Fusion of Visual and Thermal Imagery for Illumination Invariant Face Recognition System

Miloš Pavlović, Branka Stojanović, Ranko Petrović and Srđan Stanković

Abstract—**Visible light face recognition systems have been well researched and in controlled environments can reach excellent accuracy. Variation in lighting conditions results in performance degradation and illumination is the one of the major limitations in visible light face recognition systems. Using infrared facial images can provide a solution to this problem. Nearly invariant to changes in illumination, thermal IR imagery provides ability for recognition under all lighting conditions, including complete darkness. The system proposed in this paper takes advantages from both spectra and provides an effective algorithm for illumination invariant face recognition system.**

Keywords—**Face recognition, variable illumination, visible light imagery, thermal imagery**

I. INTRODUCTION

In security and safety applications, such as surveillance, access control, information security, identity control and many others, face recognition technology has great potential. In controlled environments, such as e-gates, visible light face recognition reaches accuracy of more than 99%. Various factors affect face recognition performance including facial expression changes, occlusions, pose variations, and illumination, as the one of the major limitations in visible light face recognition systems [1]. Especially in the outdoor and night vision applications illumination can significantly reduce the performance of the system. One of the solutions for the limitation caused by different illumination uses infrared facial images.

The infrared (IR) spectrum is divided into four bandwidths: Near IR (NIR) - 0.75 to 1.4 µm wavelength range, Short Wave IR (SWIR) - 1.4 to 3 µm, Medium Wave IR (MWIR) - 3 µm to 8 µm and Long Wave IR (LWIR) - located in the 8 µm to 14 µm. Thermal infrared imagery, located in the 0.35 µm to 0.74 µm wavelength range, offers a good alternative to visible imagery for face recognition because of relative insensitivity to variations in facial appearance caused by different lighting conditions. Nearly invariant to changes in illumination, thermal IR imagery provides ability for recognition under all lighting conditions including complete darkness. Emitted thermal energy from the face is less affected by dissipation and the absorption by smoke, fog or dust. Thermal IR images also contain facial anatomical information that can provide detection of faces covered by masks, scarf, etc.

Despite its advantages and robustness to illumination variations, thermal imagery has several deficiencies. IR imaging is sensitive to environment's temperature changes, variations in the heat facial patterns, such as a variations in facial expressions, physical condition, psychological conditions (e.g. excitement, fear, stress) and, as the most important, IR is opaque to glass. Glass is like a wall in thermal imagery, completely hiding everything located behind it. In situations, such as face recognition of a person wearing eyeglasses, a considerable part of the face might be occluded. Although quite sensitive to different lighting conditions, visible imagery is more robust to the temperature and glass factors.

The goal of this paper is to determine the influence of different lighting conditions on visible light and thermal imagery face recognition performance, and to find an effective algorithm to fuse information and take advantages from both spectra to improve face recognition performance.

The paper is organized as follows. Section II describes the facial image database used in the experiments. Section III presents the face recognition algorithm using Histogram of Oriented Gradients (HOG) features for thermal and visible images, Support Vector Machines (SVM) classifier and fusion of two parallel classifiers (for thermal and visible light images) into a unique system. Section IV presents evaluation methodologies, a statistical and visual comparison of the tested face recognition system on images with different illumination conditions. The Section V lists conclusions and future work in this research area.

Miloš Pavlović is with the School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mail: milos.pav1010@gmail.com).

Dr Branka Stojanović is with the Vlatacom Institute of High Technologies, Blvd. Milutina Milankovića 5, 11070 Belgrade, Serbia (e-mail: branka.stojanovic@vlatacom.com).

Ranko Petrović is with the Vlatacom Institute of High Technologies, Blvd. Milutina Milankovića 5, 11070 Belgrade, Serbia (e-mail: ranko.petrovic@vlatacom.com).

Prof. dr Srđan Stanković, is with School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia and Institute Vlatacom, Milutina Milankovića 5, 11070 Belgrade, Serbia; (phone: 381-11-3771100, e-mail: srdjan.stankovic@vlatacom.com).

978-1-5386-6974-7/18/$31.00 ©2018 IEEE

Fig. 1 Gallery set for one person from original database

II. FACIAL IMAGE DATABASE

This paper uses the face database collected by [2] with visible and the corresponding thermal infrared (LWIR) spectrum images of 30 persons. Each pair of visible light and thermal images was taken simultaneously in varying poses under different illumination (daylight, darkness and three different light sources - frontal, lateral left and lateral right) and different facial expression ('laughing', 'angry' and 'surprised') - total 176-250 images/person. The size of all images in the database (visible and thermal) is 320x240 pixels and all images have the same orientation. The gallery (training) set (Fig. 1) of this work contains 7 pairs of visible and thermal images per person captured under daylight in varying poses. For the test set in this work three different facial illumination were used – daylight, darkness and additional light source - Fig. 2. Test set for each facial illumination contains two pairs of thermal and visible images per person frontally captured.

Fig. 2 Test set for different facial illumination (from left to right: daylight, additional light and darkness)

III. SYSTEM DESCRIPTION

A. *System architecture*

Face recognition system, proposed in this paper, overcomes variable illumination problem using fusion of two different sensors (thermal and visible light cameras) and two parallel face recognition algorithms.

Generally, fusion techniques take advantage of more information sources for improving the recognition accuracy. Several fusion methods for face recognition systems have been described in literature [3], [4]. Performance improvement can be achieved by combining the visual and thermal IR imagery.

The fusion of visible and thermal imagery can be implemented on different levels - image, feature, match score and decision level.

Image fusion techniques utilize direct fusion of thermal and visible light facial images, and applying feature extraction and recognition algorithms on fused image [3].

Feature fusion approaches utilize feature extraction from two images (thermal and visible) separately and then fusion of extracted features. Recognition algorithm is applied on combined feature set [3], [4].

Match score and high-level decision fusion use the combination of decisions form multiple classifications. Decision fusion can be achieved with ranked-list combination [3].

Decision fusion classification is usually based on the average matching score or on the highest matching score obtained from individual face recognition modules [3].

This paper implements fusion of visible and thermal imagery on match score and decision level. Fig. 3 shows a block diagram of the face recognition system discussed in this paper.

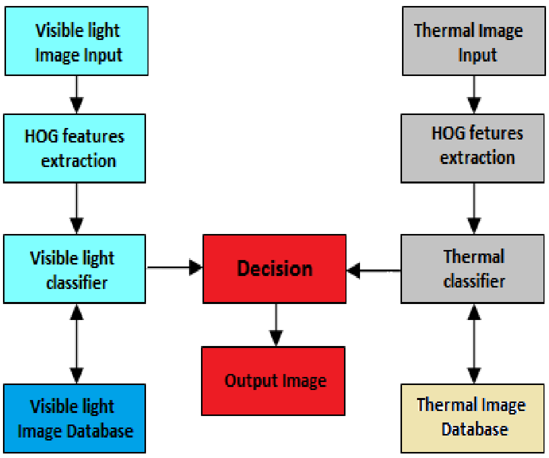

Fig. 3 Block diagram of fusion face recognition system

B. *Face recognition algorithm*

The goal of face recognition system described in this paper was accurate identification under variable illumination conditions. It requires of algorithm to find and learn features that are distinctive among people under different conditions. But it is also important to learn characteristic features that minimize differences between facial images of the same person. Choosing, preparing and getting good features is crucial for performance of all other segments of algorithm.

For robustness and simplicity, system proposed in this paper tends to use same algorithm for both thermal and visible light imagery. Face recognition algorithm used for this paper is based on Histogram of oriented gradients (HOG) [5], [6] features and Support Vector Machines (SVM) [7], [8], because it performs well regardless of imaging sensor.

Histograms of Oriented Gradients generally have good properties as a descriptor for object recognition and face recognition in particular. HOG was used in visible images to identify facial edges - an effect of unequal reflection of light from face. The HOG descriptor represents a local statistic of the orientations for the image gradients and counts occurrences of edge orientations in a local neighborhood of an image cell. HOG features vector indicating the direction of the highest increase in illumination in each image cell which contains several pixels. This method is extended and applied to detect facial edges in thermal images. Facial edges in thermal image present effect of unequal facial heat distributions.

Support Vector Machine is one of the classical machine learning method that can solve big data classification problems. SVM present the features as points in space. Features of different classes are mapped in such a way that can be efficiently separated with classification gap that have the largest possible distance to the nearest training-data point of any class [9].
The equation of this classification gap (hyper-plane) is:

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

$\mathbf{x}_i$ is element of training data set, where each element is represented by a d-dimensional vector $\mathbf{x}_i = (x_1, x_2, ..x_d)$. Equation (1) is completely determined by the parameters *w* and *b*. The parameter *w* determines the direction of the hyper-plane, while parameter *b* determines the distance of the hyper-plane from the center of the coordinate system. The classification is based on the sign of $\boldsymbol{w}^T\boldsymbol{x} + b$. SVM can also be generalized to the class of problems when data are not linearly separable.

The basic vector space in which the training set is not linearly separable can be mapped into a multi-dimensional space in which the training set is linearly separable by mapping Φ: $\mathbf{x} \rightarrow \varphi(\mathbf{x})$. Instead of a scalar product, a kernel function corresponding to a scalar product in a mapped space is introduced.

C. *Fusion of visible and thermal system*

System proposed in this paper utilizes two classifiers. Both classifiers are trained in parallel, *Thermal* with thermal images and *Visible light* classifier with visible facial images. Input in proposed face recognition system is a pair of images (thermal and visible) of one person captured in the same time, same illumination conditions and in the same facial pose. After getting the inputs, system starts parallel processing - thermal image through the thermal part and the visible image through the visible light part of the fusion system (Fig. 3). After extracting in the next block, obtained HOG features vectors are passed to the classifiers. High dimensionality of the HOG features vector justifies the use of the SVM for classifier, which eliminate the need for additional non-linear mapping into a larger space that would correspond to non-linear SVM.

After processing the obtained vectors, both classifiers give matching results. The results of each classifier are vectors with the matching scores for each person individually. Dimensions of these vectors are equal to the number of persons in the gallery database. In order to compare the results of individual classifiers and continue processing, vectors with matching scores must be normalized by choosing appropriate scaled coefficients. Proposed system utilizes the maximum of the two matching scores (thermal and visible light subsystems), that classifiers produce for each person in gallery database, as the final matching score.

IV. System Description

A. *Evaluation methodology*

This work presents results and performance statistics with three common evaluation methods:

- Receiver Operating Characteristic curve - ROC [10]
  - Accuracy is measured by area under the ROC curve (AUC)[11]
  - y axis: True Positive Rate (sensitivity)
  - x axis: False Positive Rate(1-specificity) - "false alarm"
- Cumulative Match Scores - CMC curves [12]
  - x axis: Rank - from 1 up to the number of faces in database
  - y axis: percentage of recognition accuracy at each rank
- Rank-1 accuracy [12]
  - recognition accuracy on rank 1 in CMC curve

B. *Experimental results*

Algorithm performance on test sets with visible facial images and different illumination conditions is shown by ROC curve - Fig. 4, and CMC curve - Fig. 5. These results were obtained using only subsystem that works with visible images.

Graphics of ROC and CMC curves show that the recognition is exceptionally good for facial images captured on daylight. This result is expected, because the lighting conditions on images from this test set are the most similar to those in the database. A little lower, but also very good, accuracy is obtained for facial images with an additional frontal light source. As all facial features are visible in this test set of images and the gradient of brightness is quite similar to the images in database, this is the reason for that good result in recognition accuracy. Very low performance has been obtained for faces in the darkness. In these images faces are poorly noticeable, and the gradient of brightness in this case is significantly different from the images in database, which results in high recognition error rate.

Table I supports the conclusions drawn from ROC and CMC curves with the area under the ROC curve (AUC) and Rank 1 - accuracy, as the most important indicators of the algorithm performance for variable illumination.

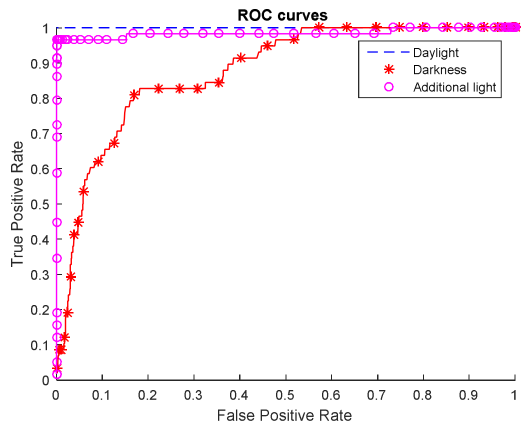

Fig. 4 ROC curves for test set with visible facial images and different illuminations

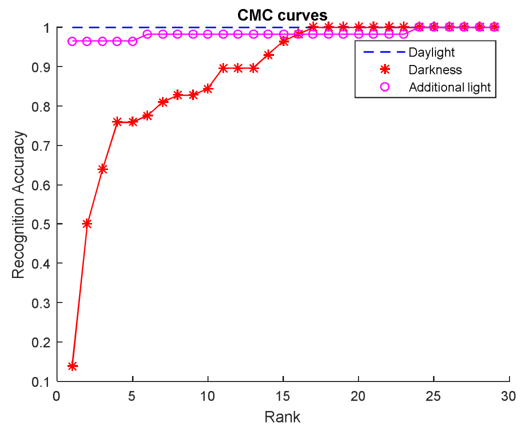

Fig. 5 CMC curves for test set with visible facial images and different illuminations

Table I AUC AND RANK-1 ACCURACY FOR TEST SET WITH VISIBLE FACIAL IMAGES AND DIFFERENT ILLUMINATION

	Daylight	Darkness	Additional light
AUC	1	0.86	0.98
Rank 1	1	0.13	0.96

Algorithm performance on test sets with thermal facial images and different illumination is shown with ROC curves - Fig. 6, and CMC curves - Fig. 7. These results were obtained using only subsystem that works with thermal facial images.

Graphics of CMC and ROC curves show that the recognition is exceptionally good for all subsets of facial images with different illumination. For all three test datasets (on daylight, in darkness and with an additional light source) obtained recognition accuracy is about and more than 90%. That shows a great advantage of using thermal images for face recognition. The biggest advantage of using this face recognition subsystem is presented in results obtained on test set of images with faces in the darkness. Compared to the subsystem that works with visible images and getting a high error rate for the images with faces in the darkness, here, for a test set of faces in the darkness, the accuracy of recognition is greater than 90%.

However, for test set of faces in daylight, *visible* subsystem achieves maximum accuracy, while the *thermal* subsystem produces a low recognition error. Slightly higher recognition error, compared to *visible* system, *thermal* subsystem produces for test set of faces with additional light source. Reason for that lies in fact that certain part of the energy that is reflected from the face in the conditions of additional lightening have influence to the appearance of thermal images.

TABLE II shows the area under the ROC curve (AUC) and Rank 1 - accuracy, as the most important indicators of the algorithm performance for different illuminations. Presented results supports the conclusions drawn from ROC and CMC curves.

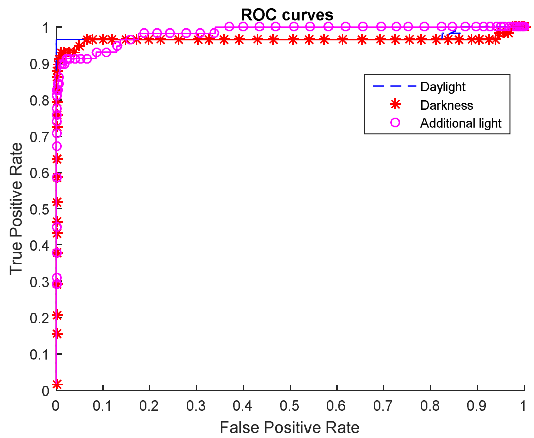

Fig. 6 ROC curves for test set with thermal facial images and different illuminations

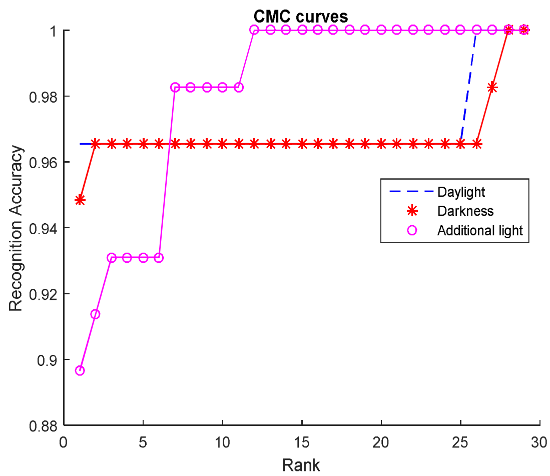

Fig. 7 CMC curves for test set with thermal facial images and different illuminations

TABLE II AUC AND RANK-1 ACCURACY FOR TEST SET WITH THERMAL FACIAL IMAGES AND DIFFERENT ILLUMINATION

	Daylight	Darkness	Additional light
AUC	0.97	0.96	0.98
Rank 1	0.96	0.94	0.89

The advantages of proposed unique system (fusion of thermal and visible subsystems) are confirmed and presented with the ROC curves - Fig. 8 and CMC curves - Fig. 9. Table III shows the area under the ROC curve (AUC) and Rank 1 for unique system. Presented results supports the conclusions drawn from ROC and CMC curves.

The accuracy of recognition for any lighting condition is over than 94%. The dominant influence of the part of the face recognition system for thermal images is shown in recognition accuracy on test images with faces in the darkness. The dominant influence of the part of face recognition system for visible images is shown in recognition accuracy on test images with faces on daylight and with additional light source. In this way systems are complementing each other which lead to maximum recognition accuracy and illumination invariant system.

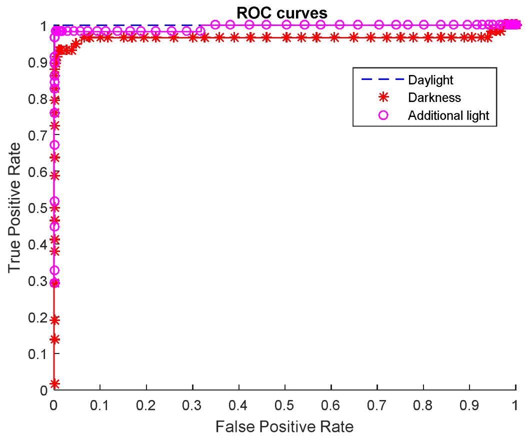

Fig. 8 ROC curves for different illuminations and fusion system

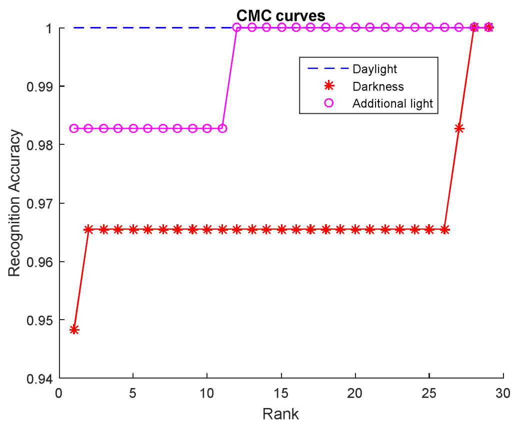

Fig. 9 CMC curves for different illuminations and fusion system

Table III AUC AND RANK-1 ACCURACY FOR DIFFERENT ILLUMINATION AND FUSION SYSTEM

	Daylight	Darkness	Additional light
AUC	1	0.96	0.99
Rank 1	1	0.9489	0.9828

V. CONCLUSION

This paper presents and compares face recognition methods for combining thermal IR and visible imagery for the illumination invariant face recognition system. Experimental results show that thermal face recognition performs better than visible face recognition in night conditions (scenes in darkness) and low light conditions, while visible face recognition system in the same lighting conditions completely fails.

In contrast to night conditions, the visible face recognition system has a very good performance tested with faces on daylight and with additional light source.

Fusion of these two systems takes the best from both systems and creates an illumination invariant system with exceptional performance. A problem with this approach is that, to the best of our knowledge, the most face recognition systems in use worldwide do not contain thermal gallery images. Adding this part to the existing systems might be challenging. On the other hand, it is easily applicable to the new face recognition systems.

In order to overcome the described challenge, and be operative with the existing visible light face recognition systems, our future research will include thermal-to-visible face recognition [13], [14], [15] algorithms.

REFERENCES

[1] M. Pavlovic, R. Petrovic, B. Stojanovic, S. Stankovic, "Facial expression and lighting conditions influence on face recognition performance", IcEtran, Palic, June, 2018.

[2] IEEE OTCBVS WS Series Bench; DOE University Research Program in Robotics under grant DOE-DE-FG02-86NE37968; DOD/TACOM/NAC/ARC Program under grant R01-1344-18; FAA/NSSA grant R01-1344-48/49; Office of Naval Research under grant #N000143010022.http://vcipl-okstate.org/pbvs/bench/

[3] K. R. Kakkirala, S. R. Chalamala, S. Jami, "Thermal Infrared Face Recognition: A review,"*UKSim-AMSS 19th International Conference on Modelling & Simulation,* 2017.

[4] S. G. Kong, J. Heo, F. Boughorbel, Y. Zheng, B. R. Abidi, A. Koschan, M. Yi, M. A. Abidi, "Multiscale Fusion of Visible and Thermal IR Images for Illumination-Invariant Face Recognition". *International Journal of Computer Vision*, *71*(2),pp. 215-233, 2007.

[5] O. Déniz, G. Bueno, J. Salido, F. De la Torre, "Face recognition using histograms of oriented gradients", *PatternRecognition Letters,*vol. 32, no. 12, pp. 1598-1603*,* 2011.

[6] C. Shu, X. Ding, C. Fang, "Histogram of the oriented gradient for face recognition," *Tsinghua Science & Technology*, vol.*16*, no. 2, pp. 216-224

[7] G. M. Fung, O. L. Mangasarian, " Multicategory proximal support vector machine classifiers," *Machine learning,* vol. 59, no. 1-2, pp. 293-300*,* 2005.

[8] J. A.Suykens, J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters,* vol. *9*, no. *3*, pp. 293-300*,* 1999.

[9] S. Čabarkapa, N. Kojić, V. Radosavljević, B. Reljin, "An implementation of SVM in CBIR system" (in Serbian) , Telfor, Belgrade, pp. 25-27, 2008.

[10] M. J. Pencina, R. B. D'Agostino, R. S. Vasan, "Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond.," *Statistics in medicine,* vol. *27*, no. 2, pp. 157-172*,* 2008.

[11] 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, pp. 171-186*,* 2001.

[12] R. M. Bolle, J. H. Connell, S. Pankanti, N. K. Ratha, A. W. Senior, "The relation between the ROC curve and the CMC," *Automatic Identification Advanced Technologies. Fourth IEEE Workshop on*, pp. 15-20, October, 2005.

[13] B. F. Klare and A. K. Jain, "Heterogeneous face recognition using kernel prototype similarities," IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 6, pp. 1410-1422, 2013.

[14] S. Hu, J. Choi and A. Chan, "Thermal-to-visible face recognition using partial least squares", Journal of the Optical Society of America A, vol. 32, no. 3, pp. 431-442, 2015.

[15] R. Petrovic, B. Stojanovic, M. Pavlovic, S. Stankovic, "Thermal-to-visible face recognition for illumination invariant systems", IcEtran, Palic, June, 2018.