Fusion of Dissimilar Features from Thermal Imaging for Improving Drunk Person Identification

Georgia Koukiou[1, 2] and Vassilis Anastassopoulos[1]
[1] University of Patras, Electronics Laboratory (ELLAB), Department of Physics, Rio Patras, Greece
[2] University of Thessaly, Department of Computer Science, Lamia, Greece
Email: {gkoukiou, vassilis}@upatras.gr

***Abstract*—In this work two different types of feature vectors are employed for drunk person identification. Both features are coming from the thermal images of the face of the inspected persons and have been derived using different image analysis techniques. Thus, they convey dissimilar information, which has to be transferred onto the same framework and fused to result into a decision with improved reliability. For this purpose conventional data association techniques are employed to explore the available information. These data association techniques are gating, normalization and weighting. After that, fusion of the information is carried out using Support Vector Machines. The resulting decision is of higher reliability compared to those achieved using the individual features separately. Experimental results are provided based on an existing sober-drunk database. The main advantage of the method is that it is not invasive and all the information is acquired remotely. In practice, an electronic system incorporating the proposed approach will point out to the police to whom an extended inspection for alcohol consumption is due.**

***Index Terms*—drunk identification, feature fusion, thermal imaging, dissimilar features, SVMs**

I. Introduction

Intoxication by means of alcohol consumption is a serious and sometimes dangerous condition that a person may fall into as far as its health, security as well as the social security is concerned. Common means of identifying drunkenness is by a breathalyzer or a blood test. Both methods require the person under test to come in touch with the device and to stand for an invasive test. Both procedures are time consuming, especially the blood test and they have a considerable cost. These techniques cannot be applied or used to monitor intoxication remotely and prevent drunk persons from being engaged in tasks that require the operator's attention and are associated with security. For example it is not efficient to perform a test with a breathalyzer before a football match if it is desirable to prevent the drunk persons entering the stadium.

The original idea on which this work is based, lies on the fact that the blood vessels' network of the face will present increased activity when the person has consumed alcohol changing in this way the temperature distribution on the person's face. This could lead to an automated system for discriminating intoxicated persons, which could be used as the first step to intoxicated person identification before proceeding to invasive biometric measurements such as breathalyzers or blood tests. Obviously, it is not possible to obtain a thermal map of the face by means of visible light. Acquiring images from faces in thermal infrared spectrum, information related to the temperature of the face is obtained which mainly depends on the physiological condition of the person (illness, exercises, drunkenness). The human face being in a mean temperature around 300^{o} K, emits according to the Wien Law [1] electromagnetic radiation as a perfect black body, with maximum at 10μm wavelength (thermal infrared). At the same time, our body absorbs thermal radiation from the environment. Therefore, thermal imaging devices such as thermal infrared cameras designed for human body inspection operate in the range of 7 to 14 μm.

Specific work on drunk identification has been carried out in the past by the authors with excellent discrimination capabilities [2]-[7]. These approaches are the only ones employing thermal information from the face from drunk discrimination. Although drunkenness is a challenging physiological condition to be investigated using infrared imagery, most of the publications in the literature refer only to automotive anti-drunk driving systems, which utilize electrical signals from the heart or brain [8].

In this paper, a fusion procedure is addressed, using features from [3] and [5], in order to increase reliability in identifying intoxication. Specifically:

1) A feature vector formed by simply taking 20 different points on the face of the person.
2) A feature describing the temperature distribution on the eyes of the sober and the drunk persons. The iris and the sclera are of the same temperature for the sober person. For the intoxicated person the iris becomes darker.

The two features are properly processed, normalized and weighted with well known data association techniques [9] and then fused by means of Support Vector Machines (SVMs) to give a decision with higher

Manuscript received August 29, 2017; revised November 23, 2017.

doi: 10.18178/ijsps.5.3.106-111

reliability. Accordingly, the first feature is firstly reduced in dimensionality so that a simpler two dimensional vector is obtained. This simpler vector is processed along with the second feature to obtain the final vector. Identification success of the fusion process outperforms the classification performance of the individual features and is higher than 95%.

The final goal of the exposed material is to achieve intoxication identification using only the thermal images of the drunk person without the need of comparisons with the corresponding images obtained when the person was sober. This is achieved by both most methods employed here and constitutes a significant challenge towards building a commercial product. Such a commercial product could scan the face of a person and in case he is identified as drunk, the system will prevent him of being engaged into critical procedures (driving or operating specialized infrastructure).

Basic elements regarding the limits in alcohol consumption posed in different countries, as well as the thermal behavior of the skin are provided in [1]. Furthermore, an analytical description of the database with sober and the corresponding drunk persons created in Electronics Laboratory Physics Department University of Patras, Greece as well as the experimental procedure carried out towards the completion of this database can be found in [4]-[7] (http://www.physics.upatras.gr/sober/).

The infrared images used in this work were acquired by means of the Thermo Vision Micron/A10 Model infrared camera (18 mm, f/1.6) of FLIR Company. Its operating range is from 7.5 to 13.0 μm, and it adjusts automatically its dynamic range (gray levels) to the minimum and maximum temperature of the scene. The resolution of the infrared images is 128x160 pixels.

Finally, it is worth mentioning that the people participated in the experiment were calm and in normal physical and psychological condition during the experiment. No illness, no psychological stress, other pathological reasons, or any kind of body exercise were recorded for any one of the participants. They were asked to be present in the room of the experiment half an hour earlier and to keep calm till the first acquisition of frames. We have to mention that our intention with this work is to distinguish drunkenness from normal condition (sober). No other kind of abnormality is considered.

In this paper the creation of the two feature vectors to be fused is presented in detail in sections 2 and 3. Data association techniques are provided in section 4. The fusion approach of dissimilar features is explained in section 5. The conclusions are drawn in section 6.

II. First Simple Feature Vector from the Face

The first simple feature vector for drunk person identification is obtained by simply taking the pixel values of 20 different points on the face of each person as shown in Fig. 1 [3]. Therefore, each face-image corresponds to a 20-dimentional feature vector:

x_i =[181 169 203 166 217 175 171 189 169 206 152 144 243 165 225 147 247 149 247 127]^t, which corresponds to a point in the 20-dimentional space. Since a set of 50 sequential images have been acquired for the same individual, a specific acquisition corresponds to a cluster of 50 points in the 20-dimentional space.

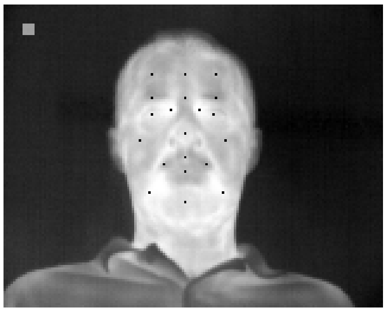

Figure 1. Twenty points were obtained on each face to monitor temperature changes with the consumption of alcohol.

It is important to find out if the cluster which corresponds to the same person moves in the feature space as the person consumes alcohol. Simultaneously, we have to examine if the cluster of each person moves towards the same direction with alcohol consumption. If the direction of movement due to alcohol consumption is different for different persons then we would have many directions in the 20-dimensional space, towards which the clusters of the drunk persons are moving. In this case it would be difficult to demonstrate the space in a simpler way (preferably in two dimensions).

It was found by means of a dimensionality reduction procedure that this problem is almost of 2-dimensions, since the solution of the generalized eigenvalue problem has given only 2-eigenvalues with significant value. Bringing all clusters in the 2-dimensional space, it is evident from a simple observation of the space that the clusters move to almost the same directions as the persons consumes alcohol. The directions of movement are estimated and it is proved that the space can easily be separated into “sober” and “drunk” regions (Fig. 2).

In our case, the feature space dimensionality was examined using the statistics of the clusters of eight persons. Consequently, the suitable directions of a 2-D space has to be found so that maximum seperability is achieved when the eight clusters for the sober persons and the corresponding eight clusters for the drunk persons are projected onto. To achieve that, the projection by means of a linear transformation W in a new space is required. The vectors w_i of W are the new directions where (each image-vector) x will be projected. This linear transformation is:

$$y_i = w^t x_i \tag{1}$$

An important criterion function that can be used for the separation of the clusters is given by

$$J = \frac{S_B}{S_W} \tag{2}$$

When J, becomes maximum the clusters are better separated. S_W is called the within-scatter-matrix while S_B is called the between-scatter-matrix. We need S_W to be small and S_B large. S_W expresses how much each separate cluster is dispersed in the space (cluster scatter) and is evaluated by summing up all cluster individual scatter matrices S_i as follows

$$S_w = S_1 + S_2 + . \quad . \quad . + S_8 \tag{3}$$

where,

$$S_i = \sum x_i * x_i^t \tag{4}$$

In the transformed space the within-scatter-matrix (S_W) is given by

$$(S_W) = w^t S_W w \tag{5}$$

The between-scatter-matrix S_B reveals how much the centers of the clusters are separated. The evaluation of the between scatter matrix S_B is realized as follows

$$S_B = \sum m_i * m_i^t, \qquad i = 1,2,\ldots 8 \tag{6}$$

where m_i corresponds to each cluster center. In the transformed space the between scatter matrix (S_B) will be given by

$$(S_B) = w^t S_B w \tag{7}$$

Therefore, we conclude that the function J in the transformed space is given by

$$J(w) = \frac{w^t S_B w}{w^t S_W w} \tag{8}$$

The transformation vectors w that maximize the function J(w) are obtained from the solution of the generalized eigenvalue problem:

$$S_B W_i = \lambda_i S_W W_i \tag{9}$$

This solution provides us the matrix W of eigenvectors w_i, which constitute the directions in the new space, on which we have to project the original image-vectors x_i. Simultaneously, it gives the eigenvalues which correspond to each of the above eigenvector. The eigenvalues express the importance of each direction-eigenvector in the feature space. The larger the eigenvalue the better the separability of the clusters we obtain towards the corresponding eigenvector. The solution of (9) is actually the Fisher Linear Discriminant (FLD) procedure [10]. It is a general procedure taking into consideration the matrices S_B and S_W with opposite effect.

The generalized eigenvalue problem was solved and the resulting 2-dimensional feature space is demonstrated in Fig. 2 where the clusters for 8 persons are depicted. A total of 16 clusters are presented in the transformed 2-D feature space, i.e. two clusters per person (sober and drunk). The sum of the two largest eigenvalues over the sum of all eigenvalues gives the quality of cluster seperability in the reduced (2D) feature space. In this experiment and for all persons of the database this ratio approaches 90%. Furthermore, in the same figure, the direction of movement of the cluster of each person is exhibited. It is obvious from Fig. 2, that there is a line in the feature space or otherwise decision boundary, which separates the space into two regions: the "drunk" and "sober". Based on this figure we can decide whether an unknown person is drunk or not, from the position of the corresponding cluster on this space, hereafter the "drunk space".

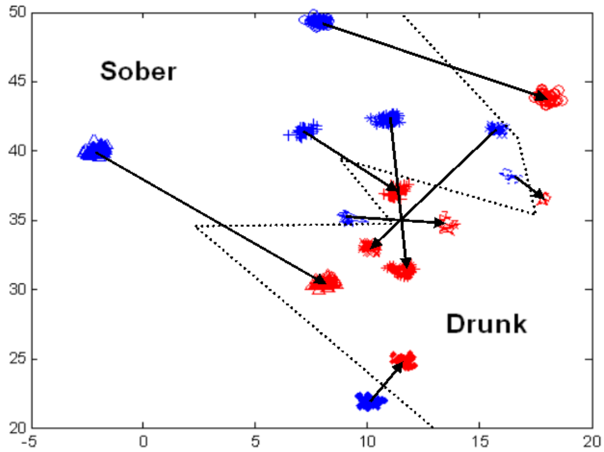

Figure 2. The 16 clusters of 8 persons in the 2-D space formed by the two most important directions (correspond to the first 2 largest eigenvalues). We call hereafter this space, the "drunk space".

The 2-D feature vector created from the Fisher Linear Discriminant analysis described previously, is to be used in conjunction (fusion) with the features obtained from the persons eyes (described in the next section) in order to create the final feature vector. Hereafter, we use the following notation for the above 2-D feature

$$x_a = \begin{bmatrix} x_{a1} \\ x_{a2} \end{bmatrix} \tag{10}$$

Additionally, it worth mentioning that the above procedure constitutes part of the data association processing which is exposed in section 4 and is needed to bring the features onto a common framework.

III. Second Simple Feature Vector from the Face

Temperature distribution on the eyes (Fig. 3) of sober and drunk persons is studied by means of thermal infrared images [5]. It is observed that the temperature difference between the sclera and the iris is zero for the sober person and increases when somebody consumes alcohol (Fig. 4). For the drunk person, iris appears darker compared to sclera which means that the sclera temperature increases. This is something expected since the sclera is full of blood vessels which present increased activity when the person consumes alcohol. Thus, in a screening procedure for drunk identification, the infrared images of the sober person are not needed. Although in most cases the sclera is brighter than the iris for the drunk persons, histogram modification algorithms can be employed when necessary to show off the gray level difference between the sclera and the iris for intoxicated persons. The discrimination capability of the procedure has been verified using the

Student t-test [5]. It was found a confidence of over 99% in drunk person discrimination.

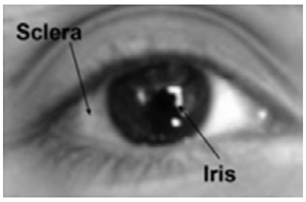

Figure 3. Sclera is surrounding iris which is actually a muscle controlled part of the eye to adjust the size of the pupil. Sclera lies on a net of blood vessels.

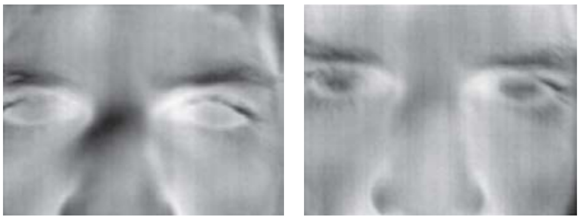

Figure 4. The mean value of 50 frames for a sober person (a) and a drunk person (b). No preprocessing has been applied on the images. It is evident for this person that the sclera becomes hotter compared to the iris when the person consumes alcohol.

Specifically, for the 28 among the forty-one people who participated in the experimental procedure, it was evident by a simple comparison of the thermal images that the sclera becomes hotter compared to the iris after alcohol consumption. As shown in Fig. 4 the iris appears darker (cooler) while the images are depicted in their original form without any kind of preprocessing. In four more people the iris appears darker than the sclera for the drunk person only after applying a histogram equalization algorithm. For all these participants, the sclera and the iris are of the same temperature in case of the sober person appearing with almost the same gray level. Moreover, another 4 persons revealed the temperature difference between the sclera and the iris for the intoxicated person only when other special histogram modification algorithms were applied. These images have been modified using a histogram modification algorithm which clips all values lower than 0.5 and higher than 0.75, while it stretches the rest histogram to occupy the whole histogram range (MATLAB imadjust ([0.5 0.75], [0 1])). For 5 persons no algorithm succeeded to show off increased temperature for the sclera in the case of drunk persons. These persons were also used to drinking wine or other spirit occasionally.

The temperature difference between the sclera and the iris was examined [4], [6] based on the statistics of the pixels in these two regions by means of two different estimation procedures which correspond to two different discrimination features. In the first procedure, the ratio of the mean value of the pixels inside the sclera's to the mean value of the pixels inside the iris, was calculated (see Fig. 5). This is the first feature x_{b1} obtained based on the statistics of the eye. This procedure was performed on the left eye of each participant, both in the case he is sober and when he has consumed alcohol. Consequently, two ratios of the mean value of the sclera to the mean value of the iris are available. It is observed that the ratio of the mean pixel value on the sclera to the mean value of the iris increases when the person has consumed alcohol. Specifically, for the 36 from the 41 cases the specific ration increases with alcohol consumption while only in 2 cases it decreases and in the rest 3 it remains almost the same. The results were analyzed using the Students-t test, in order to support statistically the drunk screening capabilities of the proposed method from eye thermal images [5].

In the second procedure, is estimated the variance of the pixels contained in the whole eye (large ellipse in Fig. 5). This evaluation was performed for the left eye of each participant when the person is sober and when he is drunk. Therefore, two variances for each participant have been calculated, corresponding to sober and drunk person respectively. This is the second feature x_{b2} obtained based on the statistics of the eye. It is observed that the variance increases in case that the person has consumed alcohol. Specifically, among the forty-one participants only 4 presented decreased variance in the region of the eye for the drunk person compared to the sober one.

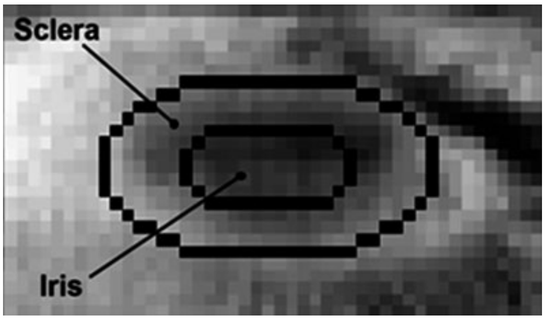

Figure 5. The two regions, small ellipse (iris) and ellipsoidal ring (sclera) are those in which the mean pixels value is evaluated separately. Simultaneously, the variance of all pixels inside the large ellipse is evaluated.

According to the material exposed in this section a 2-D feature vector can be obtained employing features x_{b1} and x_{b2} as follows

$$x_b = \begin{bmatrix} x_{b1} \\ x_{b2} \end{bmatrix} \tag{11}$$

Feature vectors described by (10) and (11) will be combined in the following for improved intoxication identification performance.

IV. Dissimilar Feature Association

In this section we present the basic aspects regarding the association between the two different types of features that are combined to give the final information. The procedure refers to manipulating the correlation between these features as well as the importance of each of the features by weighting them with proper coefficients. In a general data association procedure where the dissimilar features are presented simultaneously we can follow 4 steps to perform association [9]:

a) Retrieval of candidate features from the database
b) Gating to eliminate unwanted or unrelated features

c) Computation of an association matrix
d) Implementation of the assignment logic

The first two steps have already been carried out during the feature extraction procedure as well as throughout the feature manipulation processes such as the Fisher Linear Descriminant and the histogram modification of the thermal images of the eye.

The final feature vector to be transferred to the SVMs is as follows

$$x = \begin{bmatrix} x_a \\ x_b \end{bmatrix} = \begin{bmatrix} x_{a1} \\ x_{a2} \\ x_{b1} \\ x_{b2} \end{bmatrix} \tag{12}$$

The association matrix S_a, which reveals various kinds of correlation between the features is evaluated from the expectation

$$S_a = E\{x \quad x^t\} \tag{13}$$

and in this specific case it was found that the correlation between all 4 components is small which permits for an increased classification performance when all featured are used together.

Finally, the implementation of the assignment logic is carried out by the SVMs as described in the following section. The SVMs are implementing the proper feature normalization procedures in order to apply to the features the appropriate weights.

V. Fusion Approach-Experimental Procedure

In this section we present the results from the experimental procedure where the Support Vectors Machines (SVMs) [11] are employed. In an ordinary SVMs classification approach someone has to deal with a set of data for training the SVMs and a set of data for testing the procedure. Each vector of data sets contains the class labels (e.g. 0: for sober and 1: for drunk) and the associated features. The goal for SVMs classifier is to produce a model which is based on the training datasets and which predicts the class labels of the test datasets.

If we have a training dataset of instance, label pairs (x_i,y_i), $i=1....l$, where $x_i \in R^n$ and $y_i \in \{1,-1\}^l$, the SVMs requires the solution of the optimization problem [12]:

$$\min_{w_i b_i \vartheta} \frac{1}{2}\,\omega^T\omega + C \sum_{i=1}^{l} \theta_i \tag{14}$$

subject to

$$y_i\,(\omega^T \phi(x_i) + b) \geq 1 - \theta_i \qquad \text{with} \qquad \theta_i \geq 0 \tag{15}$$

Let the training vectors x_i which are mapped into a higher dimensional space by the function ϕ. SVM finds a linear separating hyper plane with the maximal margin in this higher dimensional space. $C > 0$ is the penalty parameter of the error term. Furthermore, $k(x_i,x_j) \equiv \phi(x_i)^T \phi(x_j)$ is called the kernel function. In the following are presented four basic kernels:

linear $\quad k(x_i,x_j) = x_i^T x_j$

polynomial $\quad k(x_i,x_j) = (\gamma x_i^T x_j + r)^d,\ \gamma > 0$

Radial Basis Function (RBF)

$$k(x_i,x_j) = exp(-\gamma \| x_i - x_j \|^2), \quad \gamma > 0$$

sigmoid $\quad k(x_i,x_j) = tanh\,(\gamma x_i^T x_j + r)$

where γ, *r and d are kernel parameters*.

In our experiment we use two procedures for achieving the best classification results. In the first procedure, we use the data without normalization. This procedure contains three steps, which are given in the following:

- Transform data to the format of SVMs
- Randomly try a new kernels and parameters
- Test

In the second procedure, we apply scaling to the data before applying SVMs. With the scaling, we achieve to avoid feature vectors in greater numeric ranges, dominating thus those in smaller numeric ranges. Also, we can avoid numerical difficulties during calculation, because kernel values usually depend on the inner products of feature vectors.

In our experimental procedure we apply three different linear scaling for each feature vector to the range [-1, 1], [-2, 2] and [-9, 9]. We have to mention here that the same method was used in order to suite both the training and the testing data.

In the second procedure the applied steps are the following:

- Transform data to the format of SVMs
- Conduct simple scaling on the data with default parameters
- Test

The final experimental results from the two procedures are presented in Table I.

TABLE I. Experimental Results

First Procedure	Second Procedure		
	Scaling Range to [-1, 1]	Scaling Range to [-2, 2]	Scaling Range to [-9, 9]
23/41	36/41	38/41	41/41
56.1%	87.8%	92.7%	100%

VI. Conclusions

A fusion approach was presented which aims at combining features obtained from thermal imagery in order to improve intoxication identification. It is actually the first approach worldwide to address drunkenness by means of fusing dissimilar features obtained from thermal infrared images of the face of the inspected person.

The two dissimilar types of features were brought onto the same unified framework following conventional data association procedures and fused by means of SVMs decision making. The combination of the two features results into a success rate which approaches 100%.

The main advantage of the method is that it is not invasive and all the information is acquired remotely. In practice, an electronic system incorporating the proposed approach will point out to the police to whom an extended inspection for alcohol consumption is due.

ACKNOWLEDGMENT

This work was supported by the "Iky Fellowships of Excelence for Postgraduate Studies in Greece-Siemens Program".

REFERENCES

[1] N. A. Diakide and J. D. Bronzino, *Medical Infrared Imaging*, 1st ed. New York: CRC Press, Taylor & Francis Group, 2008.

[2] G. Koukiou and V. Anastassopoulos, "Facial blood vessels activity in drunk persons using thermal infrared," in *Proc. 4th Int. Conf. on Imaging for Crime Detection and Prevention*, Kingston, GB, 2011.

[3] G. Koukiou and V. Anastassopoulos, "Drunk person identification using thermal infrared images," *Int. J. of Electronic Security and Digital Forensics*, vol. 4, pp. 229-243, 2012.

[4] G. Koukiou and V. Anastassopoulos, "Neural networks for identifying intoxicated persons," *Forensic Science International*, vol. 252, pp. 69-76, 2015.

[5] G. Koukiou and V. Anastassopoulos, "Drunk person screening using eye thermal signatures," *J. of Forensic Sciences*, vol. 61, pp. 259-264, 2016.

[6] G. Koukiou and V. Anastassopoulos, "Intoxicated person discrimination using infrared signature of facial blood vessels," *Australian J. of Forensic Science*, vol. 48, pp. 326-338, 2016.

[7] G. Koukiou and V. Anastassopoulos, "Local difference patterns for drunk person identification," *Multimedia Tools Appl.*, pp. 1-13, 2017.

[8] Y. C. Wu, Y. Q Xia, P. Xie, and X. W. Ji, "The design of an automotive anti-drunk driving system to guarantee the uniqueness of driver," in *Proc. Int. Conf. on Information Engineering and Computer Science*, 2009, pp. 1-4.

[9] D. L. Hall, *Mathematical Techniques in Multisensor Data Fusion*, 1st ed. Boston:Artech House, 1999.

[10] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. New York: Wiley & Sons, 2001.

[11] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: an application to face detection," in *Proc. Conf. of Computer Vision and Pattern Recognition*, Puerto Rico, 1997.

[12] C. Cortes and V. Vapnik, "Support-vector network," *Machine Learning*, vol. 20, pp. 273–297, 1995.

Georgia Koukiou (B.Sc., M.Sc., and Ph.D.) was born in November 1978, in Patras. She received the B.Sc. degree in Physics in 2004, the M.Sc. in Electronics and Computers in 2006 and her Ph.D. in Digital Image and Signal Processing in 2014 all from University of Patras.

She is a member of the research team of Digital Image and Signal Processing of Electronics Laboratory, Physics Department in the University of Patras. She is currently working in Face Identification using Thermal Infrared, Biomedical imaging, Remote Sensing, SAR imaging, Ground Penetrating Radar methods and Information Fusion specializing in Decision Fusion methods. She has 20 publications in International Journals and Conferences.

Dr. Koukiou is a Postdoc and Scholar of IKY. She is currently giving lectures on Electronics, Programming, Signal and Image Processing, Pattern Recognition and Remote Sensing at University of Patras and University of Thessaly.

Prof. Vassilis Anastassopoulos was born in Patras, Greece, in 1958. He received the B.Sc. degree in Physics in 1980 and his Ph.D. in Electronics in 1986, both from the University of Patras, Greece. He worked for two years in Canadian Scientific Institutions. He is an academic staff member in Physics Department, University of Patras, Greece from 1987. He is a Professor from 2005. He was the Head of Physics Department, University of Patras, for the biennium 2001-2003 and 2005-2007, and the Vice Rector for Strategic Research Planning and Development in the University of Patras from 2006 to 2010.

His publication record contains over 100 journal and conference papers with over than 1500 citations. His research interests are within the scope of Digital Signal Processing, Image Processing, Radar Signal Processing and Remote Sensing with emphasis in Multi-spectral, SAR and Infrared Imagery, Signal Detection, Pattern Recognition with emphasis in Handwritten Analysis, and Information Fusion including Image Fusion, Decision Fusion and Sensor Fusion Architectures. Lately, he was involved in image processing techniques for Astro-particle Physics.