FULL-LENGTH REPRINT – differs from the original in layout but not in contents
Chmielewski L.J., Orłowski A., Wieczorek G., Śmietańska K., Górski J.
Testing the limits of detection of the 'orange skin' defect in furniture elements with the HOG features.
In N.T. Nguyen et al., editors, Proc. 9th Asian Conference on Intelligent Information and Database Systems ACIIDS 2017,
volume 10192 of Lecture Notes in Artificial Intelligence, pages 276-286,
Kanazawa, Japan, 3-5 Apr 2017. doi:10.1007/978-3-319-54430-4_27

Testing the Limits of Detection of the 'Orange skin' Defect in Furniture Elements with the HOG features

Leszek J Chmielewski[1], Arkadiusz Orłowski[1], Grzegorz Wieczorek[1], Katarzyna Śmietańska[2], and Jarosław Górski[2]

[1] Faculty of Applied Informatics and Mathematics – WZIM,
[2] Faculty of Wood Technology – WTD

Warsaw University of Life Sciences – SGGW
ul. Nowoursynowska 159, 02-775 Warsaw, Poland
{leszek_chmielewski,grzegorz_wieczorek,katarzyna_laszewicz}@sggw.pl
http://www.wzim.sggw.pl, http://www.wtd.sggw.pl,

Abstract. In principle, the *orange skin* surface defect can be successfully detected with the use of a set of relatively simple image processing techniques. To assess the technical possibilities of classifying relatively small surfaces the Histogram of Oriented Gradients (HOG) and the Support Vector Machine were used for two sets of about 400 surface patches in each. Color, grey and binarized images were used in tests. For grey images the worst classification accuracy was 91% and for binarized images it was 99%. For color image the results were generally worse. The experiments have shown that the cell size in the HOG feature extractor should be not more than 4 by 4 pixels which corresponds to 0.12 by 0.12 mm on the object surface.

Keywords: orange skin, orange peel, surface defect, detection, quality inspection, furniture, Histogram of Oriented Gradients, HOG

1 Introduction

Vision systems make it possible to effectively inspect the machining quality of the selected furniture elements. They provide a possibility of quick and sufficiently precise testing of the correctness of the manufacturing of the product in terms of its relation to the prescribed tolerances [10,11]. One of the most effective ways of improving the quality of furniture elements is the system of correcting the dimensions of the element under treatment [12]. In the production of furniture, the dimensional and shape accuracy are equally important as the visual appearance of the elements, which is the aesthetic aspect of the product.

In our previous papers we assessed the viability of the image processing methods to primary measurements [4]. The 3D images from the structured light scanner, suitable for shape measurements, were used. It appeared that a common

N.T. Nguyen et al. (Eds.): ACIIDS 2017, LNAI 10192, pp. 276-286, 2017.
DOI: 10.1007/978-3-319-54430-4_27

surface defect called *orange skin* which can emerge in the painted surfaces is beyond the range of applicability of such images. In [5] we have shown that this defect should be easy to detect with simple methods.

As we have mentioned in the previous papers, in the domain of furniture elements quality control the image processing methods are very rarely if not at all mentioned in the literature. This is in contrast to the status in the timber industry, where image-based analysis of structural and anatomical defects is a well established technology with much literature (cf. the reviews [3,13]). Usually the reference to defects of furniture are only mentioned alongside with other domains of application, like for example in [8], or the quality is related more to the raw material rather than to the final product, like in [14]. The *orange skin* as a surface defect is considered in two papers. In [9] the images of *orange peel* (another name for *orange skin*) are generated and the visibility of this defect for humans in various conditions is investigated. The paper [2] deserves particular attention. A system which evokes what the authors call the *defect augmentation phenomena* with a complex, moving system of lighting synchronized with cameras, is used to substantially improve the results of quality inspection of painted surfaces in car industry. The system is aimed at a broad class of surface defects, including *orange peel*. The lighting makes it possible to use local adaptive thresholding as the detection method.

The motivation for this paper is as follows. In [5] we have shown that the *orange skin* surface defect can be successfully detected with the use of a set of relatively simple image processing techniques. We have used the gradient modulus as a single feature of the surface and we have shown that it is possible to attain a very good classification result by properly selecting a threshold on this feature. A 100% classification accuracy was attained for a broad range of thresholds. This result was obtained under an assumption that the entire tested furniture element is treated as the classified object. Therefore, a single sign of defectiveness was enough to reject an element. In this paper we try to do three things. First, we look more carefully at the fragments of the surface and try to classify them separately to see what are the limitations of classification accuracy of surface details, because the result obtained in our cited previous paper seems excessively optimistic. Second, we try to go forward in using the gradient as a feature and we apply one of the most successful while still relatively simple gradient-based feature generating method, namely the *Histogram of Oriented Gradients* (*HOG*) [7]. Third, to go beyond simple thresholding, but mainly because now we have numerous features, we apply the Support Vector Machine (SVM) [1,6] as a classifier. In this way, we show the possibilities and constraints of the method in a more reliable way and we test its limits with respect to classification accuracy of the local state of the object surface. We attempt not to use any special lighting and camera system as in [2] because of several reasons. First, the system of [2] is patented. Second, its complexity and cost would prevent it from being applied in the furniture industry, in which the cost of products is importantly lower than that in the car industry. Third, in our opinion the *orange skin* defect can be reliably detected with standard cameras and lighting.

Our results will not be compared to any reference ones. As already said, to our best knowledge there were no literature reports on automatic detection of *orange skin* in furniture elements up till now.

This paper is organized as follows. In the following Section we shall briefly characterize the defect considered. In the next Section we shall present the way we have prepared the data for finding the features in the images and for teaching and testing the classifier. We shall also give the details on the parameters of the methods we have used. In the following Section we shall report on the results obtained. The paper will be concluded in the final part.

2 The defect: *orange skin*

Orange skin or *orange peel* is a common term used to describe a surface with small, shallow hollows. In the context of furniture manufacturing it is one of the surface defects with emerges during finishing the surface by lacquering. It manifests itself with uneven structure of the hardened surface. There are various reasons for this defect: 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 surface processed is flat and fully covered with lacquer, so *orange skin* can be treated as the only reason for surface unevenness.

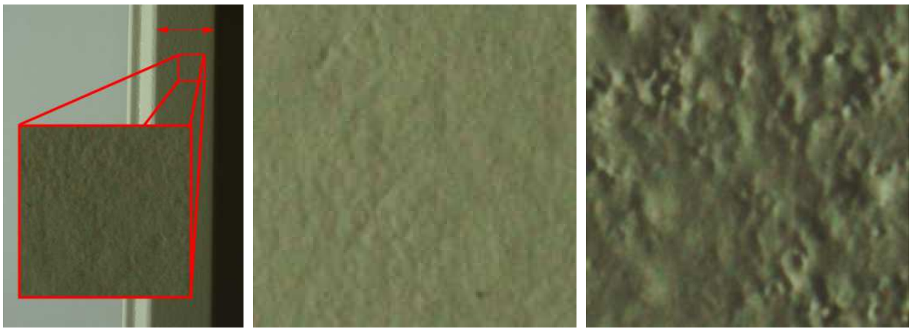
Several observations concerning this defect can be made.

1. The reasons for which a surface is considered good or defective are of esthetic nature. There are treatments and surface types for which the deviations from planarity are considered beneficial.
2. It is practically impossible to find a separate *defect* in the surface. The out-of-plane deviations or simply the small valleys and holes in the surface gradually pass to the regions of even surface. These *good* regions are smaller or larger, but irrespective of their size the whole surface should be classified as *bad*.
3. For a human eye, the smallest regions which can still be recognized as *good* or *bad* span the regions from about 0.5×0.5 to 1×1 mm.
4. The *good* surface is not free from deviations which can easily be seen in enlarged images. However, these deviations are small enough for the human to classify the whole surface as *good*.

We shall try to take these observation into account while explaining the surface classification method we have used.

3 The data and the methods

Images of elements with good surface and with *orange skin* were taken with a good quality, general-purpose color camera, at a resolution of 2.5 M pixels (1288×1936 pix). The previous study [5] indicated that the line from the light source to the center of the imaged surface should be far from perpendicular to



a **b** **c**

Fig. 1. (**a**) The way the training images were cut from the original; the length marked by an arrow is about 12 mm and corresponds to 400 pixels. (**b**) Example of a *good* region. (**c**) Example of a *bad* region. Contrast and brightness of images **b** and **c** enhanced for better visibility in print.

the surface normal. In our case the light was about 70° from the normal, so the light we have used can be considered as close to tangential. A fragment of the image is shown in Fig. 1a. The resolution at the object surface was close to 0.03 mm per pixel, so the width of the flat surface marked with an arrow corresponding to 12 mm is slightly more than 400 pixels.

From these images, the windows of size 200×200 pixels containing good and bad surfaces were cut. The width of the element and the resolution allowed for cutting two windows from the width of the element, not all the windows cut were close to the element boundary. The windows were cut at random. Each window will be treated as one object by the teaching and classification processes. The numbers of these objects will be given later on in Section 4.

In the histograms of oriented gradients according to [7] the following parameters were used: cell size 2×2, 4×4 and 8×8 pixels. The remaining parameters were not varied: block size 2×2, block overlap 1×1 pixels and number of bins in the histograms 8. The classification was performed with the SVM classifier according to [1,6]. The classification accuracy was measured as the cumulated posterior probability: the number of true positive classifications versus the number of all classified objects. The calculations were made in series: for each cell size the features were calculated from the original color image, grey level image, and additionally from the binarized image. Grey images were formed by calculating the brightness as $0.299R + 0.587G + 0.114B$. Binarization was performed with the Otsu method [15] from the grey images.

The teaching and testing were preformed in two settings. In the setting denoted as *separate* the testing objects (that is, image windows) were cut from a different set of images of furniture elements than the teaching objects were. In the setting denoted *common* the teaching and testing objects were cut from the same set of images. Obviously, in each setting, no teaching object was used in the testing, and vice versa.

Table 1. Numbers of objects in the teaching and testing sets in the two settings used.

Setting	No. teaching objects			No. testing objects			Total
	good	bad	subtotal	good	bad	subtotal	
separate	142	142	284	44	65	109	393
common	157	157	314	50	50	100	414

Table 2. Accuracy and time of teaching and testing phases together.

Setting	Cell size	Color		Grey		Binary	
		accuracy [%]	time [s]	accuracy [%]	time [s]	accuracy [%]	time [s]
Separate	2×2	**99.17**	13.04	97.25	12.66	**100.00**	9.69
	4×4	97.25	4.53	**99.08**	4.04	99.08	3.40
	8×8	94.50	2.86	97.25	2.04	97.25	2.10
Common	2×2	83.00	13.35	81.00	12.01	**99.00**	9.71
	4×4	**90.00**	4.60	**91.00**	3.80	**99.00**	3.46
	8×8	**90.00**	3.03	88.00	2.95	98.00	2.40

4 Results

The numbers of objects in the teaching and testing sets in the settings *separate* and *common* are given in Tab. 1.

The results of teaching the classifier and testing its accuracy are summarized in Tab. 2. The cumulated times of the teaching and testing phases were given only for the general information purpose. They reflect rather the teaching time, as the classification time which is more important in the case of an industrial task would be negligible with the use of suitable equipment. The calculations were carried out in Matlab® environment on a four-processor, 3.6 GHz, 64 bit personal computer.

In Figs. 2 and 3 we give the examples of a good and defective training object, its binary image and the visualization of HOG features for the 4×4 pix cell size. In Figs. 4 and 5 we give examples of negative classification results. The visualizations of a large number of features is not very informative but by showing them in this way we make an attempt to look more precisely at the classification process. The close analysis of these images indicates that for *bad* objects the HOG features for the binary image exhibit some regions of small and some regions of large values. This is due to the the lack of homogeneity of the surface of such objects. *Good* objects are more homogeneous and do not exhibit such a behavior. This can be a candidate phenomenon for explaining the success of the classifier for binary objects.

The resolution of the images indicates that the windows mentioned in observation 3, Sect. 2, corresponds to from 16×16 to 32×32 pixels. It can be disappointing that in the experiments, the errors were the smallest for windows of lesser dimensions than these, and that the error increased with the increase of the window dimension, which suggests that for larger windows the errors will not be smaller. This indicates that the way the human recognizes the defect

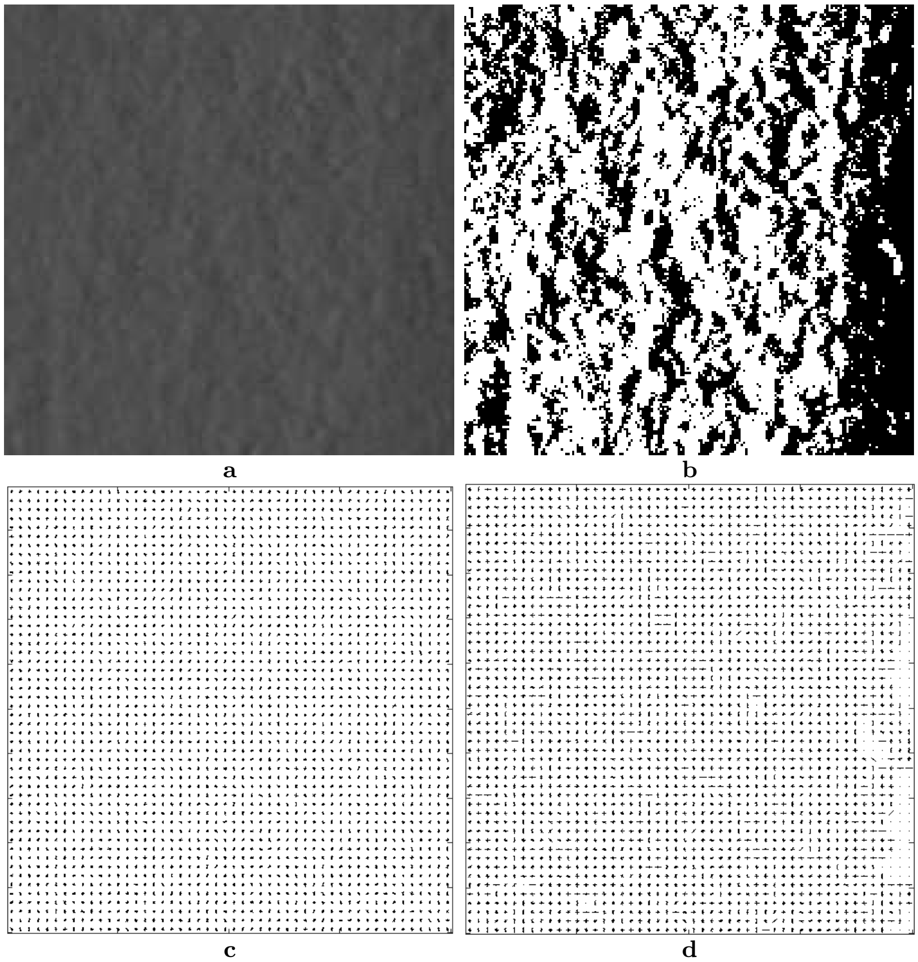


Fig. 2. An example of a *good* training object. (**a**) Grey image; (**b**) binarized **a**; (**c**) visualization of HOG features for **a** in 4×4 cell; (**d**) the same for **b**.

differs much from machine classification. The best results were obtained for the cell size not more than 4×4 pixels which corresponds to 0.12 by 0.12 mm on the object surface.

It seems reasonable that good results were received for grey images as the color information is not significant in the image we analyzed. However, it is disappointing that the best good results were obtained for binary images. This could indicate that the classical Otsu method appears to extract significant information from the images of our interest. It is also interesting that the results for the testing objects selected from the same physical objects are worse than for those from different objects.

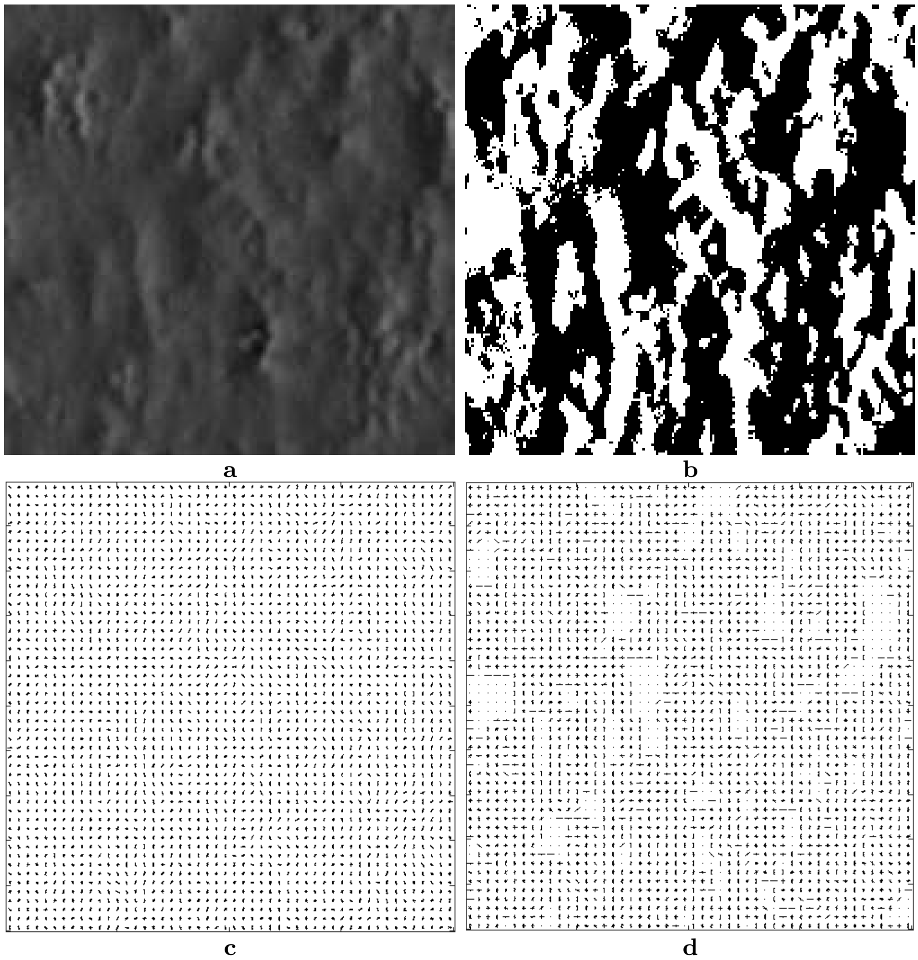


Fig. 3. An example of a *bad* training object. (**a**) Grey image; (**b**) binarized **a**; (**c**) visualization of HOG features for **a** in 4×4 cell; (**d**) the same for **b**.

The results suggest that more work should be done on choosing the features for classification, including a deeper insight into the processing of the binary images. The differences between the results for the two settings indicate that the results are not stable yet and more attention should be paid to the preparation of the training and testing sets. The possible sources of errors, like the light conditions, and the resolution and focusing of the images, should be taken into account with more care.

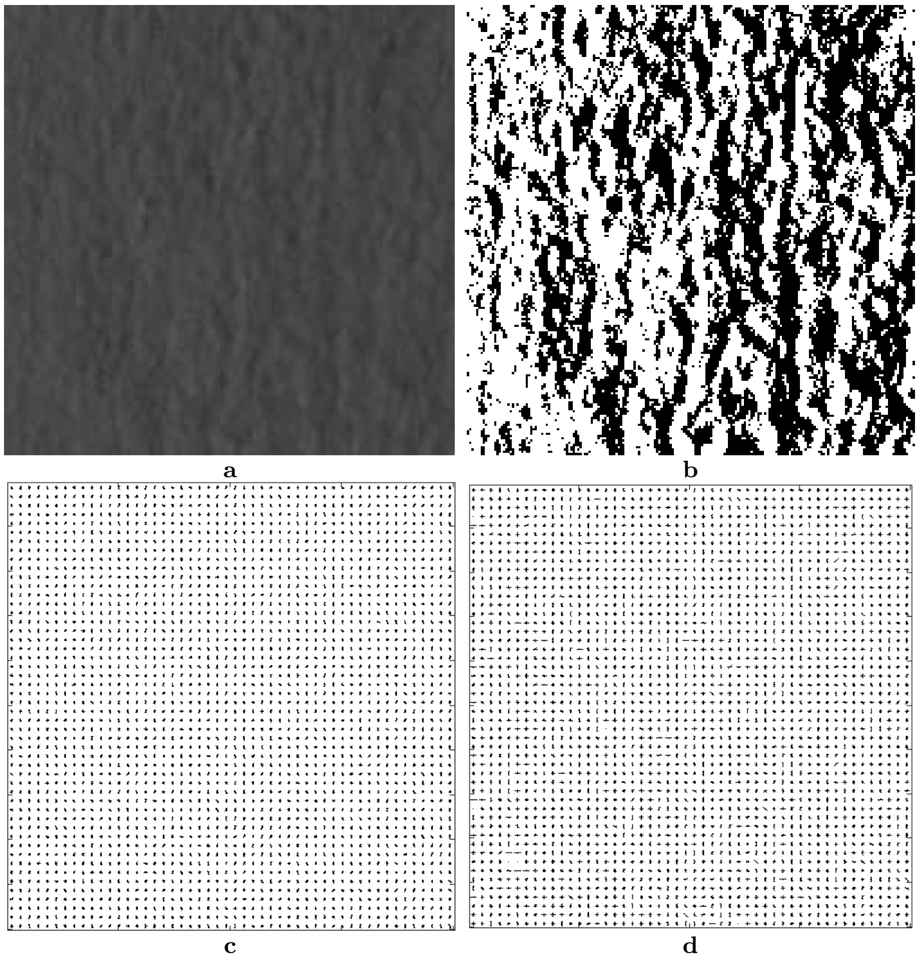


Fig. 4. An example of a false negative error: *good* object classified as *bad.* (**a**) Grey image; (**b**) binarized **a**; (**c**) HOG features for **a** in 4×4 cell; (**d**) the same for **b**.

5 Conclusion

In this report of the work in progress it has been demonstrated that the detection of the surface defect of the type *orange skin*, although in principle realizable with relatively simple image processing methods, is a demanding task in general. The well established methods of feature extraction and classification, like the histogram of oriented gradients and the support vector machine, give positive results in the majority of cases, but can fail for a significant number of images, for which a human inspector would make no error. One of the results was that the images binarized with the Otsu method carry the information on the defect of the tested type very well. Consequently, for binary images the classification

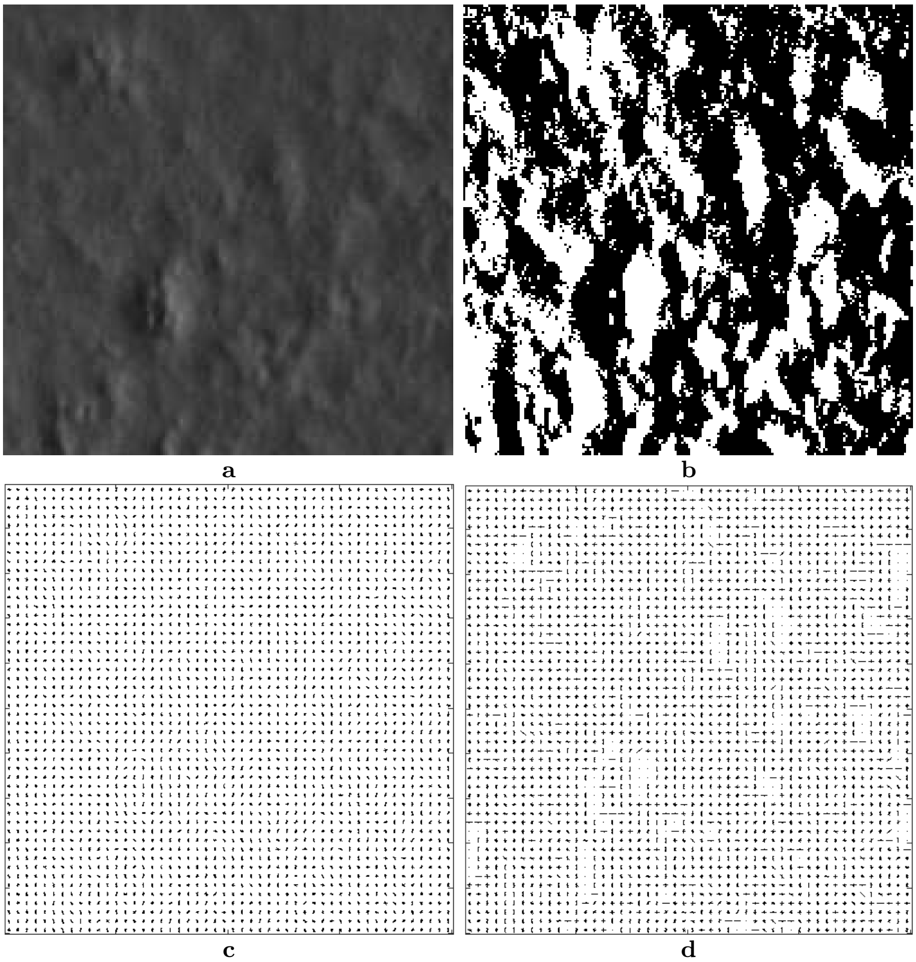


Fig. 5. An example of a false positive error: *bad* object classified as *good*. (**a**) Grey image; (**b**) binarized **a**; (**c**) HOG features for **a** in 4×4 cell; (**d**) the same for **b**.

accuracy attained 99%. This result indicates that the detection of the *orange skin* defect is tractable with the use of a standard camera and lighting system.

Having this result in mind it can be safely stated that better controlling the lighting and imaging processes and using a broader range of image features should make it possible to attain a technically acceptable level of accuracy in the task considered. The research will be continued in these directions.

References

1. Allwein, E.L., Schapire, R.E., Singer, Y.: Reducing multiclass to binary: A unifying approach for margin classifiers. Journal of Machine Learning Research 1, 113–141

(2000), www.jmlr.org/papers/v1/allwein00a

2. Armesto, L., Tornero, J., Herraez, A., Asensio, J.: Inspection system based on artificial vision for paint defects detection on cars bodies. In: 2011 IEEE International Conference on Robotics and Automation. pp. 1–4 (May 2011), doi:10.1109/ICRA.2011.5980570
3. Bucur, V.: 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. Chmielewski, L.J., Laszewicz-Śmietańska, K., Mitas, P., Orłowski, A., et al.: Defect detection in furniture elements with the Hough transform applied to 3D data. In: Burduk, R., et al. (eds.) Proc. 9th Int. Conf. Computer Recognition Systems CORES 2015. Advances in Intelligent Systems and Computing, vol. 403, pp. 631–640. Springer, Wrocław, Poland (25-27 May 2015), doi:10.1007/10.1007/978-3-319-26227-7_59
5. Chmielewski, L.J., Orłowski, A., Śmietańska, K., Górski, J., et al.: Detection of surface defects of type 'orange skin' in furniture elements with conventional image processing methods. In: Huang, F., Sugimoto, A. (eds.) Image and Video Technology – PSIVT 2015 Workshops. Lecture Notes in Computer Science, vol. 9555, pp. 26–37. Springer, 2016, Auckland, New Zealand (23-27 Nov 2015), doi:10.1007/978-3-319-30285-0_3
6. Cortes, C., Vapnik, V.: Support-vector networks. Machine Learning 20(3), 273–297 (1995), doi:10.1007/BF00994018
7. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). vol. 1, pp. 886–893 (June 2005), doi:10.1109/CVPR.2005.177
8. Karras, D.A.: 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. vol. 3, pp. 2322–2327 (July 2003), doi:10.1109/IJCNN.2003.1223774
9. Konieczny, J., Meyer, G.: 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
10. Laszewicz, K., Górski, J.: 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)
11. Laszewicz, K., Górski, J., Wilkowski, J.: 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
12. Laszewicz, K., Górski, J., Wilkowski, J., Czarniak, P.: Analysis of dimensional accuracy of milling process. Wood Research 58(3), 451–463 (2013)
13. Longuetaud, F., Mothe, F., Kerautret, B., 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
14. Musat, E.C., Salca, E.A., Dinulica, F., 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
15. Otsu, M.: 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