

Visual inspection of products: a comparison of the methods used to evaluate surface anomalies

N. Baudet*, M. Pillet**, and J.L. Maire***

Université de Savoie, Laboratoire SYMME, Annecy, France

Received: 28 February 2011 / Accepted: 3 March 2011

Abstract. This paper presents and compares different approaches currently used to assess surface anomalies identified on a product. The common point between these methods is that they are based on a document presented in the form of a table, which is to help the inspector to assess the anomaly detected in a repeatable and reproducible way. We will present three types of table: criteria/level table, tree-like presentation table and an indexed table. As each of these tables presents certain limits when applied to the inspection of a product surface, we will describe the table proposed in order to help the inspector determine the intensity to attribute to the identified anomaly.

Keywords: Surface anomaly; visual inspection; evaluation table; sensorial analysis

1 Introduction

The quality¹ of a product is generally defined as the ability of a set of product characteristics to satisfy the customer's requirements [1]. Obviously, these characteristics are not limited to the functional characteristics of the product. They also include the aesthetic characteristics identified on the product surface and which make up its appearance. The latter category is having an increasing impact on how the client perceives the quality of the product, so much so that the product appearance often dictates whether the client buys the product [2]. For example, the client no longer expects just a rolling ball cage to meet the defined functional tolerances but also to have a perfect appearance. This contributes to improving the perceived quality of the product and to assure the client that this product will meet his requirements even better. Indeed, for Dickson [3], the quality of a product is defined by something intangible, that is difficult to explain, and often the physical appearance of the product is at its source. Above and beyond buying a product of aesthetic value, the appearance of a product also indicates the quality of all its characteristics [3].

The obvious question that industry is asking is: “how do we manufacture a product which has a perfect appearance?” Another dimension to the question is also to know how to detect products with a less-than-perfect

appearance which could lower the quality of the product in the eyes of the client. The problem is difficult to resolve since surface inspection is done by inspectors who are not always in a position to detect the surface anomaly and, when they do, sometimes have great difficulty in evaluating it correctly. An additional problem is the particularity that the inspection is carried out by just one inspector. Measuring instruments for this type of control do not exist. Following the same model as the one for the inspection of the product dimensions, which methods can be used to carry out the visual inspection of the products? How can the subjectivity of the assessment of surface anomalies be reduced? Which methods guarantee the repeatability and reproducibility of the results of the visual inspection? How can we manage to reduce the variability in the judgment of the operator as he/she carries out the inspection?

The objective of our work, which is part of the INTERREG IV, a France-Switzerland 2007–2013 research programme, is to succeed in developing a method dedicated to the visual inspection of products. This programme is specially targeting products with high added value. It has brought together the University of Savoy, the Federal Polytechnic School in Lausanne, the Screw machining Technical Centre, the Watchmaking and Jewellery Technical Centre, as well as six French and Swiss companies.

This article presents and compares different approaches currently used to assess surface anomalies identified on a product.

The common objective between these approaches is to quantify the intensity of an anomaly in order to help the operator to evaluate the conformity or non conformity of a product, using the tolerances defined. The examples

* Correspondence: nathalie.baudet@univ-savoie.fr

** Correspondence: maurice.pillet@univ-savoie.fr

*** Correspondence: jean-luc.maire@univ-savoie.fr

¹ This research is supported by the European INTERREG IVA France-Switzerland programme.

presented in this article to illustrate the different approaches are taken from the work currently taking place at Dupont, a luxury product manufacturer, and Fournier, a manufacturer of bathroom and kitchen furniture.

2 Sensorial analysis for the evaluation of surface anomalies

When visual inspection is carried out, many surface anomalies can be detected. Indeed, a perfect product with zero surface anomalies does not exist because, at a given level of magnification, the controller will detect most cases of deviations from the ideal surface. Therefore he must have a method to describe and quantify the anomaly in order to decide if the defect should be judged as acceptable or unacceptable. In order to do this, the tolerance limits for each possible defect must be established. Of course these limits can vary according to where the anomaly is on the product: either on the front or on the back of the product, for example.

We have defined two types of surface anomalies. The anomalies that do not “evolve” are perceived in exactly the same way at two different moments in the life of the product. This is the case for a scratch detected on the product by the inspector, and whose length and depth will not change over the course of time and which will be perceived in exactly the same way by the client. The anomalies which evolve over time belong to the second type. Such is the case for a point on the product that is slightly oxidized. It may be considered as “not critical” by the quality inspector, but perceived as unacceptable by the client later due to the further oxidization over time. An evolving anomaly is a product defect, whatever its level of detection. In our works, we only deal with the non-evolving anomalies. From this category, and contrary to the evolving anomalies, the tolerance levels and limits can be defined and fixed.

Sensorial analysis [1, 4, 5] is an approach which helps to qualify and quantify the intensity of the sensorial perception. It has long proved its usefulness in the food industry to express the subtle differences in human perception (taste, touch, hearing, smell or sight) in an identical way as if one was using a measuring instrument. More often than not, sensorial analysis is used when perfecting a new product. Trained inspectors must describe the product samples objectively and assess the similarities and differences between many references. The results allow one to compare different products, to assess the effect of a modification of the production process on the sensorial qualities of the products, etc. Sensorial analysis is also used for the launching of a new product or when one wishes to position a product in comparison to those of the competitors. It can also be used as a quality control tool to allow one to check the sensory stability of a product over time. In contrast, we did not find any examples of the use of sensorial analysis in the visual inspection of products in previous research. However, this method seems to be perfectly suitable to describe a surface anomaly and to quantify its intensity.

With sensorial analysis, each product is first described using a set of descriptors, with each descriptor defining a perceived characteristic of the product. These descriptors must respect certain conditions [5]. First of all, they must be relevant; that is to say, relate a specific characteristic of the product. Then, they must be easily both understood and identified with the same characteristic of the product by all the inspectors. The descriptors must also cover the whole set of characteristics of the product to be examined in an exhaustive way. Finally, they must be independent of each other, each one reflecting a different characteristic of the product.

In the case of visual inspection, it is the surface anomalies of the product, and not the product itself that are described with the help of the descriptors.

As one is advised to do in sensorial analysis, the list of descriptors must be reduced to a minimum [5, 6]. It is also essential that the inspectors use the same vocabulary to describe the same anomaly. For example, the terms scratch, cut or pits can be replaced by the unique term: Mark. In the case of a visual inspection, we consider that the descriptors Mark, Distortion, Pollution (foreign body, dust, etc.) and Heterogeneity (problem of shade, of texture, etc.) suffice to cover all the surface anomalies that we could encounter on a product [5].

The evaluation criteria must then be defined for each of the descriptors, in agreement with experts. These criteria include those used to characterize the anomaly (for example, the mark descriptor can be characterized with the help of criteria such as length, depth and density) and/or criteria linked to the perception of the anomaly (for example, the Effort to make to perceive the mark) [1, 5]. Certain criteria can be more important than others, so it is necessary to carefully examine the reasoning behind the experts' choices in order to characterize an anomaly.

In sensorial analysis, each descriptor is then quantified with the help of a measure of intensity on a graded or structured scale. In our case, this means that it is necessary to define a set of levels of intensity for each criterion of an anomaly descriptor, to define the appropriate semantic expression associated with each of these levels and to determine a method of calculation to establish the overall intensity of the anomaly contingent on the values of the criteria of the descriptor.

The following paragraph lists the different approaches possible in order to quantify the intensity of a surface anomaly.

3 Different approaches to quantify the intensity of a surface anomaly

As we have mentioned previously, one of the main difficulties with visual inspection is the subjectivity that can creep into the inspector's evaluation of the anomaly. The subjectivity has an impact on the repeatability and the reproducibility of the results of these inspections. In order to overcome this difficulty, the classic approach consists in having different inspectors as well as a reference document to help them quantify the intensity of the detected

Table 1. An example of a criteria/level table.

Intensity	Scan rate	Effect of light	Tolerance of the anomaly	Contrast	Perception compared to the function
1	No anomaly visible under standard scanning				
2	Requires a thorough scan and a local effort to find the anomaly	Visible in a very particular angle (one effect of light)	The anomaly is drowning in the background, likes natural. The direction and intensity (length, depth) is comparable to the rest of the decor	No color contrast	
3	Easily visible scan local	Visible in two effects of light	The anomaly is in the sense of decor, but is not natural, does not reflect the wish of the designer	Low color contrast	
4	Easily visible in standard scanning	Visible in two effects of light + “catches the eye”	The anomaly shocked by its direction or its intensity the decor is altered	High contrast	Perception that the function or durability is maintained.
5	The defect appears instantaneously scanning without special	All effects of light	The anomaly shocked by its direction or its intensity the decor is altered	High contrast	Perception that the function or durability is altered
6	Instantly visible on a priority before everything else	All effects of light	The anomaly is extremely shocking	High contrast	Non-conformity compared to the product definition (lack of material, coating ...), deterioration making it unsuitable

anomaly in both a repeatable and a reproducible way. We will now present some of the forms this document can take:

- the criteria/level table;
- tree-like evaluation table [5];
- the indexed evaluation table [7].

These three approaches provide a quantification of the intensity perception of the surface anomaly in a scale between 0 to 6 in the two first forms and 0 to 23 for the last one.

3.1 The criteria/level table

Table 1 gives an example of an evaluation table that we call “criteria/level”. This example has been developed by the SYMME research laboratory in order to quantify a surface anomaly identified on certain watch components.

This table crosses each level of intensity for the anomaly (in this example, between 1 and 6) and the values of the criteria decided on in order to characterize the anomaly (in this example, the method of examination to perceive the anomaly, the effect of light to make it visible, the tolerance of the anomaly which characterizes the link between the anomaly and its surrounding environment, the level of contrast in colour generated by the anomaly, and the relationship between the perception of

the anomaly and the function expected from the component inspected). The inspector needs to get his perception of the anomaly to correspond to the different values described in each line of the table. If he manages to do this, he is then able to deduce the value of the intensity to assign to the anomaly.

The advantage of this table is that it clarifies what the different criteria to use to characterize the anomaly are. It also allows one to describe the different values required for each of the criterion during the period of observation. Lastly, it allows a precise link to be made between what is perceived and the level of intensity to attribute to the anomaly. It also limits the subjectivity associated with determining the value of the criterion as the latter can be confirmed by observing the values of the other criteria which are on the same line.

For example, if the inspector perceives the anomaly as instantly visible, his/her perception will be confirmed as accurate if he/she perceives the anomaly under all the different effects of light and if he/she perceives a major contrast in colour (line 6 of the table).

In contrast, this table supposes that each time the inspector perceives something, that it coincides perfectly with the description on one of the lines in the table, i.e. all the values observed for the different criteria are necessarily and only situated on the same line. It also presumes that the inspectors have a good understanding of the criteria and have had a sufficient amount of training.

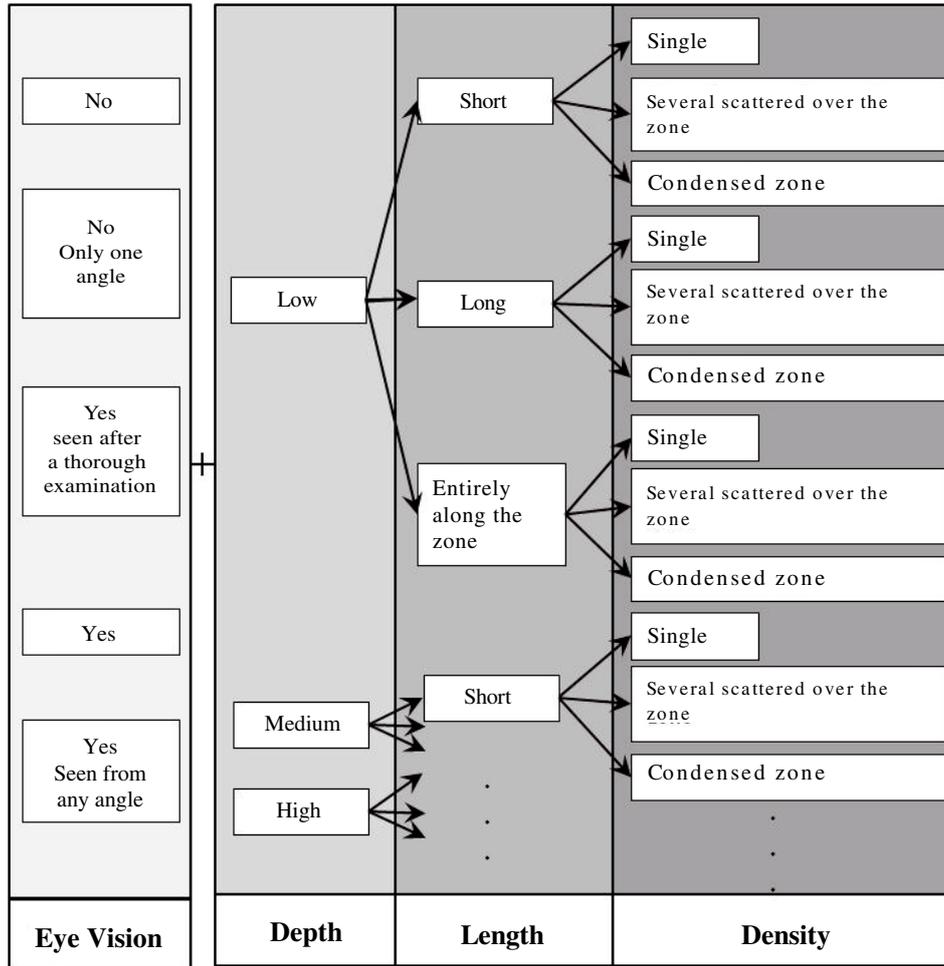


Fig. 1. Tree-like representation of the criteria for a mark-type anomaly.

3.2 Tree-like evaluation table

Figure 1 gives an example of a tree-like form of evaluation [7]. The aim of the tree-like shape is to convey the experts' hierarchical reasoning.

In order that the workers and inspectors get a better understanding we will only keep the possible combinations of criteria that will lead to a table which will define the levels of intensity for a given anomaly (Tab. 2) [7].

Like the previous table, this one creates a link between each level of intensity of the anomaly (in the example from 1 to 6) and the criteria retained in order to characterize the anomaly (the example shows the naked eye vision of the anomaly, the depth of the mark, its length and its density).

On the other hand, and contrary to the criteria/level of intensity table, this table allows several combinations of the values of the criteria for the same level of importance. In the example, it is the case for level 4, which is associated with both a shallow mark and a fairly deep mark in the case where this mark is short and located in a single place.

However, the main advantage of this form of table is to offer the inspector an operating procedure for his/her

observation. Thus, in the example, the inspector is led to consider (in the right order) the visual impact, the depth, the length and the density of the mark. Each value that is attributed to a criterion reduces the number of possibilities that the other criteria have to be used to examine.

Setting up such a table implies that a tree-like representation of the values of the different criteria has been created beforehand, and that all the values possible for these criteria have been considered.

One of the limits presented in this table comes from its tree-like representation which limits the number of criteria. Indeed, adding a criterion significantly multiplies the number of combinations of the values of the criteria to consider. In the example, we have 7 combinations with two criteria, but 22 combinations with three criteria and more than 60 combinations with four criteria.

3.3 The indexed evaluation table

Table 3 gives an example of an indexed evaluation table. The table presented and developed by QIM Eurofish [4], is used with the aim of helping the inspector to evaluate the freshness of a fish.

Table 2. An example of a table using a tree-like representation.

Intensity	Naked eye vision	Depth	Width	Mark	Density
1	No	Low	Short		Single
2	No Only one angle	Low	Short		Several scattered over the zone
3	Yes Seen after a thorough examination	Low	Long		Single
		Medium	Short		Single
		Medium	Short		Several scattered over the zone
4	Yes	Low Medium	Short Short		Condensed zone Condensed zone
5	Yes	Low	Entirely along the zone		Single
		High	Short		Single
6	Yes Seen from any angle	Medium	Long		Single
		Medium	Entirely along the zone		Single
		High	Short		Several scattered over the zone
		High	Short		Condensed zone
		High	Long		Single
		High	Entirely along the zone		Single

This approach consists in listing the different evaluation criteria, listing the different characteristics for each of these criteria and listing the different possible values for each of the characteristics. For example, the criterion eyes is characterized by a pupil being black, opaque or grey. A score is associated with each of the values for this criterion (10 criterion et 33 values, in this example) and the index attributed to the anomaly corresponds to the sum of these scores (from 0 to 23 in the example).

The advantage of this table is that it lists the whole set of possible observations that an inspector could make for a given anomaly; therefore this limits the inspector's subjectivity which appears when he/she is confronted with an anomaly that is listed.

3.4 Proposition of a corrected hierarchical evaluation table

Within the framework of the INTERREG IV A France-Switzerland project, a comparative study of different evaluation tables has been carried out using the criteria/level table, tree-like representation table and indexed table.

This study consisted in using each of the approaches in parallel and to measure their respective effects on the variability of the results of the inspections carried out by different inspectors.

This study allowed us to confirm and/or to put forward:

- the existence of a certain hierarchy in the criteria taken into account;
- the difficulty experienced by the expert in deciding on the level of intensity to attribute to an anomaly in the case of a contradiction among different criteria (for example, *Contrast* and *Effort needed to detect* in the example presented in Tab. 1);

- the difficulty in reproducing the expert's reasoning using a tree-like presentation of the criteria, as shown in Table 2;
- the sometimes tedious task of using an indexed table, as presented in Table 3.

Thanks to these points, we have deduced what we should expect from a table whose aim is to formalize the sensorial perception of the quality of a surface. The table to be used must:

- take the hierarchical character of the criteria;
- convey the reasoning behind the expert's choices (especially because some criteria can worsen the situation);
- be simple so the inspectors can use and memorise the table easily;
- be exempt from the combinations of values of criteria which could lead to a contradiction;
- be able to be translated easily into different languages.

In order to meet these requirements, Table 4 shows the table that we propose and that we call the "corrected hierarchical table".

The hierarchical nature of this table is given by the effort needed to detect the anomaly which, using the opinion of the set of experts in visual inspection, has a preponderant role in the making of the decision. First of all, the inspector evaluates the effort he/she needed to produce in order to perceive the anomaly. Therefore he/she must take a value in the first column of the table where these values match to his/her effort.

This first evaluation will then be corrected by a worsening or a reduction of the anomaly, depending on the different criteria, following assessment of all criteria, the controller found in the last column of the table, the final intensity of the anomaly or in this case the index of anomaly gravity (1 to 6 in this example).

Table 3. An example of an indexed table (www.qim-eurofish.com).

Quality parameter	Description	Score	
Appearance	Skin	Bright, indescent pigmentation	0
		Rather dull, becoming discolored	1
		Dull	2
	Stiffness	In rigor	0
		Firm, elastic	1
		Soft	2
	Very Soft	3	
Eyes	Cornea	Clear	0
		Opalescent	1
		Milky	2
	Form	Convex	0
		Flat, slightly sunken	1
		Sunken, concave	2
Gills	Pupil	Black	0
		Opaque	1
		Grey	2
	Color	Bright	0
		Less colored, becoming discolored	1
		Discolored, brown spots	2
	Brown, discolored	3	
Flesh, fillets	Smell	Fresh, seaweedy, metallic	0
		Neutral, grassy, musty	1
		Yeast, bread, beer, sour milk	2
	Mucus	Acetic acid, sulphuric, very sour	3
		Clear	0
		Milky	1
Color	Milky, dark, opaque	2	
	Translucent, bluish	0	
	Waxy, milky	1	
	Opaque, yellow, brown spots	2	
Blood	Color	Red	0
		Dark red	1
		Brown	2
Quality index		0–23	

For example, if we take the tolerance of the anomaly of the frontage of a building. The evaluation of the mark following the grain of the wood will be less reduced whereas it will worsen if this mark appears perpendicular to the grain of the wood.

The vocabulary used in the table is basic on purpose in order to satisfy the need to translate the table into different languages easily. However, each term in the table (for example, weak or strong for the size) must be explained reference documents and appropriate training. In the case we have studied, in order to evaluate the perception of the size, the experts have defined an acceptable length for a rather long mark (scratch) and an acceptable diameter for a rather circular mark, wherever the mark may be on the product.

This method is simple. It can take several combinations of criteria into account, which the previous tables could not do. Another advantage of this table is that it also allows one to follow the inspector's reasoning when he/she evaluates the anomaly. For example, if an anomaly is visible only under a particular angle, if the anomaly hidden in the background of the workpiece, if its size is important and the mark is shallow, the inspector will write: D1, T0, G1, R0, IG2 in his/her table.

With this type of approach, it is possible to both reduce the subjectivity of the inspector's evaluation and to assure oneself of the consistency of the reasoning leading to the decisions over time and between the inspectors. In order to do this, we carry out an R&R test on the products identified and for which we follow the reasoning that leads to the evaluation.

4 Discussion of the approach

In order to evaluate the result of the improvement brought about by the use of this table, we carried out a study in one of the partner companies in the INTERREG project, whose objective was to quantify the agreement among 5 experts on the evaluation of 18 products which only have one surface anomaly.

Table 5 shows the results obtained from this study before as well as after the use of this table.

Before the use of corrected hierarchical table, a test using the 18 products was carried out with the aim of verifying the agreement of the 5 experts. Each product had just one anomaly, which was clearly defined beforehand. The experts were asked to judge the conformity of the product in relation to the anomaly identified. We then calculated the percentage of products for which the responses were identical. Only 22% of the products led to an agreement among the evaluations carried out by each expert independently. This first test reveals a real problem in the harmonising of the decisions. As the decision leading to conformity is not formalized, the criteria D, T, G and R were not filled in.

Once we had defined the criteria and the levels of intensity and the corrected hierarchical table had been designed, we carried out a R&R test again. During the test, the experts were asked to be precise regarding the value associated with each of the criteria to be examined, the level of intensity they attributed to the anomaly and the decision made concerning the conformity of the product. We then calculated the percentage of anomalies which obtained the same level of intensity in the end.

The results show a significant improvement in the agreement among the experts (it increased from 22% to 46%). Going beyond the actual result of the test, it should be noted that the formalisation of the evaluation process contributed to making the discussion easier among the experts on the disagreements identified at the end of their evaluation. This last point is very important, especially within the context of a contractual framework established between customers and suppliers.

Table 4. Corrected hierarchical table.

D	T	G	R	IG
Detection effort	Tolerance of the anomaly	Size perception	Relief perception	Index of anomaly gravity
1 Visible under a particular angle and hard to find	0 (zero) if it drowned in the background, like natural 	0 (zero) if it is low 	0 (zero) if it is shallow 	1
2 Visible in several angles				2
	1 (one) decor contrast does not seem natural 	1 (one) it is high 	1 (one) if it is marked 	3
4				
3 Visible in all angles	1 (one) if it is marked 	1 (one) if it is marked 	1 (one) if it is marked 	5
				6

Table 5. Results of the degree of agreement between the experts (in percentage).

	Without the table	With the table
D	–	75%
T	–	76%
G	–	79%
R	–	87%
Total	22% ^a	46% ^a

^a Percentage calculated using the average of the agreement between five experts concerning the conformity of the products.

However, the result obtained thanks to this table must be improved or even more. This is the aim of our current research, which is especially concentrating on a better definition of the terms used in the table (shallow marked ...).

In order to satisfy the criteria of simplicity and independence in relation to the language used, we think it is not wise to increase the complexity of the table.

Each criterion to be examined is the object of specific training given to the inspectors so as to limit the difference in the way words are interpreted.

Each inspector will have to do a training course on using the table. Our objective is to reach 80% percent agreement among the five experts, which means 90% agreement for each of the criterion.

5 Conclusion

In this article, we have emphasized the necessity to rely on a formalized table in order to limit the variability of the results often observed when a visual inspection of products is carried out.

The three evaluation tables that we have presented (criteria/level table, tree-like representation table and indexed table) can be used for this type of control, and each has their own limitations.

The criteria/level table is by far the least stable of the three tables. In order to avoid confusion which could arise between the different levels of intensity possible, we especially need to remember that it requires terminology that is often too complex and therefore often very difficult to be mastered by all the inspectors.

As for the tree-like representation table, its main drawback is that it limits the number of criteria one can consider during an evaluation since the addition of just one criterion increases the number of possible combinations of the values of the criteria to be examined by the inspector far too much.

Admittedly more complete than the other two tables, in terms of the criteria to examine during the evaluation, the indexed table is however rather tedious to use and therefore not very suitable for the evaluation of the intensity of a surface anomaly.

In order to overcome these difficulties, we have developed a corrected hierarchical table. For the moment, it appears a lot more suitable for the visual inspection of products than the previous tables. Here is what the first results of this table applied at one of the companies with whom we are working seem to suggest.

These results will probably improve once the inspectors have been fully trained using this table, as well as at the end of the work which will be done so as to obtain an even more accurate definition of the criteria used in this table.

Looking to the future, it will no doubt be necessary to test the robustness of the corrected hierarchical table in other companies which carry out visual inspection by evaluating the intensity of the anomaly.

References

1. ISO 9000:2005
2. A. Lefebvre, J.F. Bassereau, L'analyse sensorielle, une méthode de mesure au service des acteurs de la conception: ses avantages, ses limites, ses voies d'amélioration. Application aux emballages, 10^e Séminaire CONFERE (Belfort, France, 2003), pp. 3–11

3. M.E.H. Creusen, J.P.L. Schoormans, The different roles of product appearance in consumer choice, *J. Prod. Innov. Manag.* **22**, 63–81 (2005)
4. E. Costell, A comparison of sensory methods in quality control, *Food Qual. Pref.* **2**, 341–353 (2002)
5. I. Etaio, M. Albisu, M. Ojeda, P.F. Gil, J. Salmeron, F.J.P. Elortondo, Sensory quality control for food certification: A case study on wine. Method development, *Food Control* **21**, 533–541 (2010)
6. T. Debrosse, M. Pillet, J.L. Maire, N. Baudet, Sensory perception of surface quality – Industrial Practices and Prospects, *Proceedings of the International Conference on Kansei Engineering and Emotion Research* (2010, Paris, France)
7. A.G. Guerra, *Métrieologie sensorielle dans le cadre du contrôle qualité visuel*, Doctoral thesis, US, University of Savoy, 2009
8. S. King, M. Gillette, D. Titman, J. Adams, M. Ridgely, The sensory Quality System: a global quality control solution, *Food Qual. Pref.* **13**, 385–395 (2002)
9. A.M. Muñoz, G.V. Civille, B.T. Carr, *Sensory evaluation in quality control* (Van Nostrand Reinhold, New York, 1992)
10. A.M. Muñoz, Sensory evaluation in quality control: an overview, new developments and future opportunities, *Food Qual. Pref.* **13**, 329–339 (2002)