# Sequential Hierarchical Image Recognition Based on the Pyramid Histograms of Oriented Gradients with Small Samples

Andrey V. Savchenko[1(✉)], Vladimir R. Milov[2], and Natalya S. Belova[3]

[1] Laboratory of Algorithms and Technologies for Network Analysis, National Research University Higher School of Economics, Nizhny Novgorod, Russia
avsavchenko@hse.ru

[2] Nizhny Novgorod State Technical University n.a. R.E. Alekseev, Nizhny Novgorod, Russia
vladimir.milov@gmail.com

[3] National Research University Higher School of Economics, Moscow, Russia
nbelova@hse.ru

**Abstract.** In this paper we explore an application of the pyramid HOG (Histograms of Oriented Gradients) features in image recognition problem with small samples. A sequential analysis is used to improve the performance of hierarchical methods. We propose to process the next, more detailed level of pyramid only if the decision at the current level is unreliable. The Chow's reject option of comparison of the posterior probability with a fixed threshold is used to verify recognition reliability. The posterior probability is estimated for the homogeneity-testing probabilistic neural network classifier on the basis of its relation with the Bayesian decision. Experimental results in face recognition are presented. It is shown that the proposed approach allows to increase the recognition performance in 2–4 times in comparison with conventional classification of pyramid HOGs.

**Keywords:** Image recognition · Hierarchical recognition · Sequential analysis · Chow's reject option · Probabilistic neural network · HOG (Histograms of oriented Gradients) · PHOG (Pyramid HOG)

## 1 Introduction

Nowadays the recognition of complex images (faces, gestures, medical objects) [1] becomes all the more acute. Several sufficiently reliable descriptors has been recently proposed, e.g., SIFT (Scale-Invariant Feature Transform) [2], HOG (Histograms of Oriented Gradients) [3], SURF (Speeded-Up Robust Features) [4], etc. However, though image recognition technology has reached a certain level of maturity [5], many researchers try to improve the recognition quality by exploiting the hierarchical processing of analyzed objects. Such processing is known to be one of the known characteristic of human intellect [6, 7]. As a result, several hierarchical image recognition


M.Y. Khachay et al. (Eds.): AIST 2015, CCIS 542, pp. 14–23, 2015.
DOI: 10.1007/978-3-319-26123-2_2

methods have been presented, e.g., pyramids of features (in particular, PHOG (Pyramid HOG) descriptor [8]), hierarchical temporal memory [6], wavelet-analysis [9], deep convolution neural networks [10], etc.

The majority of these methods include the comparison of features of query image and models from the given database at *each* level of hierarchy (pyramid). Hence, their recognition performance is much worse in comparison with conventional non-hierarchical approach. For instance, the matching of the PHOG descriptors [7] is usually 1.5–3 times slower than the matching of the state-of-the-art HOGs [3]. Thus, most popular hierarchical methods usually cannot be implemented in real-time applications [11], e.g., in video-based face recognition [12]. In this paper we propose to perform sequential analysis of query image to overcome this drawback. Namely, more detailed representation of the query image is analyzed at the next level of pyramid *only if* it is impossible to obtain a reliable solution at the current level [13, 14]. To reject unreliable solution, we use the Chow's rule and compare the maximal posterior probability with the fixed threshold [13]. This probability is estimated on the basis of the asymptotic properties of the homogeneity-testing probabilistic neural network (HT-PNN) which was proved to be a good classifier in several pattern recognition tasks [15].

The rest of the paper is organized as follows: Sect. 2 briefly presents the classification of conventional HOG and PHOG features with the HT-PNN. In Sect. 3, the sequential hierarchical approach is introduced. In Sect. 4, we present the experimental results of the proposed method in the face recognition task [12]. Finally, concluding comments are given in Sect. 5.

## 2 Classification of the PHOG with the HT-PNN

Let a training set of $R > 1$ model images $\{X_r\}$, $r \in \{1, \ldots, R\}$ with height $U_r$ and width $V_r$ be specified. It is assumed that the class $c(r) \in \{1, \ldots, C\}$ of the $r$th model image is known. Here $C$ is the total number of distinct classes. The task is to assign a query image $X$ with height $U$ and width $V$ to one of $C$ classes [16]. We assume the practically important case of small training sample $C \approx R$ [17]. Let's the objects of interest (say, faces in face recognition) are preliminary detected in either query or model images, hence, each of them contains only one object.

Let every image to be associated with a set of PHOG features [8]. At first, the whole image is divided into a regular squared grid of $K^{(1)} \times K^{(1)}$ blocks where $K^{(1)}$ is the number of rows and columns in the lowest level of pyramid with the most rough image approximation. The histogram $H_r^{(1)}(k_1, k_2) = [h_{r;1}^{(1)}(k_1, k_2), \ldots, h_{r;N}^{(1)}(k_1, k_2)]$ of gradient orientation is evaluated for each block $(k_1, k_2), k_1, k_2 \in \{1, \ldots, K^{(1)}\}$ of the $r$th model image. Here $N$ is the number of bins in the histogram. Next, according to the PHOG method [8], the image is divided into a squared grid $K^{(2)} \times K^{(2)}$ blocks where $K^{(2)} > K^{(1)}$ and its histogram $H_r^{(2)}(k_1, k_2), k_1, k_2 \in \{1, \ldots, K^{(2)}\}$ is evaluated. This procedure is repeated for $L = const$ pyramid levels. Similarly, the histograms $H^{(l)}(k_1, k_2) = [h_1^{(l)}(k_1, k_2), \ldots, h_1^{(l)}(k_1, k_2)], l = \overline{1, L}, k_1, k_2 \in \{1, \ldots, K^{(l)}\}$ for the query image are estimated.

The second part is classifier design [16]. If $C \approx R$, the nearest neighbor rule is applied as it usually outperforms other more complex state-of-the-art machine learning techniques (MLP, SVM, etc.) in this particular case [17]. Though several conventional classifiers (e.g., CNN) can be applied in this task to select the features [18], the final decision is usually done with simple nearest neighbor rule. In view of the small spatial deviations due to misalignment after object detection, the following similarity measure with mutual alignment of blocks and comparison of the histograms in $\Delta$-neighborhood of each block is used [19]

$$\rho^{(l)}(X, X_r) = \sum_{k_1=1}^{K^{(l)}} \sum_{k_2=1}^{K^{(l)}} \min_{\substack{|\Delta_1| \le \Delta, \\ |\Delta_2| \le \Delta}} \rho_H\left(H_r^{(l)}(k_1+\Delta_1, k_2+\Delta_2), H^{(l)}(k_1, k_2)\right). \tag{1}$$

Here $\rho_H$ is any distance between HOGs. In this paper, we explore the square of Euclidean distance and our HT-PNN [15]:

$$\rho\left(H_r^{(l)}, H^{(l)}\right) = \sum_{i=1}^{N} \left( h_i^{(l)} \ln \frac{2\tilde{h}_i^{(l)}}{\tilde{h}_i^{(l)} + \tilde{h}_{r;i}^{(l)}} + h_{r;i}^{(l)} \ln \frac{2\tilde{h}_{r;i}^{(l)}}{\tilde{h}_i^{(l)} + \tilde{h}_{r;i}^{(l)}} \right), \tag{2}$$

where

$$\begin{aligned} \tilde{h}_{r;i}^{(l)} &= \sum_{j=1}^{N} K_{ij} h_{r;i}^{(l)} \\ \tilde{h}_i^{(l)} &= \sum_{j=1}^{N} K_{ij} h_i^{(l)} \end{aligned}, \tag{3}$$

is the convolution of the HOGs with the Gaussian Parzen kernel $K_{ij}$ [20] and indexes $(k_1, k_2)$ are missed for clarity.

In general case ($L > 1$) the outputs (1) at each level are aggregated [8] and $X$ is assigned to the class of the closest model in terms of the following similarity measure

$$\rho_{PHOG}(X, X_r) = \sum_{l=1}^{L} w^{(l)} \cdot \rho^{(l)}(X, X_r), \tag{4}$$

Here weights of every pyramid level $w^{(l)} > 0$ are usually chosen experimentally. Unfortunately, though the accuracy of the PHOG is usually higher than the accuracy of the HOG [8], performance of the nearest neighbor rule with similarity measure (4) is quite low as it requires $\sum_{l=1}^{L} \left(K^{(l)} \cdot K^{(l)}\right) / K^{(1)} \cdot K^{(1)}$-times more calculations in comparison with conventional HOG [3] ($L = 1$). Hence, in the next section we present s way to improve the PHOG's performance by sequential analysis at each level.

## 3 Sequential Classification of the PHOG

The key idea of our paper is to perform recognition at each level independently and terminate the classification procedure if the decision was assumed to be reliable [21]. Otherwise, the next level of the hierarchy is analyzed in the same way. To verify the decision's reliability, let's use the statistical approach [16]. It is known [15, 22] that if the hypothesis $W_r^{(l)}$ for homogeneity of query $X$ and model objects $X_r$ at $l$th level are tested and it is assumed that the prior probabilities of each class are equal, then the nearest neighbor rule with the HT-PNN (1)–(3)

$$v(l) = \underset{r \in \{1,\ldots,R\}}{\arg\min}\, \rho^{(l)}(X, X_r) \tag{5}$$

is equivalent to the maximal likelihood decision. In such case, an optimal Bayesian criterion for rejection of unreliable decision is achieved with the Chow's rule [13]

$$P\left(W_{v(l)}^{(l)} \middle| X\right) \le p_0 = const, \tag{6}$$

where $P\left(W_{v(l)}^{(l)} \middle| X\right)$ is the posterior probability of hypothesis $W_{v(l)}^{(l)}$ and threshold

$$p_0 = \frac{\Pi_{10} - \Pi_{ro}}{\Pi_{01} + \Pi_{10} - \Pi_{ro}}. \tag{7}$$

Here $\Pi_{10}$ is the losses of incorrect decision (5), which has not been rejected (6), $\Pi_{01}$ is the losses of rejection (6) of correct decision (5) and $\Pi_{ro}$ is the cost of reject option (6) (obviously, $\Pi_{ro} \le \Pi_{10}$).

Hence, the task is to estimate the posterior probability $P(W_{v(l)}^{(l)} | X)$. If the classes are equiprobable, the Bayes theorem can be used

$$P\left(W_{v(l)}^{(l)} \middle| X\right) = \frac{P\left(X | W_{v(l)}^{(l)}\right)}{\sum_{r=1}^{R} P\left(X | W_r^{(l)}\right)}. \tag{8}$$

Here $P\left(X | W_r^{(l)}\right)$ is the conditional probability (likelihood) of the $r$th class. Fortunately, given the relationship of the HT-PNN and maximal likelihood [15], we get the final estimation of the posterior probability:

$$\hat{P}\left(W_{v(l)}^{(l)}\middle|X\right) = \frac{\exp\left(-UV \cdot \rho^{(l)}(X, X_{v(l)})\right)}{\sum\limits_{r=1}^{R} \exp(-UV \cdot \rho^{(l)}(X, X_r))}. \tag{9}$$

Our sequential recognition process is represented in Fig. 1. At first, the most rough approximations of images with small number of blocks $K^{(1)}$ are analyzed. If in this case it is possible to obtain a reliable solution with high posterior probability (9), the process is terminated and $c\left(v^{(1)}\right)$ (5) becomes the resulted class. Otherwise the description of the query object is detailed and the process is repeated until obtaining the reliable solution $c\left(v^{(l)}\right)$ at the $l$th level.

The last open question here is the final processing if decisions at all $L$ levels are unreliable. In this case it is necessary to obtain the best solution from the set of candidates $\left\{c\left(v^{(l)}\right)\right\}, l \in \{1, \ldots, L\}$. The most evident way here is to perform fusion of classifiers (1) for each level. The most complex fusion methods (bagging, boosting, etc.) [16] require the large training set and cannot be used if $C \approx R$. Thus, in this paper we propose to use conventional principle of maximum posterior probability [16] - the final decision is taken in favor of class $c\left(v^{(l^*)}\right)$, where

$$l^* = \underset{l \in \{1,\ldots,L\}}{\arg\max}\, \hat{P}\left(W_{v(l)}^{(l)}\middle|X\right). \tag{10}$$

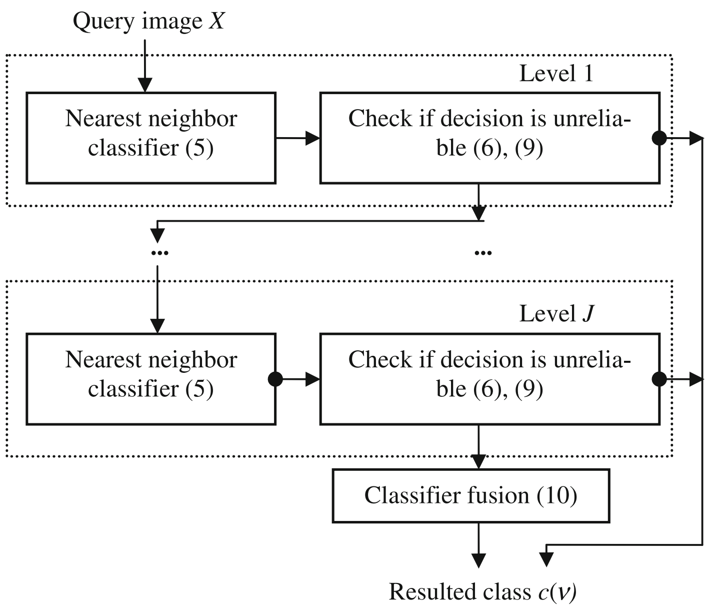


**Fig. 1.** Proposed sequential hierarchical procedure of image recognition

Thus, in the proposed method (5), (6), (9), (10) we implement an approach widely used in hierarchical image recognition [6, 16], namely, the build of the pyramid of features for various image resolution [8]. However, unlike the known similar PHOG-based methods (4) with simultaneous comparison of the union of features for all levels, we perform sequential analysis (compare with [14]) of detailed approximations of the query object and move to the next pyramid level *only if* the solution at the current level is not reliable (6). As a result, the performance of our approach should be much better in comparison with conventional nearest neighbor rule with similarity measure (4). The next section experimentally supports this claim.

## 4 Experimental Results

Our experimental study deals with the face recognition task. We combined 3 datasets, namely, AT&T [23], Yale [24] and JAFFE [25], into one set to demonstrate the flexibility of hierarchical recognition system (Fig. 1). AT&T dataset contains 400 images of 40 different people varied in pose. Yale database (165 photos of 15 persons) is used to test recognition with various light conditions. JAFFE dataset contains 213 images of 10 Japanese female persons varied in facial expressions. In total, our dataset consists of 778 photos of $C$ = 65 persons. The faces were detected with the LBP cascade classifier [26] from OpenCV library [27] which was proved to be one of the best known face detection algorithms [28]. The median filter with window size (3 × 3) was applied to remove noise in detected faces. The number of bins in the HOG $N = 8$. Threshold $p_0$ for posterior probability in the Chow's rule (6) is equal to 0.85. The neighborhood size in (1) is equal to $\Delta = 1$ [19].

We compared the performance of the proposed sequential recognition with the PHOG (4) and conventional HOG (number of levels $L = 1$). We use hierarchies with 2 and 3 levels. In the first case, $K^{(1)} = 10$ and $K^{(2)} = 20$. In the second case, $K^{(1)} = 10$, $K^{(2)} = 15$ and $K^{(3)} = 20$. Weights in (4) were found experimentally to obtain the higher accuracy.

We evaluate the error rate (in %) and the average time (in ms) to recognize one test image with a modern laptop (4 core i7, 6 Gb RAM) and Visual C++ 2013 compiler and optimization by speed. We use multithreading to make brute-force search (1), (5) faster. Each thread is implemented with Windows ThreadPool API and operates only on a subset of the database. The whole training sample is divided into 8 distinct parts, i.e., we look for the nearest neighbor (5) in 8 parallel threads. The recognition performance was estimated by the following cross-validation procedure. At first, the number of photos per one person $n_p = const$ is fixed. For each person, we randomly choose $n_p$ photos and put them into the model database $\{X_r\}$. Other photos are put into the test set. Then we estimate the error rate of test set recognition. This experiment is repeated 20 times. Finally, we estimate the mean of the error rate and recognition time for all experiments. The error rate and average recognition time for the HT-PNN (1)–(3) distance in dependence on the size of the training set $R$ are shown in Figs. 2 and 3, respectively. Here "10 × 10" bar stands for the HOG with 10 × 10 grid (i.e., $L = 1$, $K^{(1)} = 10$), "20 × 20" bar represents the results of the HOG with 20 × 20 grid ($L = 1$, $K^{(1)} = 20$), "PHOG (10 × 10 + 20 × 20)" and "PHOG

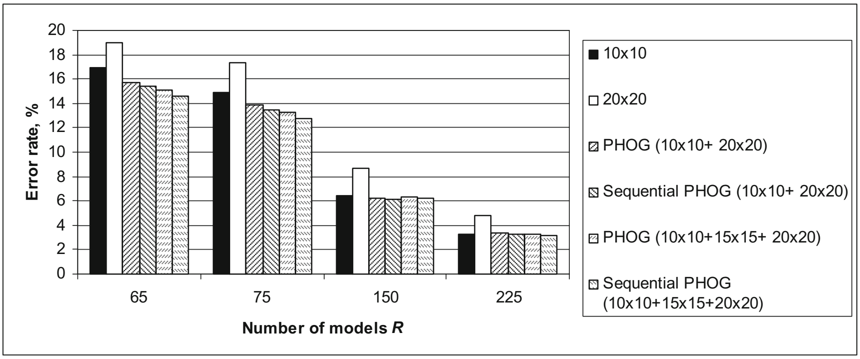


**Fig. 2.** Dependence of the error rate on the size of the training set $R$, HT-PNN (2), (3).

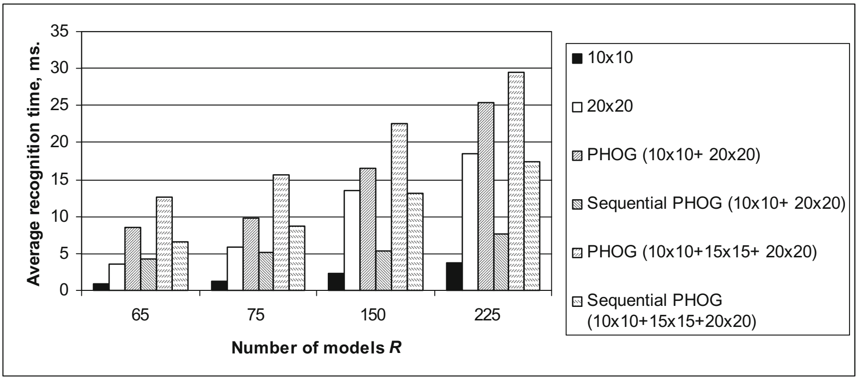


**Fig. 3.** Dependence of the average recognition time (ms.) on the size of the training set $R$, HT-PNN (2), (3).

$(10 \times 10 + 15 \times 15 + 20 \times 20)$" stand for conventional PHOG (4) with $L = 2$ (grids $K^{(1)} = 10$ and $K^{(2)} = 20$) and $L = 3$ (grids $K^{(1)} = 10$, $K^{(2)} = 15$ and $K^{(3)} = 20$) levels, respectively. The results of proposed sequential analysis (Fig. 1) are represented by the bars "Sequential PHOG $(10 \times 10 + 20 \times 20)$" and "Sequential PHOG $(10 \times 10 + 15 \times 15 + 20 \times 20)$" for $L = 2$ and $L = 3$ hierarchical levels with the same grid size, respectively.

Here one can notice that the accuracy of hierarchical approach is higher than the accuracy of the state-of-the-art HOG. It is especially true for small size $R$ of the training sample. The difference in error rates of the PHOG (4) and our approach (Fig. 1) is not statistically meaningful (Fig. 2). However, the average recognition time (Fig. 3) for the proposed method is 1.5–3.5 times lower in comparison with conventional aggregation (5) as in most cases (especially for large $R$) the reliable solution (6) was found at the first level $\left(K^{(1)} = 10\right)$.

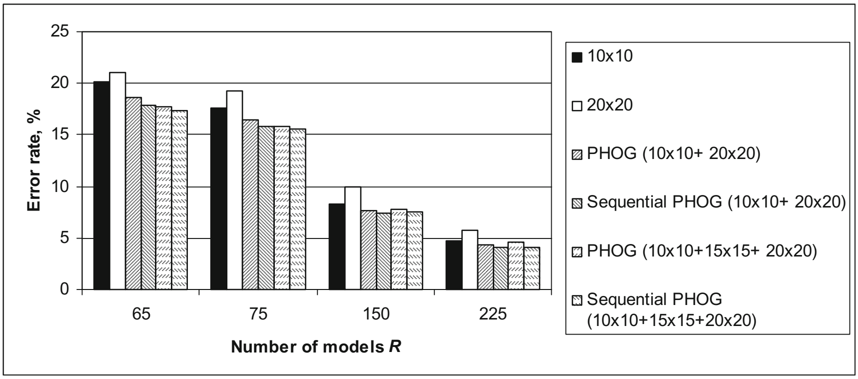


**Fig. 4.** Dependence of the error rate on the size of the training set $R$, Euclidean metric.

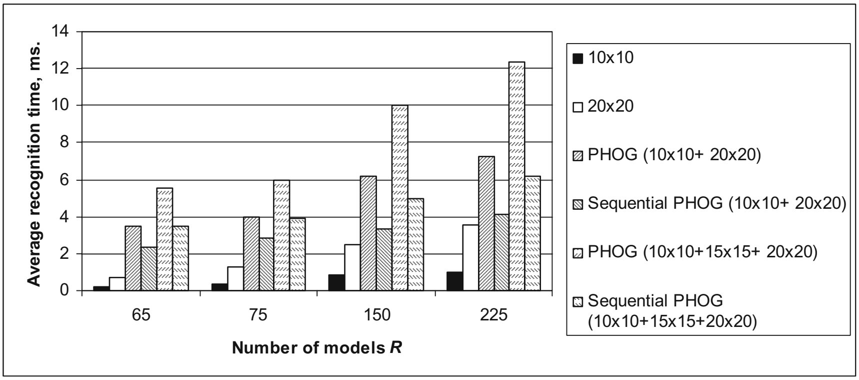


**Fig. 5.** Dependence of the average recognition time (ms.) on the size of the training set $R$, Euclidean metric.

In the second experiment we used conventional Euclidean distance to compare HOGs in (1) instead of the HT-PNN (2), (3). Error rates and average recognition time are shown in Figs. 4 and 5.

Based on these results, it is possible to draw the following conclusions. First, the error rate of Euclidean distance is 1–3.5 % higher than the HT-PNN's error rate. Second, comparing Figs. 2 and 4, the losses in the accuracy of conventional non-hierarchical approach (HOG [3]) in the experiment with Euclidean distance are much more noticeable than for the HT-PNN. Third, as the Euclidean distance is much more simple than the HT-PNN in computational sense, in some cases (with low $R$) the 1-ms gain in performance of our approach over aggregation (4) is not practically noticeable. Finally, though it is reasonable to use $L = 3$ levels for the HT-PNN with low values of the size $R$ of the training set (Fig. 2), there is no reason to prefer it to the simple case of $L = 2$ levels of the hierarchy for Euclidean metric.

## 5 Conclusion and Future Work

It is well-known that the hierarchical approach allows increasing the accuracy of image recognition (Figs. 2 and 4) [7, 8]. Unfortunately, the performance of such methods is usually insufficient for many practical applications. Hence, engineers have to apply more simple nonhierarchical methods. It seems that the improvement of performance of hierarchical methods is one of the most crucial tasks in this field.

Thus, in this paper we introduced a hierarchical image recognition algorithm based on statistical approach and the Chow's rule (6) with the estimate of the posterior probability (9) on the basis of the properties of the HT-PNN [15]. We significantly decreased the average recognition time by using the sequential analysis of the images with different level of granularity. At first, the roughest approximations are analyzed to speed-up the recognition procedure. The images are analyzed in detailed way if the current decision is not reliable (7). The proposed approach (Fig. 1) showed its efficiency in the practically important face recognition problem. In some cases (Fig. 4) our approach is even able to increase the recognition accuracy over the PHOG (4). Though the posterior probability (9) is estimated based on the HT-PNN, this expression can be used with other similarity measures. Really, expression (9) is very similar to the output of widely used probabilistic neural network [20]. For instance, in our experiment we have shown the possibility to combine our approach with the state-of-the-art Euclidean metric. It is important to emphasize that the accuracy of proposed solution is even 0.4–0.8 % higher in comparison with the PHOG.

The further research of our approach (Fig. 1) can be continued in the following directions. First, it is an application of approximate nearest neighbor methods [11] to speed-up recognition at each level of hierarchy. Another possible direction is the application of our sequential hierarchical method in various tasks of classification of complex objects, e.g., speech recognition.

**Acknowledgements.** Andrey V. Savchenko is supported by RSF (Russian Science Foundation) grant 14-41-00039 in the National Research University Higher School of Economics.

## References

1. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision, 4th edn. Cengage Learning, Boston (2014)
2. Lowe, D.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60** (2), 91–110 (2004)
3. Dalal N., Triggs B.: Histograms of oriented gradients for human detection. In: International Conference on Computer Vision and Pattern Recognition, pp. 886–893 (2005)
4. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. Comput. Vis. Image Underst. **110**(3), 346–359 (2008)
5. He, K., Zhang, X., Ren, S., Sun, J.: Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. arXiv:1502.01852 [cs], http://arxiv.org/abs/1502.01852 (2015)
6. Hawkins, J., Blakeslee, S.: On Intelligence. Times Books, New York (2004)

7. Munoz, D., Bagnell, J.A., Hebert, M.: Stacked hierarchical labeling. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part VI. LNCS, vol. 6316, pp. 57–70. Springer, Heidelberg (2010)
8. Bosch, A., Zisserman, A., Munoz, X.: Representing shape with a spatial pyramid kernel. In: 6th ACM International Conference on Image and Video Retrieval CIVR 2007, pp. 401–408 (2007)
9. Zhai, J.-H., Zhang, S.-F., Liu, L.-J.: Image recognition based on wavelet transform and artificial neural networks. In: IEEE International Conference on Machine Learning and Cybernetics, pp. 789–793 (2008)
10. Cireşan, D., Meier, U., Masci, J., Schmidhuber, J.: Multi-column deep neural network for traffic sign classification. Neural Netw. **32**, 333–338 (2012)
11. Savchenko, A.V.: Directed enumeration method in image recognition. Pattern Recogn. **45** (8), 2952–2961 (2012)
12. Chellappa, R., Du, M., Turaga, P., Zhou, S.K.: Face tracking and recognition in video. In: Handbook of Face Recognition, pp. 323–351 (2011)
13. Chow, C.K.: On optimum recognition error and reject trade–off. IEEE Trans. Inf. Theory **16**, 41–46 (1970)
14. Wald, A.: Sequential Analysis. Dover Publications, New York (2013)
15. Savchenko, A.V.: Probabilistic neural network with homogeneity testing in recognition of discrete patterns set. Neural Netw. **46**, 227–241 (2013)
16. Theodoridis, S., Koutroumbas, K.: Pattern Recognition, 4th edn. Elsevier Inc., Amsterdam (2009)
17. Tan, X., Chen, S., Zhou, Z.H., Zhang, F.: Face recognition from a single image per person: a survey. Pattern Recogn. **39**(9), 1725–1745 (2006)
18. Taigman, Y., Yang, M., Ranzato, M., Wolf, L.: DeepFace: closing the gap to human-level performance in face verification. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2014), pp. 1701–1708 (2014)
19. Savchenko, A.V.: Nonlinear transformation of the distance function in the nearest neighbor image recognition. In: Zhang, Y.J., Tavares, J.M.R.S. (eds.) CompIMAGE 2014. LNCS, vol. 8641, pp. 261–266. Springer, Heidelberg (2014)
20. Specht, D.F.: Probabilistic neural networks. Neural Netw. **3**(1), 109–118 (1990)
21. Yao, Y.: Granular computing and sequential three-way decisions. In: Lingras, P., Wolski, M., Cornelis, C., Mitra, S., Wasilewski, P. (eds.) RSKT 2013. LNCS, vol. 8171, pp. 16–27. Springer, Heidelberg (2013)
22. Kullback, S.: Information Theory and Statistics. Dover Publications, New York (1997)
23. AT&T database of faces. http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
24. Yale face database. http://vision.ucsd.edu/content/yale-face-database
25. Japanese Female Facial Expression (JAFFE) database. http://www.kasrl.org/jaffe.html
26. Liao, S., Zhu, X., Lei, Z., Zhang, L., Li, S.Z.: Learning multi-scale block local binary patterns for face recognition. In: Lee, S.-W., Li, S.Z. (eds.) ICB 2007. LNCS, vol. 4642, pp. 828–837. Springer, Heidelberg (2007)
27. OpenCV library. http://opencv.org/
28. Degtyarev, N., Seredin, O.: Comparative testing of face detection algorithms. In: Elmoataz, A., Lezoray, O., Nouboud, F., Mammass, D., Meunier, J. (eds.) ICISP 2010. LNCS, vol. 6134, pp. 200–209. Springer, Heidelberg (2010)

http://www.springer.com/978-3-319-26122-5

Analysis of Images, Social Networks and Texts
4th International Conference, AIST 2015,
Yekaterinburg, Russia, April 9–11, 2015, Revised
Selected Papers
Khachay, M.; Konstantinova, N.; Panchenko, A.; Ignatov,
D.I.; Labunets, V.G. (Eds.)
2015, XVIII, 460 p. 145 illus. in color., Softcover
ISBN: 978-3-319-26122-5