

# Behavioral Biometrics in Assisted Living: A Methodology for Emotion Recognition

Stefanos Xefteris

Distributed Knowledge and Media  
Systems Laboratory, Dept. of Electrical  
and Computer Engineering, National  
Technical University of Athens  
Athens, Greece  
xefteris@gmail.com

Nikos Doulamis

Photogrammetry and Computer Vision  
Lab, Dept. of Rural and Survey  
Engineering, National Technical  
University of Athens  
ndoulam@cs.ntua.gr

Vassiliki Andronikou

Distributed Knowledge and Media  
Systems Laboratory, Dept. of Electrical  
and Computer Engineering, National  
Technical University of Athens  
Athens, Greece  
vandro@mail.ntua.gr

Theodora Varvarigou

Distributed Knowledge and Media Systems Laboratory,  
Dept. of Electrical and Computer Engineering, National  
Technical University of Athens  
Athens, Greece  
dora@telecom.ntua.gr

George Cambourakis

Division of Communication, Electronic and Information  
Engineering, Department of Electrical and Computer  
Engineering, National Technical University of Athens  
Athens, Greece  
gcamb@cs.ntua.gr

**Abstract**—Behavioral biometrics aim at providing algorithms for the automatic recognition of individual behavioral traits, stemming from a person's actions, attitude, expressions and conduct. In the field of ambient assisted living, behavioral biometrics find an important niche. Individuals suffering from the early stages of neurodegenerative diseases (MCI, Alzheimer's, dementia) need supervision in their daily activities. In this context, an unobtrusive system to monitor subjects and alert formal and informal carers providing information on both physical and emotional status is of great importance and positively affects multiple stakeholders. The primary aim of this paper is to describe a methodology for recognizing the emotional status of a subject using facial expressions and to identify its uses, in conjunction with pre-existing risk-assessment methodologies, for its integration into the context of a smart monitoring system for subjects suffering from neurodegenerative diseases. Paul Ekman's research provided the background on the universality of facial expressions as indicators of underlying emotions. The methodology then makes use of computational geometry, image processing and graph theory algorithms for the detection of regions of interest and then a neural network is used for the final classification. Findings are coupled with previous published work for risk assessment and alert generation in the context of an ambient assisted living environment based on Service oriented architecture principles, aimed at remote web-based estimation of the cognitive and physical status of MCI and dementia patients.

**Keywords:** ambient assisted living; dementia; emotion; image processing; telemedicine

## I. INTRODUCTION

People older than 64 represent an average of 18.5% of the total European population and growing notably is the age group over 84, as reported by Eurostat in 2015[1]. Such remarkable steady

increase of older population in Europe requires innovative approaches, to make possible and maintain the independence, autonomy and quality of life of our elderly. The impact of an increasing aging population also affects greatly formal and informal care providers. An aging population is also associated with an increase in the number of people suffering from mild cognitive impairment (MCI), such as the range of dementia [2-4]. These conditions are very challenging from the outset, ranging from the problem of reaching and obtaining a diagnosis, to the tailoring of appropriate care and support for the individual and their caregivers both family and professional [5-6]. There is therefore the need for an innovative approach, to explore different ways of working with increased collaboration across disciplines: healthcare experts and technologists and greater working partnerships between older people and their formal and informal caregivers [7-9].

Progress in Information and Communications Technology (ICT) in the past years has led in exciting advances with direct effect in aspects of everyday life activities. Especially concerning modern healthcare, ICT has permeated deeply the current paradigm providing new frameworks for telecare, with comprehensive multi-stakeholder benefits, for the individual subject and direct carers, local healthcare infrastructures and healthcare system as a whole [10-12]. ICT provides frameworks that enable internet based, decentralized, cost-effective and efficient services to patients. Ambient Assisted Living (AAL) systems are being constantly on the cutting edge of ongoing research, aiming at improving home-base patients' quality of life, provide them with means to cope with their diseases, delay institutionalization and alleviate burden of family members [12-14]. The need for a paradigm shift from remedial to preventive measures in the field of assisted living is

apparent and currently in its early stages. In this context, modern AAL systems comprise significant tools, providing better information processing, extraction of patterns, analysis and comparison of data for adverse event prediction and also patient profiling capabilities to improve subject's lives delaying hospitalization/institutionalization as much as possible. A systematic flow of information to and from the subject along with identification of adverse events and assessment reports leads to a more effective care system, socially and financially beneficial to the subjects. Remote monitoring of patients under a risk prediction framework is crucial for shifting the current paradigm from remedial to preventive action [7, 15-19].

Patient profiling as well as activity recognition in the context of AAL systems, opens new routes and possibilities, aiming at actively and measurably aiding subjects to cope with their activities of daily life (ADLs) and their physical and psychological status. Supplemental to the regimented physiological measurements, profiling the subject's mood changes, emotional reactions to stimuli, attitudes and conduct can be of great aid to attending doctors and family members. Information can be acquired and processed in order to not only detect abnormalities but identify behavioral trends in order to prevent both physical and physiological adverse events [17, 20-21]. The next step in patient profiling is behavioral and activity monitoring as parts of decision support implementations in AAL context [22].

Within this context, biometrics, as a source rich in profiling information related to the physiological and behavioral characteristics of a person, form the basis for a very promising technology that can further boost and enhance patient profiling.

The strength of biometric traits lies in four fundamental characteristics [23]:

- Sufficient inter-person variability for distinctiveness
- Short-term temporal invariability
- Collectability/measurability and
- Universality

In this paper, we discuss a methodology for emotion recognition targeting applications in the framework of an ambient assisted living system following Service Oriented Architecture (SoA) principles, making use of extracted biometric facial features of the subjects. In the case of emotion recognition, the methodology makes use of the latter two characteristics of biometric traits, i.e the ease of collection and measurability and universality with the added benefit of the non-invasive nature of collection method. In our methodology, biometric information is not used to directly characterise the subject (explicit biometric profiling) but provide the background for extraction of regions of interest on a subjects' facial structure and the abstract generalisation of emotion as a geometric shape on a surface.

## II. BIOMETRIC PROFILING IN AAL SYSTEMS

ICT applications are gaining more acceptance as solutions that can act complementary to the caregiver's work by actively providing information and continuously monitoring and

assisting subjects with reduced cognitive (or even physical capacity). This is also augmented by the ease of installation in home environments of a multitude of sensors – from simple “binary” sensors, to temperature and lighting sensors [24-25] and from movement trackers on floors [26] to more complex sensors that provide audio and video recordings [27].

We use biometric profiling information as a part of the whole patient profile and assist in the detection and prediction of physical and psychological adverse events. In this context researchers usually make use of biometric technologies such as voice analysis, gesture and gait classification and emotion recognition. Emotion recognition is the task of processing a specific stream of data in order to classify the underlying emotional state. It is easier to use basic (or archetypal) emotions such as joy, sadness, anger, fear, disgust and contempt [28]. In general, emotion recognition systems are built on three distinct levels: acquisition, feature extraction and interpretation [29]. In our case, sensor technologies used in biometric data capturing are cameras which provide the advantage of being non-intrusive and cheap. Facial expressions as the basis of emotion recognition contain all the necessary information and are easily captured and analysed.

The design and implementation of monitoring systems to enable the self-management of neurodegenerative disease patients at home, impose many issues. The focus of this approach lies on prevention and prediction of adverse events through the early identification of risk factors. Thus, our methodology aims to be integrated into a risk assessment and detection component that processes patient data and if necessary, creates alerts for the clinicians. Such a component, can be based on an existing set of “Rules” and implemented to process both physical and psychological data as well as answers to cognitive and behavioral evaluation tests, to produce its warnings[30].

The main functionalities of such a component would include:

- Generate warnings for the attending physicians for situations where the patient's physiological and/or psychological measurements indicate an alarming condition for the patient's health.
- Compile automatic reports for the attending physicians on a scheduled regular basis
- Generate warnings for the subject's caregivers in case of emergencies
- Present evidence for the produced warnings

Generation of warnings is based on an easily reconfigurable rule-based schema, that attending clinicians can individualize according to each subject's personal needs, as seen in Figure 1.

## III. FACIAL EXPRESSIONS AND EMOTION

Faces are essentially very effective systems for the transmission of multiple signals and multiple messages [31]. They transmit voluntary and involuntary messages regarding emotions, mood, character, age, sex, race etc. When humans respond to a stimulus, internal or external and feel something,

their face responds accordingly. Muscles contract, wrinkles appear or disappear, mouths change shape, eyes open wide or close etc. Emotions as expressed by our faces are not static structures, but dynamic and short-lived. In this paper we concern ourselves with “emblematic” messages. The term “emblematic” is used to describe messages that have a clearly defined content, formed by specific and easily recognizable facial expressions [32]. Before venturing forth, we have to answer to 3 specific questions: a) which emotions can be expressed by the human face? b) is the initial recognition accurate? and c) can we use emotion as a behavioral biometric trait?

A. Which emotions can be expressed?

Before proceeding to the technical aspect of automatic emotion recognition from the image of a human face we must set concrete foundations, based on psychology and anthropology centred on how emotion is pictured on the human face. Research on the expression of emotion begins even from the time –and person- of Charles Darwin [33]. Since then there has been a plethora of research and hundreds of experiments whose results have been described in the works of Dr. Paul Ekman [32, 34-35]. According to this research, the human face can express (in a meaningful way for observers) six basic and archetypal emotions [31]: joy, sadness, surprise, anger, disgust and fear. There are obviously more emotions in the expression range of the human face (such as shame and enthusiasm) but these have not shown to have the same constant qualities as the 6 archetypal ones [32, 36-37].

B. Can we recognize emotions accurately?

Researchers here faced the following issue: it is not enough to just simply discern human emotion inside an expression. It is essential to discover if the interpretation by the observers is accurate or not. In Ekman’s experiments researchers who knew a priori the emotion expressed by the subject, categorized the answers of observers chosen from different cultural backgrounds, as right or wrong. The subjects ranged from healthy people who were shown a horror film/comedies etc to psychiatric hospital patients subjected to stress or relaxation experiences. The accuracy of the answers can be seen in Table I.

TABLE I. RESULTS OF EKMAN’S EXPERIMENT ON THE RECOGNITION OF EMOTION EXPRESSION BY MULTICULTURAL OBSERVERS

	Observer’s Country of Origin				
	U.S.A	Brazil	Chile	Argentina	Japan
<b>Total Observers</b>	99	40	119	168	29
<b>Emotion</b>	Percentage of Correct Answers (%)				
<b>Fear</b>	85	67	68	54	66
<b>Disgust</b>	92	97	92	92	90
<b>Joy</b>	97	95	95	98	100
<b>Anger</b>	67	90	94	90	90

C. Can emotion be used as a biometric trait?

In order to categorise emotion as a soft biometric trait – because it is obvious that emotion is a dynamic feature that cannot be used for identification or verification as other biometric traits- it has to fulfil the fundamental requirements of biometric characteristics, as they were earlier described. It is obvious that it is easily collectible, but it also needs to fulfil the universality requirement. Everybody expresses emotions through facial expressions but are emotional expressions themselves universal among people of all ages and cultural backgrounds? We have to establish here that the expressions of the six archetypal emotions are the same on people, irrelevant of cultural background, race, education, sex and age. Research has proven exactly that: the expression of emotions of the human face is the same for the six basic emotions, regardless of cultural background, age, sex and educational status. The only influence on recorded results was found when cultural background interfered on when it was appropriate to show or hide one’s emotions. The same expressions indicated the same emotions whether the observers came from the U.S, Chile or Brazil [31, 38]. In order to remove any possible objection, researchers travelled to the highlands of New Guinea to measure reactions and expressions of people totally unaffected by the western civilization and found the exact same results: The expression of emotion on the human face is universal across all cultures [31, 35, 38-41]. Thus, it can be used as a biometric trait.

IV. METHODOLOGY

In this research the authors used, after signing the relevant user agreement that gave them the right to reproduce parts of the dataset with the relevant copyright notice, the Cohn-Kanade extended dataset of facial expressions [42, 43], which

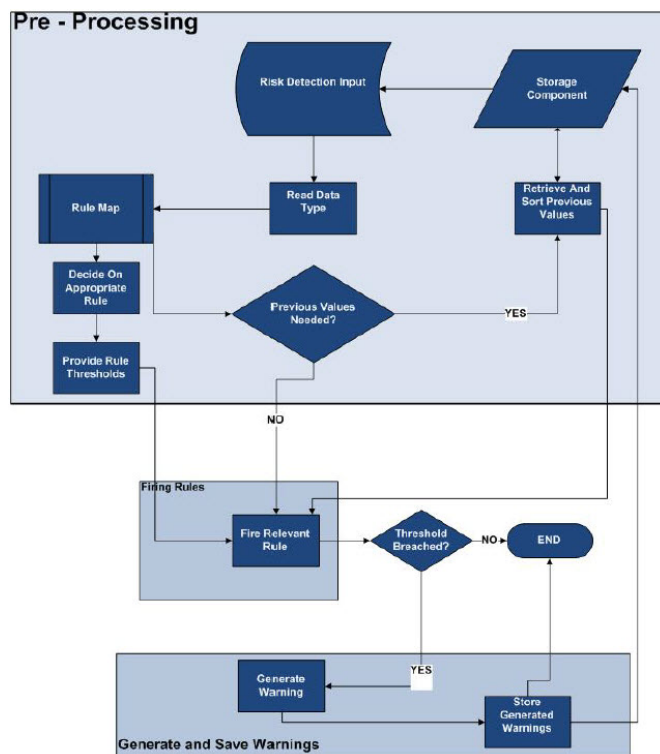


Fig. 1. Rule based risk assessment scheme

incorporates 593 sequences from 123 different people. Each sequence consists of pictures taken as the subject faces the camera, begin from neutral expression and end in a final expression depicting one of the six basic emotions along with meta-information concerning the emotion depicted and an encoding according to Ekman's FACS system. In this research only the final categorization was used, for verification of classification results.

The methodology followed, based on the fact of complete emotion expression universality aimed at providing an algorithm that needs no individual training of specific users (although training can be used in a future implementation to further improve acquired results). Also the proposed methodology aims at being easily incorporated into a SoA based telemedicine system, based on reconfigurable rules [30], or as part of a multi-algorithmic pool in a biometrics system as the one proposed in [44].

The methodology follows the classic steps of biometric traits recognition: Collection, Extraction of features, Comparison and Classification as shown in Figure 2.

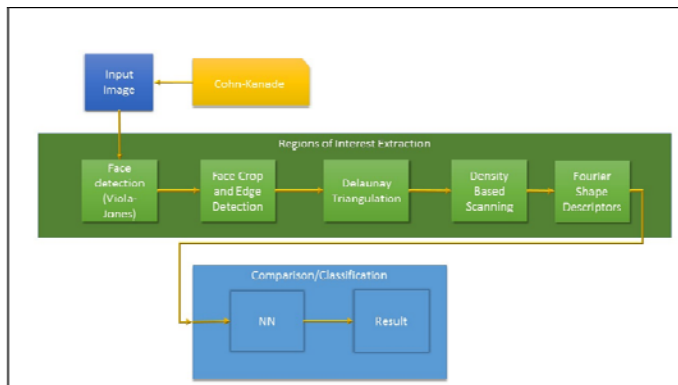


Fig. 2. The proposed methodology

#### A. Identification of Regions of Interest

The first concern in the preprocessing stage, was to crop out parts of the face that are completely irrelevant with emotion and could function as noise, such as the ears and hair, as well as totally irrelevant elements of the image itself (like timestamps). For this reason, we employed the well-known Viola-Jones face detection algorithm [45]. The Viola-Jones algorithm is a very fast and efficient solution on face detection, scale and location invariant, based on Haar-like features and the Adaboost training algorithm coupled with cascading classifiers. Moreover, it also performs extremely well in real-time applications so in a future implementation for real-time emotion recognition, it would not add delay to the methodology. The second concern is to identify which areas on the face are of interest in our case and try to separate them from the others. These regions are in order of significance: mouth, eyes, forehead and nose area.

Observing facial expressions, how they form on the human face and where the human eye focuses to perceive them, it is clear to see that where emotion appears, we have creases in the skin or facial features that define changes in luminosity. The

“background” of the face can be considered as a canvas on which harsh lines appear when we express emotion. That same “background” is not needed in the processing phase and so we employ an edge detection algorithm to remove useless parts of the phase and at the same time significantly lessen the amount of data the algorithm uses. While the employment of edge detection may seem as a vulnerability of the algorithm, due to its intolerance in high luminosity environments, we keep in mind that the target application concerns the detection inside home environments where the average luminosity is kept at regular levels.

At this stage, we must segment the regions of interest in specific sub-clusters. For this purpose we first perform Voronoi tessellation [46-47] and its dual, Delaunay triangulation [47-48] on the edge detection generated matrix of points. With this transform, the separate points in 2D space that the edge detection provided, become an undirected graph, which both preserves the geometric properties of the structure and provides an easy input for the application of clustering algorithms. Moreover, having our input in the form of a graph, makes it easily processable by a multitude of algorithmic approaches.

To define the Voronoi tessellation of a set of points on a metric space, we let  $X$  be the space and  $d$  the metric. Also, let  $K$  be a set of pointers and  $(P_k)_{k \in K}$  a tuple of non-empty subsets in space  $X$ . The Voronoi tessellation  $V_i$  of  $(P_i)$  is the set of points in  $X$  whose distance from  $(P_k)$  is less or equal from their distances of every other  $P_j$ , with  $j$  a different set of pointers. Formally:

If  $d(x, A) = \inf\{d(x, a) \mid a \in A\}$  then

$$V(p_i) = \{x \in X \mid d(x, P_i) \leq d(x, P_j) \forall j \neq i\}$$

The Voronoi tessellation diagram  $\text{Vor}(P)$  corresponding to the edge detection points in two-dimensional Euclidean space is the representation of a tuple  $(V_i)_{i \in K}$ . In our case, where the space is Euclidean and on it the regular distance metric  $l_2$ , we have a set of non-overlapping points and each Voronoi tessellation is a convex polygon. With this step we have created a polygon model of the human face with low area polygons in areas of expression. At this point in our methodology the goal is to transform this set of polygons to an undirected graph representation. To facilitate this we employ the dual method of Voronoi tessellation, the Delaunay triangulation. We define the Delaunay triangulation of a Voronoi diagram of points  $P$  as follows :

Let  $P = \{p_1, p_2, \dots, p_n\}$  a set of points on a space  $E^m$  and  $\text{Vor}(P)$  the Voronoi diagram of  $P$ , with  $V_i$  as each tessellation. The Delaunay triangulation  $\text{Del}(P)$  of  $p$  is: The complex  $\text{Del}(P)$  includes the  $k$ -simplex  $\{p_1, p_2, \dots, p_{k+1}\}$  if  $V_1 \cap V_2 \cap \dots \cap V_{k+1} \neq \emptyset$  with  $0 \leq k \leq m$ .

The complex  $\text{Del}(P)$  is the Delaunay triangulation of the convex hull of  $P$ . In essence, the Delaunay triangulation creates an edge between two points in Voronoi tessellation  $P$ , if these

two points have a common edge among their polygons. With the Delaunay triangulation we have now available an undirected graph representation of the expression of the human face, on which we can apply a graph clustering algorithm. In order to facilitate input for the neural network classification, we must first separate these clusters one by one and lessen the total number of the points that define them. As we have sets of vertexes, connected via edges in 2D space, and the basic criterion of definition of a cluster is how close these edges are, the simplest way of defining separate point clusters is the use of a density based scanning algorithm such as DBSCAN [49]. This algorithm, DBSCAN, was selected instead of other clustering algorithms due to the inherent properties of the clusters:

- there is not a priori knowledge of the number of appearing clusters so K-means clustering cannot be applied.
- Clusters are not organized in regular shapes and in some cases, one cluster can encompass another. Given the geometric nature of facial expressions, we should be aware of the single link effect (two clusters linked by a thin line of points) and negate it so there are no overlapping.
- We need an algorithm that is resilient to noise and outlier points. DBSCAN fulfils these requirements [50] with good results.

DBSCAN parses an input of a set of  $D$  points in the two dimensional Euclidean space and works on the assumption that each cluster around a point fulfils the requirement that there is a specific minimum number of points inside a given radius around that point. In this case, the normal Euclidean distance was used as a metric. The algorithm discerns between core and border points, removes outliers and noise and performs well regardless of the sequence of visiting points. Fourier shape descriptors [51, 52] on the selected clusters are then used to reduce the number of vertexes needed for the definition of a specific shape, normalizing it. This makes it easier for pattern recognition and classification since it reduces the total number of input points needed and thus provides a specific number of input points every time.

### B. Classification

Having defined the regions of interest and extracted the geometric features of the facial expression, we then need to classify the data to identify facial expressions. It is apparent from the previous sections that the geometric shapes of the regions of interest provide irregular shapes and thus a non-linear classification scheme should be adopted to improve classification accuracy. In this case we make use of a back propagating neural network for irregular shape classification [53]. The neural network classifies shapes according to their convex hulls. Since forehead and nose wrinkles are widely irregular shapes and cannot be described by a convex hull, we indicate their presence with 1 for yes's and 0 for no's. Shape of the eyes and the mouth are described by the respective significant Fourier descriptors. We have 7 different classes (happiness, sadness, fear, anger, disgust, surprise and neutral) therefore in the output layer we need 4 nodes.

To demonstrate the effectiveness of the methodology, an experiment with practical examples was performed. The Cohn-Kanade+ database contains 593 sequences from 123 different people. This subject pool was divided according to emotions depicted among different faces and part of it was used for training, part for verification. The number of available samples is shown in Table II.

TABLE II: NUMBER OF SAMPLES FOR EACH EMOTION, MANUALLY CHECKED FROM PEAK FRAMES

Emotion	Available Samples
Happiness	69
Sadness	28
Fear	25
Anger	45
Disgust	59
Surprise	83
Neutral	123

### C. Problem Formulation

Let us assume in the following that some features have been extracted from a human face analysis and included in a feature vector say  $\mathbf{f}_i$ , where index  $i$  corresponds to the  $i$ th human individual. Let us also assume that  $p$  humans' emotions are recognized by the proposed ERI computing architecture and thus each vector  $\mathbf{f}_i$  is classified to one of the  $p$  available classes  $\omega_j$ ,  $j=1,2,\dots,p$ , i.e., emotions. As it can be shown in several publications, six emotions are basic; the happiness, sadness, fear, anger, surprise, and disgust [31]. It is not in doubt that these basic emotions are key points of reference and can be described as 'archetypal emotions', which reflects the fact that they are undeniably the obvious examples of an emotion state. New studies attempt to define a set of terms covering a wider range of emotions without becoming unmanageable. Let us denote in the following as  $\mathbf{y}(\mathbf{f}_i)$ , a  $p$ -dimensional vector, the elements of which  $p_{e_k}^{(i)}$  contains the degrees of coherence of vector  $\mathbf{f}_i$  to one of the  $k$ th emotional state among the  $p$  available,

$$\mathbf{y}(\mathbf{f}_i) = [p_{e_1}^{(i)} p_{e_2}^{(i)} \dots p_{e_p}^{(i)}]^T \quad (1)$$

Let us assume in the following, for simplicity purposes and clarification reasons, a two-class classification problem, where the two classes  $\omega_1$ ,  $\omega_2$  refer, for example, to two emotional states, such as happiness and sadness. Extension to multiple emotional states can be performed in a similar way. In this case the model output is scalar, say  $y(\mathbf{f}_i)$  instead of a vector (output values close to one (1) correspond to the first emotional state, while values close to zero to the second emotional state.).

Then, using concepts derived from functional analysis, it can be proved that any continuous non-linear function can be expressed as a parametric relation of known functional components  $\Phi_l(\cdot)$  within any degree of accuracy. Then, the model output can be written as

$$y(\mathbf{f}_i) \approx \sum_{l=1}^L v_l \cdot \Phi_l \left( \sum_{k=1}^P w_{k,l}(f_{i,k}) \right) \quad (2)$$

where  $f_{i,k}$  refers to the  $k$ th element of feature vector  $\mathbf{f}_i$ , the  $v_l$  and  $w_{k,l}$  to the model parameters (coefficients), and  $Q$  the size of the feature vector  $\mathbf{f}_i$ . Variable  $L$  corresponds to the approximation order of the model. In particular, the larger the number of  $L$  is the higher the approximation accuracy becomes at the expense of the model complexity. The most familiar class of functional components  $\Phi_l(\cdot)$  is the sigmoid functions, which are equal to

$$\Phi_l(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

The model parameters  $v_l$  and  $w_{k,l}$  are estimated using a set of representative samples and then by applying a training algorithm. Let us denote as  $S = \{(\mathbf{f}_1, d_1), \dots, (\mathbf{f}_N, d_N)\}$  this set, where  $N$  is the total number of samples and  $d_i$  the desired output vector for the  $i$ th human individual. The values of  $d_i$  is equal to 1 for the first emotional state and 0 for the second emotional state. Let us also denote as  $\mathbf{W}$  a vector, which contains all parameters  $v_l$  and  $w_{k,l}$  of (2), so  $\mathbf{w} = [\dots v_l \dots w_{kl} \dots]^T$ .

#### D. Adaptable Emotionally Rich Architecture

Although the general expression of archetypal emotions is the same, we nevertheless have to take into account what would be considered as noise: More specifically, since we are targeting an elderly user-base, the most important noise inducing factor would be the appearance of wrinkles. Moreover, we should take into account also micro-changes in the expressions, owed to individual quirks. Thus, direct application of the model of (2), in which the model parameters are considered constant, would not work well, since the model cannot be adapted to the slight changes in the way that humans use to express their emotions. To overcome this difficulty, we equip the proposed architecture with an adaptable mechanism, as seen in **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.** Figure 3, which can update the system response with respect to the current way that the individual use to express their emotions.

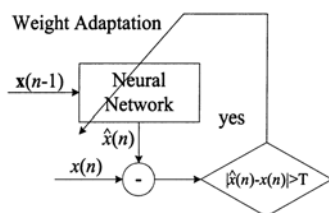


Fig. 3. Weight adaptation mechanism

In this way, a new set of parameters, say  $\mathbf{w}_a$ , should be created for each emotional state, which are capable of updating the response of the ERI architecture to the current actual way that the users express their emotions. To perform the adaptation, a new set of representative samples, say  $S_c = \{(\mathbf{f}'_1, d'_1), \dots, (\mathbf{f}'_M, d'_M)\}$  should be created, which has a similar form to the set  $S$ . Then, the new parameters  $\mathbf{w}_a$  are estimated by minimizing the following error criterion,

$$\mathbf{w}_a = \arg \min_{\mathbf{w}} E = \frac{1}{2} \sum_{i=1}^N (y_{\mathbf{w}}(\mathbf{f}_i) - d_i)^2 = \frac{1}{2} \sum_{i=1}^N E_i \quad \text{for all data in } S \quad (4)$$

and

$$y_{\mathbf{w}_a}(\mathbf{f}_i) = d'_i \quad i = 1, \dots, M, \quad \text{for all data in } S_c \quad (5)$$

In (4, 5) we added the dependence of the model output  $y_{\mathbf{w}}(\mathbf{f}_i)$  on the model parameters  $\mathbf{w}$ . Equations (4,5) indicate that the new model parameters are estimated so that the current data, as expressed by the set  $S_c$ , are satisfied as much as possible, while simultaneously a minimal degradation of the previous obtained knowledge is derived

#### E. Optimal Estimation of the Model Parameters

In this section, we describe a novel, fast and reliable algorithm for estimating the new model parameters  $\mathbf{w}_a$ . Initially, we assume that a small perturbation of the model parameters before the adaptation is enough to achieve good classification performance. Then,

$$\mathbf{w}_a = \mathbf{w} + \Delta \mathbf{w} \quad (6)$$

where  $\Delta \mathbf{w}$  is a small incremental vector. This assumption leads to an analytical and tractable solution for estimating  $\mathbf{w}_a$ , since it permits linearization of functional components  $\Phi_l(\cdot)$  of (2, 5) using a first order Taylor series expansion.

It can be shown [54] that linearization of (5) with respect to the weight increments  $\Delta \mathbf{w}$  is equivalent to a set of linear equations

$$\mathbf{c} = \mathbf{A} \cdot \Delta \mathbf{w} \quad (7)$$

where vector  $\mathbf{c}$  and matrix  $\mathbf{A}$  are appropriately expressed in terms of the previous model parameters  $\mathbf{w}$ . In particular,

$$\mathbf{c} = [y_{\mathbf{w}_a}(\mathbf{f}_1) \Delta y_{\mathbf{w}_a}(\mathbf{f}_M)]^T - [y_{\mathbf{w}}(\mathbf{f}_1) \Delta y_{\mathbf{w}}(\mathbf{f}_M)]^T \quad (8)$$

indicates the difference between model outputs after and before the adaptation for all input vectors in  $S_c$ .

The size of vector  $\mathbf{c}$  is smaller than the number of unknown weights  $\Delta\mathbf{w}$ , since in general a small number  $M$ , of current data are available. Thus, many solutions exist for (7), since the number of unknowns is much greater than the respective number of equations. Uniqueness, however, is imposed by an additional requirement, which takes into consideration the previous network knowledge. Among all possible solutions that satisfy (7), the one which causes a minimal degradation of the previous model knowledge is selected as the most appropriate. This is expressed by equation (4).

It can be shown [54] that (4) takes the form of

$$E = \frac{1}{2}(\Delta\mathbf{w})^T \cdot \mathbf{K}^T \cdot \mathbf{K} \cdot \Delta\mathbf{w} \quad (9)$$

where the elements of matrix  $\mathbf{K}$  are expressed in terms of the previous network weights  $\mathbf{w}$  and the data in the set  $S$ . Thus, the problem results in the minimization of (9) subject to constraints of (7).

$$\text{Minimize } E = \frac{1}{2}(\Delta\mathbf{w})^T \cdot \mathbf{K}^T \cdot \mathbf{K} \cdot \Delta\mathbf{w} \quad (10)$$

$$\text{Subject to } \mathbf{c} = \mathbf{A} \cdot \Delta\mathbf{w} \quad (11)$$

The error function defined by (10) is convex since it is of squared type, while the constraints of (11) are linear equalities. Thus, the solution should lie on the hyper-surface defined by (11) and simultaneously minimize the error function given in (10). The gradient projection method is used in this paper to solve this problem. The gradient projection method starts from a feasible point and moves in a direction, which decreases  $E$  and simultaneously satisfies the constraints; a point is called feasible, if it satisfies all constraints. The weights are adapted as follows:

$$\Delta\mathbf{w}(n+1) = \Delta\mathbf{w}(n) + \mu(n)\mathbf{h}(n) \quad (12)$$

where  $n$  is the iteration index and  $\mu(n)$  is a scalar that determines the rate of convergence. Using the methodology of **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.** we can estimate vector  $\mathbf{h}(n)$  as

$$\mathbf{h}(n) = -\mathbf{P}\nabla E = -\mathbf{Q}\Delta\mathbf{w} \quad (13)$$

$$\text{with } \mathbf{P} = \mathbf{I} - \mathbf{A}^T(\mathbf{A} \mathbf{A}^T)^{-1}\mathbf{A} \quad \text{and } \mathbf{Q} = \mathbf{P}\mathbf{K}^T\mathbf{K} \quad (14)$$

using  $\nabla E$  computed from (13). The computational complexity required to independently update each network weight is proportional to the number of network weights.

#### F. Computational Complexity

The computational complexity of the adaptation algorithm, in contradiction to the generally long training periods of the initial model parameters, is very small. In particular, the adaptation process is separated in two phases; the initialization phase where matrix  $\mathbf{Q}$ , which projects the negative gradient of function  $E$  onto the surface defined by the constraints of (11),

is estimated, and the iteration phase where the weight increments  $\Delta\mathbf{w}$  are updated. The iteration phase involves a simple multiplication of matrix  $\mathbf{Q}$ , of size  $N_w \times N_w$ , by vector  $\Delta\mathbf{w}$ , of size  $N_w \times 1$ , where  $N_w$  denotes the number of model parameters. For a typical value of  $N_w$  equal to 500, the product  $\mathbf{Q} \cdot \Delta\mathbf{w}(n)$  in (13) requires few msec to be executed (~10 ms on a core i7 PC). The number of iterations  $n$  that the gradient projection method requires to derive the optimal solution is in general small (8-12 iterations are usually sufficient to obtain a weight increment close to the optimal one). Thus, the computational cost of the iteration phase is about half a sec.

In the initialization phase, matrix  $\mathbf{Q}$  is estimated as a product of matrix  $\mathbf{P}$  and matrices  $\mathbf{K}^T \cdot \mathbf{K}$ . The main computational load of  $\mathbf{P}$  is the estimation of the inverse matrix  $(\mathbf{A} \cdot \mathbf{A}^T)^{-1}$ . However, the size of  $\mathbf{A} \cdot \mathbf{A}^T$  is equal to the number of constraints, which is usually very small since it refers to the number of representative data of the current environment. Consequently, the computational complexity for matrix  $(\mathbf{A} \cdot \mathbf{A}^T)^{-1}$  is significantly low (in the order of few msec). Finally, computation of  $\mathbf{P} \cdot \mathbf{K}^T \cdot \mathbf{K}$  requires about few seconds for a typical network size of 500 weights. Thus, the total cost of retraining is small (of order of few seconds), allowing the efficient use of the proposed scheme to real-life interactive multimedia systems.

After the classification, results of the process can be stored in xml format as seen in Figure 4 **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.** For the evaluation of the classification we used ten-fold cross validation. Available data are divided into ten non-overlapping sets of roughly equal size, and nine of them are used for training, one for validation. The classification accuracy of the BPN ranges from ~47% to 98,4%. Lowest success rates were noted in the expressions of anger and fear, where facial expressions are of similar geometric attributes, while sadness, happiness and surprises were met with the highest success rates.

```

<?xml version="1.0" encoding="UTF-8"?>
<PatientStatus PatientId="123456"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="PatientStatus.xsd">
  <PatientId>123456</PatientId>
  <TimeStamp>2015-05-30T09:30:10.5</TimeStamp>
  <Detected>
    <emotion>sadness</emotion>
    <attribute>negative</attribute>
  </Detected>
</PatientStatus>

```

Fig. 4. Detection report

#### V. CONCLUSION

The implementation of the proposed methodology was specifically designed to be flexible and scalable so as to fit into telemedicine based AAL systems. Providing a ready XSD

schema, the results of the methodology can be easily integrated into a reconfigurable rule based risk assessment module, as part of a telemedicine system [30]. The results from this methodology are promising and useful, especially in the aforementioned context, and can be used to evaluate –and even predict– the overall health status of subjects suffering from neurodegenerative diseases. Depressive symptoms are widely present in people with the onset of MCI and can even further aid at the deterioration of their mental and cognitive status [55-58]. It is thus crucial for the attending clinician to be able to monitor not only the physical but also the psychological status of subjects.

The proposed algorithm, although having some shortcomings, such as lack of real-time properties and strong illumination sensitivity, is a first step towards the full emotional status estimation and prediction of adverse events in the framework of an internet-based AAL system for people with MCI or other neurodegenerative disease. These disadvantages though, are not deemed deteriorating for the overall efficacy of the methodology: The methodology is able to grab image frames from a web-camera at regular intervals (several times a minute), which can be adjusted as deemed necessary and provide the relevant results to a rule-based implementation of risk detection [30]. Moreover, since this methodology is aimed at home based AAL systems, it is considered a given that subjects are lit by a home's natural lighting which in everyday life activities are not usually extreme so they don't induce significant errors in the process.

Finally, the proposed methodology is ready to be incorporated into a metadata-aware biometric architecture as the one proposed in [44] specifically targeted in ambient assisted living systems. Future improved implementations of the methodology aim to encompass real time capabilities, especially based on Constrained Local Models [59-60] and deep learning networks for better hierarchical multimodal classification of the subject's emotional status [61-63].

#### Acknowledgment

The authors would like to acknowledge the ALADDIN project, funded under the first call of AAL Joint Programme, AAL-2008-1, as well as the people of the Distributed Knowledge and Media Systems lab of the Electrical and Computer Engineering School of the National Technical University of Athens. Finally, this work would not have been possible without the work of Cohn-Kanade and their database of facial expressions, the images of which provided the basis for our research.

#### REFERENCES

- [1] Eurostat, "Population structure and ageing," 2015, available at: [http://ec.europa.eu/eurostat/statistics-explained/index.php/Population\\_structure\\_and\\_ageing](http://ec.europa.eu/eurostat/statistics-explained/index.php/Population_structure_and_ageing)
- [2] A. L. Busse, G. Gil, J. M. Santarém, W. J. Filho, "Physical activity and cognition in the elderly: A review", *Dement. Neuropsychol.*, Vol. 3, No. 3, pp. 204–208, 2009
- [3] A. Di Carlo, M. Baldereschi, L. Amaducci, S. Maggi, F. Grigoletto, G. Scarlato, D. Inzitari, "Cognitive impairment without dementia in older people: prevalence, vascular risk factors, impact on disability. the italian longitudinal study on aging", *J. Am. Geriatr. Soc.*, Vol. 48, No. 7, pp. 775–782, 2000
- [4] D. Laurin, R. Verreault, J. Lindsay, K. MacPherson, K. Rockwood, "Physical activity and risk of cognitive impairment and dementia in elderly persons", *Arch. Neurol.*, Vol. 58, No. 3, pp. 498–504, 2001
- [5] A. Bradford, M. E. Kunik, P. Schulz, S. P. Williams, H. Singh, "Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors", *Alzheimer Dis. Assoc. Disord.*, Vol. 23, No. 4, pp. 306–314, 2009
- [6] T. Kitwood, *Dementia Reconsidered: the Person Comes First*, Open University Press, 1999
- [7] M. Haritou, Y. Glickman, A. Androulidakis, S. Xefteris, A. Anastasiou, A. Baboshin, S. Cuno, and D. Koutsouris, "A technology platform for a novel home care delivery service to patients with dementia", *J. Med. Imaging Heal. Informatics*, Vol. 2, No. 1, pp. 49–55, 2012
- [8] M. Torkamani, L. McDonald, I. S. Aguayo, C. Kanios, M.-N. Katsanou, L. Madeley, P. D. Limousin, A. J. Lees, M. Haritou, M. Jahanshahi, "A randomized controlled pilot study to evaluate a technology platform for the assisted living of people with dementia and their carers", *J. Alzheimers Dis.*, Vol. 41, No. 2, pp. 515–523, 2014
- [9] P. Rashidi, A. Mihailidis, "A survey on ambient-assisted living tools for older adults", *IEEE J. Biomed. Heal. Informatics*, Vol. 17, No. 3, pp. 579–590, 2013
- [10] C. Frantzidis, C. Bratsas, C. L. Papadelis, E. Konstantinidis, C. Pappas, P. D. Bamidis, "Toward emotion aware computing: an integrated approach using multichannel neurophysiological recordings and affective visual stimuli", *Inf. Technol. Biomed. IEEE Trans.*, Vol. 14, No. 3, pp. 589–597, 2010
- [11] F. Bellotti, B. Kapralos, K. Lee, P. Moreno-Ger, R. Berta, "Assessment in and of serious games: an overview," *Adv. Human-Computer Interact.*, Vol. 2013, pp. 1–11, 2013
- [12] E. I. Konstantinidis, A. S. Billis, C. Mouzakidis, V. Zilidou, P. E. Antoniou, P. D. Bamidis, "Design, implementation and wide pilot deployment of FitForAll: an easy to use exergaming platform improving physical fitness and life quality of senior citizens", *IEEE J. Biomed. Heal. Informatics*, Vol. 20, No. 1, pp. 189 – 200, 2014
- [13] A. S. Billis, E. I. Papageorgiou, C. Frantzidis, E. I. Konstantinidis, P. D. Bamidis, "Towards a hierarchically-structured decision support tool for improving seniors' independent living", 6th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA'13), pp. 1–4, Rhodes, Greece, May 29-31, 2013
- [14] A. Artikis, P. D. Bamidis, A. Billis, C. Bratsas, C. Frantzidis, V. Karkaletsis, M. Klados, E. Konstantinidis, S. Konstantopoulos, D. Kosmopoulos, H. Papadopoulos, S. Perantonis, S. Petridis, C. S. Spyropoulos, "Supporting tele-health and AI-based clinical decision making with sensor data fusion and semantic interpretation: The USEFIL case study", International Workshop on Artificial Intelligence and NetMedicine, Montpellier, France, August 27, 2012
- [15] S. Xefteris, M. Haritou, K. Tserpes, A. Serretti, J. R. Llopart, R. Calati, T. Varvarigou, "Analysis of requirements and specifications for a monitoring system to support the self-management of dementia patients at home", 3rd International Conference on Pervasive Technologies Related to Assistive Environments (PETRA'10), Samos, Greece, June 23-25, 2010
- [16] S. Xefteris, A. Baboshin, K. Tserpes, A. Androulidakis, Y. Glickman, T. Varvarigou, M. Haritou, F. D'Andria, "Enabling risk assessment and analysis by event detection in dementia patients using a reconfigurable rule set", 4th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA'11), Crete, Greece, May 25-27, 2011
- [17] P. D. Bamidis, P. Fissler, S. G. Papageorgiou, V. Zilidou, E. I. Konstantinidis, A. S. Billis, E. Romanopoulou, M. Karagianni, I. Bearatis, A. Tsapanou, G. Tsilikopoulou, E. Grigoriadou, A. Ladas, A. Kyriolidou, A. Tsolaki, C. Frantzidis, E. Sidiropoulos, A. Siountas, S. Matsi, J. Papatriantafyllou, E. Margiotti, A. Nika, W. Schlee, T. Elbert, M. Tsolaki, A. B. Vivas, I.-T. Kolassa, "Gains in cognition through combined cognitive and physical training: the role of training dosage and severity of neurocognitive disorder", *Front. Aging Neurosci.*, Vol. 7, Art. No. 152, pp. 1-15, 2015.



- [18] C. Styliadis, P. Kartsidis, E. Paraskevopoulos, A. A. Ioannides, P. D. Bamidis, "Neuroplastic effects of combined computerized physical and cognitive training in elderly individuals at risk for dementia: an eLORETA controlled study on resting states", *Neural Plast.*, Vol. 2015, Article ID 172192, 2015
- [19] A. S. Billis, N. Katzouris, A. Artikis, P. D. Bamidis, "Clinical decision support for active and healthy ageing: an intelligent monitoring approach of daily living activities", 17th Portuguese Conference on Artificial Intelligence (EPIA 2015), Coimbra, Portugal, September 8-11, 2015
- [20] D. N. Monekosso, P. Remagnino, "Behavior analysis for assisted living", *IEEE Trans. Autom. Sci. Eng.*, Vol. 7, No. 4, pp. 879–886, 2010
- [21] R. A. Stucki, P. Urwyler, L. Rampa, R. Mürli, U. P. Mosimann, T. Nef, "A web-based non-intrusive ambient system to measure and classify activities of daily living", *J. Med. Internet Res.*, Vol. 16, No. 7, Article ID e175, 2014
- [22] A. S. Billis, E. I. Papageorgiou, C. A. Frantzidis, M. S. Tsatali, A. C. Tsolaki, P. D. Bamidis, "A decision-support framework for promoting independent living and ageing well", *IEEE J. Biomed. Heal. Informatics*, Vol. 19, No. 1, pp. 199–209, 2015
- [23] A. K. Jain, A. Ross, S. Prabhakar, "An introduction to biometric identification", *IEEE Trans. Circuits Syst. Video Technol. Spec. Issue Image-and Video-based Biometrics*, Vol. 14, No. 1, pp. 4-20, 2004
- [24] E. M. Tapia, S. S. Intille, K. Larson, *Activity recognition in the home using simple and ubiquitous sensors*. Springer, 2004
- [25] M. Mozer, "Lessons from an adaptive house", in *Smart Environments: Technologies, Protocols, and Applications*, Hoboken, NJ, pp. 273–294, 2004
- [26] A. Steinhage, C. Lauterbach, "Monitoring movement behavior by means of a large area proximity sensor array in the floor", 2nd Workshop on Behaviour Monitoring and Interpretation (BMI'08), Kaiserslautern, Germany, September 23, 2008
- [27] M. Muhlenbrock, O. Brdiczka, D. Snowdon, J. -L. Meunier, "Learning to detect user activity and availability from a variety of sensor data", 2nd IEEE International Conference on Pervasive Computing and Communications (PerCom'04), pp. 13-22, Florida, USA, March 14-17, 2004
- [28] C. Darwin, P. Ekman, P. Prodger, *The expression of the emotions in man and animals*. Oxford University Press, USA, 1998
- [29] B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey," *Pattern Recognit.*, Vol. 36, No. 1, pp. 259–275, 2003
- [30] S. Xefteris, A. Baboshin, K. Tserpes, A. Androulidakis, Y. Glickman, T. Varvarigou, M. Haritou, F. D'Andria, "Enabling risk assessment and analysis by event detection in dementia patients using a reconfigurable rule set", 4th International Conference on Pervasive Technologies Related to Assistive Environments, (PETRA'11), Crete, Greece, May 25-27, 2011
- [31] P. Ekman, W. V Friesen, *Unmasking the face: A guide to recognizing emotions from facial clues*. Ishk, 2003
- [32] P. Ekman, R. J. Davidson, *The nature of emotion: Fundamental questions*. Oxford University Press, 1994
- [33] C. Darwin, *The Expression of the Emotions in Man and Animals*. 1872.
- [34] P. Ekman, W. V Friesen, P. Ellsworth, *Emotion in the human face: Guidelines for research and an integration of findings*, Oxford Pergamon Press, 1972
- [35] P. Ekman, W. V Friesen, M. O'Sullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider, R. Krause, W. A. LeCompte, T. Pitcairn, P. E. Ricci-Bitti, and others, "Universals and cultural differences in the judgments of facial expressions of emotion", *J. Pers. Soc. Psychol.*, Vol. 53, No. 4, pp. 712-717, 1987
- [36] P. Ekman, "An argument for basic emotions", *Cogn. Emot.*, Vol. 6, No. 3–4, pp. 169–200, 1992
- [37] P. Ekman, *Darwin and Facial Expression: A century of Research*. New York: Academic Press, 1973
- [38] C. Izzard, *The Face of Emotion*. New York: Appleton-Century-Crofts, 1971.
- [39] P. Ekman, W. V. Friesen, P. Ellsworth, *Emotion in the human face. Guidelines for research and an integration of findings*, Pergamon Press, USA, 2013
- [40] H. Hwang, D. Matsumoto, "Evidence for the Universality of Facial Expressions of Emotion", in *Understanding Facial Expressions in Communication*, Springer India, pp. 41–56, 2015,
- [41] I. E. Eiblfeldt, *Ethology, the Biology of Behavior*. 1970
- [42] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression", 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 94–101, 2010
- [43] T. Kanade, J. F. Cohn, Y. Tian, "Comprehensive database for facial expression analysis", 4<sup>th</sup> IEEE International Conference on Automatic Face and Gesture Recognition, pp. 46–53, 2000
- [44] V. Andronikou, S. Xefteris, T. Varvarigou, "A novel, algorithm metadata-aware architecture for biometric systems", IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS), pp. 1–6, 2012
- [45] M. Jones, P. Viola, "Fast multi-view face detection", *Mitsubishi Electr. Res. Lab TR-20003-96*, Vol. 3, 2003, available at <http://www.merl.com/publications/docs/TR2003-96.pdf>
- [46] A. Okabe, B. Boots, K. Sugihara, S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*, Vol. 501. John Wiley & Sons, 2009
- [47] F. P. Preparata, M. Shamos, *Computational geometry: an introduction*. Springer Science & Business Media, 2012.
- [48] B. Aronov, M. Dulieu, F. Hurtado, "Witness (Delaunay) graphs", *Comput. Geom.*, Vol. 44, No. 6–7, pp. 329–344, 2011
- [49] J. Gan, Y. Tao, "DBSCAN Revisited: Mis-Claim, Un-Fixability, and Approximation", *ACM SIGMOD International Conference on Management of Data*, pp. 519–530, 2015
- [50] R. Xu, D. Wunsch, and others, "Survey of clustering algorithms", *IEEE Transactions on Neural Networks*, Vol. 16, No. 3, pp. 645–678, 2005
- [51] I. Kunttu, L. Lepistö, A. Visa, "Efficient Fourier shape descriptor for industrial defect images using wavelets", *Opt. Eng.*, Vol. 44, No. 8, p. 80503-1 - 80503-3, 2005
- [52] F. J. van Rensburg, J. Treurnicht, C. J. Fourie, "The Use of Fourier Descriptors for Object Recognition in Robotic Assembly", 5th CIRP ISME Intelligent Computation in Manufacturing Engineering (CIRP ISME), Italy, July 25-28, 2006
- [53] S.-W. Lin, S.-Y. Chou, S.-C. Chen, "Irregular shapes classification by back-propagation neural networks", *Int. J. Adv. Manuf. Technol.*, Vol. 34, No. 11–12, pp. 1164–1172, 2007
- [54] A. D. Doulamis, N. D. Doulamis, S. D. Kollias, "An adaptable neural-network model for recursive nonlinear traffic prediction and modeling of MPEG video sources", *IEEE Transactions on Neural Networks*, Vol. 14, No. 1, pp. 150–166, 2003
- [55] P. Delivery, "Home Precautions For A Family Member With Dementia", *Adolesc. Heal.*, Vol. 56, No. 5, pp. 502–507, 2015
- [56] E. H. Rickenbach, K. L. Condeelis, W. E. Haley, "Daily Stressors and Emotional Reactivity in Individuals With Mild Cognitive Impairment and Cognitively Healthy Controls", *Psychol. Aging*, Vol. 30, No. 2, pp. 420-431, 2015
- [57] S. Mussele, K. Bekelaar, N. Le Bastard, Y. Vermeiren, J. Saerens, N. Somers, P. Mariën, J. Goeman, P. P. De Deyn, S. Engelborghs, "Prevalence and associated behavioral symptoms of depression in mild cognitive impairment and dementia due to Alzheimer's disease", *Int. J. Geriatr. Psychiatry*, Vol. 28, No. 9, pp. 947–958, 2013
- [58] E. Richard, C. Reitz, L. H. Honig, N. Schupf, M. X. Tang, J. J. Manly, R. Mayeux, D. Devanand, J. A. Luchsinger, "Late-life depression, mild cognitive impairment, and dementia", *JAMA Neurol.*, Vol. 70, No. 3, pp. 383–389, 2013
- [59] D. Cristinacce, T. F. Cootes, "Feature detection and tracking with constrained local models", *BMVC*, Vol. 3, pp. 929-938, Edinburgh, September 4-7, 2006

- [60] T. Baltrušaitis, P. Robinson, L. -P. Morency, "3D constrained local model for rigid and non-rigid facial tracking", 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2610–2617, 2012
- [61] Y. Bengio, I. J. Goodfellow, A. Courville, Deep Learning, MIT Press, 2016 (in preparation)
- [62] H. Zhang, X. Xiao, O. Hasegawa, "A Load-Balancing Self-Organizing Incremental Neural Network", IEEE Trans. Neural Networks Learn. Syst., Vol. 25, No. 6, pp. 1096–1105, 2014
- [63] H.-I. Suk, S.-W. Lee, D. Shen, "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis", Neuroimage, Vol. 101, pp. 569–582, 2014