Using Soft Computing for a Prototype Collagen Plate Inspection System

M. Köppen, A. Soria-Frisch and Th. Sy
Fraunhofer IPK
Dept. Security and Inspection Technologies
Pascalstr. 8-9, 10587 Berlin, Germany
{mario.koeppen|aureli|thorsten.sy@ipk.fhg.de}

Abstract- A framework for the automated visual inspection of collagen plates was implemented and used in an industrial application for the evaluation of fault perceptual relevance. The here presented paper analyzes the methodologies used in this framework in order to predict the end user's opinion, especially the usage of genetic algorithm to adapt the fuzzy weights for making the final classifications.

Keywords: Automated Visual Inspection, Fuzzy Aggregation, Order Weighted Averaging, Choquet Fuzzy Integral, Pattern Recognition, Genetic Algorithm.

1 Introduction

The control of the industrial mass production of raw collagen plates by freeze-dry technology from bovine skins has improved strongly within the last years. However, the production line still needs a lot of personal experience and proper system settings in order to produce collagen plates of a quality satisfying customer demands. While the number of technical and functional faults (like ruptures, inclusions) decreased steadily, a new quality of customer demands came more and more into focus. This new quality demand is related to the inhomogenuous appearance of a collagen plate. The analysis of this quality is specially complex due to the fact that the collagen plates originate from an organic raw material and that they are finally going to be used as end products in the cosmetics and pharmaceutical industry (e.g. tailored to face masks or used as human skin replacement). So, in that scope also the subjectively determined aesthetical quality matters, and a quality inspection procedure is needed regarding this fact. For the production companies, this means to hire staff for manual visual inspection of the plates after production and before packaging, and to attempt a rough specification of perceptible plate features that may lead a customer to the final decision that a certain collagen plate looks somehow "ugly". Once having such features, they can be used to train the human inspectors.

Fig. 1 gives some examples for such perceptual faults. Under normal lightning conditions, collagen plate look similar to paper, so the contrast of the images in the figure was strongly enhanced by histogram equalization procedure to make the features more clearly visible. While subfigure (a) shows a plate that will be accepted by the customer, the other show features, the presence of which often leads to a customer return, as there are:

- non-persistent holes in the material (b), often only visible on one side of the plate. The size, count and positions of all such holes goes into the perceptual decision of whole plate appearance: while one bigger hole at the border may not count much, multiple small holes in the center gives a strong reject. This is also related to the further processing of the plates, e.g. to the placement of cutting face mask template on it that circumvent hole positions. Persistent holes are usually considered a technical fault.
- thawing structures (c), resulting from small fluctuations of freeze-drying control variables, yield a large-scale structure on the plate that normally gives a customer rejection. The contrast enhancement acts more strongly here, since it affects larger parts of the whole image.
- Eddies (d) are always present on the plate, but they are of lower contrast to the background and of fuzzy regional appearance if compared to holes. Also, count, size, position, but also distinctness and contrast of all eddies influence the whole imprecision of the plates complexion.

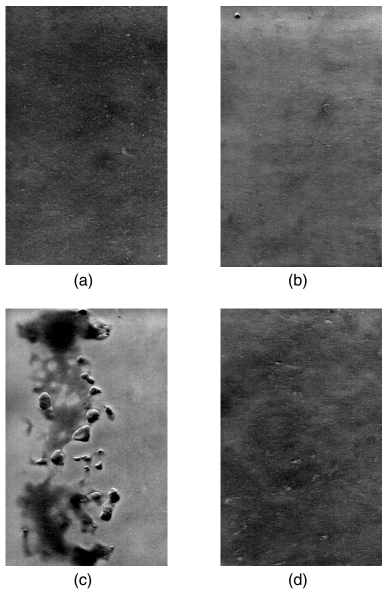

Figure 1: Contrast enhanced images of collagen plate with some typical perceptual faults; (a) no faults; (b) hole in upper left corner; (c) thawing structures; (d) high number of eddies distributed over the whole area.

There are further classes of such features, referred to as perceptual faults in the following. However, their general impact on the final appearance decision is doubtfully, and the subjective base of the decision makes it even more harder to establish an objectively working framework, with resulting customer returns of collagen plates being as minimal as possible.

This works presents a prototype system for the visual inspection of collagen plates. This system is in a test phase at a collagen producer's site. Despite of the many facetts of such a system (lighting installation, safe transport of the plates, task synchronization), the emphasis here is on the evaluation of perceptual faults that are featurized by appropriate image analysis algorithms, and their fusion into an anticipated customer decision. So, the point of departure of the approach is the acquired image of a collagen plate, and the reference is given by two-class decisions of the inspecting personal, which is trained in the procedure to reflect customer demands by their own decisions. In between, procedures are needed that may model the subjective decisions on whole-plate appearance. Summarizing it can be stated that the prototype system attains the simulation of the end user's aesthetic evaluation of the plates, which is undertaken in the production line by the inspecting personal.

The underlying modular framework of the system will be shortly presented in section 2. The basic idea behind it is to employ the stacking of fuzzy computations from low-level image features to a final class decision. The components of the framework treat specific texture detection problems with different binarization approaches. The evaluation of the binary images delivered by the binarization modules is realized through a hierarchical network of fuzzy fusion operators. More details on the framework itself and the employment of fuzzy concepts can be found in [9] [11].

The main achievment presented here targets the problem of adaption of the final evaluation to the decisions of a human making the same inspection. In the presented framework this problem maps to the problem of adapting the fuzzy densities, which are used in the fuzzy integral giving the final decision. How this task is achieved by using genetic algorithm will be shown in section 3 of this paper. Finally, the reader can find the conclusions in section 4.

2 Framework for Perceptual Relevance Evaluation

The purpose of the visual inspection system prototype is the evaluation of the perceptual relevance of different fault types (e.g. holes, eddies) on collagen plates (see Fig. 1). The framework is basically composed of a processing chain of alternating Binary Pattern Processing Modules ($BPPM_i$), each of them having the same internal structure. These modules reduce the grayvalue domain of the input images to a common binary one by detecting a specific type of faults. Thus all fault types can be further analyzed in the same way. Some additional testing modules can be found between such modules (see Fig. 2). The testing modules, where a fast testing routine based on reduced information is undertaken, are optional design components. Their main purpose is to bypass a BPPM (possibly time-costly) computations, if there is no evidence for the faults, which are processed by that BPPM.

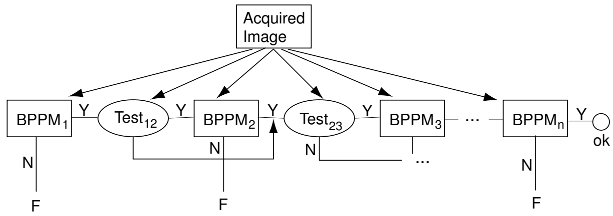

Figure 2: Framework for perceptual relevance evaluation. Different fault types, whose detection presents increasing complexity from the left to the right, are detected through different Binary Pattern Processing Module ($BPPM_i$). Different Test modules ($Test_{ij}$) are interpolated among them in order to break the analysis and thus to optimize the performance of the system in terms of detection time. The BPPMs deliver a decision on the relevance of the analyzed fault (F).

Each BPPM has access to the acquired image and to the evaluation of the foregoing testing module. Hence, the processing of each BPPM is independent of the processing of the others, but may refer to the results of the foregoing modules. Modules for the detection of the more frequent faults, or of the more simple to detect faults should come first in the chain. Since the appearance of a sufficient relevant fault in foregoing modules can interrupt the more time consuming computation in later ones, the system is computationally optimized.

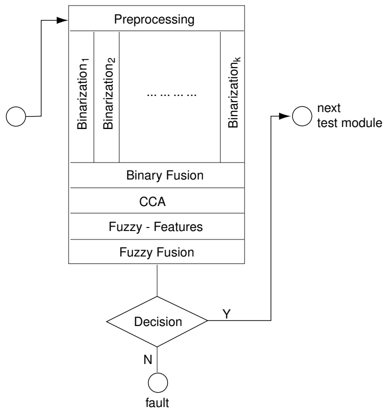

Figure 3: Structure of a Binary Pattern Processing Module. Although the faults to be treated are different, the BPPM maintain the same structure up to the number of parallel binarizations. CCA: Connected Component Analysis.

Each single BPPM is designed for a particular fault type, which is defined with respect to the fault's morphological and contrast structure. Each module is composed by a preprocessing, a binarization, a fuzzy feature extraction and a fuzzy evaluation stages (see Fig. 3). In the binarization stage, a set of k binarizing procedures is performed in parallel, giving k resulting binary images. These images act as complementary pieces of evidence in the analysis of the object under inspection. So far, four basic designs of a BPPM for four different fault types have been implemented (for a detailed explanation the reader is referred to [9]). A taxonomy of these types is given as follows:

Strong-contrast localized faults appear as an area presenting a grayvalue very different from the image background. Only one binarization operation is applied, which may be an interval thresholding.

Long-range faults are not related to a strong local contrast, but to a distortion within the global distribution of grayvalues within the image. Such faults may be detected by employing the auto-lookup procedure [9] based on the co-occurrence matrix [7] of a subset of pixels from the image. The procedure acts as a novelty filter (see Fig. 4).

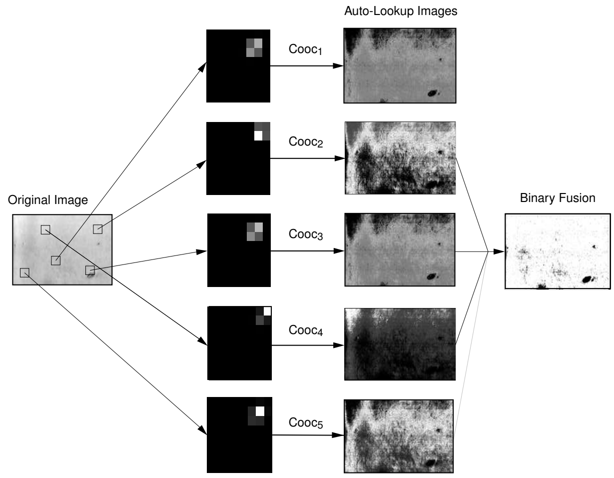

Figure 4: Auto-Lookup based procedure. From the acquired image, five co-occurrence matrices [7] are derived at five randomly located windows. Then, for each co-occurrence matrix, the auto-lookup procedure [9] is performed on the whole image. As a result dark regions stand for "atypical" regions.

Low-contrast faults are characterized by its low contrast to the background. For their treatment, the framework called Lucifer2 [8] may be used. The purpose of the framework is the automated generation of texture filters given the original and the expected goal images.

Frequency related faults appear in form of a disruption of the uniformity resulting from the repetition of a basic element. For its detection a Gabor image decomposition [2] for different spectral bands is calculated following the schemata presented in [10]. Finally, the total image energy in the analyzed spectral bands is computed and thresholded.

A fusion procedure derives a binary image by applying logical operators on the k images delivered by the binarization stage. The foreground (or Black pixels) of that binary image are candidates for perceptual faults which will be fuzzy processed in the following stages. The CCA module connects components from the binary fused image, which are perceptual fault regions, and removes noise that remained after the binary fusion.

The fuzzy evaluation of the binary image is undertaken by a network of fusion operators (see below). Thus, the network closes the gap between binary pixel information and the relevance of the defect under inspection. As a result each BPPM delivers a pair of fuzzy membership degrees upon two classes ("no relevant faultiness", "relevant faultiness"). For instance, the vector $(0.3, 0.7)$ would indicate an inspected objet where the analyzed texture fault are more perceptually relevant than irrelevant.

The final decision for the rejection of a piece upon each defect type is undertaken by the production system based on this information.

2.1 Fuzzy Aggregation for Perceptual Relevance Evaluation

The binary images extracted by the parallel binarization submodules (see Fig. 3) have to be fuzzy processed in order to better approximate the desired subjective evaluation.

The analysis of the binary images obtained is undertaken at this point by a network of fuzzy aggregation operators (see Fig. 5), where the complexity of the problem attached in each following stage needs the usage of operators of also increasing complexity. This concept has been analyzed in detail in [12]. With each fusion operator in the network a new abstraction level in the way from pixel information to the quantification of the perceptual relevance of the defects is achieved. The different stages will be analyzed in the following.

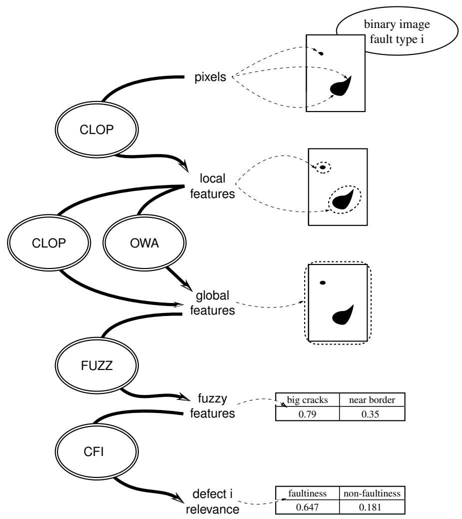

Figure 5: Hierarchically organized network of fusion operators used in the here presented paper for the fuzzy evaluation of binary images. The operators of increasing softiness [12] are in charge of tasks of increasing complexity. These correspond to increasing levels of abstraction in the evaluation from the consideration of pixel information, $value(x, y)$, to the result in form of a double fuzzy membership degree expressing "faultiness". CLOP: Classical Operators; OWA: Ordered Weighted Averaging operators; FUZZ: fuzzification stage; CFI: Choquet Fuzzy Integral.

At first the remaining connected components are measured (see Fig. 6) through classical operators as the statistical first moment of the black pixel positions or the sum of pixel values (1 or 0). As a result, local geometric features for each detected fault are extracted. E.g. height, width, area, perimeter, roundness.

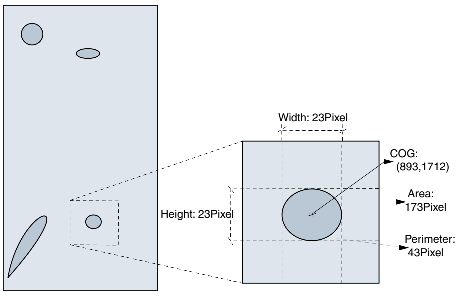

Figure 6: Measuring of binary patterns after CCA. Geometric features for each detected fault are extracted. COG: Center of gravity.

Then, a holistic description of the local features is needed, in order to get global description of the elements under analysis. For that end, traditional fusion operators on the one hand and Ordered Weighted Averaging (OWA) operators on the other are applied on them. The employment of traditional fusion operators is trivial, e.g. computation of the number of faults in an item. On the other hand it is worth detailing the employment of OWAs. The global descriptor obtained up to them can be considered as a meta-feature, resulting from the fuzzy aggregation of the local fault features.

The weighting configuration of the OWAs, which are softer aggregation operators than traditionally used ones [1], increase the flexibility. Since the weighting is done by taking into consideration the numerical ranking of the features, the result can be biased for giving preference to a determined range of values or for reinforcing the presence of coincident ones [17]. Furthermore, the usage of the OWAs [16] allows the comparison of vectors with different number of components without suffering the low-pass filtering effect of traditionally used ones, e.g. average. An example of the usage of the OWA weights and its effect on the result is shown in Fig. 7b.

The global descriptors till this stage are fuzzified when necessary by defining linguistic terms with trapezoidal fuzzy membership functions (see Fig. 7c). The fuzzification of the *meta-features*, which apply a linguistic descriptor on them, eases the conceptual development and parameterization of the following stages. Moreover the usage of two opposed fuzzy features increase the robustness of the evaluation system.

The last stage of the hierarchical network of fuzzy aggregation operators (see Fig. 5) is implemented through a Choquet Integral. This operator attains binding the different *fuzzy meta-features* in order to make a decision over the goodness of the item. Thus for each perceptual fault class the *fuzzy meta-features* are fused into a value from $[0, 1]$ by using the Choquet fuzzy integral in order to characterize the perceptual relevance of the fault in form of a membership degree. The application of an aggregation operator can produce such a membership function.

The most important reason for the application of the fuzzy

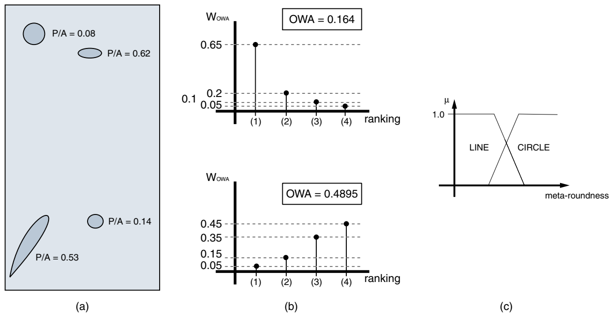

Figure 7: Example of the employment of OWAs in the computation of *fuzzy meta-features*. The image shows the exemplary computation of the meta-roundness of a plate based on the fuzzy aggregation of the local features Perimeter/Area (P/A) (a). (b) A weight configuration very sensitive to the presence of just one fault with circular form is displayed at the top. On the contrary if a detection of more than one elongated fault is desired, a configuration like the one at the bottom would be used. (c) Example of fuzzification functions for the computation of *fuzzy meta-features* (in this case, meta-roundness). The result of the OWAs is fuzzified with trapezoidal fuzzifying functions.

integral is the capability of this fuzzy operator for fusing information taking into consideration the *a priori* importance of both individual and groups of attributes. The fusion of the different fuzzy features is needed in order to find the joint perceptual relevance of the faults in the object under inspection. In such a process the interaction between the different fuzzy features has to be considered. The fuzzy integral is the only fuzzy fusion operator known so far to allow such a characterization [5].

The kind of analysis undertaken reflects in the result the different possibilities of interaction. E.g., if the presence of defects on the plate border is very important, the result of the relevance quantification should increase; if such a presence is not so important but coincides with a very big defect, the relevance should also increase; if the defects are small and there are not so much of them, the relevance decreases. Such a characterization could have been undertaken also with a system of fuzzy rules. However, the fuzzy integral approach is more synthetic and comprehensive from the developer point of view. When the kind of interactions to be characterized are very complex or numerous, the number of rules increase so much that the problem is no longer tractable. Furthermore, in many cases the descriptions delivered by the inspection experts are difficult to collect, if not even full of contradictions. The automated finding of the fuzzy measures helps avoiding this problem and allows an easier redesign of the feature extraction stage.

It should be noted here that the Choquet Fuzzy Integral has been already used as classificator in pattern recognition problems, where its performance was superior to that of other classificators, e.g. multilayer perceptron, bayesian independent classification [4]. Moreover the parameterization of the fuzzy integral was taken into consideration. Diverse algorithms have been presented for the parameterization of

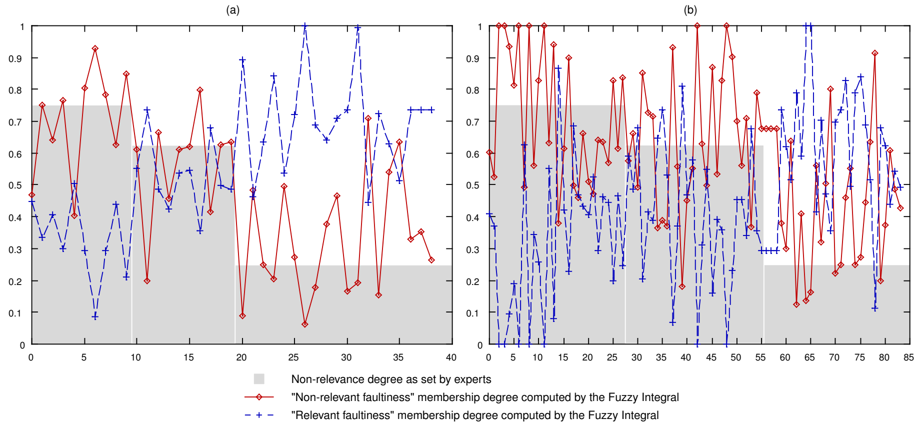

Figure 8: (a) Results obtained on a 39 plates training set. (b) Results obtained on a 84 plates test set with the automatically determined fuzzy measure coefficients.

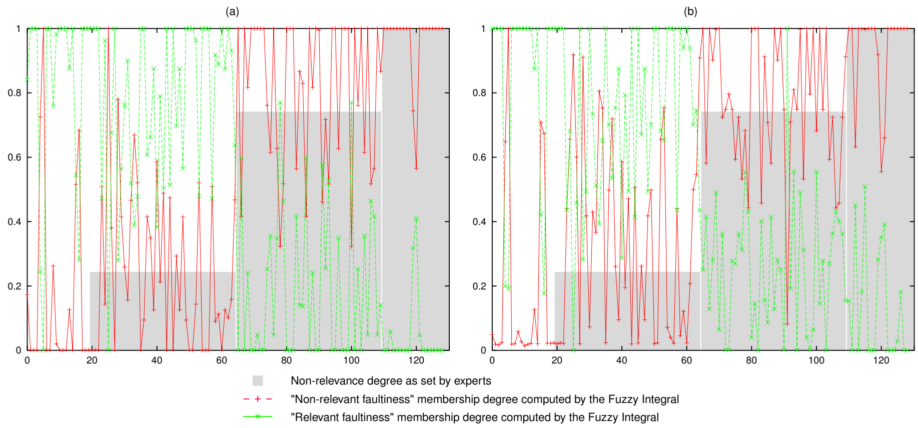

Figure 9: Results obtained on a 129 plates training set (a) using general fuzzy measures and (b) $\lambda-$fuzzy measures.

the Choquet integral [4][5][15], while the Sugeno integral [13] lacks of such a diversity.

3 Automated Parameterization of the Fuzzy Integral

As already mentioned, the prototype system attains the evaluation of

the perceptual relevance of each type of fault taken individually. Thus the final output of the system is a two-component vector for each plate and fault, which expresses the membership degree of the plate to the "accepted" and the "rejected" classes. This vector can be understood as a two-component feature for each fault type, upon which an end decision of the commercialization channel of the corresponding plate is to be made, e.g. some plates are sold as first choice quality, other ones as second choice,

The results of the system are analyzed by taking into consideration just the image understanding part of the system. Thus the binary images are considered already computed. Up to this point the features are extracted and fuzzified by undergoing the network of fuzzy fusion operators (see Fig. 5). It is worth mentioning that the fuzzy membership functions are heuristically defined on hand of the classification results. This process is not trivial and the achieved classification results demonstrated to depend on this stage. Therefore its automation can systemize the future implementations of the framework. The results of the stage, where the resulting two sets of fuzzy meta-features are fused with a Choquet Fuzzy Integral [6] with respect to two different fuzzy measures, are analyzed in the following.

The fuzzy measures are constructed through genetic algorithms. The genetic algorithm implemented is a steady-state genetic algorithm, which guarantees not losing the better result in each generation. The genetic algorithms employs: random initialization, a rank selection operator, a two-point crossover operator with probability of 0.7, mutation probability of 0.01, and replacement probability of 0.95 [3]. The settings of the genetic algorithm were found heuristically. Finally the perceptual relevance of the fault is delivered in form of a two-component vector, where each component is defined in $[0, 1]$.

The generalization capability of the system is first analyzed. λ-fuzzy measures are used for that purpose. The BPPM taken into consideration analyzes the relevance of textured areas on the plates. The training of the fuzzy integral was undertaken for this BPPM based on a set of 39 plates.

Table 1: Statistical comparison of results obtained through the application of the fuzzy integral with respect to general (μ) and $\lambda-$fuzzy measures (μ_λ). ρ: Crisp recognition rate of the system by taking into consideration the larger membership degree as a crisp result. MSE: Mean square error between the "relevance" membership established by the experts and the one obtained. σ_{MSE}: Variance of the mean square error. Values are taken as average over different simulations and computed for three different data sets.

	Set I			Set II			**Set I+II**		
	ρ	MSE	σ_{MSE}	ρ	MSE	σ_{MSE}	ρ	MSE	σ_{MSE}
μ	92.2	0.15	0.01	86.7	0.167	0.041	86.82	0.184	0.028
μ_λ	83.0	0.15	0.01	83.0	0.16	0.04	86.82	0.184	0.034

Table 2: Statistical analysis of the results obtained by the automated industrial system on 11 evaluation data sets from production line. EFT: Error in fault type. FR: False rejectance. FA: False acceptance. ϵ_m: Minimum. $\bar{\epsilon}$: Average. σ_ϵ: Variance. ϵ_M: Maximum.

Set	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	ϵ_m	$\bar{\epsilon}$	σ_ϵ	ϵ_M
EFT	1	1	7	13	9	2	5	11	13	4	4	1	6.4	4.5	13
FR	1	3	1	0	3	9	2	0	4	4	4	0	2.8	2.6	9
FA	24	9	21	17	11	3	5	5	6	5.3	11	3	10.7	7.1	24

The obtained results on the training set after determining the λ-fuzzy measure are depicted in Fig. 8a. The consideration of the trend condition in the fitness function as described in the former section improved the results in terms of false accepted/rejected rates and of approximation to the relevance established by the experts. Taking into consideration the larger membership degree as a crisp result the system achieved a recognition rate of 84.61% on the training set.

The results obtained on a test set of 84 plates are depicted in figure 8b. The recognition rate with the test set was of 70.23%. On hand of these results the generalization capability of the evaluation framework can be analyzed.

In another simulation the influence of type of fuzzy measures to be used in the classification result is analyzed. Thus the training phase is completed using a data set of 129 items. The simulation compares the classification results obtained by using $\lambda-$ and general fuzzy measures (see Fig. 9). Although the recognition rate is the same (86.82%) for both types of fuzzy measure in the here presented simulation, the greater flexibility of the general fuzzy measures in front of the λ ones improve the performance of the system (see Tab. 1). Moreover the system presents more errors with the λ-fuzzy measures (12.4%) in the set of items qualified as "accepted" by the experts than with the general ones (8.5%). These type of items are the most common ones in the production line.

Furthermore the fuzzy integral with respect to the general fuzzy measures seems to fit better the evaluation of the inspector personal. This lightly better performance can be observed on hand of the variance of the mean square error (σ_{MSE}) between the relevance curve determined by the experts (see gray surface in Fig. 9) and the results of the fuzzy integral. The MSEs and their variances are summarized in Tab. 1. Although the general fuzzy measures achieve a better performance for all data sets, the difference between the results obtained with both types of measures strongly depends on the considered data set. For instance the fuzzy integral with respect to the general fuzzy measures outperfoms the results obtained with respect to the λ-fuzzy measures in the first data set (see Tab. 1). Taking into consideration all these facts general fuzzy measures are chosen for the implementation of the final system.

The installed prototype system was tested with 11 different evaluation data sets, which are composed by 100 plates taken directly from the production line. The results obtained on these evaluation sets can be observed in Tab. 2. The system had been previously trained and tested with other data sets. Thus the results can be used for the evaluation of the system's generalization as well. Three parameters characterize the goodness of the system. First the percentage of plates with relevant faultiness that present a false fault type as the principal factor influencing its final classification (EFT). Finally the false rejectance (FR) and acceptance (FA) rates are given.

4 Conclusions

The here presented framework has been successfully implemented. The slight differences between the recognition rates of the experimental results and those of the industrial system lays on the fact that the training and test sets just take faulty plates into consideration.

After analyzing the results on hand of the binary images entering the fuzzy evaluation subsystem it could be stated that a great part of the errors were due to external factors. Among them is worth mentioning the presence of contradictions in the expected output membership degrees established by the experts. The interactive assessment of these output membership degrees taking into consideration the results of the automated system should be undertaken in order to minimize such errors.

Bibliography

[1] D. Dubois, H. Fargier, H. Prade (1997). Beyond min aggregation in multicriteria decision: (ordered) weighted min, discri-min, lexmin. In: Yager, R. R., Kacprzyk, J. (eds.), *The ordered weighted averaging operators — Theory and applications*, Kluwer Academic Publishers, Dordrecht.

[2] D. Gabor (1946). Theory of Communication, *J. Inst. Elec. Eng.*, vol. 93, pp. 429-457.

[3] D.E. Goldberg (1989). *Genetic algorithms in search, optimization & machine learning*, Addison-Wesley, Reading, MA.

[4] M. Grabisch, J.-H. Nicolas (1994). Classification by fuzzy integral: Performance and tests, *Fuzzy Sets and Systems* (65), pp.255-271, 1994.

[5] M. Grabisch (1997). Fuzzy Measures and Integrals for Decision Making and Pattern Recognition. In *Tatra Mountains: Fuzzy Structures Current Trends*, vol.13, pp. pp. 7-34.

[6] M. Grabisch, T. Murofushi, M. Sugeno (1999). *Fuzzy Measures and Integrals*, Physica-Verlag.

[7] R. Haralick, K. Shanmugam, I. Dinstein (1973). Textural features for image calssification, *IEEE Trans. SMC*, 3, (6), pp. 610-621.

[8] M. Köppen, B. Nickolay (2001). Genetic Programming based Texture Filtering Framework, *Nikhil R. Pal (Ed.), Pattern Recognition in Soft Computing Paradigm, FLSI Soft Computing Series - Vol. 2, Chapter 12*, World Scientific, pp. 275–304.

[9] M. Köppen, A. Soria-Frisch, Th. Sy (2001). Binary Pattern Processing Framework for Perceptual Fault Detection, *Proc. 12th Scandinavian Conf. Image Analysis, SCIA 2001*, Bergen, Norway, pp.363-370.

[10] O. Nestares, *et al.* (1998). Efficient spatial-domain implementation of a multiscale image representation based on Gabor functions, *J. Electronic Imaging*, vol. 7, pp. 166-173.

[11] Soria-Frisch, M. Köppen, Th. Sy (2003). Is she gonna like it? Automated Inspection System Using Fuzzy Aggregation. To appear in *Intelligent Systems for Information Processing: From Representation to Applications, Elsevier Science (in print)*

[12] A. Soria-Frisch (2002). Soft Data Fusion in Image Processin. In *Soft-Computing and Industry: Recent Advances*, R. Roy et al. eds., pp. 423-444, Springer-Verlag.

[13] M. Sugeno (1974). *Theory of fuzzy integrals and its applications*, Ph.D. thesis, Tokyo University.

[14] K. Tanaka, M. Sugeno (1991). A study on subjective evaluation on printed color images, *Int. J. Approximate Reasoning*, vol. 5, No. 3, pp 213-222.

[15] D. Wang *et al.* (1997). Determining Fuzzy Integral densities using a Genetical Algorithm for Pattern Recognition, *Proc. 1997 NAFIPS Annual Meeting*, pp. 263-267.

[16] R.R. Yager (1988). On ordered weighted averaging aggregation operators in multi-criteria decision making, *IEEE Trans. SMC*, 18, pp.183-190.

[17] R.R. Yager, A. Kelman (1996). Fusion of Fuzzy Information with considerations for compatibility, partial aggregation and reinforcement,*International Journal of Approximate Reasoning*, 1996.