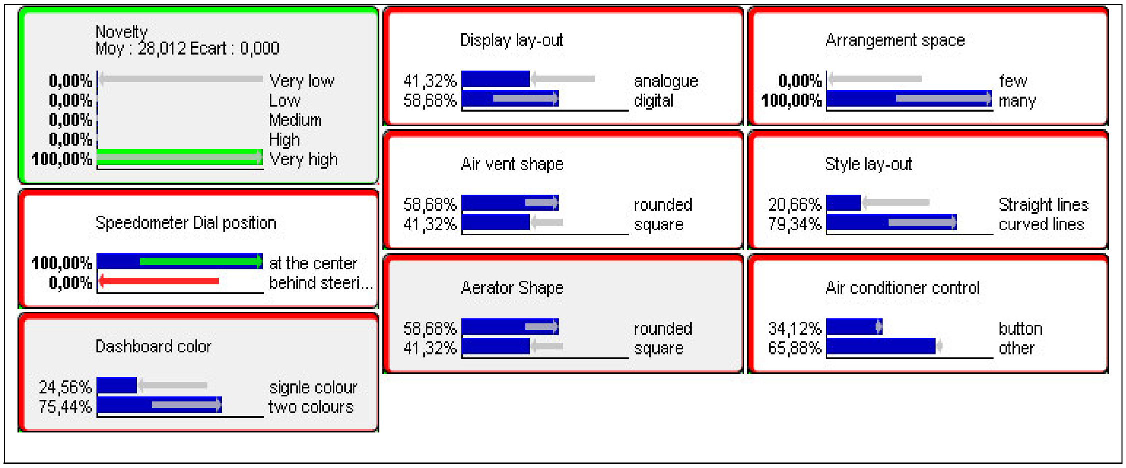

Figure 5. The optimal technical choices designer should carry out in order to improve the novelty perception of a dashboard

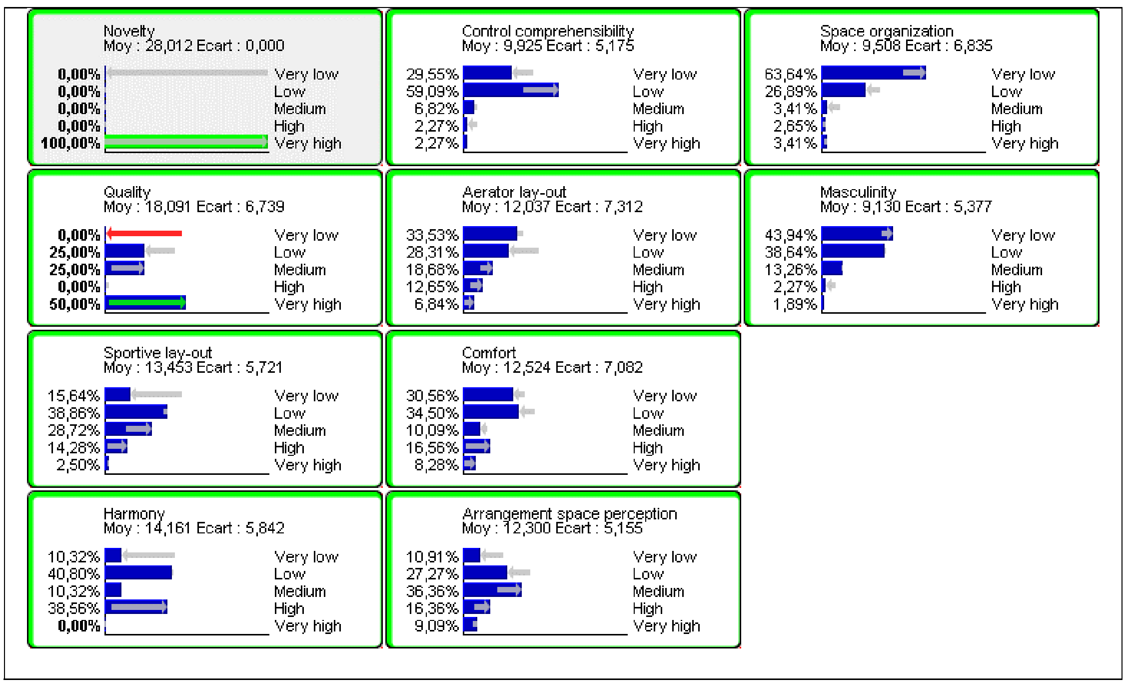

Figure 6. The impact of improving the novelty perception of a dashboard. Improving the novelty perception of a dashboard may be coherent with the improvement of other perceptual attributes (e.g. quality, harmony), but it may also deteriorate some other perceptual attributes (e.g. control comprehensibility, space organization).

5.3. A compound scenario for designing under constraints

In a design process, a designer is confronted to many constraints, which make the control of some technical attributes hard (or impossible) because of restrictions related to the cost and/or time and/or technology. In that case, the designer has to freeze some technical choices and tune others. Suppose a designer looks for the different technical possibilities to improve the *novelty perception* of a dashboard he is developing and suppose he has no other choice but putting the *speedometer dial behind the steering wheel*. According to the previous sections (cf. Figure 5), this constraint does not fit with his objective because a dashboard whose speedometer dial position is at the center looks more novel than a dashboard whose speedometer is behind the steering wheel. Thereby the designer has to identify the other levers (i.e. the other technical choices) he can perform in order to improve the novelty perception of the dashboard he is designing. The unsupervised model we presented in Figure 2 allows handling that issue.

In a sense, this is a mixed scenario of the two previous ones: at a first step a decision about a technical choice, so a direct scenario, is performed (*"speedometer dial position*" = *"behind the steering wheel"*). Then an inverse scenario is carried out to know what are the other technical attributes the designer can tune to optimize the perceptual attribute (i.e. to maximize the *novelty perception*). A simulation of a compound scenario is represented in Figure 7.

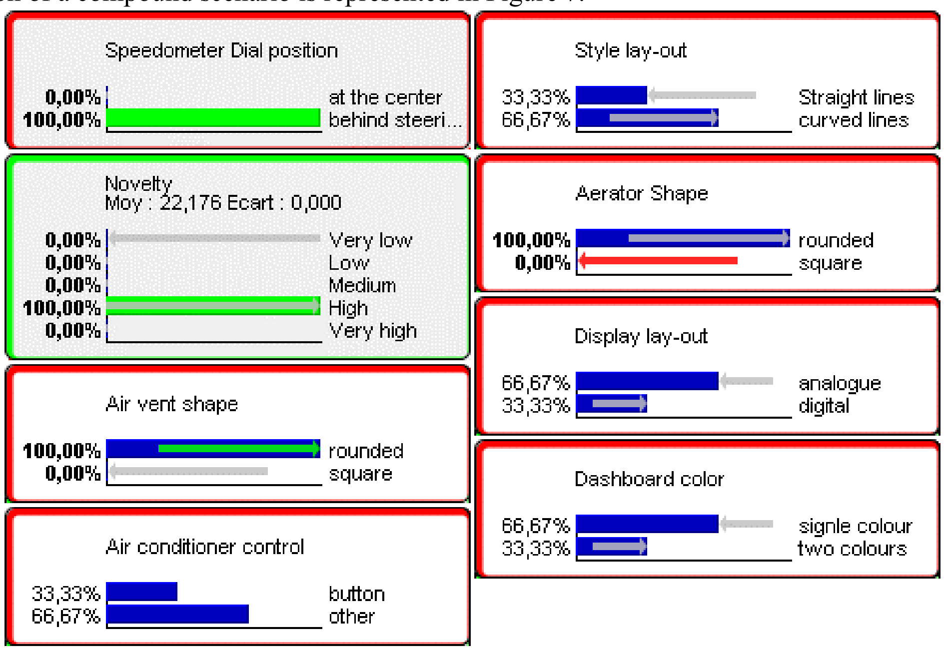

Figure 7. A compound scenario. A direct scenario is performed ("speedometer dial position" = "behind the steering wheel") and then an inverse scenario is carried out ("novelty perception" = "High") to know what are the other technical attributes the designer can tune to optimize the perceptual attribute (i.e. to maximize the novelty perception)

6. CONCLUSION AND PERSPECTIVES

In a design process, designers need tools to help them understanding customers' needs and thereby predicting their appreciation level of a new product. In this paper, we propose unsupervised learnt kansei models to represent the whole set of probabilistic relationships existing within the data. They are very useful to carry out a global optimization of the design. We defined three scenarios to use these models:

- **A direct/analysis scenario** in which the input is a technical attribute and the output consists of an analysis of its impact on the other design attributes (technical and perceptual). This scenario is typically used when a designer wants to compare different technical solutions.
- **An inverse/synthesis scenario** in which the input is a perceptual attribute and the output

consists of a list of technical choices the designer can perform to optimize the perceptual attribute. This scenario is typically used when a designer wants to identify all technical solutions to optimize a given performance of his design. He can then choose the most relevant and the least constraining technical attributes to tune.

- **A compound scenario** in which both technical and perceptual attributes are observed and the impact on the other attributes is analyzed. This scenario is typically used when a designer wants to optimize a given performance of his design under design constraints (cost, time, technology, etc.).

Through these three use scenarios, we showed that Bayesian Networks are a very flexible and powerful technology in preliminary perceptual (or emotional) design in terms of simulation and prediction capacities.

In a future work, we will propose other types of bayesian kansei models intended to help designers performing a local optimization of their design. In fact, the unsupervised model we propose in this paper is more adapted to carry out a global optimization of the design in that it allows learning all the probabilistic relations that hold between all attributes. However, in a local optimization, a designer looks generally for characterizing a target attribute (technical or perceptual). Thereby, he needs to focus the learning on identifying relations between the target attribute he is optimizing and the other attributes instead of learning all the probabilistic relations that hold in the data set. We will show that supervised learning is more accurate with respect to the target attribute characterization than unsupervised learning.

REFERENCES

[1] Akaike, H. (1974) 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control*, Vol. 19, No. 6, pp.716–723

[2] BayesiaLab, (2006). *Bayesia tutorial Book*. www.bayesialab.com

[3] Bouckaert, R. R. (1993) 'Probabilistic network construction using the minimum descriptions length principle', *Proceedings of the European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*.

[4] Buntine, W. (1996) 'A guide to the literature on learning graphical models', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 8, pp.195–210.

[5] Chickering, D. M. (2002a) 'Learning equivalent classes of bayesian network structures', *Journal of Machine Learning Research*, Vol. 2, pp.445-498.

[6] Chickering, D. M. (2002b) 'Optimal structure identification with greedy search', *Journal of Machine Learning Research*, Vol. 3, pp.507-554

[7] Cooper, G. F. and Herskovits, E. (1992) 'A bayesian method for the induction of probabilistic networks from data', *Machine Learning*, Vol. 9, No. 309.

[8] Friedman, N. (1998) 'The bayesian structural EM algorithm', *Proceedings of the Fourteenth Uncertainty in Artificial Intelligence Conference (UAI-1998)*, *San Francisco, CA*, Morgan Kaufmann Publishers, pp.129–138.

[9] Friedman, N. and Koller, D. (2003) 'Being bayesian about network structure. A bayesian approach to structure discovery in bayesian networks', *Machine Learning*, Vol. 50, pp.95–125.

[10] Harvey A., (2005), *Application of an integrated method to a study of the consumer perceptions of automobile dashboards*, Research Internship report in Ecole Centrale Paris, University of Bath.

[11] Heckerman, D., Geiger, D. and Chickering, D. M. (1995) 'Learning Bayesian networks: the combination of knowledge and statistical data', *Machine Learning*, Vol. 20, No. 3, pp.197–243

[12] Hsiao S.W., Wang H.P., 1998. Applying the semantic transformation method to product form design. Design Studies 19, 309-330.

[13] Huang C., A. Dawiche (1996). Inference in Belief Networks : A Procedural Guide. *International Journal of Approximate Reasoning*, 15, p225-263.

[14] Jensen F.V. (1996). *An Introduction to Bayesian Networks*. (UCL Press (Ed)), London.

[15] Jindo T., Hirasago K., 1997. Application studies to car interior of Kansei engineering. *International journal of industrial ergonomics* 19, 105-114.

[16] Jordan, M. I. (1998) *Learning in Graphical Models,* MIT Press, Cambridge, Massachusetts.

[17] Jouffe L., Munteanu P. *(2001),* New Search Strategies for Learning Bayesian Networks, *Proceedings of Tenth International Symposium on Applied Stochastic Models, Data Analysis*,

Compiègne, France.
[18] Jouffe L.,(2002), Nouvelle classe de méthodes d'apprentissage de réseaux bayésiens, *Journées francophones d'Extraction et de Gestion des Connaissances (EGC),* janvier 2002, Montpellier, France.
[19] Lam, W. and Bacchus, F. (1994) 'Learning bayesian belief networks: An approach based on the MDL principle', *Computational Intelligence*, Vol. 10, pp.269–293.
[20] Limayem F., Yannou B., (2004), Generalization of the RCGM and LSLR Pairwise Comparison Methods. *Computers and Mathematics with Applications*, vol. 48: p. 539-548.
[21] Meek, C. (1997), '*Graphical Models: Selecting causal and statistical models*', PhD thesis, Carnegie Mellon University.
[22] Mehmet, K. and Gregory, C. (2002) 'A bayesian network scoring metric that is based on globally uniform parameter priors', *Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence (UAI-02)*, *San Francisco, CA*, Morgan Kaufmann Publishers, pp. 251–258
[23] Munteanu P., Bendou M. (2001) The EQ Framework for Learning Equivalence Classes of Bayesian Networks, *First IEEE International Conference on Data Mining (IEEE ICDM),* San José.
[24] Nagamachi M., 1995. Kansei engineering: a new ergonomic consumer-oriented technology for product development. *International Journal of Industrial Ergonomics* 15, 3-11.
[25] Nagamachi M., 2002. Kansei engineering as a powerful consumer-oriented technology for product development. *Applied Ergonomics* 33, 289–294.Osgood C.E., Suci G.J., Tannenbaum P.H., 1957. *The measurement of meaning*, Illinois press.
[26] Naïm, P., Wuillemin, P. H., Leray, P., Pourret, O. and Becker, A. (1999) *Les Réseaux Bayésiens,* Eyrolles, Paris
[27] Neapolitan, R. E. (2003) *Learning Bayesian networks,* Prentice Hall, Upper Saddle River, NJ.
[28] Petiot J.-F., Yannou B., (2004), Measuring consumer perceptions for a better comprehension, specification and assessment of product semantics. *International Journal of Industrial Ergonomics*, vol. 33(6): p. 507-525.
[29] Rissanen, J. (1986) 'Stochastic complexity and modeling', *The Annals of Statistics*, Vol. 14, No. 3, pp.1080–1100.
[30] Robinson, R. (1977) 'Counting unlabeled acyclic diagraphs' In *Combinatory Mathematics V*(Ed, Little, C.) Springer, Berlin, pp.28-43
[31] Spirtes, P., Glymour, C. and Scheines, R. (2000) *Causation, Prediction, and Search*, MIT Press, 2nd edition, Cambridge, MA.
[32] Schwarz, G. (1978) 'Estimating the dimension of a model', *Annals of Statistics*, Vol. 6, pp.461–464.
[33] Suzuki, J. (1993) 'A construction of bayesian networks from databases based on an MDL principle', *Proceedings of the 9th Annual Conference on Uncertainty in Artificial Intelligence (UAI-93)*, *San Francisco, CA*, Morgan Kaufmann Publishers.
[34] Yannou B., Petiot J.-F., (2004), A methodology for integrating the customers' assessments during the conceptual design. *in DETC/DTM: ASME Design Engineering Technical Conferences / Design Theories and Methodologies,* Sept. 28 - Oct. 2, Salt Lake City, Utah, USA.
[35] Yannou B. (2007). Easy and flexible specifications and product evaluations by expert and customer comparisons with existing products, *submitted to ICED'07*

Contact:	Bernard Yannou
institution/university:	Ecole Centrale Paris
department:	Laboratoire Genie Industriel (LGI)
street:	Grande Voie des Vignes
PO Box, City:	92295 Châtenay-Malabry Cedex
Country:	France
Phone:	(33) 1 41 13 15 21
Fax:	(33) 1 41 13 12 72
e-mail:	bernard.yannou@ecp.fr