Textural Features Based on Run Length Encoding in the Classification of Furniture Surfaces with the Orange Skin Defect

Jakub Pach[1], Leszek J. Chmielewski[1], Arkadiusz Orłowski[1], Michał Kruk[1], Jarosław Kurek[1], Bartosz Świderski[1], Izabella Antoniuk[1], Grzegorz Wieczorek[1], Katarzyna Śmietańska[2], Jarosław Górski[2]
[1]*Institute of Information Technology*
[2]*Institute of Wood Sciences and Furniture*
Warsaw University of Life Sciences – SGGW, Poland

Abstract. Textural features based upon thresholding and run length encoding have been successfully applied to the problem of classification of the quality of lacquered surfaces in furniture exhibiting the surface defect known as *orange skin*. The set of features for one surface patch consists of 12 real numbers. The classifier used was the one nearest neighbour classifier without feature selection. The classification quality was tested on 808 images 300 by 300 pixels, made under controlled, close-to-tangential lighting, with three classes: *good*, *acceptable* and *bad*, in close to balanced numbers. The classification accuracy was not smaller than 98% when the tested surface was not rotated with respect to the training samples, 97% for rotations up to 20 degrees and 95.5% in the worst case for arbitrary rotations.

Key words: quality inspection, furniture surface, orange skin, textural features, run length coding, thresholded image, one nearest neighbour, leave-one-out testing.

1. Introduction

In various sectors of the manufacturing industry, the computer vision methods are used to perform inspection and measurement tasks, while in the furniture industry it seems that the inspection with the unarmed human eye is the most commonly used tool. In this way, the tasks like the evaluation of esthetic quality of an object or a surface can be performed in a straightforward way and at a moderate cost. This state does not exclude the research on computerized image-based inspection in the application to furniture.

The application of our interest will be the classification of a painted surface with respect to the defect known as *orange skin* or *orange peel*. Up till now, the research on this defect was performed to a limited extent. In [10] the images of *orange peel* were generated and its visibility for humans was studied, but neither automatic detection nor classification was considered. In [2] a complex system of moving lights and cameras was investigated which improved the visibility of a variety of defects, including the *orange peel*, so that the local adaptive thresholding could be used as the defect detection method. The defects on furniture were mentioned within the general domain of defect

detection in [9]. The quality of raw materials was considered in [17]. In many patents the methods to perform the painting process so that orange skin is removed or avoided, for example [1], but with no reference to image analysis. This state of the lack of interest in automatic image-based inspection in the furniture industry is interesting in view of that in the timber industry the analysis of a whole range of raw wood defects is widely applied and described [3, 16]).

The quality of wood products has been the domain of our interest in a number of papers (shape and dimensional accuracy in [5, 13, 14, 15] and surface quality in [5, 6, 7, 11, 21]). In particular, in [5] we have demonstrated that structured light-based scanning of the surface can not reveal the orange skin defect reliably. In [6] we have tested the applicability of such conventional image processing techniques like special lighting, differentiation and thresholding. The use of conventional and novel textural features was tested in [7] and [11]. A set of seven feature selection methods were used to find the best features for this task [21], with the best attained classification precision of 95.8%.

Finding the best features in the sense of precision as well as calculation speed is still an interesting problem. As far as moderately expensive equipment is considered, this can not be solved with such general but hardware-intensive and costly solutions like deep learning. In this paper we present yet another set of features which are relatively simple to calculate to be used in the orange skin detection and classification problem. In the algorithm a set of features based upon the classic run length encoding is applied. The features similar to those presented here have been successfully used in the problem of finding the handwriting area in a manuscript [19]. The accuracy of the classification expressed with its accuracy, can attain values over 98%, despite of the relatively simple structure of the features.

The remaining part of the paper is organized as follows. In the next section the images to be analysed will be described. In Section 3 the algorithm of finding the features will be proposed. The classifier will be briefly introduced and the attained classification quality will be presented in Section 4. The results will be discussed in Section 5. The paper will be concluded in Section 6.

2. Images

Orange skin is a defect of lacquered and hardened surfaces which can be seen as shallow hollows which form a characteristic pattern. The defect can emerge due to insufficient quantity or bad quality of dilutent, excessive temperature difference between the lacquer and the surface, bad pressure or distance of spraying, excessive air circulation during spraying or drying, and insufficient air humidity. The defect can be the reason for treating the surface as unacceptable for vending, or as acceptable as a lower quality product. Therefore, the surface fragments considered have been divided into three classes: *good*,

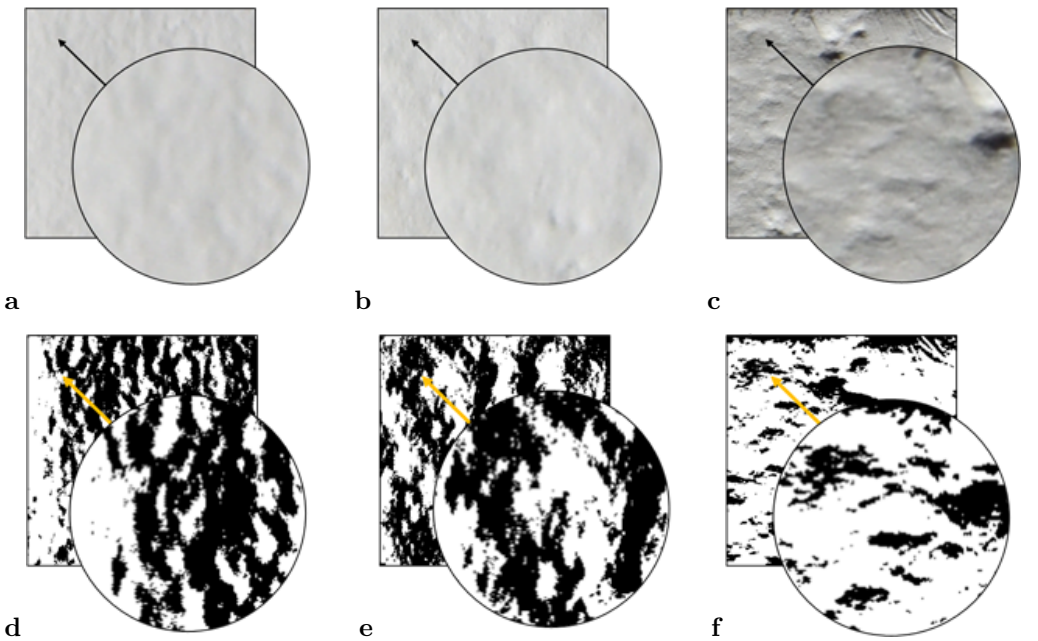


Fig. 1. Examples of images of the surface without and with the orange skin defect. (**a**) Class *good*, defect absent; (**b**) class *acceptable*, weak defect; (**c**) class *bad*, defective. Binarized images: (**d**) *good*; (**e**) *acceptable*; (**f**) *bad*.

acceptable and *bad*. The examples of such surfaces with fragments enlarged are shown in Figure 1a, b and c, respectively.

The set of images, described already in [11,21], contained images 300×300 pixels each, cut from a smaller number of images taken with the Nikon D750 24 Mpix camera with the Nikon lens F/2.8, 105 mm. The distance from the focal plane to the object surface was 1 m and the optical axis of the camera was normal to the surface. The surface was illuminated with a flash light located at 80 cm from the object, with the axis of the light beam inclined by 70° from the normal to the surface (lighting close to tangential). The images were made in colour mode, with lossless compression, transformed into CMY colour model, and the Y component was taken for further processing.

There were $N = 808$ images in total: 282 *good*, 278 *acceptable*, and 248 *bad* (92 images were excluded from the set considered in [21] due to that it was very problematic to which class they should be assigned). All these images were used in the assessment of the classification quality, due to that the cross-validation procedure was used, as it will be described further in Section 4.

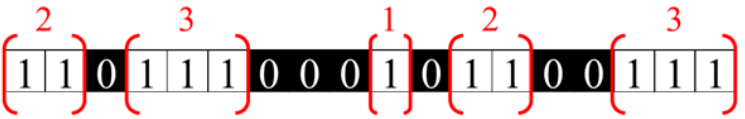


Fig. 2. Illustration of a run length code of a line of a binary image. Ones represent white pixels, zeros represent black ones.

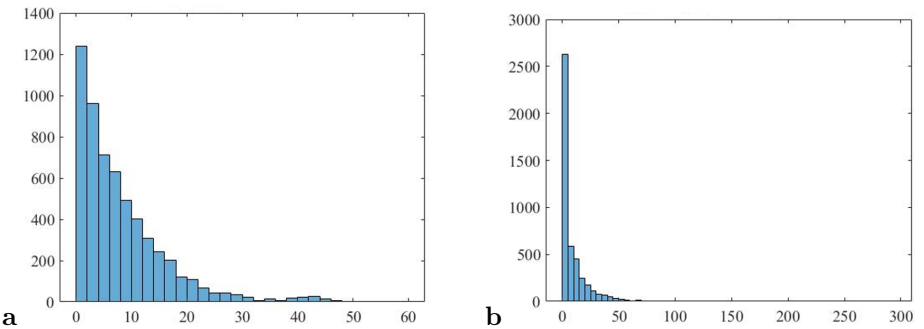


Fig. 3. Histograms or run lengths of ones in lines of a binary image of the orange skin defect. (**a**) Lines taken horizontally; (**b**) lines taken vertically. The histograms shown have bins containing more than one length of the runs (for example, histogram **a** has bins containing two consecutive lengths each: 1 and 2, 3 and 4, etc.); in the features, the histograms containing 10 consecutive lengths in each bin are used.

3. Features

The features used previously to successfully differentiate the regions with handwritten text from other regions [19] will be used. These were the components of the histogram of runs of ones, in run length encoded binarized image of the sheet of a manuscript. A hypothetic line of a binary image is shown in Fig. 2 in the form of the run length code. It occurs that the longer the run, the less frequent it is in an image. Histograms of run lengths of ones in an image coded by rows (horizontally) and by columns (vertically) are shown in Fig. 3.

In the histograms, each bin contains L consecutive lengths, and the lengths up to $l_{\max}$ are considered. To find the features, $L = 10$ lengths are counted in each bin, and lengths up to $l_{\max} = 60$ are considered, for horizontal as well as for vertical lines. This gives $n = l_{\max}/L = 6$ features for each direction, which makes $m = 2n = 12$ features for an image. The numbers of the runs are normalized to give the sum equal to one,

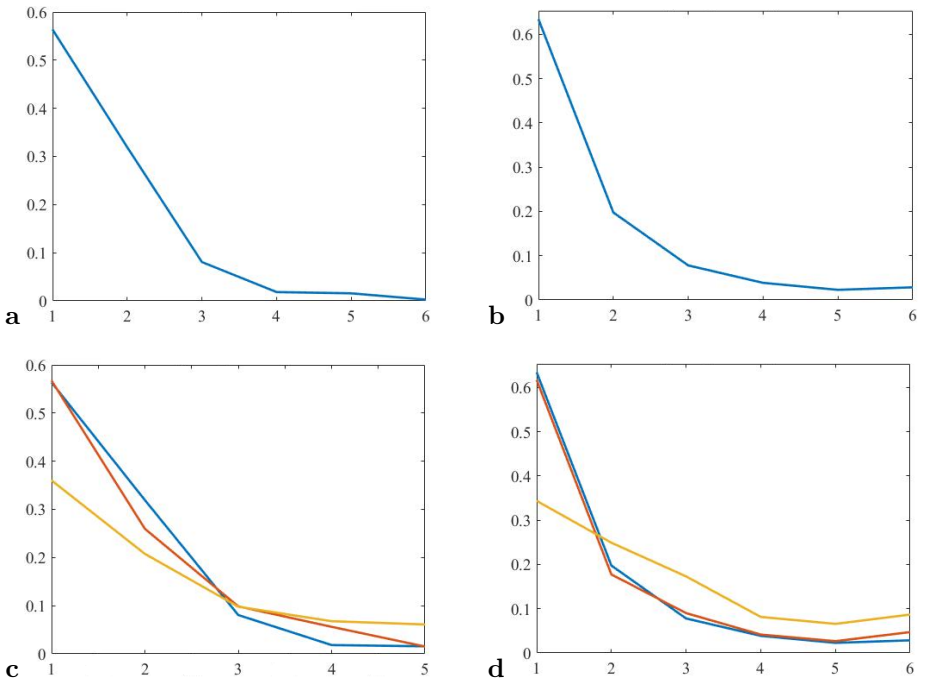


Fig. 4. Features, that is, normalized histograms for horizontally nad vertically coded images. (**a**) Example features of an image coded horizontally, corresponding to the raw histogram of Fig. 3a. (**a**) Example features of an image coded vertically, corresponding to the raw histogram of Fig. 3b. (**c**) Features of an image coded horizontally, belonging to three classes: blue – class *good*, red – class *acceptable*, and yellow – class *bad*. Zero values of feature 6 not shown. (**d**) Same as **c**, but the image coded vertically. Lines joining the values for the integer indexes of features have no meaning; shown only to indicate the relations between the points.

so the features are real numbers (rational fractions). Normalization makes the features insensitive to the size of an image, but they remain sensitive to its scale.

In Fig. 4 the features, that is, the normalized histograms for a horizontally and vertically encoded image, corresponding to the histograms in Fig. 3, and the features for the three classes of the surface are shown.

To binarize the images, the classical Otsu method [18] was used due to that this classic parameter-less method still has many citations and applications and its performance is good in cases where the numbers of dark and bright pixels do not differ much [20].

4. Classification

4.1. Classifier

The classical nearest neighbour classifier was used, like in [19]. The simplest one nearest neighbour (1-NN) version was applied. The more complex nearest neighbours schemes (see for example [8]) were not tested due to that in this paper the effectiveness and simplicity of the features, not the classifier, is of the greatest interest, and the results appeared to be satisfactory.

As the distance function between objects e_i and e_j represented with features $F(e) = \{f_k(e), k = 1, \ldots, m\}$ the Manhattan distance was used

$$d(f_i, f_j) = \sum_{k=1}^{m} |f_k(e_i) - f_k(e_j)| \ . \tag{1}$$

4.2. Errors and classification quality

4.2.1. Data without transformations

The quality of the classification in the setting defined above was tested with the cross-validation method in its extreme form, that is, the leave-one-out method. Besides some well theoretically justified criticism [12], this method is still treated as attractive and applicable (let us cite [4] as just one example of many possible ones). This method consists of taking one pattern away from the set of known N patterns, training the classifier with the remaining $N - 1$ ones, and then classifying this left pattern with the so trained classifier. After cycling through all the N patterns in the set, the errors made in the N classifications are calculated. In the case of 1-NN classifier, there is actually no training, and in our case the set of features is fixed so the feature selection process is absent either. Therefore, this otherwise time-consuming procedure is effective in this case and is an attractive solution.

The errors found with the cross-validation method are shown in Table 1. As the global measure of classification quality, the accuracy expressed as the ratio of the number of

Tab. 1. Confusion matrix C_{ij} for the three classes: *good* $c = 1$, *acceptable* $c = 2$ and *bad* $c = 3$.

	c	Actual class		
Class no. $= c$		1	2	3
Classification result	1	275	2	0
	2	7	274	3
	3	0	2	245

proper classifications to the number of all the patterns was used

$$A = \frac{\sum_{i=1}^{3} C_{ii}}{\sum_{i,j=1}^{3} C_{ij}}\,, \tag{2}$$

with the attained value $A = 98.29\%$.

4.2.2. Testing the sensitivity to rotation

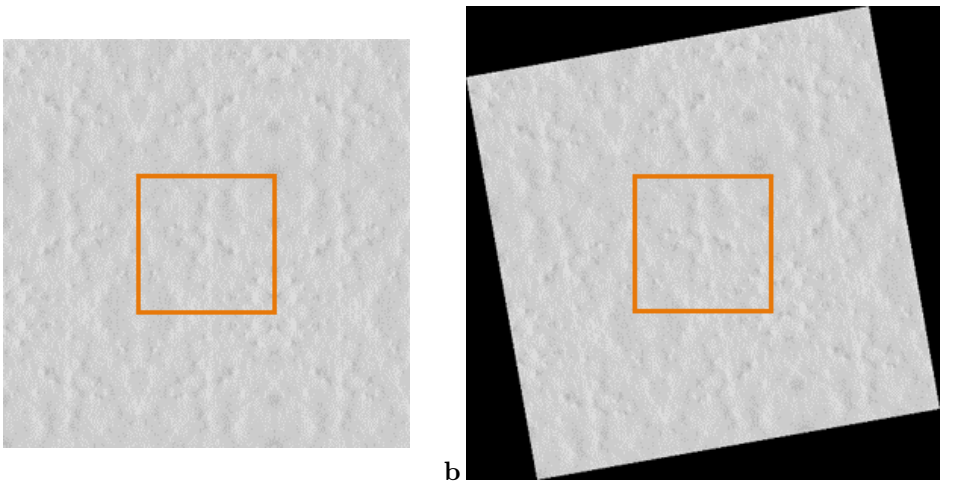
In an industrial quality inspection process the location of the tested object can be easily restricted to that which corresponds to the conditions in which the teaching of the classification system was performed. However, the sensitivity of the testing system to such transformations as rotation and scale should be assessed. From the two transformations mentioned, the scale seems to be easier to control by setting the distance from the focal plane of the camera to the object surface to the desired value. The influence of rotation, which can result from that the direction of cutting the manufactured object from the painted surface can be arbitrary, is more difficult to restrict, so the sensitivity of the proposed features to rotation has been tested.

To achieve this goal, each of the images used in the cross-validation process was rotated, by $k = 360$ angles from 1° to 359°. The images should have the same dimensions and scale as those without rotation, so before the rotation each image was complemented at its edges and corners with its six mirror images, such extended image was rotated, and its center 300×300 was cut, as shown in Fig. 5. In this way, the appearance of the texture within the image was maintained.

For each of the rotated images, the cross-validation procedure was repeated, with the $N - 1$ reference images without rotation, and the classified image with rotation, by each of the k angles. The accuracy A obtained in this way, versus the rotation angle, is shown in the graph in Fig. 6. The difference between the best and worst accuracy is close to 3.5%. The accuracy is the worst (95.5%) for the angle close to $310^\circ = 360^\circ - 50^\circ$, and is the best at 270° and happens to surpass 99%. The shape of the graph follows the commonsense expectation that the rotations by angles close the multiples of 90° deteriorate the classification quality to the smallest extent (or even improve it), while the rotations by 45° reduce it the most. The variability of inter-class errors is illustrated in Table 2 with the standard deviations of the errors.

5. Discussion

The classification accuracy obtained with the features presented has been measured with a different methodology than in the previous papers [11, 21], where cross-validation with the set of images was partitioned into nine subsets, 90 images each. Also the set of images was better prepared in this paper than in [21], in that the images for which the class was



a **b**

Fig. 5. Illustration of rotating an image with the scale and dimensions maintained. (**a**) Image 300×300 (colour frame) complemented at the edges and corners with its mirror images, forming a 900×900 extended image. (**b**) Window 300×300 (colour frame) cut from the rotated extended image.

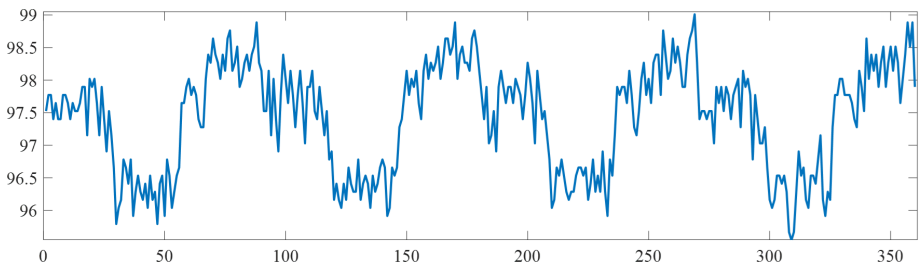


Fig. 6. Global classification accuracy [%] versus rotation of the classified image [°], in the range $[1°, 359°]$.

Tab. 2. Variability of the confusion matrix for the experiment with rotation illustrated by standard deviations of the inter-class errors for three classes.

	c	Actual class		
Class no. $= c$		1	2	3
Classification result	1	2.21	1.05	0.72
	2	2.13	3.28	2.56
	3	0.44	2.95	2.88

difficult to assess by a human eye were excluded. Nevertheless, it can be observed that the single set of features, relatively simple to calculate, appeared to perform better, or at least not worse, than the large set of state-of-the-art features selected with advanced methods. The best average accuracy attained in [21] was 95.88%. The comparison with the accuracy of 98.29% found in this paper indicates that the features proposed have a potential in the application of our interest. It should be also noted that here there was entirely no confusion between the classes *good* and *bad*, which is important from the point of view of the application to quality inspection.

The sensitivity of the classification accuracy to the rotation of the image was not tested previously. The variability shown in Table 2 can be considered relatively small with respect to the global accuracy. From the graph shown in Fig. 6 it follows that the accuracy should not fall below 95.5% obtained for the rotation of the image close to 45°, which is the worst case if it is taken into account that the original features were found for angles close to zero. For the angles differing from those close to the multiple of 90° by not more than ±20° the accuracy is not worse than 97%. This result indicates that the decrease of classification accuracy related to small rotations is negligible, and that the sensitivity of the classification quality to arbitrary rotations can be considered small.

6. Conclusion

Quality of the surface of manufactured elements is usually inspected by a naked eye of human inspectors in the furniture industry practice. As a continuation of our previous research aimed at changing this practice, a set of textural features based upon thresholding and run length encoding, positively tested in other applications, have been successfully applied to the problem of classification of the quality of lacquered surfaces in furniture exhibiting the surface defect known as *orange skin*. The set of proposed features was formed by 12 easy to calculate real-valued features for each object. The classifier used was the simple one nearest neighbour classifier, with the Manhattan distance, and without feature selection. The classification method was tested on 808 images 300 × 300 pixels, made under controlled, close-to-tangential lighting, and classified by an expert in the way of inspection by an unarmed eye into three classes: *good*, *acceptable* and *bad* surface. The numbers of examples from each class were close to each other. The classification quality was assessed by cross-validation with the leave-one-out method. The attained accuracy was not smaller than 98% when the tested surface was not rotated with respect to the training samples, 97% for rotations up to ±20° and 95.5% in the worst case for arbitrary rotations. In the case without rotation, there were no classification errors between the classes *good* and *bad* surface.

The proposed features seem to be a simple and effective alternative to the advanced features tested previously for this defect.

References

[1] M. Allard, C. Jaecques, and I. Kauffer. Coating material which can be thermally cured and hardened by actinic radiation and use thereof, September 27 2005. US Patent 6,949,591.

[2] L. Armesto, J. Tornero, A. Herraez, and J. Asensio. Inspection system based on artificial vision for paint defects detection on cars bodies. In *2011 IEEE International Conference on Robotics and Automation*, pages 1–4, May 2011. doi:10.1109/ICRA.2011.5980570.

[3] V. Bucur. Techniques for high resolution imaging of wood structure: a review. *Measurement Science and Technology*, 14(12):R91, 2003. doi:10.1088/0957-0233/14/12/R01.

[4] G. C. Cawley and N. L. C. Talbot. Efficient leave-one-out cross-validation of kernel Fisher discriminant classifiers. *Pattern Recognition*, 36(11):2585 – 2592, 2003. doi:10.1016/S0031-3203(03)00136-5.

[5] L. J. Chmielewski, K. Laszewicz-Śmietańska, P. Mitas, A. Orłowski, et al. Defect detection in furniture elements with the Hough transform applied to 3D data. In R. Burduk et al., editors, *Proc. 9th Int. Conf. Computer Recognition Systems CORES 2015*, volume 403 of *Advances in Intelligent Systems and Computing*, pages 631–640, Wrocław, Poland, 25-27 May 2015. Springer. doi:10.1007/978-3-319-26227-7_59.

[6] L. J. Chmielewski, A. Orłowski, K. Śmietańska, J. Górski, et al. Detection of surface defects of type 'orange skin' in furniture elements with conventional image processing methods. In F. Huang and A. Sugimoto, editors, *Image and Video Technology – PSIVT 2015 Workshops*, volume 9555 of *Lecture Notes in Computer Science*, pages 26–37, Auckland, New Zealand, 23-27 Nov 2015. Springer, 2016. doi:10.1007/978-3-319-30285-0_3.

[7] L. J. Chmielewski, A. Orłowski, G. Wieczorek, K. Śmietańska, and J. Górski. Testing the limits of detection of the 'orange skin' defect in furniture elements with the HOG features. In N.T. Nguyen, S. Tojo, et al., editors, *Proc. 9th Asian Conference on Intelligent Information and Database Systems ACIIDS 2017, Part II*, volume 10192 of *Lecture Notes in Artificial Intelligence*, pages 276–286, Kanazawa, Japan, 3-5 Apr 2017. Springer. doi:10.1007/978-3-319-54430-4_27.

[8] A. Jóźwik, S. Serpico, and F. Roli. A parallel network of modified 1-NN and k-NN classifiers – Application to remote-sensing image classification. *Pattern Recogn. Lett.*, 19(1):57–62, January 1998. doi:10.1016/S0167-8655(97)00155-4.

[9] D. A. Karras. Improved defect detection using support vector machines and wavelet feature extraction based on vector quantization and SVD techniques. In *Proc. Int. Joint Conf. on Neural Networks, 2003*, volume 3, pages 2322–2327, July 2003. doi:10.1109/IJCNN.2003.1223774.

[10] J. Konieczny and G. Meyer. Computer rendering and visual detection of orange peel. *Journal of Coatings Technology and Research*, 9(3):297–307, 2012. doi:10.1007/s11998-011-9378-2.

[11] M. Kruk, B. Świderski, K. Śmietańska, J. Kurek, L. J. Chmielewski, J. Górski, and A. Orłowski. Detection of 'orange skin' type surface defects in furniture elements with the use of textural features. In K. Saeed, W. Homenda, and R. Chaki, editors, *Proc. 16th IFIP TC8 Int. Conf. Computer Information Systems and Industrial Management Applications CISIM 2017*, volume 10244 of *Lecture Notes in Computer Science*, pages 402–411, Białystok, Poland, 16-18 Jun 2017. Springer, Cham. doi:10.1007/978-3-319-59105-6_34.

[12] W. J. Krzanowski and D. J. Hand. Assessing error rate estimators: The leave-one-out method reconsidered. *Australian Journal of Statistics*, 39(1):35–46, 1997. doi:10.1111/j.1467-842X.1997.tb00521.x.

[13] K. Laszewicz and J. Górski. Control charts as a tool for the management of dimensional accuracy of mechanical wood processing (in Russian). *Annals of Warsaw University of Life Sciences-SGGW, Forestry and Wood Technology*, 65:88–92, 2008.

[14] K. Laszewicz, J. Górski, and J. Wilkowski. Long-term accuracy of MDF milling process–development of adaptive control system corresponding to progression of tool wear. *European Journal of Wood and Wood Products*, 71(3):383–385, 2013. doi:10.1007/s00107-013-0679-2.

[15] K. Laszewicz, J. Górski, J. Wilkowski, and P. Czarniak. Analysis of dimensional accuracy of milling process. *Wood Research*, 58(3):451–463, 2013.

[16] F. Longuetaud, F. Mothe, B. Kerautret, et al. Automatic knot detection and measurements from X-ray CT images of wood: A review and validation of an improved algorithm on softwood samples. *Computers and Electronics in Agriculture*, 85(0):77–89, 2012. doi:10.1016/j.compag.2012.03.013.

[17] E.-C. Musat, E.-A. Salca, F. Dinulica, et al. Evaluation of color variability of oak veneers for sorting. *BioResources*, 11(1):573–584, 2016. doi:10.15376/biores.11.1.573-584.

[18] M. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man and Cybern.*, 9(1):62–66, 1979. doi:10.1109/TSMC.1979.4310076.

[19] J. L. Pach. *Identification of the author of Latin manuscripts with the use of image processing methods*. Ph.d. thesis, Warsaw University of Technology, Faculty of Electronics and Information Technology, Warsaw, 2019.

[20] M. Sezgin and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146–165, 2004. doi:10.1117/1.1631315.

[21] B. Świderski, M. Kruk, G. Wieczorek, J. Kurek, K. Śmietańska, L. J. Chmielewski, J. Górski, and A. Orłowski. Feature selection for 'orange skin' type surface defect in furniture elements. In L. Rutkowski et al., editors, *Proc. Int. Conf. on Artificial Intelligence and Soft Computing ICAISC 2018*, volume 10842 of *Lecture Notes in Artificial Intelligence*, pages 81–91, Zakopane, Poland, 3-7 Jun 2018. doi:10.1007/978-3-319-91262-2_8.