Krzysztof SIWEK[1], Stanisław OSOWSKI[1,2], Jacek JAKUBOWSKI[2]

Warsaw University of Technology (1), Military University of Technology (2)

doi:10.15199/48.2015.07.08

Face Recognition in Visible and Infra-Red Imagery - Comparison of Methods

Abstract. *The paper is concerned with the recognition of faces represented by the visible and infra-red images. Different methods of image feature generation at application of different classifiers will be studied and compared for both types of face imagery. The investigated approaches include the linear and nonlinear methods of transformation: principal component analysis (PCA), Kernel PCA, Sammon transformation and stochastic neighbor embedding with t-distribution (tSNE). The representation of the image in the form of limited number of main components of transformation is applied to the input of support vector machine classifier and random forest. The numerical results of experiments will be presented and discussed.*

Streszczenie. *Praca przedstawia porównanie metod rozpoznawania twarzy na podstawie dwu rodzajów obrazów: widzialnego oraz w podczerwieni. Zbadano kilka metod przetwarzania obrazu w cechy diagnostyczne: metodę opartą na PCA, nieliniową metodę KPCA, odwzorowanie Sammona oraz transformację stochastyczną tSNE. Każda z tych metod generuje inny zestaw cech diagnostycznych użytych jako atrybuty wejściowe dla klasyfikatora. W pracy zastosowano zespół klasyfikatorów stosujących sieć SVM oraz las losowy Breimana . Przedstawiono wyniki rozpoznania każdego z tych klasyfikatorów współpracujących z odpowiednim zestawem atrybutów wejściowych oraz wynik fuzji poszczególnych rezultatów. Jako jednostkę integrującą zespół zastosowano las drzew losowych. Wyniki pokazują, że zastosowanie wielu metod przetwarzania obrazu w cechy diagnostyczne i równoległego obrazowania twarzy w postaci widzialnej i w podczerwieni pozwala zwiększyć efektywność rozpoznania o około 30%.(**Rozpoznawanie twarzy w podczerwieni i w świetle widzialnym – porównanie metod**).*

Słowa kluczowe*:* obrazy widzialne i w podczerwieni, rozpoznawanie twarzy, transformacje danych, klasyfikacja.
Keywords: visible and infra-red imagery, face recognition, transformation of data, classification

Introduction

There is an increased interest in application of face recognition in biometrics nowadays, since this is non-invasive and relatively non-expensive way of person identification and verification. However, it has been shown [Philips] that even the well known face recognition systems may perform poor in less controlled situations due to the large variations of external illumination, imaging systems, etc.

Recently infrared (IR) images have been used for face recognition [3,4,5]. While visual cameras measure the electromagnetic energy in the visible spectrum range of 0.4-0.7µm, sensors in the IR camera respond to thermal radiation in the infrared spectrum range at 0.7-14µm . The light in the thermal IR range is emitted rather than reflected. Thermal emissions from skin are an intrinsic property, independent on illumination. This is one of the most important advantage of IR images. Moreover, the IR face images contain more information about the natural anatomical face characteristics. At the same time the tasks, such as face detection, location and segmentation are relatively easier and more reliable than these in visual images []. The thermal imaging has also some limitations, such as poor recognition of persons wearing glasses, since they block a large portion of thermal energy near the eyes, resulting in some loss of information.

The paper will be concerned with the face recognition techniques based on different preprocessing methods of the image and application of two types of the most efficient classifiers: the support vector machine (SVM) and random forest of decision tree(RF). We compare them on the basis of their visible and infra-red imagery.

In the first step the face should be characterized by the set of numerical descriptors, which are the most representative for the image. The important task is to find the transformation method of the highest compression ability of the image, which is able to pack the most characteristic features of the faces into smallest possible number of the significant descriptors. These descriptors will be used as the input attributes for the classifiers, performing the final role of recognizing tool. The numerical experiments will be directed to find the strongest points in both types of face imagery.

Characterization of images

In our work we will compare two types of face imagery: the visible and IR images. The typical examples of these types of face images in different illumination conditions are presented in Fig. 1

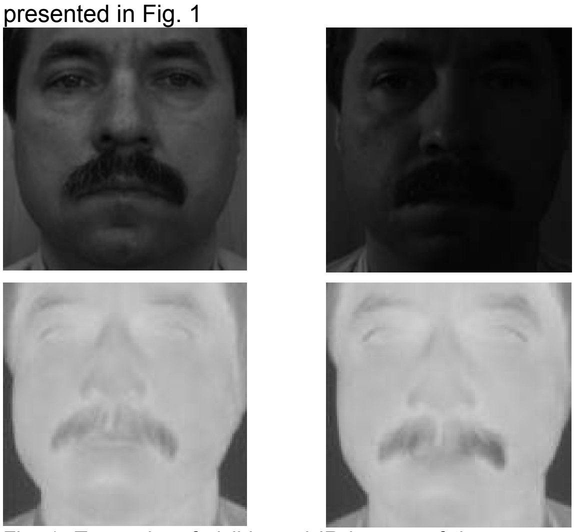

Fig. 1. Two pairs of visible and IR images of the same person at different illuminations

Thermal IR imagery is nearly invariant to changes in ambient illumination. Consequently, no light compensation is necessary and the within-class variability is lower than in a visible imagery.

This is well seen on the example of histograms of the images presented in Fig. 2. The histograms of the visible imagery at different illumination have changed a lot. Not only the shape but also the maximum values and distribution of bins of histograms are different at changed illumination. At the same time the difference in histograms of the infra-red imagery taken at different illumination conditions are almost the same. We can see very similar shape of the histogram and the range of values. It means much easier work for the classifier in the case of IR imagery. At the same time we should note, that IR imagery results in loss of many significant details of the image. This might be significant in recognition of images belonging to different classes.

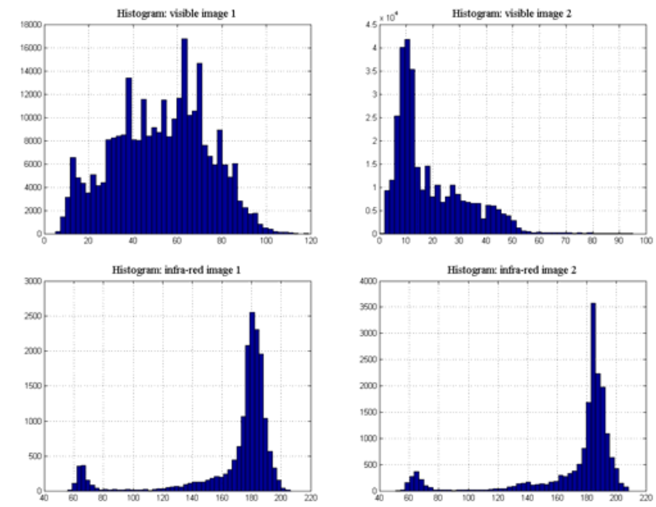

Fig. 2. Two pairs of visible and IR images of the same person with different illuminations

Feature generation methods

In this paper we will study few methods of descriptors generation, including PCA, Kernel PCA, Sammon transformation and stochastic neighbor embedding. Their use in visual image description were found as valuable [1,3,5]. This paper will study their extension to infra-red imagery of the face.

The original face image in the matrix form will be represented by the vector $\mathbf{x} = [x_1, x_2, ..., x_N]^T$ formed by the succeeding rows of the matrix. In this way each image is represented by the vector.

The PCA represents linear transformation **y=Wx** of the transformation matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_K]^T$ defined on the basis of eigen-vector decomposition.

From the nonlinear transformations we have applied Kernel PCA (KPCA) representing the linear PCA defined on the nonlinear mapping of the vectors **x** [2].

The next nonlinear method is the Sammon transformation designed to minimize the differences between corresponding inter-point distances in the original and transformed spaces. These distances are determined in Euclidean measure.

The last of feature generation is the stochastic neighbor embedding with a Student distribution [3]. It tries to find the map points (vectors $\mathbf{y}_i$ and $\mathbf{y}_j$) of the high-dimensional data points ($\mathbf{x}_i$ and $\mathbf{x}_j$) in a way to minimize a Kullback-Leibler divergence between the joint probability distribution p_{ij} in high-dimensional space and a joint probability distribution q_{ij} in the transformed (lower dimensional) space.

All these preprocessing methods characterize the images by the limited number of features, which should represent them in a way providing the highest uniformity within the same class and highest differences for images representing different classes.

The original images transformed into numerical features form the input attributes to the classifiers, which are responsible for the final recognition of images . As the classifiers we have used Support Vector Machine of Gaussian kernel [2] and Breiman random forest of the decision tress [7]. Both types of classifiers belong to the most efficient classifiers. However, both solutions apply different mechanisms of decision taking.

Applied classifiers

The final recognition is done by two types of the classifiers: SVM and RF, regarded now as the most efficient in pattern recognition problems. The applied SVM [] is a linear machine, working in the high dimensional feature space created by the non-linear mapping of the N-dimensional input vector **x** into a L-dimensional feature space (L>N) by using the kernel function $K(\mathbf{x},\mathbf{x}_i)$. The learning problem of SVM is defined as the task of separating the learning vectors into two classes of the destination values: d_i = 1 (one class) or d_i = -1 (the opposite class), with the maximal separation margin. The SVM of the Gaussian kernel has been used in our application as the most universal and efficient kernel. The hyperparameters (the regularization constant C and Gaussian kernel width) have been adjusted by repeating the learning experiments for the set of their predefined values and choosing the best one on the validation data sets.

The second classifier, Breiman random forest is an ensemble learning method for classification that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes outputs by individual trees. The randomness in selecting the learning data is implemented to improve the generalization property of the forest. The trees are grown using also randomly selected input variables in each node. The number of trees may be arbitrarily high, increasing in this way the generalization ability of it.

Numerical results

The numerical experiments of face recognition have been performed for the set of face images of 51 classes of people. Each class was represented by 20 individuals in different poses and illumination conditions. The same images have been acquired simultaneously in visible and infra-red forms. The size of original images in both cases was the same and equal 100×100.

The typical examples of images under recognition are presented in Fig. 3 in visual and infra-red imageries. As we can see the images differ by the pose, illumination, glasses.

Fig. 3. The examples of diversities of face images of one person taking part in experiments: the upper row – the visual images, the lower row – infra red images

The numerical results comparing the accuracy of recognition for both forms of imagery will be based on 10-fold cross validation approach. In this approach the whole set of data is split into 10 equal parts. Nine parts are used in learning and the last one for testing the learned classifiers. The experiments are repeated 10 times. Each time exchanging the testing part. In this way all samples take part in learning and testing phases of classification. Only the results of testing will be presented and discussed.

Table 1. The misclassification rate (mean+/-std) of 51 classes of faces committed by SVM

Number of classes	PCA [%]	KPCA [%]	Sammon [%]	tSNE [%]
Visual image	13.5±1.53	13.16±1.44	23.71±2.62	16.04 ±1.63
Infra-red image	14.8+/-1.7	14.72±1.82	24.43±3.14	16.74±1.91

The results of comparison of different methods of feature generation combined with the SVM classifier are presented in Table 1. They present the average misclassification rate and the standard deviation obtained in all cross validation experiments.

The results show that PCA (linear and nonlinear) belonged to the most efficient. The tSNE methods of feature generation was only slightly worse. However, the Sammon transformation seems to be useless for feature generation. This conclusion is true in both types of face imagery.

Application of random forest as the classifier has resulted in only slightly different level of these mean errors. The numerical results in this case are presented in Table 2. The tSNE method of feature generation is the most stable in all forms of imagery. The PCA and KPCA produce different results for both forms of imagery.

Table 2. The misclassification rate (mean+/-std) of 51 classes of faces committed by RF

Number of classes	PCA [%]	KPCA [%]	Sammon [%]	tSNE [%]
Visual image	13.28±1.66	13.72±1.82	21.42±2.49	16.44 ±1.36
Infra-red image	17.74+/-1.68	18.83±1.97	23.93±1.99	16.48±1.34

The important point in application of many classifiers is possibility of integrating them into one final system. We have made additional experiments combining many individual results into the final one by applying random forest as an integrator. In these experiments we have excluded the Sammon transformation, as the least efficient in both forms of imagery. The experiments have been done first by combining the results of SVM and RM classifiers separately for 3 forms of data preprocessing (PCA, KPCA and tSNE) resulting in ensemble of 3 classifiers. In the second case we have combined SVM and RF results together (6 classifiers in ensemble). These forms have been repeated separately for visual images and for infer red images.

Table 3. The misclassification rate of 51 classes of faces after fusing the results of many classifiers

	SVM	RF	SVM+RF
Number of classifiers	3	3	6
Visual image	12.65%	12.98%	12.96%
Infra-red image	14.12%	16.37%	16.23%

The results of integration are depicted in Table 3. In all forms of integration we observe the improvement of accuracy. The best results correspond to support vector machine classifier. The random forest is slightly worse. Combining SVM and RF together did not bring the improvement of results.

The last experiments have been performed for fusing both forms of imagery: visual and infra red. We have combined the results of different data preprocessing and different classifiers into one system. Once again the random forest was used as the integrating unit. This time the average error of classification has been reduced to 11.02%. This is the evidence that different forms of imagery cooperating together lead to increase of recognition accuracy of faces.

Conclusions

The paper has shown the comparison of two types of imagery in face recognition. Four different methods of feature generation combined with two classifiers (SVM and RF) have been tried. The results of numerical experiments of recognition of 51 classes of faces have shown, that the IR imagery is of comparable (although slightly worse) efficiency to the visible imagery. Furthermore, we demonstrated that combined use of both imaging modalities results in even higher performance, with identification error dropping by more than 20%.

REFERENCES

[1] Belhumeur P., Hespanha J., Kriegman D.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, *IEEE Trans. PAMI*, 19 (1997), no 7, 711–720
[2] Phillips P. J., Blackburn D, Bone M., Grother P., Michaels R., Tabassi E., Face recognition vendor Test 2002, available at: *http://www.frvt.org/FRVT2002/default.htm*
[3] Wu S., Lin W., Xie S., Skin heat transfer model of facial termograms and its application in face recognition, *Pattern Recognition*, vol. 41 (2008), pp. 2718-2729
[4] Kong S., Heo J., Abidi B.R, Paik J. Abidi M. A., Recent advances in visual and infrared face recognition – a review, *Comput. Vision Image Understanding*, 97 (2005), 103-135
[5] Socolinsky D. A., Selinger A., Neuheisel J. D., Face recognition with visible and thermal infrared imagery, *Comput. Vision Image Understanding*, 91 (2003), 72-114
[6] Siwek K., Osowski S., Comparison of methods of feature generation for face recognition, *Przegląd Elektrotechniczny*, 90 (2014), nr. 4, 206-209
[7] Schölkopf B., Smola A., "Learning with kernels", Cambridge, MIT Press, (2002), MA
[8] Van der Maaten L., "Matlab toolbox for dimensionality reduction", v0.8.1, (2013), *Delft University of Technology*
[9] "Matlab user manual – image toolbox", (2012), MathWorks, Natick
[10] Sammon J. W., A nonlinear mapping for data structure analysis, *IEEE Trans. Computers*, 18 (1969), no 5, 401–409
[11] Breiman L., Random forests, *Machine Learning*, 45 (2001),. 5–32

***Authors**: dr hab. inż. Krzysztof Siwek, Warsaw University of Technology, Institute of the Theory of Electrical Engineering, Measurement and Information Systems, E-mail: ksiwek@iem.pw.edu.pl.*
prof. dr hab. inż. Stanisław Osowski, Warsaw University of Technology, Institute of the Theory of Electrical Engineering, Measurement and Information Systems, Military University of Technology, Institute of Electronic Systems, E-mail: sto@iem.pw.edu.pl.
dr hab. inż. Jacek Jakubowski, Military University of Technology, Institute of Electronic Systems, E-mail: jjakubowski@wat.edu.pl