SERBIAN JOURNAL OF ELECTRICAL ENGINEERING
Vol. 16, No. 3, October 2019, 387-403

UDC: 621.38:681.586]:536.5, 004.93 **DOI: https://doi.org/10.2298/SJEE1903387P**

Optimal HOG Cell to Image Ratio for Robust Multi-Sensor Face Recognition Systems

Miloš Pavlović[1], Branka Stojanović[1], Ranko Petrović[1], Snežana Puzović[1], Srđan Stanković[2]

Abstract: The main problem for modern visible light face recognition has been accurate identification under variable environmental conditions. Thermal infrared facial images utilization in face recognition systems can provide a solution for problems related to uncontrolled environmental conditions, especially to those caused by illumination limitations. This paper compares the results of the use of visible light and thermal infrared imagery for face recognition based on the HOG feature descriptor. In particular, the paper suggests an optimal HOG cell to image size ratio in order to improve recognition accuracy and reduce computational complexity. Performance statistics are presented on facial images with different facial expressions. The obtained results support the conclusion that recognition with thermal infrared images is more robust and that fusion of sensors should be included for improving recognition accuracy.

Keywords: Face recognition, Visible light imagery, Thermal imagery, Image scaling, Facial expression, HOG.

1 Introduction

Despite a significant level of recognition accuracy (even more than 99% in controlled conditions) face recognition is still a highly challenging and the most attractive task in pattern recognition and computer vision. In uncontrolled operating conditions, face recognition based only on the visible light facial images degrades recognition accuracy very quickly.

Illumination dependency is one of the major disadvantages of visible light face recognition systems and the primary reason why current face recognition technology cannot be used in outdoor and night vision applications. Other

[1]Vlatacom Institute of High Technologies, Blvd. Milutina Milankovića 5, 11070 Belgrade, Serbia;
E-mails: milos.pavlovic@vlatacom.com; branka.stojanovic@vlatacom.com; ranko.petrovic@vlatacom.com; snezana.puzovic@vlatacom.com
[2] School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11200 Belgrade, Serbia and Vlatacom Institute of High Technologies, Blvd. Milutina Milankovića 5, 11070 Belgrade, Serbia;
E-mail: srdjan.stankovic@vlatacom.com

factors that affect recognition performance include facial pose variations, facial expression changes and occlusions [1].

In order to overcome these problems, introducing thermal infrared sensors in current systems has become an area of growing interest [2, 3]. Thermal IR spectrum comprising Medium Wave IR (MWIR) and Long Wave IR (LWIR), has been suggested as an alternative source of information for face detection and recognition [4].

The face recognition testing environment proposed in this paper includes both visible and thermal imagery, and uses a face recognition algorithm based on Histogram of Oriented Gradients (HOG) [5, 6] as feature descriptors and the Support Vector Machine (SVM) classifier [7, 8].

The common problem in both visible light and thermal face recognition systems is robustness – the system needs to catch as many features as possible, but those features need to be a pose or environmental independent. This indicates that there is a need to find a proper balance between sensitivity of feature descriptor and robustness of the system.

The other common problem, when it comes to online system applications, is processing speed – the system needs to process facial images in real-time.

The goal of this paper is to find the right balance between face recognition system robustness and processing speed on the one side, and accuracy on the other side by proposing the optimal ratio between cell size (for calculating HOG features) and image size (scale factor) for both image sensors. Scaling images to dimensions that correspond to the optimal ratio between the HOG cell and the image size, the paper discusses the influence of facial expression on recognition performance in the existing face recognition systems using visible light and thermal IR facial images. The obtained results are compared with those obtained using the original size facial images for both sensors. Also, based on a comparison of environmental influence, a system that uses a fusion of thermal and visible imagery can be developed with maximum robustness.

The paper is organized as follows. Section II gives the system description. It presents techniques for classification using the Support Vector Machine (SVM) method with the Histogram of Oriented Gradients extraction procedure, as well as the system architecture and facial image database used for this work. Section III presents evaluation methodologies, as well as statistical and visual comparison of *HOG cell size to image size ratio* and facial expression influence on face recognition performance in the proposed testing environment. Section IV lists conclusions and presents future work in this research area.

2 System Description

This paper proposes an environment for testing performance-robustness relation of the face recognition system, in a variety of conditions for both thermal and visible light imagery. The testing system for classification is based on Histogram of Oriented Gradients (HOG) as feature descriptors and Support Vector Machine (SVM) classifier.

2.1 SVM

Support Vector Machine (SVM) [7, 8] is a classical machine learning method for data classification and regression, developed by Vapnik and associates in 1995. SVM has been established as one of the state-of-the-art tools for machine learning and enjoys great popularity in this area due to very good results that are similar or better than those obtained by other soft computing techniques for big data classification problems, such as neural networks. The application of SVMs to face detection and face recognition problem have been proposed recently. In [9] SVM is used for face recognition with the component-based method and two global methods. The SVMs with a binary tree recognition strategy are used to solve the face recognition problem in [10]. In [11] SVM is presented as a very effective method for face detection.

Here is an explanation of the basic theory of SVM for binary classification. SVM is a supervised learning model that generates input-output mapping functions from a set of labeled training data. Training examples are mapped from the input space of values of the attributes into the new feature space with larger dimensionalities, in which training examples can be potentially linearly split. The task of the training phase is to find an optimal separating hyperplane in the feature space where the data is presented, in such a way that all data from the one class can be on the same side of the separating hyperplane.

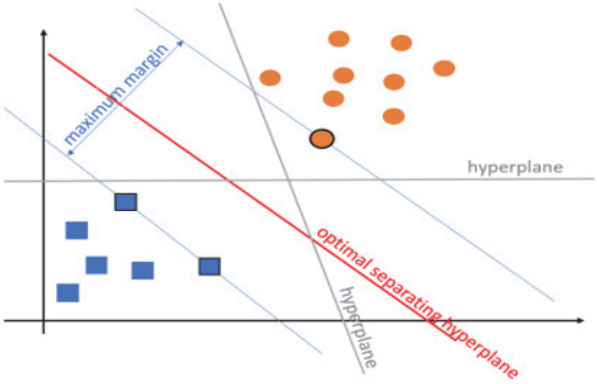

Fig. 1 – *Optimal separating hyper-plane with maximum margin.*

The optimal separating hyperplane is the one which separates training examples with the maximum margin (distance between the separating hyperplane and each training data) – Fig. 1.

Equation (1) presents equation of separating hyperplane

$$w^T x + b = 0 \ . \tag{1}$$

The equation of this separating hyperplane represents the model on the basis of which classification is performed. The SVM determines the optimum separating hyperplane by maximizing the distance between the separating hyperplane and the features that are closest to the potential separating line - called support vectors. The separating hyperplane is completely determined only by the support vectors, by which the method is named. In (1), vector *w* determines the direction of the optimal separating hyperplane, while the parameter *b* determines the distance of separating hyperplane from the center of the coordinate system.

The classification of the test examples is based on the sign of (1). For each test point $\{x_i, y_i\}$ the separation can be formulated by the following conditions:

$$\begin{aligned} w^T x_i + b \geq 0 \quad y_i = 1, \\ w^T x_i + b < 0 \quad y_i = -1. \end{aligned} \tag{2}$$

In order to maximize the margin, the objective function – Equation (3) needs to be minimized, subject to the condition in (4):

$$J(w) = \frac{1}{2} w^T w \, . \tag{3}$$

$$y_i \left(w^T x_i + b \right) \geq 1 \, . \tag{4}$$

Equation (4) says that distance between the separating hyperplane and each training data (margin for each training data) should be at least 1. This is a quadratic optimization problem with linear conditions, and one way for solving it is the method of Lagrange's multipliers [12]. Solving this problem, optimal vector *w** and parameter *b** are obtained and the decision function is:

$$h_{w,b} = \operatorname{sgn}\left(w^*, x + b^* \right) . \tag{5}$$

SVM can also be generalized to the class of problems when data is not linearly separable. The first way is the introduction of variable ξ_i that would tolerate “small” errors on train data. Now, the objective function is (6), subject to the condition (7):

$$J(w) = \frac{1}{2} w^T w + C \sum \xi_i \, , \tag{6}$$

$$y_i\left(w^T x_i + b\right) \geq 1 - \xi_i, \quad \xi_i \geq 0, \tag{7}$$

where C is constant of regularisation which makes a compromise and balances between the margin size and the learning error (minimizing the error with the cost of not getting the maximum margin of separation). Regularisation allows that margin can be less than 1 for the value of ξ_i.

Another way is based on mapping the input vector space *x*, in which the training set is not linearly separable, to a higher-dimensional space $\Phi(x)$, in which the training set is linearly separable – graphically presented in – Fig. 2

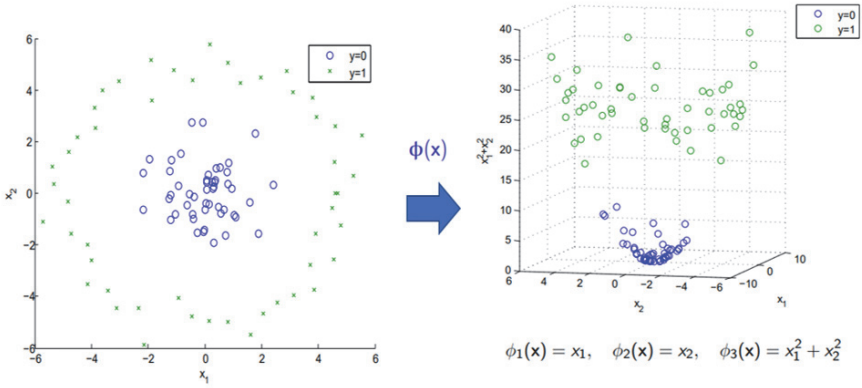

Fig. 2 – *Mapping input space to a higher- dimensional space.*

Instead of a scalar product of high dimensional vectors in a mapped space, a *kernel function K* can be introduced, corresponding to the scalar product in the mapped space:

$$K\left(x_i, x_j\right) = \Phi\left(x_i\right)^T \Phi\left(x_j\right). \tag{8}$$

For the problem of choosing and constructing the kernel K corresponding to the scalar product in the new space for a given task, a special mathematical theory has been developed [13].

2.2 HOG features

Choosing, preparing and getting good features is crucial for the overall performance of face recognition systems. Histogram of Orientated Gradients (HOG) [5, 6] is a feature descriptor used in computer vision and image processing primarily for object detection. The concept behind HOG was first described by Robert K. McConnell in 1986 [14]. However, the use of HOGs has become widespread since 2005, when Navneet Dalal and Bill Triggs, researchers from the French National Institute for Research in Computer

Science and Automation (INRIA), presented their work [15] based on HOG descriptors for detection of pedestrians in static images.

Distribution of the intensity of brightness or directions of the local edges in the image can be described by the HOG feature descriptor. Unequal reflection of light from face presents facial edges in visible light images, while unequal emitted heat energy from face presents facial lines in thermal images. HOG features can be efficiently used for the identification of these edges, independently of the image sensor. In [6] is presented that the HOG feature for face recognition achieves better accuracy than the LBP descriptor and almost the same recognition rate with much lower computational time than the widely used Gabor feature. HOG works on local cells, so it's invariant on geometric and photometric transformations, except for the orientation of the object and these are key advantages over other descriptors.

Gradient in digital image processing presents direction of the highest increase in the illumination in each image pixel. Combined changes in the illumination along the horizontal and vertical directions in the observed pixels correspond to the gradient values, represented by the matrix of gradient modules G – Equation (10). The gradient orientation in each pixel is contained in the matrix of angles θ – Equation (11).

$$\nabla f(x,y) = \left[\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\right]^T = \left[G_x\, G_y\right]^T, \tag{9}$$

$$G = |\nabla f| = \sqrt{G_x^2 + G_y^2}, \tag{10}$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right). \tag{11}$$

The goal of calculating the HOG features is to obtain discrete gradient histograms that correspond to cells. Therefore, it is first necessary to divide an image into cells. The size of a HOG cell is specified in pixels. To capture large-scale spatial information, the cell size should be increased, but by increasing the cell size, small-scale details might be lost. The cell size is conditioned by the block dimension, in which the analysis is implemented. A large block dimension reduces the ability of suppressing local illumination changes. These changes may be lost by averaging the values of pixels in a large block. Reducing the block dimension leads to capturing the significance of local pixels. A smaller block dimension can suppress illumination changes of the HOG features.

For calculating HOG features, it is necessary to discretize possible gradient values. After defining allowed directions that gradient can occupy in each pixel,

the gradient is projected in two closest, adjacent directions. The projection implies the redistribution of the gradient module value to the allowed, discrete directions in a relationship that corresponds to the angular distance of the gradient from each of them. Directions can be signed ($0\ ^\circ \leq \theta \leq 360\ ^\circ$) and unsigned ($0\ ^\circ \leq \theta \leq 180\ ^\circ$). In the case of unsigned directions, there is no need for more than nine discrete directions. The obtained histograms for one face from the facial image database are shown graphically in Fig. 3.

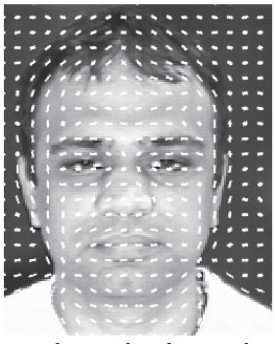

Fig. 3 – *HOG features for one facial image in database.*

Histograms are calculated by counting the possible directions in each cell – summing values of the gradient modules corresponding to the observed direction, across all pixels in the cell for each of the discrete directions. The histogram values describe the presence of the dominant directions in the observed cell, and in this way, locally, the presence of edges in the image for multiple adjacent cells.

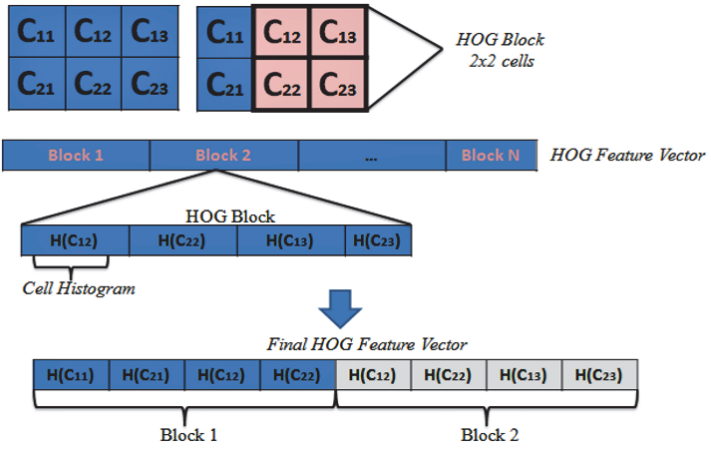

Fig. 4 – *Calculating HOG features vector.*

Local normalization of the histogram value for each cell is necessary because histogram values between spatially close cells can vary significantly. Normalization is done by blocks that include several cells. Overlapping between blocks is important and individual cells can enter the composition of multiple blocks. With overlapping blocks, each histogram enters the final HOG features vector several times. Large overlap values can capture more information, but they produce a larger feature vector size. Fig. 4 shows the described procedure for calculating HOG features.

2.3 System architecture

A block diagram of the proposed system is shown in Fig. 5. The same algorithm was used for both thermal and visible light sensors. HOG features are calculated on resized input thermal or visible facial image. The algorithm for HOG features extraction is implemented with the cell size of 8×8 pixels. Dimensions for the normalization block are 2×2 cells with 2 overlapping cells between the adjacent blocks. Unsigned directions are implemented with 9 discrete directions. Then HOG feature vectors are passed to the classifier. The classifier is trained with images from the database – for visible light test images with visible light images and for thermal test images with thermal database facial images. The Multiclass SVM [16] in the one-vs-all scheme with Gaussian Kernel trick [17] has been used for classification. After processing the obtained feature vectors, SVM classifier gives matching results. The result of SVM classifiers is a vector with the matching scores for each person individually. The proposed system utilizes the maximum of the matching scores vector that classifier produces as the final matching score for the recognized face.

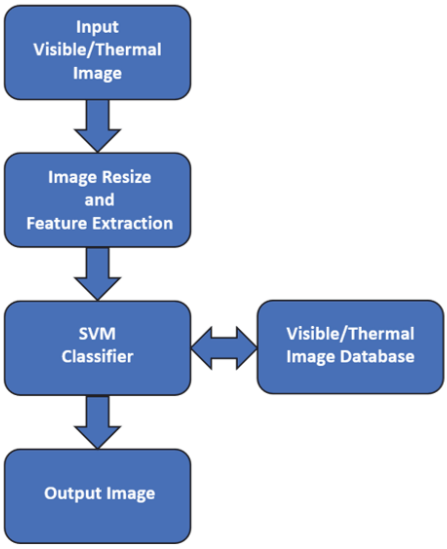

Fig. 5 – *Block Diagram of proposed face recognition system.*

2.4 Facial Image Database

Database used for this paper contains visible light and the corresponding thermal infrared (LWIR) images of 29 persons under different lighting conditions, different facial pose, and different facial expression. The database is collected by the authors of [18]. Each pair of visible light and thermal images was taken simultaneously, so using this database, facial expression influence on recognition performance between two imaging sensors can be easily compared.

Training set for this work contains 7 pairs of visible light and thermal images per person, captured from 7 different angles, on daylight and with neutral facial expression – Fig. 6.

Fig. 6 – *Training set of one person in database.*

Test set for this work contains four subsets of facial images for testing different facial expression influence – *neutral*, *smile*, *boring* and *open mouth* - Fig. **7**. All images have the same orientation and the size of all images is 320×240 pixels. Each test subset contains 2 images per person frontally oriented.

Fig. 7 – *Test set for different facial expression*
(from left to right: neutral, smile, boring and open mouth).

3 Results and Discussion

3.1 Evaluation methodology

Results and performance statistics for visible and thermal face recognition have been presented with common evaluation methods: *Receiver Operating Characteristic* (ROC) and *Cumulative Match Characteristic* (CMC) curves.

Receiver Operating Characteristic curve [19] presents the diagnostic ability of a classification system when its discrimination threshold is varied, where each point on the curve defines the sensitivity pair corresponding to a particular decision threshold. Accuracy is measured by the area under the ROC curve [20], in such a way that greater area means better recognition accuracy. The ROC space is defined by True Positive Rate (*sensitivity*) – the number of well-recognized faces and False Positive Rate (*1-specificity*) – the number of wrong recognized faces – "false alarm'.

The CMC curve [19] presents the percentage of recognition accuracy on specified rank. Rank is plotted on the horizontal axis and runs from 1 up through the number of persons in the database. For each test face, there is a corresponding database face of the same person. Generating a CMC curve is based on comparing the matching result for each facial test image against all database sorted by decreasing similarity. Test face is correctly recognized at rank k if the database face of the same person is among the first k faces in the sorted database. Recognition accuracy on rank 1 is the most important performance indicator on CMC curves.

3.2 Experimental results

Figs. 8 and 9 show the effect of image scale factor (image resolution) on recognition performance for thermal and visible face recognition.

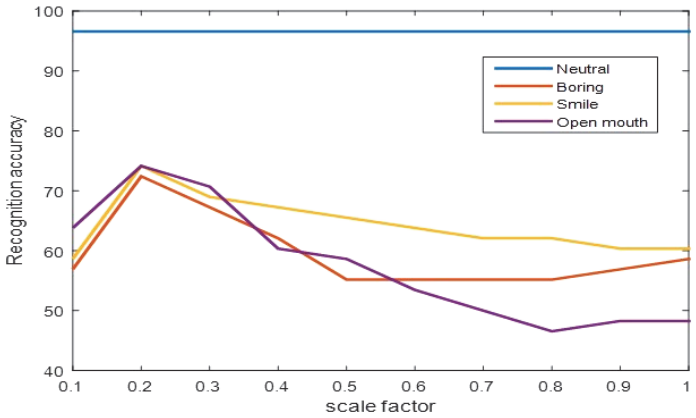

Fig. 8 – *Scale size effect on recognition performance for thermal images (different facial expressions).*

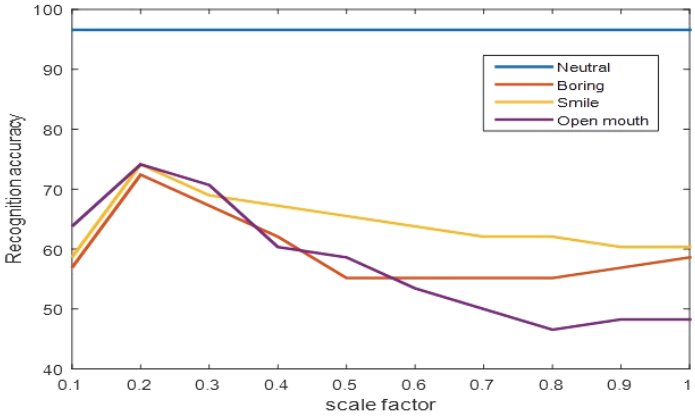

Fig. 9 – *Scale size effect on recognition performance for thermal images (different facial expressions).*

Experimental results show that both thermal and visible face recognition test environments provide better results for images with resized dimensions. Although some information from the original image is lost with rescaling, experiments indicate that the lost information is related to distinctions between the test and database images affected by different transformations, which makes the system more robust. Removing pixels that make differences in the facial appearance of the same person on different images leads to a greater similarity between the test and database images. Graphics in Fig. 8 indicates that the optimal scale factor for the test set of thermal images with different facial expressions is equal to 0.2. Consequently, it can be concluded, based on results of all test sets, that the optimal scale factor for thermal sensors is 0.2. Fig. 9 shows that, for visible facial images, the highest recognition accuracy for the test set of different facial expressions is obtained for the scale factor of 0.1. Table 1 shows image size, HOG feature vector length, the time needed for single-frame processing and recognition and time needed for training SVM classifier depending on different scale factors. For resized images, HOG features vectors have a shorter length, which is also an advantage from the implementation point of view – the processing of these vectors is less computationally expensive. Processing and recognition time per frame and time needed for training SVM classifier is measured offline in Matlab software package. It can be concluded from the obtained recognition results and processing time that the optimal image scale factor for a common visible and thermal system should be equal to 0.2. Implementation in Matlab is slower than it would be in the real-time application, but it can be concluded that with the resized image with scale factor 0.2, processing per frame is faster (1.26 times)

than with the original image size, with the similar expectation in real-time applications.

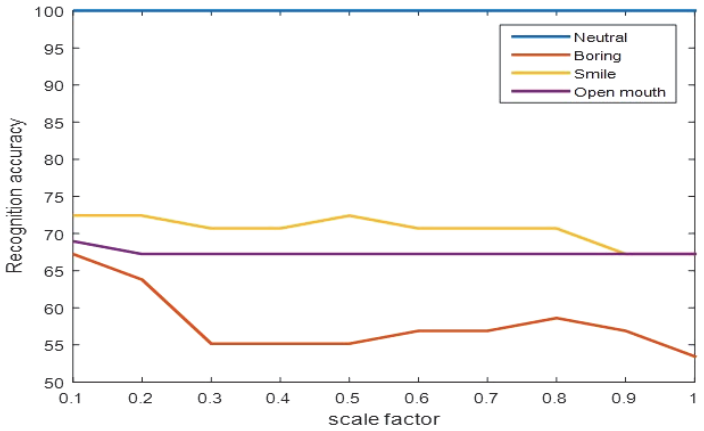

Fig. 10 – *Scale size effect on recognition performance for visible light images (different facial expressions).*

For the chosen scale factor of 0.2 for both sensors, the scaled images have dimensions of 24×32 pixels. Cell size for calculating HOG features is 8x8 pixels, and the optimal hog cell to image ratio is 12.

Table 1

Image size, HOG features vector length, processing time per frame and training SVM time depending on scale factor.

Scale factor	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
Image size [pixels]	240×320	216×288	192×256	168×224	144×192	120×160	96×128	72×96	48×64	24×32
HOG features	40716	32760	25668	19440	14076	9576	5940	3168	1260	216
Processing time per frame [s]	0.282	0.279	0.273	0.263	0.260	0.256	0.240	0.228	0.224	0.216
Training SVM time [s]	27.839	22.589	17.934	14.261	11.391	8.516	6.572	5.065	3.806	3.161

The results of tests described in following text are obtained with images scaled with factor 0.2 and 8×8 pixels HOG cell size.

Performance of the algorithm on test sets with visible light facial images for different facial expression is shown using ROC and CMC curves – Figs. 10 and 11.

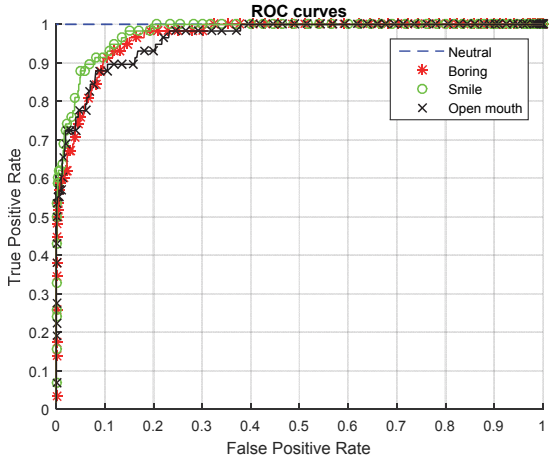

Fig. 11 – *ROC curves; visible facial images, different facial expressions.*

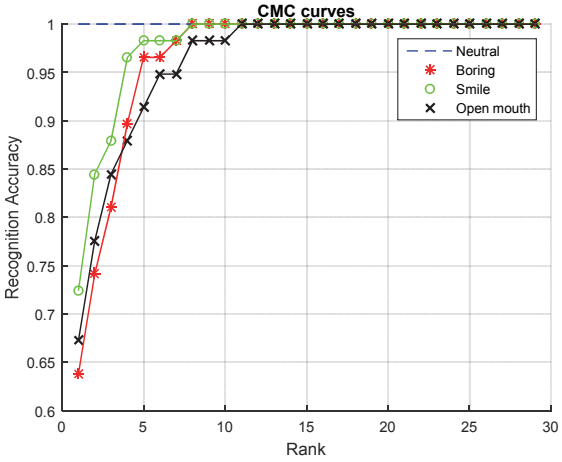

Fig. 12 – *CMC curves; visible facial images, different facial expressions.*

The presented graphics of ROC and CMC curves show that the system obtains exceptionally good results for images of people with *neutral* facial expression, which is expected, as the persons in that test set have the closest facial expression to those who are in training database. Lower accuracy is obtained for boring, smiley and faces with open mouth because the faces at these images have much more changes relative to the images in the database.

Changes in those images affect the gradient of brightness and HOG features and lead to performance degradation.

Algorithm's performance on test sets with thermal facial images for different facial expression is shown with ROC and CMC curves – Fig. 13 and Fig. 14.

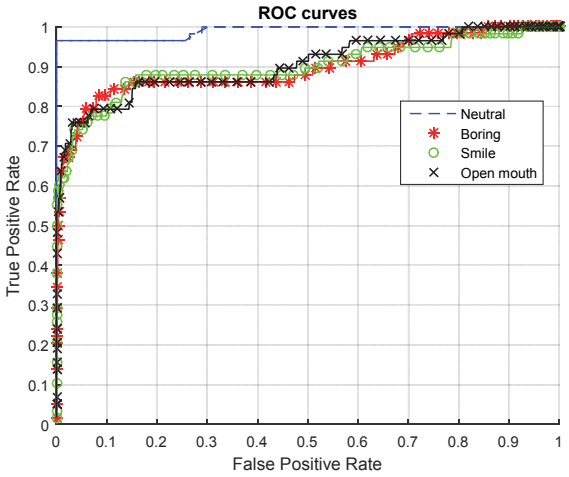

Fig. 13 – *ROC curve; facial images, different facial expressions.*

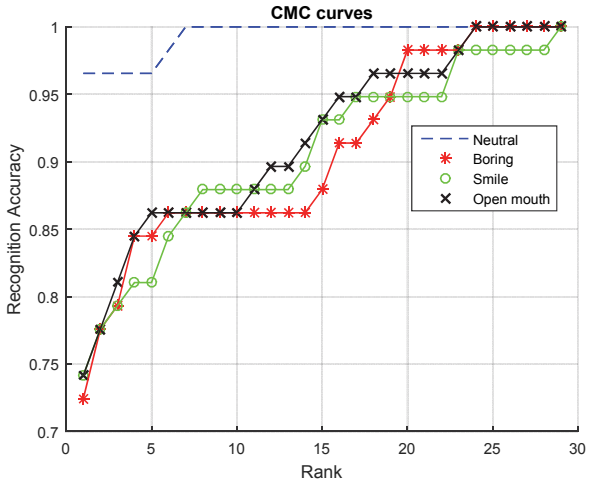

Fig. 14 – *CMC curves; thermal facial images, different facial expressions.*

Compared to visible imagery, better results (by several percent) are obtained for all test sets of thermal images which shows that the thermal infrared imagery is less dependent on the face changes caused by facial

grimaces and more robust to facial expression changes. Also, the advantage of thermal spectrum for face recognition lies in the fact that thermal sensors collect the heat energy emitted from the face which is not affected by illumination changes that have the most important influence on visible light imagery [1].

Table 2 presents the area under the ROC curve (AUC) and Rank 1 accuracy for both sensors, as the most important indicators of the algorithm performance and support of the conclusions drawn from the ROC and CMC curves. Compared to the results obtained for images with original dimension presented in [1, 21], **Table 2** shows that the recognition accuracy is significantly higher for images scaled using factor 0.2 to the dimensions of 24×32 pixels. Both indicators (Rank 1 and AUC) show the advantage of scaling images for both sensors and all facial expressions. The biggest improvement for visible light images is obtained for boring facial expression – 10%. For the boring look of the face, the whole surface of the face has appearance changes compared to the neutral facial expression, and, therefore, pixels in this test set of visible images are changed in a much wider range compared to those in the images in database. Scaling images results in losing some pixels, and in this case, the pixels that make the difference between neutral and boring look of the face are lost, resulting in a greater similarity between the test and database images. Improvement in recognition accuracy for thermal facial images is noticed for all facial expressions in the range between 14% and 26%.

Table 2
Rank 1 Accuracy and AUC for different facial expression.

Facial expression	Visible light				Thermal			
	Rank 1		AUC		Rank 1		AUC	
	Resized	Original	Resized	Original	Resized	Original	Resized	Original
Neutral	1	1	1	1	0.96	0.98	0.99	0.97
Smiley	0.72	0.67	0.97	0.95	0.74	0.60	0.90	0.94
Boring	0.63	0.53	0.96	0.92	0.72	0.57	0.89	0.83
Open mouth	0.67	0.67	0.96	0.91	0.74	0.48	0.90	0.84

4 Conclusion

This paper presents visible light and thermal face recognition performance in an operational scenario, including facial expression changes. It has been shown that visible light face recognition works very well, with exceptional

performance, for facial images with neutral facial expression. This situation represents controlled conditions in which the recognition accuracy is usually good. However, changes even in a small part of the face, such as the area around the mouth, lead to a significant degradation in the recognition accuracy. On the other side, thermal infrared systems are less dependent on the angle of capturing images, and small facial changes caused by facial grimaces which results in much better recognition accuracy than visible light systems. However, in spite of the fact that they are more robust to facial expression and illumination changes, thermal images are sensitive to changes in environment's temperature, variations in thermal emission from the skin and opaque to glass, what is particularly evident in images of people wearing eyeglasses.

Most importantly, experimental results show that both, thermal and visible face recognition systems perform better results using images with downsized dimensions. Although some information from the original images is lost with scaling, experiments indicate that the lost information is related to distinctions between the test and database images affected by different transformations (pose, facial expression…), making the system more robust. It can be concluded from the obtained results in the described experiments that the optimal scale factor for both visible and thermal images should be equal to 0.2.

Also, the processing of images with resized dimensions is less computationally expensive, which is another advantage of resizing dimensions. For cell size for calculating HOG features of 8x8 pixels, based on presented research, optimal HOG cell to image ratio should be equal to 12.

Presented results also indicate that, in order to take advantages from both spectra, systems that use fusion of thermal and visible imagery [21, 22] should be developed for maximum recognition accuracy and robustness in uncontrolled environmental conditions.

5 References

[1] M. Pavlovic, B. Stojanovic, R. Petrovic, S. Stankovic: Facial Expression and Lighting Conditions Influence On Face Recognition Performance, Proceedings of the 5th International Conference IcETRAN 2018, Palic, Serbia, June 2018, pp. 777 – 781.

[2] K. R. Kakkirala, S. R. Chalamala, S. K. Jami: Thermal Infrared Face Recognition: A Review, Proceedings of the UKSim-AMSS 19th International Conference on Modelling & Simulation (UKSim), Cambridge, UK, April 2017, pp. 55 – 60.

[3] G. Bebis, A. Gyaourova, S. Singh, I. Pavlidis: Face Recognition by Fusing Thermal Infrared and Visible Imagery, Image and Vision Computing, Vol. 24, No. 7, July 2006, pp. 727 – 742.

[4] J. Choi, S. Hu, S. S. Young, L. S. Davis: Thermal to Visible Face Recognition, Proceedings of the SPIE Digital Library, Vol. 8371, May 2012, pp. 83711L-l – 83711L-l0.

[5] O. Deniz, G. Bueno, J. Salido, F. De la Torre: Face Recognition Using Histograms of Oriented Gradients, Pattern Recognition Letters, Vol. 32, No. 12, September 2011, pp. 1598 – 1603.

[6] C. Shu, X. Ding, C. Fang: Histogram of the Oriented Gradient for Face Recognition, Tsinghua Science and Technology, Vol. 16, No. 2, April 2011, pp. 216 – 224.

[7] G. M. Fung, O. L. Mangasarian: Multicategory Proximal Support Vector Machine Classifiers, Machine Learning, Vol. 59, No. 1-2, May 2005, pp. 77 – 97.

[8] J. A. K. Suykens, J. Vandewalle: Least Squares Support Vector Machine Classifiers, Neural Processing Letters, Vol. 9, No. 3, June 1999, pp. 293 – 300.

[9] B. Heisele, P. Ho, T. Poggio: Face Recognition with Support Vector Machines: Global versus Component-Based Approach, Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV 2001), Vancouver, Canada, July 2001, pp. 688 – 694.

[10] G. Guo, S. Z. Li, K. Chan: Face Recognition by Support Vector Machines, Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France, March 2000, pp. 196 – 201.

[11] E. Osuna, R. Freund, F. Girosit: Training Support Vector Machines: An Application to Face Detection, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR IEEE'97), San Juan, Puerto Rico, June1997, pp. 130 – 136.

[12] R. T. Rockafellar: Lagrange Multipliers and Optimality, SIAM Review, Vol. 35, No. 2, June 1993, pp. 183 – 238.

[13] B. Schölkopf: The Kernel Trick for Distances, Proceedings of the 13th International Conference on Neural Information Processing Systems, Denver, Colorado, May 2000, pp. 283 – 289.

[14] R. K. McConnell: Method of and Apparatus for Pattern Recognition, Patent No. US 4567610, USA, January 1986.

[15] N. Dalal, B. Triggs: Histograms of Oriented Gradients for Human Detection, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR IEEE'05), San Diego, USA, June 2005, pp. 1 – 8.

[16] C.- W. Hsu, C.- J. Lin: A Comparison of Methods for Multiclass Support Vector Machine, IEEE Transactions on Neural Networks, Vol. 13, No. 2, March 2002, pp. 415 – 425.

[17] S. S. Keerthi, C.- J. Lin: Asymptotic Behaviors of Support Vector Machines with Gaussian Kernel, Neural Computation, Vol. 15, No. 7, July 2003, pp. 1667 – 1689.

[18] OTCBVS Benchmark Dataset Collection - Dataset 02: IRIS Thermal/Visible Face Database, Available at: http://vcipl-okstate.org/pbvs/bench/Data/02/download.html

[19] R. M. Bolle, J. H. Connell, S. Pankanti, N. K. Ratha, A. W. Senior: The Relation between the ROC Curve and the CMC, Proceedings of the 4th IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05), Buffalo, USA, October 2005, pp. 15 – 20.

[20] D. J. Hand, R. J. Till: A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems, Machine Learning, Vol. 45, No. 2, November 2001, pp. 171 – 186.

[21] M. Pavlović, B. Stojanović, R. Petrović, S. Stanković: Fusion of Visual and Thermal Imagery for Illumination Invariant Face Recognition System, Proceedings of the 14th Symposium on Neural Networks and Applications (NEUREL), Belgrade, Serbia, November 2018, pp. 1 – 5.

[22] R. Petrović, B. Stojanović, M. Pavlović, S. Stanković: Thermal-to-Visible Face Recognition for Illumination Invariant Systems, Proceedings of the 5th International Conference IcETRAN 2018, Palic, Serbia, June 2018, pp. 766 – 771.