# Sensory Quality Management and Assessment: from Manufacturers to Consumers

Ludovic Koehl*, Xianyi Zeng*, Bin Zhou*,#, and Yongsheng Ding#

*GEMTEX Lab., ENSAIT 9, rue de l'Ermitage, 59070 Roubaix
ludovic.koehl@ensait.fr
#College of Information Sciences and Technology Donghua University,
1882 Yan-An West Road, Shanghai 200051, P.R. China

**Abstract.** This paper presents an intelligent technique based method for analyzing and interpreting sensory data provided by multiple panels for the evaluation of industrial products. In order to process the uncertainty existing in these sensory data, we first transform all sensory data into fuzzy sets on a unified scale using the 2-tuple fuzzy linguistic model. Based on these normalized data sets, we compute the dissimilarities or distances between different panels and between different evaluation terms used by them, defined according to the degree of consistency of data variation. The obtained distances, expressed with crisp numbers, are turned into fuzzy numbers for a better physical interpretation. Thus, these fuzzy distances permit to characterize in an easier way the evaluation behaviour of each panel and the quality of the evaluation terms used. Also, based on soft computing techniques and the dissimilarity between terms, we develop procedures for interpreting terms of one panel using those of another panel and a model for setting the relationships between the physical product features and the evaluation terms. Then, we introduce a new method to forecast the consumer preference from the sensory evaluation provided by an expert panel. This general approach has been applied to two kinds of industrial products concerning both cosmetic and textile industries.

**Key words:** Sensory Evaluation, quality, assessment, fuzzy linguistic model, dissimilarity, distance, fuzzy distance, interpretation

## 1 Introduction

In many industrial sectors such as food, cosmetic, medical, chemical, and textile, sensory evaluation is widely used for determining the quality of end products, solving conflicts between customers and manufacturers, developing new products, and exploiting new markets adapted to the consumer's preference [1, 2, 3]. In the [2], sensory evaluation is defined as a scientific discipline used to evoke, measure, analyze, and interpret reactions to the characteristics of products as they are perceived by the senses of sight, smell, taste, touch,

Ludovic Koehl et al.: *Sensory Quality Management and Assessment: from Manufacturers to Consumers*, Studies in Computational Intelligence (SCI) **5**, 375–400 (2005)
www.springerlink.com 

and hearing. In general, sensory evaluation can be described as "under predefined conditions, a group of organized individuals evaluate some products with respect to certain given target". Consequently, there are four basic factors in sensory evaluation: evaluation product, evaluation panel, evaluation target and evaluation environment. According to the difference cases of these factors, we can divide sensory evaluation into two levels [4]: (1) design-oriented sensory evaluation; and (2) market-oriented sensory evaluation. Design-Oriented Sensory Evaluation (DOSE) is done by a trained panel composed of experienced experts or consultants inside the enterprise for judging industrial products on a number of analytical and non-hedonic linguistic descriptors in a controlled evaluation environment, such as an evaluation laboratory. The evaluation target of design-oriented sensory evaluation is to obtain the basic sensory attributes of products to improve the quality of product design and development. Market-Oriented Sensory Evaluation (MOSE) is given by untrained consumer panels using analytical and hedonic descriptors according to their preference on the products to be evaluated in an uncontrolled evaluation environment, such as supermarkets. The evaluation target of market-oriented sensory evaluation is to obtain the preference degree of consumers in order to forecast the market reaction to the evaluated product. Afterwards, the key issue is to compute data provided by a DOSE in order to forecast the consumers' preference (B2C: Business to Consumer) assuming the data are obtained for a precise context and/or end-use for the industrial product, since sensory evaluation is context dependant.

Sensory evaluation of industrial products leads to a set of linguistic terms, named subjective evaluation, strongly related to consumer's preference but difficult to be normalized due to their uncertainty and imprecision (see Fig. 1). As such, this evaluation restricts the scientific understanding of product characteristics for those who wish to design high quality product by engineering means. Hence, a great number of researchers tried to develop objective evaluation systems by physical measurements in order to replace sensory evaluation, e.g. [5, 6]. In practice, these objective evaluation systems are often expensive and lead to precise numerical data describing indirectly products but their

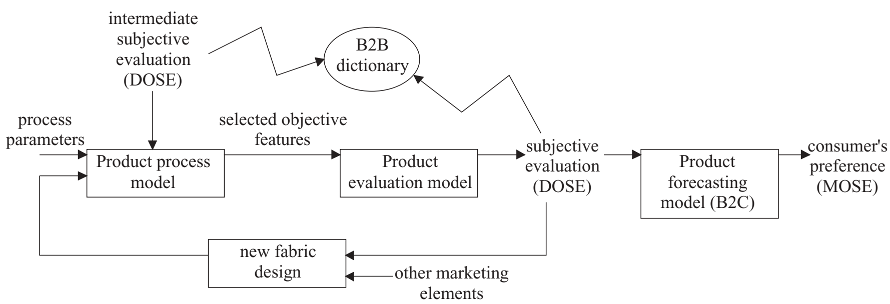


**Fig. 1.** Different steps for the industrial product sensory evaluation

interpretation for the product quality related to consumer's preference has to be exploited. In this chapter (Sect. 4.4), we present a fuzzy based method for modeling the relationships between the sensory evaluation provided by panels and the physical features measured on appropriate apparatus, which is called objective evaluation. The published model is named 'Product evaluation model'. Compared with physical measures, sensory evaluation as a result of a human measurement is more efficient for quality determination and it can not be, for a long term, completely replaced by objective evaluation. This method can be applied to set up the product process model (see Fig. 1).

In sensory evaluation, the main difficulties can be summarized as follows:

1. For an individual, the evaluation of a sample (in numerical score or linguistic expression) gives a relative result depending on the comparison with the other samples. This score is significant only for one specific collection of products and for one particular individual. It is not normalized in a general background.
2. The terms used by different individuals in an evaluation are not normalized. Even if they use a common term, its significance is not necessarily the same for them.
3. The scales and the upper and lower bounds used by different individuals are often different, which should be unified to the same scale so that the aggregation of all sensory data can be done.

In our previous work [7], we propose an approach based on a linguistic 2-tuple model [8] for the formalization and the analysis of sensory data. This approach permits to initially normalize and aggregate sensory data inside each panel, i.e. a group of evaluators and compute the dissimilarities or distances between different panels and between different terms used according to the degree of consistency of relative data variation. The quality of each panel and evaluation terms can be estimated from these dissimilarity criteria. However, the physical meaning of the corresponding results is not easy to be interpreted and distances for different panels are not easy to be compared because the definition of the distance is related to a specific evaluation space, whose number of evaluation terms is not the same with respect to another evaluation space.

We propose here two procedures in order to provide a systematic interpretation to sensory data obtained by different panels and to forecast the consumer preference using a training set of products with known DOSE. The first procedure permits to interpret the values of the dissimilarities between panels and between evaluation terms. In this procedure, each evaluation score is generated by a random variable distributed between its lower and upper bounds uniformly and then the scores of each panel on all terms for all products to be evaluated constitute a random evaluation matrix. The statistical distribution of the dissimilarity between two panels or between two terms can be obtained from the corresponding random evaluation matrices using the equations given in Sect. 4.1. For specific panels or terms, the fuzzy values or linguistic values of distances between them can be calculated according to these statistical distributions. This interpretation will be very important in the understanding of the behavior of panels and terms used in the evaluation.

The second procedure permits to interpret the relationship between terms used by different panels. This relationship is recurrently determined using a genetic algorithm with penalty strategy. It can be considered as a dictionary for the understanding between different panels. Using this dictionary, an evaluation term used by one panel can be transformed into one or several terms used by another panel. It will be very helpful for solving commercial conflicts between producers and consumers at the level of understanding of evaluation terms (B2B: Business to Business). The general proposed approach tries to allow manufacturers to reduce cost and time for designing new products and thus to become more reactive and competitive for the market demands and requirements. Using suitable tools, the manufacturers will be able to predict the ability for a product to become a success on a specified market. The final challenge is to be able to tune the process parameter for producing the appropriate product which will fit the market demand. Additional components, such as marketing elements (price, retailer location, and so on, ...), have also to be taken into account.

In order to illustrate the effectiveness of our proposed approach, we apply it to sensory data provided by two sets of industrial products. The first set corresponds to fabrics designed for apparel: T-shirts. The hand evaluation for those 43 knitted cotton samples are obtained from 4 sensory panels in France and China. Each panel uses its own terms and evaluation bounds different from the others. Based on the proposed method, we compare the behaviors on fabric hand evaluation between textile professionals and students and between French and Chinese consumers in order to make the adaptive design of textile products to consumer's preference. The second set of data is related to the cosmetic industry. It includes 8 lotions with varying performance according to their interaction with the skin or the human feeling they express at the first contact.

## 2 Description and Formalization of Sensory Data

The concepts of sensory evaluation used in this chapter can be formalized as follows.

$P = \{P_1, P_2, \ldots, P_r\}$ : the set of $r$ panels, each panel, $P_i = \{I_{i1}, I_{i2}, \ldots, I_{i,h(i)}\}$, being composed of $h(i)$ individuals evaluating the hand feeling of fabric samples.

$A = \{a_{ik} | i = 1, 2, \ldots, r;\ k = 1, 2, \ldots, m(i)\}$ : the set of linguistic terms used for the fabric hand evaluation. For the panel $P_i$, it uses $m(i)$ terms, i.e. $A_i = \{a_{i1}, a_{i2}, \ldots, a_{i,m(i)}\}$. For one or several panels, different linguistic terms can be correlated between them, but in general they can never be replaced one by another. It is also possible for two terms used by different panels to be identical, but they are not necessarily equivalent at semantic level.

$T = \{t_1, t_2, \ldots, t_w, \ldots, t_n\}$ : the set of $n$ industrial products to be evaluated. The relationships among different terms of the evaluation, the behaviors of different panel members can be studied from these samples.

$E_i = \{A_i; E_{i1}, E_{i2}, \ldots, E_{i,h(i)}\}$ : the evaluation space for the panel $P_i/(i \in \{1, 2, \ldots, r\})$.

$E_{ij}$: the evaluation matrix $n \times m(i)$ of the individual $I_{ij}$. Each element of this matrix, $e_{ij}(k, 1)$ represents a relative numerical score or a granular linguistic expression given by $I_{ij}$ for evaluating the sample $t_k$ on the term $a_{il}/\left(k \in \{1, 2, \ldots, n\}, l \in \{1, 2, \ldots, m\left(i\right)\}\right)$. It is obtained by a classification procedure for the whole samples of $T$.

The sensory data, provided by the consumers, are obtained through a survey and the consumers' panel can be considered as a free profiling panel. The consumers fill a questionnaire form. For each question, they select the most appropriate grade (intensity) according to their feeling. For example, if the question about the product concerns its 'softness', the consumer's answer can be chosen from the following grades of softness: {not Soft, slightly Soft, Soft, very Soft, extremely soft}. Since the survey is repeated for many people, the consumer evaluation matrix is computed so that each element denotes the population percentage who thinks the product satisfies the considered grade compared to the term and for all the products of $T$. Thus, a consumer panel (MOSE) can be regarded as a regular panel of $P$, but with only one member ($h(i) = 1$), with a number of terms that corresponds to the number of terms used in the survey and with an evaluation matrix (scores) which corresponds to the population percentage. We also define the vector $X = (x_1, \ldots, x_w, \ldots, x_n)^T$ which concerns the consumers' preference about all the products which belong to $T$. Through a survey, they express their preference for one product over the all set. Thus, the component $x_w$ of $X$ corresponds to the consumers' rate who appreciate the $w$-*th* product over the $(n-1)$-*th* other products. As a mother of fact, the more consumers are enrolled in the survey, the more precise is the market preference evaluation. The sum for all the $x'_w s$, as the expression of the consumers rate who rank the $w$-th product at the top level, is then equal to 1: $\sum_{w=1}^{n} x_w = 1$

## 3 Linguistic 2-Tuple for Finding the Optimal Unified Scale

In sensory evaluation, results given by different individuals on different attributes or terms may have different scales. This is because the sensitivity of each individual to the samples to be evaluated, strongly related to his personal experience and the corresponding experimental conditions, is often different from others. Moreover, these sensory data may be in a numerical form or a granular linguistic form. So it is necessary to develop a suitable unified scale in order to normalize and aggregate these data.

The 2-tuple fuzzy linguistic model [8] can be used for unifying multi-granular linguistic information without loss of information. Using this model, sensory data provided by different individuals on different attributes can be normalized on the common optimal scale. However, in [8], the transformation is carried out between levels of a linguistic hierarchy only. Here, we generalize the 2-tuple model to be used in transforming among arbitrary scales and several quantitative criteria permitting to select the most suitable scale for all individuals and all attributes or terms.

### 3.1 2-Tuple Fuzzy Linguistic Model

In the panel $P_i$, for each individual $I_{ij}$ and each term $a_{il}$, the corresponding sensory data varying between 0 and $g(i,j,l) = \max\{e_{ij}(k,l)|k = 1,2,\ldots,n\}$ can be transformed into a fuzzy set of $g(i,j,l) + 1$ modalities denoted by $U_{ijl} = \{u_1, u_2, \ldots, u_{g(i,j,l)}\}$ as shown in Fig. 2.

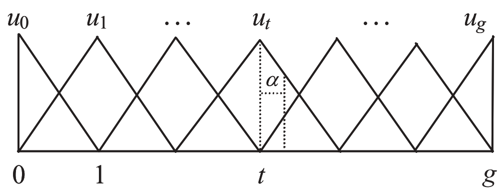


**Fig. 2.** Fuzzy set $U_{ijl}$ composed of $g + 1$ modalities

For simplicity, $g(i,j,l)$ is denoted by $g$ for unambiguous cases. We consider that any evaluation score of $I_{ij}$ for the term $a_{il}$ is included between 0 and $g$ and it can be represented by a 2-tuple $(u_t^g, \alpha^g)$ with $\alpha^g \in [-0.5, 0.5]$.

For any $\beta \in [0, g]$, the 2-tuple that expresses the equivalent information is obtained using the following function:

$$\Delta(\beta) = \begin{cases} u_t & t = \mathrm{round}(\beta) \; . \\ \alpha = \beta - t\alpha \in [-0.5, 0.5) \end{cases} \text{ and } \Delta^{-1}(u_t, \alpha) = t + \alpha = \beta \; .$$

For the panel $P_i$, the evaluation results of the individuals $I_{ij}$'s can be aggregated by transforming all the corresponding fuzzy sets to be on a unified scale.

Let $ug(i,l)$ be the value of the unified scale for all the individuals of $P_i$ on the term $a_{il}$. For each individual $I_{ij}$, any evaluation score $(u_t^g, \alpha^g)$ on this term can be transformed into a new 2-tuples:

$$(u_s^{ug}, \alpha^{ug}) = \Delta\left(\frac{\Delta^{-1}(u_t^g, \alpha^g) \cdot ug}{g}\right) \tag{1}$$

This transformation can be denoted by the function $s = Tr(t,\ g,\ ug)$.

### 3.2 Obtaining the Optimal Common Scale

In order to best aggregate the sensory evaluation of multiple individuals in the same panel, we have to find an optimal value of the unified scale $ug$ for all the individuals of $P_i$.

For the panel $P_i$, its optimal unified scale can be calculated according to the two following principles:

1. The sensory data given by the individuals $I_{ij}$ should cover all the modalities of the unified scale, i.e., any $u_s^{ug} (s \in \{0, 1, 2, \ldots, ug\})$ should correspond to at least one data.
2. The variation or the trend of the sensory data should not change very much with the transformation of the scale.

The sensory data of $I_{ij}$ for evaluating $n$ samples on the term $a_{il}$ before the transformation are $\{e_{ij}(1, l), e_{ij}(2, l), \ldots, e_{ij}(n, l)\}$. After the transformation, these data become $\{s_{ij}(1, l), s_{ij}(2, l), \ldots, s_{ij}(n, l)\}$, where

$$s_{ij}(k, l) = Tr(e_{ij}(k, l), g, ug) \quad \text{for } k = 1, 2, \ldots, n \, .$$

According to the first principle, we first calculate the number of data for each modality $q$ of the unified scale $ug$, i.e.,

$$N\_\mathrm{mod}_i(l, q) = \sum_{j=1}^{h(i)} \sum_{k=1}^{n} equ(s_{ij}(k, l), q) \, ,$$

with

$$equ(p, q) = \begin{cases} 1 & \text{if } p = q \\ 0 & \text{otherwise} \end{cases}$$

The criterion of coverage of the unified scale is then defined by

$$\mathrm{Cover}_i(l) = \min \{ N\_\mathrm{mod}_i(l, q) \, | q = 0, 1, 2, \ldots, ug \} \, .$$

According to this criterion, $ug$ should be selected so that $\mathrm{Cover}_i(l)$ is as big as possible. If this value is 0, it means that there exists at least one non-significant modality $u_s^{ug}$ on the unified scale $ug$.

According to the second principle, the difference of the trend between two data sets $\{e_{ij}(1, l), e_{ij}(2, l), \ldots, e_{ij}(n, l)\}$ and $\{s_{ij}(1, l), s_{ij}(2, l), \ldots, s_{ij}(n, l)\}$ should be as small as possible. So, the corresponding criterion is defined by

$$\mathrm{Trend}_i(l) = \min \{ \mathrm{trend}_{ij}(l) \, | j = 1, \ldots, h(i) \} \, ,$$

with

$$\mathrm{trend}_{ij}(l) = \frac{2}{n(n-1)} \sum_{k_1 < k_2} c_{ij}(k_1, k_2, l),$$

and

$$c_{ij}(k_1,k_2,l)=\begin{cases}1 & \text{if } (e_{ij}(k_1,l)-e_{ij}(k_2,l))(s_{ij}(k_1,l)-s_{ij}(k_2,l))>0\\ 0 & \text{otherwise}\end{cases}.$$

According to this criterion, $ug$ should be selected so that the value of $\text{Trend}_i(l)$ is as big as possible. If two data sets $\{e_{ij}(1,l), e_{ij}(2,l),\ldots,e_{ij}(n,l)\}$ and $\{s_{ij}(1,l),\ s_{ij}(2,l),\ldots,s_{ij}(n,l)\}$ for any individual $I_{ij}$ vary completely with the same trend, the value of $\text{Trend}_i(l)$ will be maximized and there is no information lost in the data set $\{s_{ij}(k,l)\}$ on the unified scale $ug$.

The optimal value of $ug$ can be obtained by maximizing the linear combination of these two criteria as follows:

$$\max\{\text{Cover}_i(l)+\rho\cdot\text{Trend}_i(l)\} \tag{2}$$

where $\rho$ is a positive constant adjusting the ratio of these two criteria.

As the optimal value of $ug$ is obtained, for each individual $I_{ij}$ and each term $a_{il}$, the optimal unified evaluation score $(u_s^{og},\alpha^{og})$ can be obtained by transforming sensory data into a new 2-tuples by using (1). The sensory data for all individuals of this panel can be aggregated on this optimal scale.

### 3.3 Aggregating Sensory Data on the Desired Domain

On the common optimal scale, the transformed evaluation scores on term $a_{il}$ for all individuals can be aggregated using an averaging operator. The aggregated evaluation result of $P_i$ for one sample on the term $a_{il}$ can be calculated by

$$(\bar{u}_s^{ug},\bar{\alpha}^{ug})=\Delta\left(\frac{\sum_{j=1}^{h(i)}\Delta^{-1}(u_{s_j}^{ug},\alpha_{s_j}^{ug})}{h(i)}\right)$$

where $(u_{s_j}^{ug},\alpha_{s_j}^{ug})$ is the transformed 2-tuples of the evaluation scores of the individual $I_{ij}$ for the same sample on the term $a_{il}$.

In the same way, all the aggregated evaluation scores of $P_i$ for different terms of $A_i$ are transformed to be on the unique unified scale and can be expressed by a matrix of $n\times m(i)$ 2-tuples, denoted by $S_i$ (evaluation space for $P_i$). Each element of $S_i$ is a 2-tuples denoted by $(u_{s_i(k,l)},\alpha_{s_i(k,l)})$ for $k=1,2,\ldots,n$ and $l=1,2,\ldots,m(i)$. For simplicity, this 2-tuples is replaced by $s_i(k,l)$ if the weak influence of $\alpha_{s_i(k,l)}$ can be neglected. The matrix $S_i=(S_{i1}\ S_{i2}\cdots S_{i,n})^T$ includes $n$ vectors, each of them representing the evaluation results for one sample. These vectors will be used in the next section for the analysis and interpretation of panels and term.

## 4 Analysis and Interpretation of Sensory Data

In this section, we first define a number of criteria, according to the normalized matrices $S_i=(s_i(k,l)), i\in\{1,\ldots,r\}$ obtained from Sect. 3, in order to analyze the performance of the evaluators and the used description terms.

### 4.1 Dissimilarity between Individuals and between Terms

The sensory data of two panels $P_a$ and $P_b$ constitute two evaluation spaces $S_a$ and $S_b$. The dissimilarity between $P_a$ and $P_b$ cannot be defined using classical methods, which compute distances between the vectors $S_{ak}$ and $S_{bk}$ ($k \in \{1, \ldots, n\}$) because these two vectors are not in the same space. So a new dissimilarity criterion between two individuals $P_a$ and $P_b$ has been defined in [7].

In this definition, the dissimilarity criterion takes into account the degree of consistency of relative variations of two different sensory data sets. If the internal relative variations of these two data sets are close each other, and then the dissimilarity between the corresponding panels is small. Otherwise, this dissimilarity is great. Formally, this dissimilarity is defined by

$$D_{ab} = \frac{2}{n(n-1)} \sum_{i<j} d_{ab}(i,j) \tag{3}$$

It depends on the following elements:

1. The dissimilarity between $P_a$ and $P_b$ related to the relative variation between fabric samples $t_i$ and $t_j$: $d_{ab}(i,j) = |vr_a(i,j) - vr_b(i,j)|$.
2. The relative variations between $t_i$ and $t_j$ for $P_a$ and $P_b$:

$$vr_a(i,j) = \frac{1}{\sqrt{m(a)}} \left\| S_{ai} - S_{aj} \right\|$$

$$vr_b(i,j) = \frac{1}{\sqrt{m(b)}} \left\| S_{bi} - S_{bj} \right\| .$$

The definition of $D_{ab}$ permits to compare, between these two panels, the relative variations on the samples of $T$. The dissimilarity between two panels reaches its minimum only when the internal variations of their sensory data are identical.

The dissimilarity $D_{ab}$ can be considered as a distance between $P_a$ and $P_b$ because it satisfies the following three conditions:

(i) $D_{aa}$=0
(ii) $D_{ab} = D_{ba}$
(iii) $D_{ab} + D_{bc} \geq D_{ac}$

The two first conditions (i)&(ii) can be easily proved from the definition of the dissimilarity. The proof of the third condition (iii) is given as follows.

If $vr_a(i,j) \geq vr_b(i,j) \geq vr_c(i,j)$, then $d_{ab}(i,j) + d_{bc}(i,j) = vr_a(i,j) - vr_c(i,j) = d_{ac}(i,j)$.

If $vr_a(i,j) \geq vr_c(i,j) \geq vr_b(i,j)$, then $d_{ab}(i,j) + d_{bc}(i,j) = vr_a(i,j) - vr_b(i,j) + vr_c(i,j) - vr_b(i,j) \geq vr_a(i,j) - vr_c(i,j) = d_{ac}(i,j)$.

For any other conditions, the inequality $d_{ab}(i,j) + d_{bc}(i,j) \geq d_{ac}(i,j)$ also holds and then we have $D_{ab} + D_{bc} \geq D_{ac}$. So, the three conditions of distance are completely satisfied by the dissimilarity defined previously.

Another criterion is developed in order to compare two panels according to the sensitivity of data for the evaluation of the samples of $T$ [7]. The sensitivity of $P_a$ is defined by

$$SSB_a = \frac{2}{n(n-1)} \sum_{i<j} vr_a(i,j) \tag{4}$$

where, $vr_a(i,j)$ characterizes the relative variation of the sensory data given by $P_a$ from the sample $t_i$ to $t_j$. If the value of $SSB_a$ is bigger than that of $SSB_b$, then we consider that $P_a$ is more sensitive to the samples of $T$ than $P_b$. However, this does not mean that $P_a$ is more efficient than $P_b$.

In the same way, we also define the dissimilarity or distance between terms used by the same panel and by different panels [7]. This criterion permits to study the redundancy of the terms used by each panel. In general, the bigger the dissimilarity between any two terms used by a panel is, the more efficient the evaluation results are.

### 4.2 Fuzzy Dissimilarity or Fuzzy Distance

The criteria of dissimilarity and sensitivity for panels and terms defined in Sect. 4.1 are significant only for comparison because we do not know how to physically interpret the absolute values of these criteria. We do not know if a slight variation of such a criterion is physically important or not. In order to give a physical interpretation to the results calculated from dissimilarity criteria and sensitivity criteria, we transform these numerical values into fuzzy numbers, whose membership functions are generated according to the probability density distributions of the corresponding random matrices. The detailed procedure is given as follows and, according to the above section, $P_a$ and $P_b$ denote two panels $a$ and $b$:

Step 1: For fixed values $n$, $m(a)$ and $m(b)$, generating two random matrices $S_a$ (dimension: $n \times m(a)$) and $S_b$ (dimension: $n \times m(b)$), whose elements obey the uniform distribution between lower and upper bounds of normalized evaluation scores, i.e. 0 and *ug*.

Step 2: Computing the values of dissimilarity and sensitivity $D_{ab}$, $SSB_a$ and $SSB_b$ according to the equations in Sect. 4.1.

Step 3: Repeat Step 1 and Step 2 several times in order to obtain the probability density distributions for $D_{ab}$, $SSB_a$ and $SSB_b$(see Fig. 3).

Step 4: We then divide equally the area of each distribution into 5 parts. According to these divided areas, we generate 5 fuzzy sub-sets for each of $D_{ab}$, $SSB_a$ and $SSB_b$: {*very small, small, medium, large, very large*}. The corresponding membership functions can be determined from these 5 fuzzy numbers.

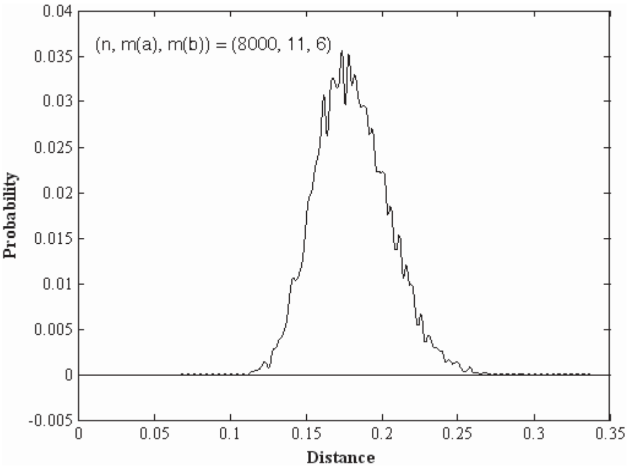


**Fig. 3.** Distance Distribution Function ($D_{ab}$) with uniformly selected random evaluation scores

Figure 3 gives the probability distribution for the dissimilarity $D_{ab}$ between two panels $P_a$ and $P_b$ using 11 and 6 terms respectively as Fig. 4 shows how we turn a crisp number of dissimilarity (or sensitivity) into a fuzzy one,which makes easier the understanding. The membership functions corresponding to the five fuzzy values equally dividing the area of this distribution are given in Fig. 4. From these membership functions, we can see that the dissimilarity $D_{ab}$ is sensitive only in the interval of [0.11, 0.30], in which three fuzzy values small (S), medium (M) and large (L) are asymmetrically distributed. A value of $D_{ab}$ smaller than 0.161 is considered as very small (VS) and a value of $D_{ab}$ larger than 0.207 as very large (VL).

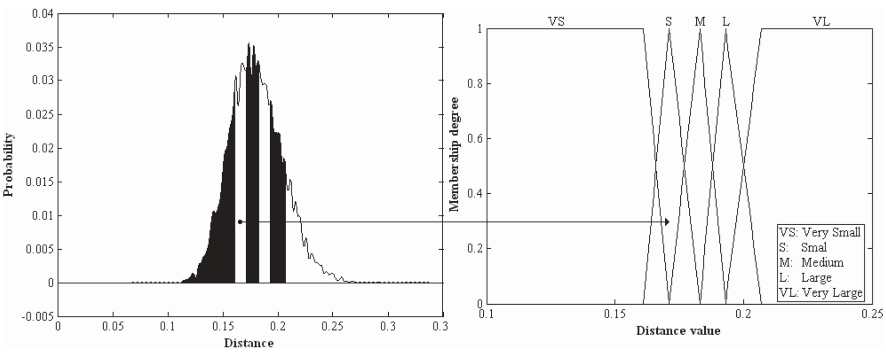


**Fig. 4.** From a distance crisp number to a fuzzy distance number

In this way, each numerical value of dissimilarity criteria and sensitivity criteria, calculated from (3) and (4) can be transformed into a fuzzy number whose value includes the linguistic part taken from the previous 5 terms and the corresponding membership degree. This fuzzy number permits to interpret the dissimilarity or the sensitivity with respect to the whole distribution of random values. The evaluation behaviors of different panels can be effectively analyzed and compared from these fuzzy numbers.

Moreover, according to our experiments, the distributions of the dissimilarity and the sensitivity for different values of $n, m(a)$ and $m(b)$ are rather similar. This is because the normalization with respect to these parameters has been taken into account in the corresponding equations.

The interpreted results of the dissimilarity and the sensitivity are strongly related to their probability distributions. In this paper, we suppose that there does not exist any restriction in evaluation scores and values of the elements of $S_a$ and $S_b$ and then they are selected randomly from the uniform distribution. If some restriction exists in evaluation scores, the probability distributions of the dissimilarity and the sensitivity will change accordingly and new membership functions of the corresponding fuzzy values should be generated in order to guarantee the correctness of the interpreted results.

This principle of interpretation using fuzzy distances can also be applied to the analysis of terms used by the same panel and different panels.

### 4.3 Relationships Settings between Linguistic Terms

In industrial applications, there exists a strong need for interpreting evaluation terms of one panel using those of another panel. In this paper, we propose a genetic algorithm based procedure to do so. This procedure can be considered as a dictionary of terms for different panels and it is helpful for solving commercial conflicts between sensory panels related to the understanding of quality criteria. The details of this procedure are given as follows.

The sensory data of two panels $P_a$ and $P_b$ are obtained by evaluating the same set of representative samples denoted by $T$. The terms sets of $P_a$ and $P_b$ are denoted by $A_a = \{a_{a1}, a_{a2}, \ldots, a_{a,m(a)}\}$ and $A_b = \{a_{b1}, a_{b2}, \ldots, a_{b,m(b)}\}$ respectively. For each term $a_{ak}$ of $P_a$ $(k \in \{1, \ldots, m(a)\}$, we try to find the optimal linear combination of the terms $a_{b1}, a_{b2}, \ldots, a_{b,m(b)}$ to generate a new term denoted by $a(P_a, P_b, k)$ which is the closest to $a_{ak}$ in semantics, i.e. $a(P_a, P_b, k) = w_1^k \cdot a_{b1} + w_2^k \cdot a_{b2} + \cdots + w_{m(b)}^k \cdot a_{b,m(b)}$ with $\sum_{i=1}^{m(b)} w_i^k = 1$. The corresponding weights $\{w_1^k, w_2^k, \ldots, w_{m(b)}^k\}$ are determined using a genetic algorithm with penalty strategy [9] so that the distance between $a_{ak}$ and $a(P_a, P_b, k)$ is minimal. This optimization procedure is realized by performing the following steps:

Step 1: Finding the term of $P_b$ the closest to $a_{ak}$

Computing the distance between $a_{ak}$ and each term of $P_b$: $a_{b1}, a_{b2}, \ldots, a_{b,m(b)}$ using the method presented in Sect. 4.1. The corresponding values are denoted by $\{D_1, D_2, \ldots, D_{m(b)}\}$. Selecting the term $a_{bx}$ so that $D_x$ (the distance between $a_{bx}$ and $a_{ak}$) is the smallest of $\{D_1, D_2, \ldots, D_{m(b)}\}$.

Step 2: Building the support set of terms of $P_b$ related to $a_{ak}$

Building a set of new terms of $P_b$: $\{{a_{b1}}', \ldots, {a_{b,x-1}}', {a_{b,x+1}}', \ldots, {a_{b,m(b)}}'\}$ by adding the normalized evaluation scores of $a_{bx}$ for all samples of $T$ to those of each term $a_{bi}$ ($i \in \{1, \ldots, b(m)\}$ and $i \neq x$). Computing the distance between $a_{ak}$ and each of these new terms and denoting the corresponding results as $\{D_1^{'}, \ldots, D_{x-1}^{'}, D_{x+1}^{'} \cdots D_{m(b)}^{'}\}$. If $D_j^{'} < D_x$ ($j \in \{1, \ldots, m(b)\}$ and $j \neq x$), we consider that $a_{bj}$has a contribution to the decrease of the distance between $a_{bx}$ and $a_{ak}$ and then the corresponding weight $w_j^k > 0$. If $D_i^{'} > D_x$, we consider that $a_{bj}$ has no contribution to the decrease of the distance between $a_{bx}$ and $a_{ak}$ and then the corresponding weight $w_j^k = 0$. Therefore, we define the support set of terms of $P_b$ related to $a_{ak}$ by

$A_{_b}^k = \{a_b h_1, a_{bh_2}, \ldots, a_{bh_q}\}$ with $a_{bh_1} = a_{bx}$ and $a_{bh_j} \in A_b$ and $w_{h_j}^k > 0$ for $j \in \{2, \ldots, q\}$.

In this case, all terms of $P_b$ satisfying the condition $w_j^k$=0 are deleted from $A_b$ and only the relevant terms having contributions to the construction of the new term $a(P_a, P_b, k)$ are preserved in the support set $A_{_b}^k$. This step can largely reduce the computing complexity of Step 3.

Step 3: Building the term $a(P_a, P_b, k)$ that is the closest to $a_{ak}$

By applying a genetic algorithm, we compute the optimal weights of the relevant terms of $A_{_b}^k$ in order to construct the term $a(P_a, P_b, k)$ that is the closest to $a_{ak}$ in semantics. It is an optimization problem with constraints because the sum of the weights should be equal to 1 and each weight should be no smaller than 0. In this case, we use the penalty strategy [9] in the genetic algorithm. The detail for this algorithm is given as follows.

Procedure for computing the weights of the relevant terms using a Genetic Algorithm:

Begin

Coding and initializing the population of weights $W_{_b}^k(t) = (w_{_{h_1}}^k(t) w_{_{h_2}}^k(t) \cdots w_{_{h_q}}^k(t))$ $(t \leftarrow 0)$

IF $W_b^k(t)$ satisfies the constraints ($\sum_{j=1}^{q} w_{h_j}^k = 1$ and $w_{h_j}^k \geq 0$ for $j \in \{1, \ldots, q\}$)

THEN Evaluate the fitness by fitness function A
ELSE Evaluate the fitness by fitness function B
End_IF

While Not satisfying stop conditions

Do
Random Selection Operation
Crossover Operation

Mutation Operation
Updating $W_b^k(t)$ $(t \leftarrow t+1)$ for generating the next population of weights

IF $W_b^k(t)$ satisfies the constraints

THEN Evaluate the fitness by fitness function A
ELSE Evaluate the fitness by fitness function B
End_IF
End_While
End

The fitness function A is defined by $D_{ab}(a(P_a, P_b, k), a_{ak}) = \frac{2}{n(n-1)} \sum_{i<j} d_{ab}(a(P_a, P_b, k), a_{ak}, i, j)$ with $a(P_a, P_b, k) = (1 - \sum_{j=2}^{q} w_j^k) \cdot a_{bh_1} + \sum_{j=2}^{q} (w_j^k \cdot a_{bh_j})$ $\left(a_{bh_j} \in A_b^k\right)$.

under the constraints $\begin{cases} 1 - \sum_{j=2}^{q} w_j^k > 0 \\ w_j^k \geq 0 \quad (2 \leq j \leq q) \end{cases}$.

The concepts $D_{ab}$ and $d_{ab}$ are computed according to the definitions in Sect. 4.1.

We also define the fitness function B with penalty factor $\rho$ as follows:

$$D_{ab}(a(P_a, P_b, k), a_{ak}) = \frac{2}{n(n-1)} \sum_{i<j} d_{ab}(a(P_a, P_b, k), a_{ak}, i, j) + \rho$$

where $\rho = \gamma \sum_{j=2}^{q} w_j^k$. $\gamma$ is the parameter of penalty.

Figure 5 gives a practical example which recurrently computes the weights of the relevant terms of $P_b$ related to $a_{ak}$ using the genetic algorithm running for 100 generations. The evolution of the best value and the averaged value of the fitness function shows that the algorithm converges to its optimum after 20 populations. The best linear combination of the terms of $P_b$ related to the term $a_{ak}$ of $P_a$ is then obtained.

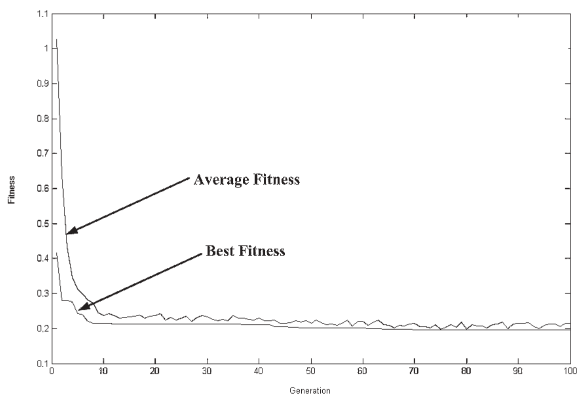


**Fig. 5.** Evolution of the fitness function for computing the weights (one example)

### 4.4 Forecasting the Consumers' Preference

Usually, the number of descriptors used by experts is greater than those of consumers. This lies in the fact the consumers' knowledge is more basic. But even if their experience on the product quality evaluation is poorer, their feeling can be included in the results of experts or trained panels. Thus, it means that relationships can be found between linguistic terms used by experts or trained panels and terms used by the consumers. Since the consumer evaluation matrix is computed so that each element denotes the population percentage, we need to adapt the relative variation between fabric samples $t_i$ and $t_j$ to compute the distances between DOSE and MOSE panels. The relative variation $vr_{\mathrm{MOSE}}(i, j)$ is now estimated by considering the averaged scores, computed with the grades values for a specific descriptor (see Table 1). For each consumer's term and each grade, we assign an absolute score according to the number of grades and the range [0, 1], since the sensory data are all normalized between 0 and 1. The average is then obtained by multiplying the population rate and the absolute score. For the considered term, we obtain a score which tends toward 1, if the major part of the consumers agrees with the descriptor. Then, using the procedure described in Sect. 4.3, we compute the relationships between the consumers' sensory evaluation and the terms used by DOSE panels. In the same way, we reproduce the same technique for extracting the relationships between the consumers' preference (vector $X$) and the DOSE-linguistic descriptors. These last relationships lead to a vector of weights, which characterizes the optimal linear combination of the $m(i)$ DOSE-terms for explaining the consumers' preference $X$: $(w_1^k, w_2^k, \ldots, w_{m(i)}^k)^T$ where $k$ equals 1, because the consumers' preference is there considered as only one descriptor and no more as a set of linguistic terms. Our target is to estimate what could be the consumers' preference for an additional $(n + 1)$-*th* product when its sensory evaluation is performed by the DOSE. Assuming the $x_i$'s vary in the range [0, 1], to find the preference $x_{n+1}$, we construct an array of all the preference vectors with an additional component which lies in [0, 1] and represents the possible preference. Then, we compute all the distances between the preference vector with additional components and the optimal linear combination of DOSE-terms for the new product. The predicted preference corresponds to the minimum of those distances, because it is the nearest distance between the optimal linear combination which models the preference and the forecasted preference. The Fig. 6 shows the curve obtained with varying consumers' preference. In this example, the nearest distance between the DOSE and the preference is about 13, which means that 13% of the consumers are expected to appreciate the product.

## 5 Product Evaluation and Process Models

In the product evaluation model (see Fig. 1), we have numerical input variables (selected physical features). Those input parameters, measured with

**Table 1.** Example of consumer sensory evaluation for a set of samples $T$

| | | Not 'soft' | Very Little 'soft' | Slightly 'soft' | 'soft' | Very 'soft' | Extremely 'soft' |
|---|---|---|---|---|---|---|---|
| | Absolute scores | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 1 |
| Product #1 | Population rate (%) | 9 | 16 | 18 | 27 | 27 | 3 |
| | Score ([0, 1]) | 0.09 | 0.16 | 0.18 | 0.27 | 0.27 | 0.03 |
| average = 0.09x0 + 0.16x0.2 + 0.18x0.4 + 0.27x0.6 + 0.27x 0.8 + 0.03x1 = **0.512** | | | | | | | |
| … | … | | | … | | | |
| Product #n | Population rate (%) | 1 | 7 | 5 | 45 | 32 | 10 |
| | Score ([0, 1]) | 0.01 | 0.07 | 0.05 | 0.45 | 0.32 | 0.10 |
| average = 0.01x0 + 0.07x0.2 + 0.05x0.4 + 0.45x0.6 + 0.32x 0.8 + 0.10x1 = **0.66** | | | | | | | |

appropriate devices, are chosen for their influence on the product quality. The Abe's method is then applied for extracting fuzzy rules directly from these numerical data [10] and to build the model between the physical measurements on the products and the sensory evaluation. This method permits to obtain a good compromise between precision, robustness and interpretability.

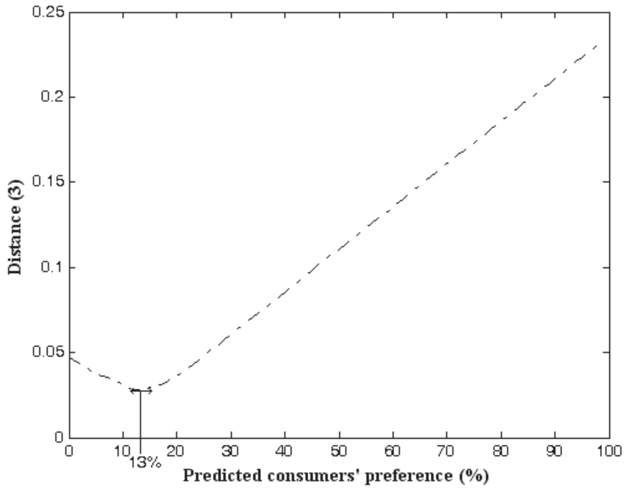


**Fig. 6.** Distance curve with varying consumers' preferences (percentage)

The Abe's method is briefly described below. At first, the universe of the output is divided into a number of intervals. By putting the input data into different classes according to the output intervals, we define two kinds of regions in the input space: activation hyperboxes and inhibition hyperboxes. For a given class of input data, an activation hyperbox contains all data belonging to this class and an inhibition hyperbox inhibits the existence of data for this class. Inhibition hyperboxes can be located by finding overlaps between neighboring activation hyperboxes. In these located inhibition hyperboxes can be defined new activation and inhibition hyperboxes for the next level. This procedure is repeated until overlaps are solved (see Fig. 7).

In this procedure, the fuzzy rules are defined by activation and inhibition hyperboxes (see Fig. 7). We select a suitable Gaussian function as membership function and calculate the output value using Sugeno's defuzzification method.

By comparison with other methods of fuzzy rules extraction which assume the space of input variables is partitioned into a number of fixed regions, this procedure generates more accurate fuzzy partition and fuzzy rules.

In practice, the fuzzy rules extracted using Abe's method are less efficient when the number of input variables is too great with respect to the quantity of available data. It is the case in many industrial applications. For solving this problem, we use Principle Component Analysis (PCA) [11] to reduce the number of input variables before starting the procedure of fuzzy rules extraction. By using this technique, the lower dimensional input space used in the Abe's method is obtained from the projection of the original high dimensional space. Its principle is given below.

PCA performs a linear transformation of an input variable vector for representing all original data in a lower-dimensional space with minimal information lost.

The principle of the product evaluation model for one descriptor is shown in Fig. 7. A very slight model internal parameters adjustment is required for any other descriptor and the same procedure can be repeated in order to build the product process model described in Fig. 1.

In the example described below (Sect. 6), experts obtain 7 levels for "Soft" and for the whole fabric samples. Their evaluation scores are taken as output data of the model. The data measured on 11 selected physical features are taken as input data after the projection of the original space into the two-dimensional subspace using PCA. Fuzzy rules are then directly extracted from these input-output learning data (see Fig. 8).

## 6 Application

In order to highlight and to illustrate the effectiveness of the above approach, we apply it to sensory data on fabric hand evaluation provided by 2 sensory panels in France, including a fashion design expert (FE) and a group of trained

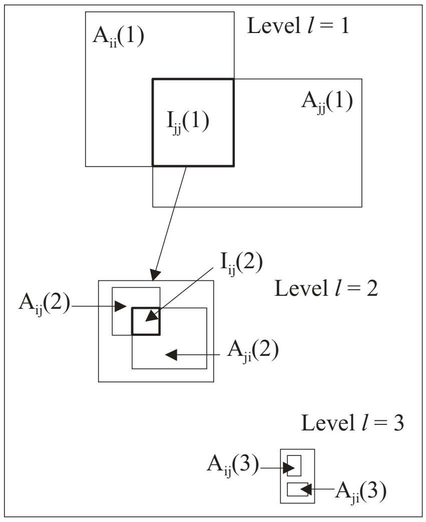


**Fig. 7.** Fuzzy rules extraction by generating Activation and Inhibition hyperboxes in the input space (2-D)

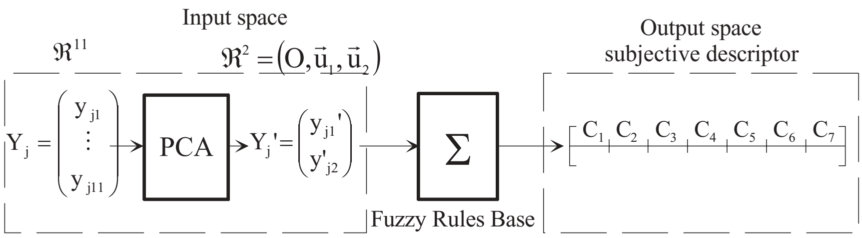


**Fig. 8.** Product evaluation model for one descriptor with 11 product features

students (FTS) and 2 sensory panels in China, including a group of textile experts (CE) and a group of trained students (CTS). The fabric set is composed of 43 knitted cotton samples produced using 3 different spinning processes. These samples can be then classified into 3 categories: Carded, Combed and Open-End, corresponding to different touch feelings. For FE, 11 evaluation terms have been used and they include “soft”, “smooth”, “tight”, “slippery”, “floppy”, “compact”, “hollow”, “pleasant”, “fresh”, “dense”, “flexible”. For FTS, they use 4 terms, including “smooth”, “slippery”, “soft”, “tight”. For CE and CTS, the terms used are not normalized and they vary with evaluator. In general, each CE uses 6 or 7 terms such as “soft, “slippery”, “flexible”, “texture”, “elasticity”, “thickness” and each CTS uses 3 or 4 terms such as “soft”, “slippery”, “flexible”.

### 6.1 Analysis at the Level of Evaluators

Applying the methods presented in Sects. 4.1 and 4.2 to analyze and interpret the evaluation scores given by these 4 evaluation panels, we obtain the corresponding crisp and fuzzy values of dissimilarity or distance and sensitivity in Table 1.

From Table 2, we can notice that the averaged distances between French and Chinese experts and between professional experts and students are very small. It means that the general evaluation on fabric hand is related to neither the cultural background nor the professional background. Moreover, since the experts' sensitivities are medium or large and the students' sensitivities are very small or small, it means that experts are more sensitive in the evaluation of products.

### 6.2 Analysis in the Level of Terms

Using the methods in Sects. 4.1 and 4.2, we also calculate the crisp and fuzzy distances between evaluation terms used by each panel. The averaged results are given in Table 3. From Table 3, we can notice that the crisp values of distances between different terms are very large or large for French experts (FE) and very small or small or large for the other panels. This means that French experts define more suitable terms for describing fabric hand and understand better their meaning in fabric hand evaluation. The overlap between two different terms is not important.

The results of crisp and fuzzy distances between different panels on the term "soft" are given in Table 4 and Table 5 respectively.

**Table 2.** Averaged distances and sensitivities for different panels

| | | FE | FTS | CE | CTS | |
|---|---|---|---|---|---|---|
| Distances | *Crisp#* | 0 | 0.1563 | 0.1284 | 0.1717 | FE |
| | *Fuzzy set* | VS | VS | VS | VS | |
| | *Crisp#* | 0.1563 | 0 | 0.1456 | 0.1692 | FTS |
| | *Fuzzy set* | VS | VS | VS | VS | |
| | *Crisp#* | 0.1284 | 0.1456 | 0 | 0.1622 | CE |
| | *Fuzzy set* | VS | VS | VS | VS | |
| | *Crisp#* | 0.1717 | 0.1692 | 0.1622 | 0 | CTS |
| | *Fuzzy set* | VS | VS | VS | VS | |
| Sensitivities | *Crisp#* | 0.4219 | 0.3635 | 0.4187 | 0.3977 | |
| | *Fuzzy set* | S 0;2833 | VS | S 0.5459 | VS 0.4900 | |
| | | M 0;7167 | | M 0.4531 | S 0.5100 | |

Table 4 and Table 5 show that the distances between different panels on the term "soft" are sometimes rather important (M, L, VL) although most of the fuzzy values of these distances are very small. The same phenomenon

**Table 3.** The averaged crisp and fuzzy distances between terms

<table>
<tr><th>PROCESS</th><th></th><th>FE</th><th colspan="2">CE</th><th>FTS</th><th colspan="2">CTS</th></tr>
<tr><td rowspan="3">Carded</td><td>Crisp#</td><td>0.2937</td><td colspan="2">0.2546</td><td>0.1844</td><td colspan="2">0.289</td></tr>
<tr><td rowspan="2">Fuzzy Set</td><td rowspan="2">VL</td><td colspan="2" rowspan="2">VS</td><td rowspan="2">VS</td><td>L</td><td>0.792</td></tr>
<tr><td>VL</td><td>0.208</td></tr>
<tr><td rowspan="3">Combed</td><td>Crisp#</td><td>0.2911</td><td colspan="2">0.291</td><td>0.1585</td><td colspan="2">0.2869</td></tr>
<tr><td rowspan="2">Fuzzy Set</td><td rowspan="2">VL</td><td>L</td><td>0.167</td><td rowspan="2">VS</td><td>L</td><td>0.967</td></tr>
<tr><td>VL</td><td>0.833</td><td>VL</td><td>0.033</td></tr>
<tr><td rowspan="3">Open-End</td><td>Crisp#</td><td>0.3708</td><td colspan="2">0.3779</td><td>0.2774</td><td colspan="2">0.3289</td></tr>
<tr><td rowspan="2">Fuzzy Set</td><td rowspan="2">VL</td><td colspan="2" rowspan="2">VL</td><td rowspan="2">VS</td><td>S</td><td>0.144</td></tr>
<tr><td>M</td><td>0.856</td></tr>
</table>

can be observed for the other terms such as "slippery" and "smooth". This means that one evaluation term is often semantically interpreted in different ways by different panels. This remark has been validated by some industrial companies. In these companies, there exist conflicts between suppliers and consumers on quality criteria expressed in linguistic terms. A dictionary is then needed for the understanding of evaluation terms between different professional populations.

The performance of each evaluator or panel can be characterized by the following criteria: 1) dissimilarity between terms he/she uses; 2) total number of terms used; 3) sensitivity to the products be evaluated and 4) stability or capacity of reproduction of evaluation scores. According to the previous analysis on the sensory evaluation data provided by our panels, we can see that the experts specialized in textile technology, especially the French expert (FE) are more efficient than the other panels in the evaluation of fabric hand. A sensory panel can also be trained according to these 4 criteria.

**Table 4.** The values of crisp distances between different evaluators on the common term "soft"

| Process | Carded | | | Combed | | | Open-End | | |
|---|---|---|---|---|---|---|---|---|---|
| Evaluators | CE | FTS | CTS | CE | FTS | CTS | CE | FTS | CTS |
| FE | 0.1343 | 0.2123 | 0.1743 | 0.2262 | 0.2311 | 0.2661 | 0.102 | 0.21 | 0.1803 |
| CE | | 0.2061 | 0.201 | | 0.1325 | 0.1643 | | 0.1803 | 0.1973 |
| FTS | | | 0.2189 | | | 0.1574 | | | 0.2874 |

### 6.3 Interpretation of the Relationship between Terms used by Different Panels

We use the method in Sect. 4.3 to interpret each term used by the panel of French trained students (FTS), the panel of Chinese experts (CE) and the

**Table 5.** The values of fuzzy distances between different evaluators on the common term "soft"

| | PROCESS | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Carded | | | Combed | | | Open-End | | |
| Evaluators | CE | FTS | CTS | CE | FTS | CTS | CE | FTS | CTS |
| FE | VS | VS 0.813<br>S 0.187 | VS | S 0.417<br>M 0.583 | S 0.417<br>M 0.583 | VL | VS | VS 0.031<br>S 0.969 | VS 0.969<br>S 0.031 |
| CE | | VS 0.813<br>S 0.187 | VS | | VS | VS | | VS 0.969<br>S 0.031 | VS 0.344<br>S 0.656 |
| FTS | | | VS 0.125<br>S 0.875 | | | VS | | | L 0.1<br>VL 0.9 |

panel of Chinese trained students (CTS) by those of the French Expert (FE). For simplicity, we only discuss the case of $P_a$ = FTS and $P_b$ = FE in this section. The 11 terms used by FE corresponds to $A_b = \{a_{b1}, a_{b2}, \ldots, a_{b,11}\}$ and the 4 terms of FTS to $A_a = \{a_{a1}, a_{a2}, a_{a3}, a_{a4}\}$. After applying the genetic algorithm with penalty strategy presented previously, we obtain the optimal linear combination of the terms of FE related to each term of FTS. The corresponding weights $w_j^k$'s of these linear combinations are shown in Table 6.

**Table 6.** The weights of optimal linear combinations of terms of FE related to those of FTS

| | **Terms of FE** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Terms Of FTS** | Soft | Smooth | Tight | Slippery | Floppy | Compact | Hollow | Pleasant | Fresh | Dense | Flexible |
| Smooth | 0.6 | 0 | 0 | 0 | 0.1 | 0 | 0.3 | 0 | 0 | 0 | 0 |
| Slippery | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Soft | 0.3 | 0 | 0 | 0 | 0.3 | 0 | 0.4 | 0 | 0 | 0 | 0 |
| Tight | 0 | 0.1 | 0 | 0 | 0 | 0 | 0.1 | 0 | 0 | 0.4 | 0.4 |

Table 6 permits to interpret the relationship between terms used by FTS and FE. Under this relationship, the terms used by FTS can be approximately expressed by linear combinations of the terms of FE. For example, the term *"soft"* used by FTS can be approximately expressed as a linear combination of three terms of FE: *"soft", "floppy" and "hollow"*, i.e.

$$soft\_FTS \approx 0.33 \cdot soft\_FE + 0.32 \cdot floppy\_FE + 0.35 \cdot hollow\_FE$$

Table 7 gives the distances between the terms of FTS and their corresponding optimal linear combinations of terms of FE, obtained using the procedure

**Table 7.** Evolution of distances between terms of FTS and FE

| | Distance between the term of FTS and the closest term of FE | | | Distance between the term of FTS and the optimal linear combination of terms of FE | |
|---|---|---|---|---|---|
| Term of FTS | *Crisp#* | *Fuzzy set* | | *Crisp#* | *Fuzzy set* |
| Smooth | 0.211 | **VS** | 0.431 | 0.156 | **VS** |
| | | **S** | 0.569 | | |
| Slippery | 0.161 | **VS** | | 0.161 | **VS** |
| Soft | 0.212 | **VS** | 0.367 | 0.131 | **VS** |
| | | **S** | 0.638 | | |
| Tight | 0.237 | **M** | 0.501 | 0.165 | **VS** |
| | | **L** | 0.499 | | |

in Sect. 4.3. These results show that the obtained optimal linear combinations are very close to the terms of FTS related to the original terms of FE. Each term of FTS corresponds to several terms of FE.

For the term "slippery" of FTS, there is only one term in the support set of terms of FE. So, the procedure described in Sect. 4.3 can not decrease the closest distance between terms of FTS and FE. For the other three terms, the procedure implemented in Sect. 4.3 can effectively decrease the closest distance between terms of FTS and FE. From the optimal linear combinations, we obtain new terms of FE much closer to those of FTS than the original terms of FE.

### 6.4 Interpretation of the Relationship between Terms used by Different Panels

In the modeling of fabric evaluation, we first use Principle Component Analysis (PCA) to obtain the reduced data and then extract fuzzy rules using the Abe's method. The corresponding results are given as follows.

The two components obtained from PCA are $\lambda_1 = 4.43$ and $\lambda_2 = 1.16$ with 46.9% and 16.7% of explanation respectively. For testing the effectiveness of the model, we remove at each time one sample from the learning base and we apply the Abe's algorithm with the remaining data for extracting the fuzzy rules base. By taking the removed sample as testing data, we compare the output estimated from the model and the real output of the removed sample. The corresponding results are shown in Tables 8 & 9.

From Table 8, we can see that the difference between the real output and the output estimated by the fabric evaluation model is rather small. For the descriptor "Soft", the maximum evaluation error doesn't exceed 2 marks over 7. For example, "0" means that the model gives the same score as the experts and "1" means that there exists only one level of difference between the model's evaluation and the expert's evaluation. The model also gives good results for the other samples and the other descriptors. In average,

**Table 8.** Fabric evaluation model results for the descriptor "soft"

| Removed Sample | # Rules | # Level(s) | Real Output | Estimated Output | Error |
|---|---|---|---|---|---|
| 4_1 | 8 | 2 | 7 | 5.83 | 1 |
| 10_1 | 8 | 2 | 7 | 4.95 | 2 |
| 14_1 | 8 | 2 | 6 | 3.50 | 2 |
| 14_2 | 7 | 1 | 4 | 4.50 | 1 |
| 16_1 | 8 | 2 | 6 | 4.25 | 2 |
| 16_2 | 8 | 2 | 4 | 3.74 | 0 |
| 22_1 | 8 | 2 | 6 | 3.50 | 2 |
| 22_2 | 8 | 2 | 4 | 4.02 | 0 |
| 24_1 | 7 | 2 | 5 | 3.27 | 2 |
| 24_2 | 8 | 2 | 4 | 3.51 | 0 |
| 26_1 | 8 | 2 | 3 | 3.44 | 0 |
| 26_2 | 8 | 2 | 3 | 3.71 | 1 |
| 28_1 | 8 | 2 | 3 | 3.85 | 1 |
| 28_2 | 8 | 2 | 2 | 2.89 | 1 |
| 31_1 | 8 | 2 | 2 | 2.49 | 0 |
| 31_2 | 8 | 2 | 2 | 2.71 | 1 |
| 34_1 | 8 | 2 | 1 | 1.65 | 1 |
| 34_2 | 8 | 2 | 1 | 1.20 | 0 |
| | | | | *average* | ***0.94*** |

by considering 9 linguistic terms all together for describing the touch handle of the whole samples set, the evaluation error is about 1 point (1.11 – Table 9). This represents the fabric evaluation model accuracy.

### 6.5 An Example of Consumers' Preference Forecasting

To illustrate the consumers' preference prediction, we collected one set of sensory data related to the cosmetic industry. It includes 8 lotions with varying performance according to their interaction with the skin or the human feeling they express at the first contact. 19 terms are required by the experts for describing the lotions quality. The consumers expressed their feeling with two linguistic terms: the softness and the touch feeling. They ranked also the lotions according to their preference. The running of the procedure described above in Sect. 4.3 leads to the weights set of Fig. 9. Then, applying the method of Sect. 4.4., we are able to predict the consumers' preference (see Table 10). The forecasting works quite well even if sometimes the error raises 15% over 100%. This drawback is due to the size of the training data set base which contains only 8 samples. For testing the methods, we used 7 samples and tried to predict the preference of the 8-th.

**Table 9.** Results of fabric evaluation model for 9 linguistic terms used by a French Experts panel

| | Tight | Shiftless | Smooth | Compact | Weak | Pleasant | Fresh | Heavy | Pliant |
|---|---|---|---|---|---|---|---|---|---|
| Removed Sample | | | | | Class Error | | | | |
| 4_1 | 1 | 1 | 2 | 1 | 4 | 2 | 2 | 0 | 1 |
| 10_1 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 2 |
| 14_1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| 14_2 | 1 | 1 | 1 | 1 | 4 | 1 | 1 | 1 | 2 |
| 16_1 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 0 | 2 |
| 16_2 | 2 | 2 | 1 | 3 | 0 | 3 | 1 | 1 | 2 |
| 22_1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| 22_2 | 1 | 1 | 1 | 1 | 3 | 2 | 1 | 0 | 2 |
| 24_1 | 0 | 3 | 1 | 0 | 0 | 2 | 1 | 0 | 1 |
| 24_2 | 2 | 1 | 1 | 3 | 0 | 1 | 1 | 0 | 2 |
| 26_1 | 1 | 1 | 1 | 0 | 2 | 0 | 1 | 2 | 0 |
| 26_2 | 1 | 2 | 1 | 0 | 1 | 0 | 1 | 0 | 3 |
| 28_1 | 0 | 2 | 1 | 0 | 0 | 0 | 1 | 1 | 0 |
| 28_2 | 2 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 3 |
| 31_1 | 2 | 3 | 1 | 1 | 0 | 0 | 1 | 2 | 1 |
| 31_2 | 1 | 1 | 0 | 2 | 3 | 0 | 1 | 2 | 3 |
| 34_1 | 1 | 1 | 1 | 2 | 0 | 0 | 1 | 2 | 1 |
| 34_2 | 2 | 1 | 1 | 2 | 1 | 0 | 1 | 1 | 3 |
| Total average: | | | | | errors (average): | | | | |
| ***1.11*** | **1.17** | **1.33** | **1.00** | **1.11** | **1.11** | **0.83** | **1.06** | **0.72** | **1.66** |

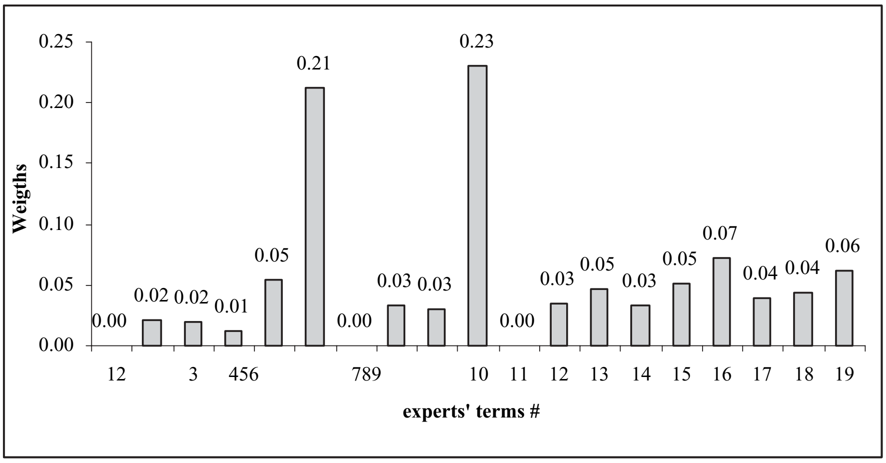


**Fig. 9.** Weights values for explaining the relationships between consumers' preference about the softness and experts' terms

**Table 10.** Difference between forecasted and real consumers' preference for 8 lotions

| Expressed (%) | Forecasted (%) | Error (%) |
|---|---|---|
| 14 | 17 | 3 |
| 9 | 10 | 1 |
| 16 | 7 | 9 |
| 16 | 19 | 3 |
| 18 | 7 | 11 |
| 4 | 20 | 16 |
| 10 | 9 | 1 |
| 13 | 18 | 5 |

## 7 Conclusions

This paper presents a general method for analyzing and interpreting sensory data given by different panels. The 2-tuple linguistic model is used for normalizing and aggregating sensory data of different individuals inside each panel on an optimal unified scale. The dissimilarity criteria and the sensitivity criteria are transformed into fuzzy numbers in order to obtain a suitable physical interpretation, leading to a better understanding of the quality of panels and evaluation terms. Also, we propose a procedure permitting to interpret terms of one panel using the linear combination of terms of another panel. The optimal weights of this linear combination are obtained using a genetic algorithm with penalty strategy. This procedure is particularly significant for solving commercial conflicts related to the understanding of product quality criteria expressed in linguistic terms. The proposed method has been successfully applied to the analysis and the interpretation of the sensory data on fabric hand evaluation provided by four panels.

This paper deals mainly with the introduction of intelligent methods for both formalizing sensory data, which are expressed by human being, and modeling the relationships between these sensory data and objective measures operated on the fabrics. It gives promising results for assessing the sensory quality of industrial products from manufacturers to consumers. Another contribution in this paper is that we used PCA to project original higher dimensional data into a lower dimensional subspace before starting the procedure of fuzzy rules extraction. In general, a fuzzy model is efficient only when the number of input variables is small enough with respect to the number of learning data. We have to reduce the number of input variables if we can not measure more numerical data.

## References

1. H. Stone, J.L. Sidel, Sensory Evaluation Practice, Academic Press Inc., 1993. 375
2. G.B. Dijksterhuis, Multivariate Data Analysis in Sensory and Consumer Science, Food & Nutrition Press Inc., Trumbull, Connecticut, USA, 1997. 375
3. J.R. Piggot, E.A. Hunter, Evaluation of assessor performance in sensory analysis, Italian Journal of Food Science, vol. 11, no. 4, pp. 59–64, 1999. 375
4. Bin Zhou, Xianyi Zeng, Ludovic Koehl, Yongsheng Ding, A 2-level Model for Description of Fabric Hand Sensory Evaluation, Int. Conf. World Textile Conference – 4th AUTEX Conference, Roubaix, France, June 22–24, 2004. 376
5. Kawabata S. and Niwa M., Objective measurement of fabric hand, Modern Textile Characterization Methods (Eds. M. Raheel and M. Dekker), 1996, 329–354. 376
6. Hu J., Vhen W., and Newton A., A psychophysical model for objective fabric hand evaluation: An application of Steven's law, *J. the Textile Institute*, 1993, 84(3): 354–363. 376
7. X. Zeng, Y. Ding and L. Koehl, A 2-tuple fuzzy linguistic model for sensory fabric hand evaluation, in Intelligent sensory evaluation, D. Ruan and X. Zeng, Eds. Berlin: Springer, 2004, pp. 217–234. 377, 383, 384
8. F. Herrera, L. Martizez, A model based on linguistic 2-tuples for dealing with multigranular hierarchical linguistic contexts in multi-expert decision-making, IEEE Trans. Systems, Man, and Cybernetics – Part B: Cybernetics, vol. 31, no. 2, pp. 227–234, 2001. 377, 380
9. F. Back, Evolutionary Algorithms in Theory and Practice, Oxford University Press, New York, 1996. 386, 387
10. S. Abe and M. Lan (1995). Fuzzy rules extraction directly from numerical data for function approximation, IEEE Trans. SMC 25(1), 1995, pp. 119–129.
11. K. Fukunaga (1990), Introduction to Statistical Pattern Recognition, 2nd ed. San Diego, CA: Academic 391