EMOTHAW: A Novel Database for Emotional State Recognition from Handwriting and Drawing

Laurence Likforman-Sulem, *Senior Member, IEEE,* Anna Esposito, Marcos Faundez-Zanuy, Stéphan Clémençon, and Gennaro Cordasco

***Abstract*—The detection of negative emotions through daily activities such as writing and drawing is useful for promoting well-being. The spread of human-machine interfaces such as tablets makes the collection of handwriting and drawing samples easier. In this context, we present a first publicly available database which relates emotional states to handwriting and drawing, that we call EMOTHAW. This database includes samples of 129 participants whose emotional states, namely anxiety, depression and stress, are assessed by the Depression Anxiety Stress Scales (DASS) questionnaire. Seven tasks are recorded through a digitizing tablet: pentagons and house drawing, words copied in handprint, circles and clock drawing, and one sentence copied in cursive writing. Records consist in pen positions, on-paper and in-air, time stamp, pressure, pen azimuth and altitude. We report our analysis on this database. From collected data, we first compute measurements related to timing and ductus. We compute separate measurements according to the position of the writing device: on paper or in-air. We analyse and classify this set of measurements (referred to as features) using a random forest approach. This latter is a machine learning method [1], based on an ensemble of decision trees, which includes a feature ranking process. We use this ranking process to identify the features which best reveal a targeted emotional state. We then build random forest classifiers associated to each emotional state. We provide accuracy, sensitivity and specificity evaluation measures, obtained from cross-validation experiments. Our results show that anxiety and stress recognition perform better than depression recognition.**

***Index Terms*—Affective database, Random Forests, Handwriting, Emotional state, DASS scales, Depression, Anxiety, Stress.**

I. Introduction

HEALTH care mainly relies on the early detection of illnesses, and clinical tests have been developed to diagnose diseases and follow their evolution [2], [3]. Among tests, those based on human activity (speech, handwriting, body movements) have the advantage of being non invasive and are valuable tools for complementing laboratory analyses and clinical examination. In particular, simple pen and paper tests can detect cognitive impairments through handwriting: lack of legibility, jagging and perseveration of letters are well-known effects of Parkinson (PD) and Alzheimer (AD) diseases [4].

The importance of detecting early signs of illnesses can be extended to the detection of negative emotions since emotions such as depression, anxiety and stress influence health. Depression is a complex and heterogeneous mood disorder that is expressed by behavioral disinterest and sad feelings and it may cause serious social, occupational and cognitive impairments [5] [6]. As defined by Eysenck et al. [7](pp.336) *"Anxiety is an aversive emotional and motivational state occurring in threatening circumstances"*. It affects cognition and reduce the individual's effectiveness and efficiency in performing cognitive tasks. The most comprehensive definition of stress is *"a [negative] emotional experience accompanied by predictable biochemical, physiological and behavioral changes"* cited in [8]. The causes of stress are extremely diverse, ranging from difficulties to handle everyday experiences and changes to traumatic events such as surviving to a natural disaster. Depression, anxiety and stress, are natural responses to changes and challenges of everyday life. However, when persisting over a long time period, they can produce serious illnesses such as Major Depression Disorders (MDD), Generalized Anxiety Disorder (GAD) and Chronic Stress [6] [9] [10].

Mundt et al. [11] and Yang [12] have exploited the speech signal to detect depression. They report significant differences in the acoustic biomarker values (such as F0 and F0-derived measures) of clinical/non-clinical subjects. However, as pointed out in Esposito & Esposito [13] *"[voice acoustic measures] appear to be affected at various degrees by many sources of variability that causes distortions and modifications in the original signal, thus modifying the acoustic features useful for its [automatic] recognition"*. In handwriting acquisition however, the signal is easily acquired without error propagation.

In the present study we propose to detect negative emotions such as depression, anxiety and stress through handwriting, a human daily activity. Our approach consists in collecting an individual's handwriting through a computerized platform and predict his/her emotional state through a machine-learning approach.

Machine learning approaches such as Support Vector Machines, Neural Networks, Bayesian Networks have been used successfully in related domains such as affective computing and personality computing [14] where measurements (referred to as features) are extracted from behavioral signals such as face expressions, speech or body gestures, and fed to a classifier. The outcome of the classification in affective computing is one simple emotion among the set of basic ones: happiness, surprise, sadness, anger, fear, disgust. The goal is to improve human-machine interfaces by adequately reacting from users' inputs. For personality computing, the outcomes are personality traits: openness, agreeableness, conscientious-

L. Likforman and S. Clémençon are at Télécom ParisTech and Université Paris-Saclay, Paris, France.

A. Esposito and G. Cordasco are at the Second University of Naples, in Caserta, Italy.

M. Faundez-Zanuy is at Escola Superior Politecnica, TecnoCampus Mataro-Maresme, Spain.

Manuscript received

ness, extraversion, neuroticism. Machine-learning approaches require labeled databases for training the classifier [15]. The personality trait labels are most often assessed through the Big-Five questionnaire [16]. Similarly to BigFive and personality, negative emotions such as depression, anxiety and stress can be scored through the DASS (Depression-Anxiety-Stress Scales) scales [17]. These scales are now well assessed [18] and also use a self-reported questionnaire. Administering the DASS questionnaire to each participant, we have built "EMOTHAW" (EMOTion recognition from HAndWriting and drAWing) a first publicly available database[1] relating emotional states to handwriting. Other handwriting databases devoted to biometry [19] or recognition [20] have been developed, and do not include emotional labels.

In addition to the database, we present a non-parametric classifier based on the random forests machine learning approach [1]. We extract features related to timing and ductus from the collected data. We compute separate features according to the position of the writing device: on paper or in-air. Random forest training provides ranking of the input features according to their importance. We develop an analysis of the extracted features based on these rankings. From this analysis, we deduce which tasks and which features better characterize a targeted emotional state. We also build random forest classifiers based on the extracted features, and provide recognition results for the three emotional states covered by the DASS scales.

Our paper is organized as follows. In Section II, we present previous computerized studies devoted to handwriting analysis for diagnosis purposes. We formulate the assumptions of our analysis in Section III. Section IV describes the process of data collection as well as the cohort recruitment. The set of features extracted on raw handwriting data is described in Section V-A. Section V-B presents our feature ranking and classification of emotions, based on random forests. In Section VI, we apply this analysis to the collected data and provide recognition results.

II. Handwriting analysis for diagnosis purposes

Handwriting collected on paper has long been used as a means for authentication, cognitive impairment detection and personality trait assessment. Indeed, collecting handwriting is not invasive, simple and cheap and requires little expertise from the operator. Various pen and paper tests have been developed and used to complement laboratory data, physician examination or face-to-face interviews.

From handwritten samples, it has been observed that Alzheimer and Parkinson patients [21] deteriorate character shapes, add or omit characters within words and have less fluent movements. In the biometry domain, off-line signatures have long been used as a means of authentication as well as writing samples collected with or without the knowledge of the writer [22]. Handwriting analysis have also lead to graphology-related applications such as detection of lies. In this context, Tang [23] assumes that when honest people lie, their writing is modified due to the associated cognitive stress. Writing may be less fluent and margins may be modified at places where the lie is expressed. Well-assessed tests in the medical domain are:

- The clock drawing test (CDT) where participants are required to draw a clock, including all digits, and setting the hands to 11:50. Patients with Alzheimer disease may not space digits evenly and sometimes make mistakes when setting the clock hands. A maximum score of 15 points is awarded for shape and spatial arrangement of the clock numbers, and hands.
- The MMSE test (Mini Mental State Examination) [24] is a 30-point test that evaluates cognitive functions related to registration, attention and calculation, recall, language, ability to follow simple commands and orientation. Among other tasks, the subject is asked to write a sentence of his own, draw interlinking pentagons and memorizing a sequence of 3 words.
- The HTP (House-Tree-Person) test, designed by J. Buck [25] is a clinical test where participants are asked to draw a house, a tree and a person. The drawing tasks are accompanied by questions to infer personality. This test is sometimes extended to evaluate brain damage.

Due to the development of scanners and tablets, the following studies aim at converting pen and paper tests to computerized platforms. The collected data within computerized platforms are on-line in contrast to off-line [22]. The advantage of on-line data is that pen positions are recorded along with non-directly visible features such as time stamp, pressure and stylus inclination. The offline case lacks the temporal dimension since only scanned images of the ink trace are available. This is why the analysis of on-line signatures is more efficient than the analysis of off-line signatures [26].

The CDT test has been computerized into the ClockMe system [27] which automatically evaluates the CDT score, following the paper-based protocol. No difference can be observed between the paper and tablet based scores, the tablet offering supplemental data awaiting to be processed and interpreted.

Recently, the 'strokes against stroke' system has been proposed to detect brain stroke risk from stroke measurement achieved on a tablet from stimuli displayed on a screen [28]. A previous version of this system [29] was based on signatures and triangle drawings.

The CompPET system [30] (Computerized Penmanship Evaluation Tool) is a computerized platform which collects on-line handwritten data, including pressure, in air and on tablet points. From data collected with ComPET, changes in handwriting have been measured according to age, and various major health disorders such as depression (MDD: Major Disorder Depression). Recently, this platform has been used to automatically detect lies in handwriting [31], using speed and size of handwriting units.

The previous studies deal with handwriting analysis but none of them deal with emotion recognition. Our present contribution provides a database and an analysis of that database relating emotional states to handwriting/drawing tasks and pen movements.

[1] https://sites.google.com/site/becogsys/emothaw

III. Setup for our analysis

We assume that writing and drawing are related to behavior, and are influenced by one individual's emotional state. We focus here on negative emotional states such as depression, anxiety and stress, which are seen as distinct states. However, their clinical symptoms largely overlap [17]. The DASS scales (Depression-Anxiety-Stress Scales) proposed by Lovibond and Lovibond (1995) [17] have been designed to provide a maximum discrimination between these negative emotional states. The items included in each scale have been chosen in order to produce three orthogonal axes from a factor analysis: DASS-Depression focuses on items related to low motivation and self esteem, DASS-anxiety to fear and perceived panic, and DASS-stress to tension and irritability.

Following this classification, we assume that each emotion can be separately recognized through measurements on handwriting and drawing samples, and that each emotional state is likely to be related to specific writing/drawing tasks and measurements.

Starting with little knowledge on which state influences which measurement, we propose a machine learning approach, namely random forests [1], which automatically ranks the measurements associated to each recognition task. The random forest approach also provides an estimate on the participants' emotional state from these measurements.

The starting point of this study is the collection of a database from 129 participants, which we refer to as the EMOTHAW database. On-line data are collected through a computerized platform. This database is in itself a useful tool due to the lack of publicly available labeled data for this domain. The writing/drawing tasks from which the measurements are extracted are well assessed tasks used in medical diagnosis or in scoring handwriting/drawing proficiency (house drawing, text copying, ...). The emotional state of participants is assessed by the DASS scale. Measurements related to the kinematic and ductus of the writing and likely to change under the mentioned emotional states, are then extracted from the database. These measurements (referred to as *features* in the machine learning terminology) are fed to classifiers in order to automatically recognize the above emotional states.

IV. Data collection

We describe here the data collection process of the EMOTHAW database: raw measurements provided by the digitizing tablet, tasks completed by the participants, and ground truth assessment.

A. Computerized platform

Data have been registered thanks to an INTUOS WACOM series 4 digitizing tablet and a special writing device named Intuos Inkpen. This device provides high spatial and pressure accuracies and can be considered as a state-of-the-art tablet. Participants were required to write on a sheet of paper (DIN A4 normal paper) laid on the tablet. Figure 1 shows the sample acquired from one participant. No time restrictions were provided to participants. While the digital signal is visualized on the screen, it is also visible on the paper (due to the inkpen), and the participant normally looks at the paper. There is a human supervisor sitting next to the participant, controlling a specifically designed acquisition software on a computer linked to the tablet. When the supervisor pushes the "end" button after a given task, a new svc file is created for the next task. But the software does not register any in-air point before the pen touches the paper (first on-paper stroke), and after the last on-paper stroke. Svc files thus start and end with on-paper strokes and there is no bias related to transition between one task to the other. The same procedure was used in the MCYT and BIOSECURID databases [19]. Data samples

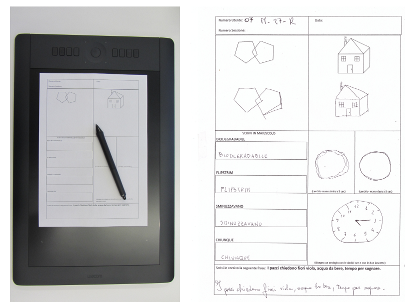


Fig. 1. Left: Acquisition tablet with ready to fill A4 sheet and writing device. Right: sheet with the whole set of tasks filled by one participant.

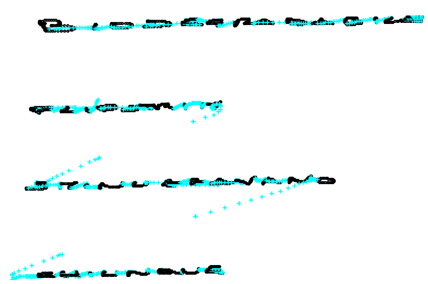


Fig. 2. Task 3 sample with on-paper points (black dots, pen status down=1) and in-air points (blue +, pen status up=0). Additional in-air movements (writing device too high above the tablet) are not registered.

are acquired in real time. The resulting files are svc files, svc being the file extension provided by Wacom. Svc files are ASCII files that can be opened with WORD, NOTEPAD and other standard editor applications.

The following information is captured:

(i) Position in x-axis.
(ii) Position in y-axis.
(iii) Time stamp
(iv) Pen status (up=0 or down =1)
(v) Azimuth angle of the pen with respect to the tablet (see Fig. 4).
(vi) Altitude angle of the pen with respect to the tablet (see Fig. 4).
(vii) Pressure applied by the pen.

Using this set of dynamic data, further information such as velocity features (acceleration, velocity), instantaneous trajectory angle, instantaneous displacement, time features and ductus-based features (see Section V-A) can be inferred. In addition, the system has the nice property to capture in-air movements, which are lost using the on-paper ink. The in-air information has proven to be as important as the on-surface information [32] [33]. However, when the writing device is too far from the tablet (at a distance greater than 1 cm), in-air points are not registered. Subjects were not informed of this distance issue from tablet to stylus. Figure 2 shows pen-down and pen-up points acquired at 100 points per second, corresponding to handwriting in capital letters. When the speed of writing is low, registered points are close to each other and strokes seem darker than when the speed is high. In our study, we have used for each sample a vector of seven parameters, which consists of the previous list of measurements provided by the WACOM pen tablet. Figure 3 shows a short extract of an svc file. The extract includes both on-paper points (pen status equal to 1) and in-air points (pen status equal to 0). It can be noted that timestamp values always increase and that pressure (column 7) is equal to zero for in-air points. For instance, the point drawn at time stamp 17606786 is an in-air point at x position 50621 and y position 33860. Its pen status and pressure are both equal to 0, and azimuth and altitude values are equal to 1900 and 540, respectively. Azimuth and altitude values can be normalized using maximum values as normalization factors. The normalized values are in degrees: azimuth=$1900 \times 360/4095 = 167°$ and altitude=$540 \times 90/1023 = 47.5°$.

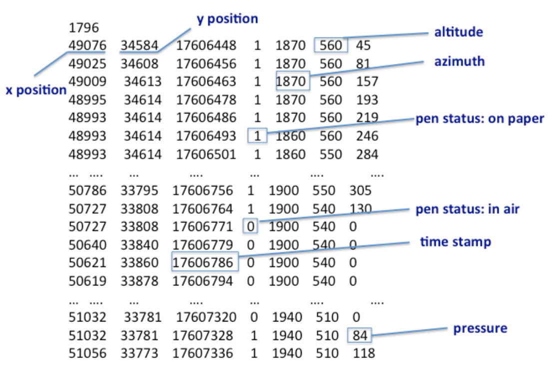


Fig. 3. Extract of an svc file corresponding to the pentagon drawing task. The file includes 1796 points, each one having 7 measurements (x position, y position, time stamp, pen status, azimuth, altitude and pressure).

The tasks acquired by the tablet are the following (see Fig. 5):

I. Copy of a two-pentagon drawing
II. Copy of a house drawing
III. Writing of four Italian words in capital letters (BIODEGRADABILE (biodegradable), FLIPSTRIM (flipstrim), SMINUZZAVANO (to crumble), CHIUNQUE (anyone))
IV. Loops with left hand
V. Loops with right hand
VI. Clock drawing
VII. Writing of the following phonetically complete Italian sentence in cursive letters (I pazzi chiedono fiori viola, acqua da bere, tempo per sognare : *Crazy people are seeking for purple flowers, drinking water and dreaming time*).

The two-pentagon drawing is part of the minimental test (MMSE test), while drawing a clock is part of the Clock Drawing Test (CDT). House drawing is part of the HTE (House-Tree-Person) projective test of personality (see Section II). Copying text has been also used for assessing writing abilities [34]. Thus two of the proposed tasks consist in copying a sequence of words, according to two types of script: cursive and handprint.

The handprinted words (Biodegradabile,...) have been selected based on our previous Spanish database BIOSECURID [19]. They are neutral words that do not have positive or negative connotations. The cursive sentence (I pazzi....) is a phonetically complete word sequence which was reviewed by PhD experts at the Psychology department of the Seconda Università di Napoli. No priming effect [23] related to the meaning of the sentence was considered here, though a possible negative effect could be related to the word ”pazzi” which means ”fool”. But in italian the word ”pazzi” is commonly used for kidding, poetry and in the daily conversational language as synonym of ”special persons” particularly when they were asking for ”purple flowers, drinking water, and dreaming time” (lyrics of a famous song from F. de Gregori entitled ”I matti”). Thus, in Italian the connotation of this sentence is smoother than it could be in another language.

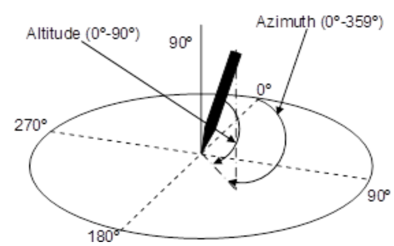


Fig. 4. Azimuth and inclination angles of the pen with respect to the plane of the graphic card.

B. Participants and Emotion Scoring

For the data collection, 129 subjects aged between 21 and 32 years, (mean age 24.8 years, standard deviation SD= 2.4 years) were recruited, all Master and BS students at the Seconda Università di Napoli, Department of Psychology, located in Caserta, Italy. It can be noted that in medical and psychological studies a database of 129 subjects with a specific background and age is a noteworthy research accomplishment. In these fields, acquiring a large database is not simple and even case studies are of great interest. There were 71 Female and 58 Male participants. The range of years has been limited in order to reduce the inter-subject variability of the experiment. Otherwise it would be difficult to attribute differences to years, or psychological conditions. There can be differences in handwriting features because of ages also and this would have been interfered with differences due to the emotional states. Each participant first filled in and signed a consent form providing her/his general demographic information. Measuring emotion is a well known difficult problem [35]. We opted to measure emotional/affective states through a questionnaire as it is commonly done in affective computing studies. Participants' emotional state was determined by using the Italian version of the Depression Anxiety Stress Scales (I-DASS-42). The construct validity of DASS-42 was assessed by Henry and Crawford in 2003 [18] for the English version and by Severino (http://www2.psy.unsw.edu.au/dass/Italian/Severino.htm) for the Italian one. The I-DASS-42 is a 42-item questionnaire including three self-report scales, each containing 14 items, designed to measure depression, anxiety and stress. These affective states differ from transient emotional states that may appear and vanish rapidly with good/bad weather, good/bad news etc. For each of these negative emotional states, subjects are asked to rate on a 4-point severity/frequency scales (from 0 to 3) the extent to which they have experienced each item[2] listed in the questionnaire *over the past week*. The severity-rating index assigned to each subject is made according to Table I. The distributions of DASS scores in the EMOTHAW database are shown in Fig. 6. The consistency and temporal stability of the DASS scales have been assessed through several studies (f.i. [36]). In the present study, each participant performed the writing/drawing tasks as soon as he/she completed the DASS questionnaire. We are thus confident that the participant performed the tasks under the measured states.

[2]Such items include : "I found myself getting upset rather easily", and "I had a feeling of shakiness (eg, legs going to give way)."

	Depression	Anxiety	Stress
Normal	0-9	0-7	0-14
Mild	10-13	8 - 9	15 - 18
Moderate	14 - 20	10 - 14	19 - 25
Severe	21 - 27	15 - 19	26 - 33
Extremely Severe	28+	20+	34 +

TABLE I
DASS SCORE RANGE ACCORDING TO EMOTIONAL STATE LEVEL.

Scoring details are described in the document found at http://www.aasw.asn.au/document/item/2794. It can be noted that DASS questionnaire is intended to screen normal adolescents and adults for detecting psychological mood disorders: severe disorders should be assessed through clinical examination.

Cross-tables in Fig. 7 show that depression, anxiety and stress can be observed separately or in conjunction. The scores have been dichotomized as explained in the following (see Section V). From the matrices in the 2nd row of Fig. 7, we observe that

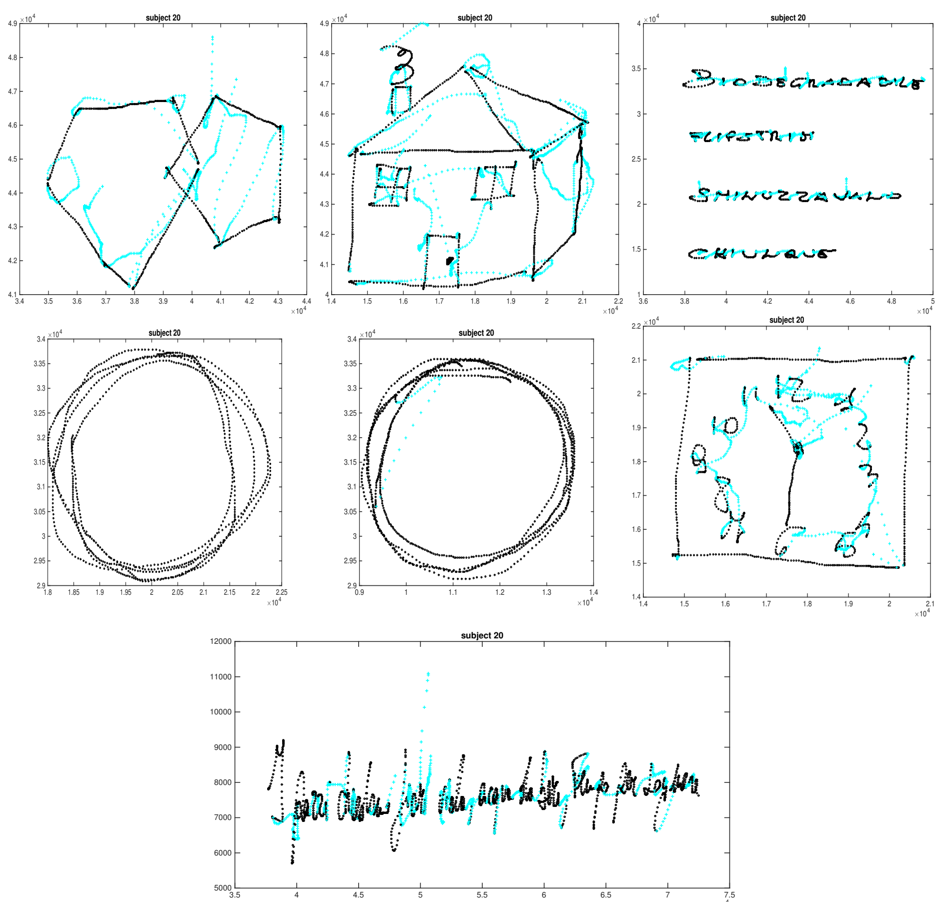


Fig. 5. Writing and drawing samples collected from all tasks: pentagons and house drawings, handprint writing, loops (left hand and right hand), clock drawing and cursive writing. Pen-down and pen-up data points are in black and blue, respectively.

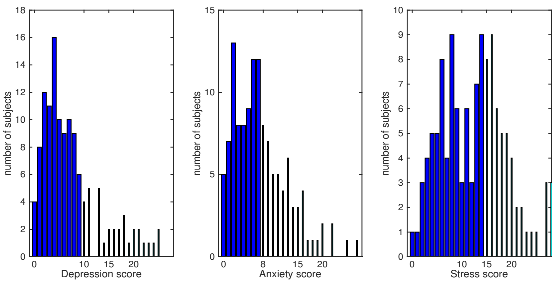


Fig. 6. DASS score distributions in the EMOTHAW database. Normal scores are in dark blue.

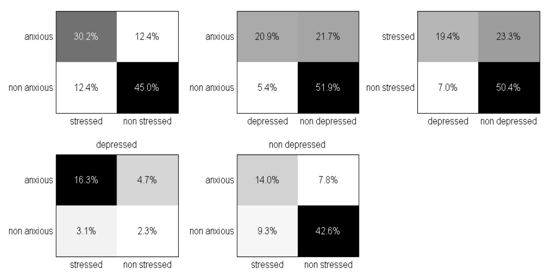


Fig. 7. Cross-tables showing the percentage of co-occurrence of emotional states in the EMOTHAW database.

for about 20% of the participants (19.4%=2.3%+7.8%+9.3%), a single negative state is observed (f.i. such as anxious/non stressed/non depressed). For about the same percentage of participants (21.8%=3.1%+4.7%+14.0%), two negative emotional states are observed in conjunction. However, due to the scales' construct, each of these emotional states can be predicted separately. A Pearson's χ^2 test conducted on the anxious/stressed, stressed/depressed and anxious/depressed cross-tables, shows that the qualitative variables anxiety and stress, as well as stress and depression and depression and anxiety, are linked (p-values below 0.01). The strongest link is between anxiety and stress. Then the link is stronger between anxiety and depression, than between stress and depression.

V. Emotional State Recognition

In the following experiments, DASS scores have been dichotomized in order to predict the state of an individual as a two-class problem: anxious or not anxious, stressed or not, or depressed or not. This is motivated by the fact that approaches such as ordinal regression would need more sample points to build the regression function. The non-anxious state refers to normal anxiety level (score less than 7), while the anxious state refers to levels ranging from mild to high (scores up to 20). The stressed emotional state is scored more than 14 and the depression state, more than 9 (see Table I).

A. *Proposed features and tasks*

Feature extraction is the first step of an automatic recognition system. It consists in representing the raw data signal in a more concise and accurate way. Indeed, the registered signal includes raw features such as x-y positions of the writing device, as well as absolute time and pressure at each of x-y position (see Section IV-A). From these raw features, accurate features must be extracted. Such features should be independent of absolute pen position and time, as well as being efficient in solving the targeted state recognition task.
The definition of features for recognizing emotional states from handwriting or drawing is not straightforward. To our knowledge, there is no assessed features for such tasks. Thus our approach consists in proposing a number of features, extracted from the tasks described above (see Section IV), then ranking and analyzing their importance through a random forest approach (see Section V-B).

The first features we propose are timing-based features which have proven to be efficient for assessing writing proficiency of pupils and elderly people copying texts [33] [37]. Both time spent on paper and in air were found efficient. We extend such timing-based features to cursive writing and drawing tasks. We have also observed that ductus (ductus is the way how strokes are drawn including stroke order, direction and speed) presents great variations across data. Ductus may be related to individual differences but it may also be related to the emotional state of an individual. We have thus added a feature related to ductus. For each task, we thus propose the following timing-based and ductus-based features:

- F_1: time spent in-air while completing the task
- F_2: time spent on-paper while completing the task
- F_3: time to complete the whole task
- F_4: number of on-paper strokes while completing the task

In this context an on-paper (resp. in-air) stroke corresponds to consecutive drawing points achieved without lifting (resp. dropping) the pen. Tracking pen status changes in svc files allows us to segment data into on-paper and in-air strokes (see column 4 in Fig. 3). Starting and ending with on-paper strokes, a drawing or writing task consists of a sequence of strokes denoted by $\{s^{(1)}, s^{(2)}, \ldots s^{(2K+1)}\}$. Since on-paper and in-air strokes alternate, $\{s^{(2k+1)}\}_{k=0,\ldots K}$ is the set of on-paper strokes and $\{s^{(2k)}\}_{k=1,\ldots K}$ the set of in-air strokes. The number of on-paper strokes is thus $F_4 = K + 1$. Given a stroke $s^{(i)}$, we denote $\{s_1^{(i)}, s_2^{(i)}, \ldots, s_{|s^{(i)}|}^{(i)}\}$ the sequence of its compound points. $| s^{(i)} |$ is the number of points in $s^{(i)}$ registered at time stamps $t_1^{(i)}, t_2^{(i)}, \ldots, t_{|s^{(i)}|}^{(i)}$ (see column 3 in Fig. 3). The duration of stroke $s^{(i)}$ is thus $d_i = t_{|s^{(i)}|}^{(i)} - t_1^{(i)}$ and the values for features F_1 to F_2 are obtained by summing over the appropriate set of strokes:

$$F_1 = \sum_{k=1}^{K} d_{2k} \quad F_2 = \sum_{k=0}^{K} d_{2k+1} \quad F3 = t_{|s^{(2K+1)}|}^{(2K+1)} - t_1^{(1)}$$

From the seven registered tasks, the two loop drawings have been discarded since such tasks include almost no pen-up movement. Thus we extract the above features on the five remaining tasks, cumulating 20 measurements for each subject. In the following, we analyze these measurements by a machine-learning approach in order to identify those that are important for a targeted emotional state.

B. *Feature analysis with random forests*

We have proposed above (see Section V-A) a set of 20 measurements, referred to as features, to extract from five handwriting/drawing tasks. Our objective is now to find which tasks and which features are relevant in order to recognize an emotional state.

Feature ranking consists in scoring features either in isolation or in the context of other features. It is an essential component of *filter* and *embedded* feature selection approaches which aim at removing redundant and irrelevant features [38], [39]. In embedded approaches, the ranking is performed within the training of a classifier (decision trees, SVMs). Features are seen in the context of other features, taking into account their interaction. In contrast, filter-based approaches do not rely on classifiers, and rank features from the data only, as individual variables. Impurity measures (such as Gini index) and weights provided by the Relief algorithm are examples of popular feature scoring measures [40].
Feature selection is often necessary in high-dimensional feature spaces. Since our feature space is relatively small, we are interested in understanding the impact of our variables on the proposed classification tasks. The random forest approach [1] satisfies this objective by providing several importance measures useful for ranking features, along with training.

Training a random forest consists in building $ntree$ decision trees which are combined at decision level. Each tree is built from a subset of the training data and from a subset of the features or variables. A random process occurs for the selection of data used for training the t^{th} tree and another one for the selection of features at each split node of this tree. At each node, the candidate set of features used for splitting the node into child nodes, is restricted to $mtry$ features. This restricted set is obtained by randomly sampling the original feature set. The training data, not selected by the first random process, are used for testing the efficiency of this tree. Such data are the so-called OOB (Out-Of-Bag) data [41]. Since each data point may be an OOB for several trees in the forest, all decisions for this OOB can be recorded and accumulated. The final decision for an OOB is provided by a majority vote. OOB decisions permit to compute the so-called OOB error rate which is a valuable estimation of the forest accuracy.

Another advantage of the random forest approach is that variable relative importance measures are provided through the training process. There are four importance measures for a given feature f_i:

- Measure 1: is the amount of decrease of the OOB error rate. It is computed as follows. The value of feature f_i is randomly changed for each out-of-bag data point. The error rate is thus modified compared to using the original data. If the error rate significantly increases, this means that feature f_i is important. Conversely if it decreases, the importance value is negative and this means that feature f_i is not reliable.
- Measure 2: is the average margin decrease. The so-called margin of an OOB point is the difference between the proportion of trees in the forest which classify correctly this OOB point minus the proportion of trees which misclassify this point. When the value of feature f_i is randomly changed, the margin is modified (it typically decreases). The decreases are averaged across OOB points to provide measure 2.
- Measure 3: is the normalized difference between the number of margins which have decreased and the number of margins which have increased when applying the previous process (f_i randomly changed for OOBs).
- Measure 4: is the mean decrease of the Gini criterion. The Gini criterion is an impurity measure used to choose the best split variable at each tree node. For a given feature f_i, one can sum up the gains in impurity obtained when using this feature in the forest trees. This sum is then normalized by the number of trees in the forest to provide measure 4 for f_i.

Features are then ranked according to each importance measure and each targeted emotional state. The automatic ranking process is the following. For a given emotional state, the ranks of each feature according to each importance measure are summed up. The lowest the sum, the best the feature. However, since building a forest includes a part of randomness, the ranking of features may slightly vary from one forest to another. Thus, an ensemble of T forests is built and the ranks are cumulated from each forest of the ensemble.

Random forests are mostly used for classification purposes. Thus, we also train three random forest classifiers from the previous set of measurements in order to recognize depression, anxiety and stress.

VI. Experiments

We apply our analysis based on random forests (see Section V) to the recognition of anxiety, depression and stress.

A. *Feature analysis*

Our analysis consists in ranking the proposed timing and ductus-based features (see Section V-A). The process starts by building a random forest ensemble from the set of extracted features (20 features: 5 tasks, 4 features per task). For each emotion, a forest is built according to Section V-B, using $ntree$ trees, and at each tree node, $mtry$ variables are considered for a split. In this study, forests are built with a large number of trees ($ntree$=100). Since we have $nfeat = 20$ features, $mtry = \sqrt{nfeat} \sim 5$ is a popular tuning for the number of variables considered at each tree node.
Random forest training provides both a forest of $ntree$ trees and four importance measures for each feature (variable) considered. Features are ranked according to the automatic ranking process described above, using an ensemble of $T = 50$ forests (see Section V-B). The ten top-ranked features provided by each random forest model are shown in Table II. These top-10 features belong to specific writing or drawing tasks. These tasks differ according to the emotion under consideration (see Fig. 8). Figure 8 is a task-centered representation of Table II which shows which tasks are privileged for the recognition of a given emotion.

We observe from Fig. 8 that depression is expressed through drawing tasks only (clock, pentagons, house). In contrast,

Random Forest Model	Features
Depression	in-air duration (clock), on-paper duration (clock)
	total duration (clock), in-air duration (pentagons),
	total duration (pentagons), in-air duration (house)
	on-paper duration (pentagons), total duration (house)
	on-paper duration (house), number of pen-down strokes (house)
Anxiety	on-paper duration (clock), total duration (cursive)
	on paper duration (pentagons), in air duration (cursive)
	in air duration (house), number of pen-down strokes (house)
	in air duration (pentagons), total duration (handprint)
	total duration (house), total duration (pentagons)
Stress	on paper duration (clock), in air duration (cursive)
	total duration (clock), on-paper duration (pentagons)
	in-air duration (clock) , number of pen-down strokes (cursive)
	on-paper duration (house), on-paper duration (handprint)
	total duration (cursive), in air duration (house)

TABLE II
TOP RANKED FEATURES ACCORDING TO EACH MODEL.

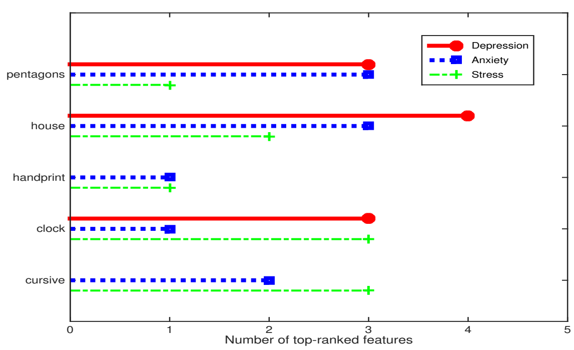


Fig. 8. Number of top-ranked features according to writing/drawing tasks and emotional state.

anxiety and stress use both handwriting (cursive, handprint) and drawing cues.

Table II indicates that both features related to in-air and on-paper motions are useful for characterizing emotions: all models use three in-air cues (in air duration) picked from the proposed handwriting/drawing tasks. This shows the importance of in-air cues which cannot be observed in the ink trace. In contrast, the proposed ductus-related cues (number of pen-down strokes) are used only once for each model.

B. Emotion recognition

To provide emotion recognition results, we conduct $C = 10$ repetitions of leave-one-out cross-validation experiments with $K = 129$ folds. Each leave-one-out cross-validation experiment consists in building a forest from all but one data point (the $K - 1$ folds), then using this forest to test the remaining data point (the Kth fold). This process is repeated K times. From C repetitions of this process, C cross validation results are collected and the mean accuracies are shown for each model in Table III, along with confidence intervals. However, the overall accuracy of a constant-decision classifier based on the highest class prior could be high, especially if classes are unbalanced. ROC-based evaluation measures such as true positive rates (TPR) and false positive rates (FPR) are thus necessary to complement accuracy values. We provide sensitivity and specificity values in Table III which are equal to TPR and 1-FPR respectively. The ROC for all systems is shown in Fig. 9 and the diagonal corresponds to random guesses. Systems are ranked in this space, according to their distance to the top-left corner: anxiety and stress recognition (with drawing, or drawing and writing features) perform better than depression recognition. Although the depression recognition systems have the best overall accuracies, they are the worst in ROC space. It can be noted that random forest systems using drawing features, or writing and drawing features, are all above the diagonal. All random forest experiments have been conducted using the R language, and the *random forest* and *ipred* packages [42]. Fig. 10 shows clock drawings from

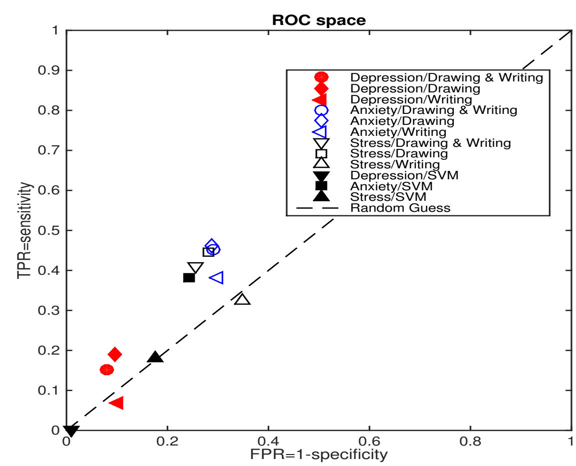
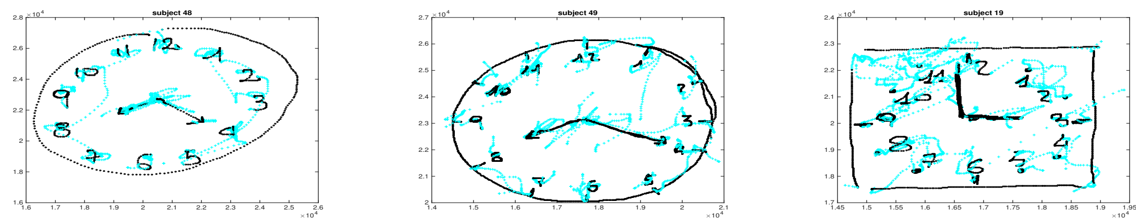


Fig. 9. Emotion recognition systems plotted in the ROC space.

several participants: participant 48 is normal (non-depressed, non stressed), participant 49 is depressed, and participant 19, stressed. The duration on paper is much longer for the depressed (36,8 seconds) and the stressed (36,3 seconds) participants than for the normal one (8,7 seconds). This is also true for the duration in-air: longer for the depressed (39 seconds) and stressed (46 seconds) participants than for the normal one (22 seconds). This can be related to studies on speech processing where it has been shown that speech pauses

Random Forest Model	Feature type	Accuracy (in %)	Sensitivity (in %)	Specificity (in %)
Depression	Writing	67.8 [66.5 69.2]	6.7 [4.3 9.2]	89.7 [88.1 93.6]
Depression	Drawing	71.6 [70.9 72.3]	19.1 [17.6 20.6]	90.4 [89.4 94.4]
Depression	Drawing & Writing	71.2 [69.9 72.5]	12.6 [10.6 14.6]	92.2 [90.7 93.6]
Anxiety	Writing	56.3 [55.2 57.5]	38.2 [35.2 41.2]	69.9 [68.5 71.2]
Anxiety	Drawing	60.5 [59 62]	46.2 [44.1 48.2]	71.2 [68.5 73.9]
Anxiety	Drawing & Writing	60 [58.2 61.8]	45 [42.9 47.2]	71 [68.5 73.9]
Stress	Writing	51.2 [49.9 52.5]	32.3 [30.6 34]	65.3 [63.1 67.4]
Stress	Drawing	60.1 [59.2 61.1]	44.5 [42.4 46.7]	71.7 [71 72.5]
Stress	Drawing & Writing	60.2 [59.6 60.8]	41 [39.9 42.8]	74.5 [73.5 75.4]

TABLE III
MEAN RECOGNITION ACCURACY, SENSITIVITY AND SPECIFICITY VALUES, FOR EACH MODEL PROVIDED BY $C = 10$ REPETITIONS OF LEAVE-ONE-OUT CROSS-VALIDATION EXPERIMENTS, ALONG WITH 95 % CONFIDENCE INTERVALS. SEPARATE RESULTS ARE PROVIDED FOR WRITING AND DRAWING FEATURE SUBSETS.



(a) (b) (c)

Fig. 10. Sample clock drawings. a) non stressed/non depressed participant 48: DASS-depression 3, DASS-anxiety 6, DASS-stress 6. b) depressed participant 49: DASS-depression 11, DASS-anxiety 2, DASS-stress 5. c) stressed participant 19: DASS-depression 9, DASS anxiety 4, DASS-stress 15.

are longer for depressed participants [43]. From the blue points of Fig. 10-b and c, we observe that in-air movements of stressed and depressed participants differ: they are erratic for the stressed participant, smooth for the depressed one. Thus a long in air duration is indeed the sign of a disorder but more features are necessary to distinguish between a stressed participant and a depressed one. This shows the interest of using several features for emotion recognition.

Table II shows that top-ranked features for depression recognition are all extracted from drawing tasks. Furthermore we consider the subset of pure drawing features (extracted from tasks 1, 2 and 6), and compare it to the subset of pure writing features (extracted from tasks 3 and 7). Recognition accuracies for these subsets are given in Table III. We observe that using the subset of writing features performs worse than using the subset of drawing features. This can be explained by the fact that there are less features in the writing set (8 features) than in the drawing set (12 features). Interestingly we also observe that the performance using the drawing feature subset is quite similar to that of the set of all features (drawing and writing). Using drawing features only seems to be sufficient to recognize the targeted emotions. However the writing feature subset has an intrinsic predictive power, though reduced compared to the drawing feature subset (the decrease in accuracy is about 4%). This experiment shows that the writing feature subset is redundant with the drawing feature subset when both sets are combined. Finally, we conduct an experiment where the random forest classifier is replaced by the SVM classifier (Support Vector Machine). We set up the parameters as follows: radial kernel, $C = 1$ and $\gamma = 0.5$. Mean cross validation accuracies are shown in Table IV, along with sensitivity and specificity values. Similar to random forest classifiers, Fig. 9 shows that the SVM classifier for anxiety recognition performs best. The SVM classifier for depression recognition performs poorly since it is under the diagonal of the ROC space. The SVM classifier for stress recognition performs better since it is slightly above the diagonal.

SVM	Depression	Anxiety	Stress
Accuracy (in %)	72.8	59.7	55
Sensitivity (in %)	0	38.2	18.2
Specificity (in %)	98.9	75.7	82.4

TABLE IV
MEAN RECOGNITION ACCURACY, SENSITIVITY AND SPECIFICITY VALUES WITH SVM CLASSIFIER PROVIDED BY LEAVE-ONE-OUT CROSS-VALIDATION EXPERIMENTS.

VII. CONCLUSION

We have presented EMOTHAW, a novel database devoted to a new task: the recognition of emotional states from handwriting and drawing. The affective database includes samples from 129 participants whose emotional states, namely depression, anxiety and stress, are assessed by the DASS scales. Seven tasks are recorded through a digitizing tablet: pentagons and house drawings, words copied in handprint, circles and clock drawings, and one sentence copied in cursive writing. Records consist in pen positions, time stamp, pressure, pen azimuth and altitude. These samples are registered as on-line data including both on paper and in air points. Each participant also completed the self-reported DASS questionnaire from which DASS scores were computed.

Recognition experiments have been conducted using the random forest machine learning approach. In this context, we have computed features related to timing both from in air and on paper movements, and ductus (number of strokes). We have developed an analysis based on random forests in order to highlight important features associated to a targeted emotional state. Our results show that both in-air and on-paper features are important for predicting emotional states. However each state is characterized by its own set of relevant in-air and on-paper features and tasks.

In particular, depression is characterized by cues extracted from drawing tasks while anxiety and stress use both writing and drawing cues. We have shown however that the subset of writing cues may be redundant with the subset of drawing cues when they are both combined in a single set. Comparing all systems in the ROC space, depression was found more difficult to recognize than anxiety and stress. The present research involved a specific class of subjects (students). This was done on purpose in order to reduce the variability in the collected data. However, this is also a limitation since it does not allow for a generalization of the identified features and the obtained results to a larger sample of individuals. In the future we plan to analyze other population groups such as kids, mature and elder people, looking for commonalities and differences in order to delineate the proposed procedure and identify a general methodology for detecting emotional disorders from handwriting/drawing data. We consider that the first step must be a study in each group because a lot of indicators are related to age (for instance, it is believed that happiness in function or years has a U shape with a minimum around 45 years old). Nevertheless, the strengths of the proposed work are:

- Proposing, for the first time, the exploitation of handwriting/drawing data for the detection of emotional disorders
- Proposing and experimental design for the data collection.
- Proposing an automatic procedure for the ranking of handwriting/drawing features useful for the task at hand; these features are easy to collect and do not need manual measurements.
- Providing a handwriting/drawing database to the scientific community for further testing.
- Opening new investigation approaches to help psychologists and clinicians to assess emotional disorders.

In this first study we have considered depression, anxiety and stress recognition as independent two-class problems. However such emotions can be observed in conjunction. We plan to investigate the field of multitask learning [44] where several classifiers are learnt in conjunction, taking into account their interactions.

The set of extracted features considered in this work is restricted to timing and ductus-based ones. It could be extended to other ductus-based features such as for example speed. The stress state which induces modifications in the muscle tonus may benefit from new features linked to pressure [45]. Higher level features such as fluidity of movements could be added by an analysis of speed changes. Model-based features such as the parameters provided by the kinematic theory of rapid human movements [46] could also be extracted. At last, the features we have dealt with in this paper are global features that do not reveal possible local changes in the tracing due to an emotional state. In order to improve classification results, future work will consist of integrating local features. Such features (f.i. velocity, curvature, straightness index) can be directly extracted on sample points or on specific areas or shapes (corners, straight lines). Extracting these areas and shapes could be done by extending mathematical morphology tools to on-line data.

We hope that the database release will provide possibilities for new and extended features and recognition approaches.

Acknowledgments

This work has been supported by FEDER and Ministerio de ciencia e Innovacion, TEC2012-38630-C04-03, and by the industrial chair "Machine Learning for Big Data" from Telecom ParisTech.

References

[1] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[2] L. Bender, "A visual motor gestalt test and its clinical use," *Research Monographs, American Orthopsychiatric Association*, vol. 3, 1938.

[3] M. Shin, S. Park, S. R. Park, S. Seol, and J. S. Kwon, "Clinical and empirical applications of the Rey-Osterrieth complex figure test," *Nature protocols*, vol. 1, 2006.

[4] J. Neils-Strunjas, J. Shuren, D. Roeltgen, and C. Brown, "Perseverative writing errors in a patient with Alzheimer's disease," *Brain and Language*, vol. 63, pp. 303–320, 1998.

[5] K. Scherer, "What are emotions ? and how can they be measured ?" *Social Science Information*, pp. 695–729, 2005.

[6] L. Rehm, *Depression*, ser. Advances in Psychotherapy. Hogrefe Publishing GmbH, 2010.

[7] M. W. Eysenck, N. Derakshan, R. Santos, and M. G. Calvo, "Anxiety and cognitive performance: attentional control theory." *Emotion*, vol. 7, no. 2, p. 336, 2007.

[8] A. Baum, "Stress, intrusive imagery, and chronic distress," *Health Psychology*, vol. 6, pp. 653–675, 1990.

[9] S. Cohen, D. Janicki-Deverts, and G. E. Miller, "Psychological stress and disease," *Jama*, vol. 298, pp. 1685–1687, 2007.

[10] K. Mogg and B. P. Bradley, "A cognitive-motivational analysis of anxiety," *Behaviour research and therapy*, vol. 36, no. 9, pp. 809–848, 1998.

[11] J. C. Mundt, A. P. Vogel, D. E. Feltner, and W. R. Lenderking, "Vocal acoustic biomarkers of depression severity and treatment response," *Biological psychiatry*, vol. 72, no. 7, pp. 580–587, 2012.

[12] Y. Yang, C. Fairbairn, and J. F. Cohn, "Detecting depression severity from vocal prosody," *IEEE Transactions on Affective Computing*, vol. 4, no. 2, pp. 142–150, 2013.

[13] A. Esposito and A. M. Esposito, "On the recognition of emotional vocal expressions: motivations for a holistic approach," *Cognitive Processing*, vol. 13, no. 2, pp. 541–550, 2012.

[14] N. Sebe, I. Cohen, T. Gevers, and T. S. Huang, "Multimodal approaches for emotion recognition: a survey," in *Electronic Imaging 2005*. International Society for Optics and Photonics, 2005, pp. 56–67.

[15] N. Fourati and C. Pelachaud, "Collection and characterization of emotional body behaviors," in *International Workshop on Movement and Computing, MOCO '14*, 2014.

[16] L. R. Goldberg, "Personality structure: emergence of the five factors model," *Annual Review of Psychology*, vol. 41, pp. 417–440, 1990.

[17] S. Lovibond and P. Lovibond, *Manual for the Depression, Anxiety, Stress Scales*, ser. Psychology Foundation monograph. School of Psychology, University of New South Wales, 1995.

[18] J. R. Crawford and J. D. Henry, "The depression anxiety stress scales (DASS): Normative data and latent structure in a large non-clinical sample," *British Journal of Clinical Psychology*, vol. 42, no. 2, pp. 111–131, 2003.

[19] J. Fierrez, J. Galbally, J. Ortega-Garcia, M. Freire, F. Alonso-Fernandez, D. Ramos, D. Toledano, J. Gonzalez-Rodriguez, J. Siguenza, J. Garrido-Salas, E. Anguiano, G. Gonzalez-de Rivera, R. Ribalda, M. Faundez-Zanuy, J. Ortega, V. Cardeñoso Payo, A. Viloria, C. Vivaracho, Q. Moro, J. Igarza, J. Sanchez, I. Hernaez, C. Orrite-Uruñuela, F. Martinez-Contreras, and J. Gracia-Roche, "Biosecurid: a multimodal biometric database," *Pattern Analysis and Applications*, vol. 13, no. 2, pp. 235–246, 2010.

[20] I. Guyon, L. Schomaker, R. Plamondon, R. Liberman, and S. Janet, "Unipen project of online data exchange and recognizer benchmarks," in *International Conference on Pattern Recognition ICPR*, 1994, pp. 29–33.

[21] M. Faúndez-Zanuy, A. Hussain, J. Mekyska, E. Sesa-Nogueras, E. Monte-Moreno, A. Esposito, M. Chetouani, J. Garre-Olmo, A. Abel, Z. Smékal, and K. López-de-Ipiña, "Biometric applications related to human beings: There is life beyond security," *Cognitive Computation*, vol. 5, no. 1, pp. 136–151, 2013.

[22] R. Plamondon and S. Srihari, "Online and offline handwriting recognition: A comprehensive survey," *IEEE PAMI*, vol. 22, no. 1, 2000.

[23] T. L.-P. Tang, "Detecting honest people's lies in handwriting," *Journal of Business Ethics*, vol. 106, no. 4, pp. 389–400, 2012.

[24] M. F. Folstein, S. E. Folstein, and P. R. McHugh, *Journal of Psychiatric Research*, vol. 12, no. 3, 1975.

[25] P. Kline, *The Handbook of Psychological Testing*. Routledge, 1999.

[26] R. Plamondon and G. Lorette, "Automatic signature verification and writer identification: the state of the art," *Pattern recognition*, vol. 22, no. 2, pp. 107–131, 1989.

[27] H. Kim, "The clockme system: computer-assisted screening tool for dementia," Ph.D. dissertation, Georgia Institute of Technology, 2013.

[28] R. Plamondon, C. O'Reilly, and C. Ouellet-Plamondon, "Strokes against stroke - strokes for strides," *Pattern Recognition*, vol. 47, no. 3, 2014.

[29] C. O'Reilly and R. Plamondon, "Design of a neuromuscular disorders diagnostic system using human movement analysis," in *ISSPA*, 2012.

[30] J. Heinik, P. Werner, T. Dekel, I. Gurevitz, and S. Rosenblum, "Computerized kinematic analysis of the clock drawing task in elderly people with mild major depressive disorder: an exploratory study," *Int Psychogeriatr.*, no. 3, pp. 479–488, 2010.

[31] G. Luria, A. Kahana, and S. Rosenblum, "Detection of deception via handwriting behaviors using a computerized tool: Toward an evaluation of malingering," *Cognitive Computation*, vol. 6, no. 4, 2014.

[32] E. Sesa-Nogueras, M. Faúndez-Zanuy, and J. Mekyska, "An information analysis of in-air and on-surface trajectories in online handwriting," *Cognitive Computation*, vol. 4, no. 2, pp. 195–205, 2012.

[33] S. Rosenblum, S. Parush, and P. Weiss, "The in air phenomenon:temporal and spatial correlates of the handwriting process," *Perceptual and Motor Skills*, 2003.

[34] S. Rosenblum, B. Engel-Yeger, and Y. Fogel, "Age-related changes in executive control and their relationships with activity performance in handwriting," *Human Movement Science*, 2013.

[35] P. Petta and C. Pelachaud, Eds., *Introduction to Emotion-Oriented Systems - The HUMAINE Handbook*. Springer, 2011.

[36] T. Brown, B. Chorpita, W. Korotitsch, and D. Barlow, "Psychometric properties of the depression anxiety stress scales (DASS) in clinical samples," *Behav. Res. Ther.*, vol. 35, pp. 79–89, 1997.

[37] S. Rosenblum, P. Werner, T. Dekel, I. Gurevitz, and J. Heinik, "Handwriting process variables among elderly people with mild major depressive disorder: a preliminary study," *Aging Clinical and Experimental Research*, pp. 141–147, 2009.

[38] H. Liu and H. Motoda, *Computational Methods of Feature Selection (Data Mining and Knowledge Discovery Series)*. Chapman & Hall/CRC, 2007.

[39] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, pp. 1157–1182, 2003.

[40] I. Kononenko and M. Kukar, *Machine Learning and Data Mining: Introduction to Principles and Algorithms*. Horwood Publishing Limited, 2007.

[41] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.

[42] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, 2008.

[43] A. Esposito, A. M. Esposito, L. Likforman-Sulem, N. Maldonato, and A. Vinciarelli, "On the significance of speech pauses in depressive disorders: results on read and spontaneous narratives," in *Advances in non linear speech processings*. Springer, 2016.

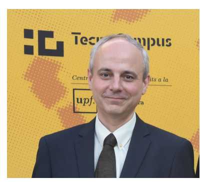
[44] R. Caruana, "Multitask learning," *Machine Learning*, vol. 28, no. 1, pp. 41–75, 1997.

[45] G. Keinan and S. Eilatgreenberg, "Can stress be measured by handwriting analysis? the effectiveness of the analytic method," *Applied Psychology: An International Review*, 1993.

[46] R. Plamondon, C. O'Reilly, J. Galbally, A. Almaksour, and É. Anquetil, "Recent developments in the study of rapid human movements with the kinematic theory: Applications to handwriting and signature synthesis," *Pattern Recognition Letters*, vol. 35, pp. 225–235, 2014.

Laurence Likforman-Sulem received the engineering degree from ENST-Bretagne (Ecole Nationale Supérieure des Télécommunications) in 1984 and the PhD degree from ENST-Paris in 1989. She is an Associate Professor at Telecom ParisTech where she serves as an instructor in pattern recognition. She chaired the program committee of CIFED held in Fribourg, Switzerland, in 2006 and the program committees of two DRR Conferences (Document Recognition and Retrieval) held in 2009 and 2010 in San Jose, California.

Anna Esposito received the PhD Degree in Applied Mathematics and Computer Science from the University of Naples "Federico II". She is currently associate professor in the Department of Psychology, at Seconda University of Naples (IT) and director of the Behavioural Cognitive Systems (BeCogSys) laboratory. She is research affiliate of the International Institute for Advanced Scientific Studies (IIASS). She has published 160 peer reviewed papers and is editor/coeditor of 24 international books.



Marcos Faundez-Zanuy received the B.Sc. degree in telecommunication in 1993 and the Ph.D. degree in 1998, both from the Polytechnic University of Catalunya. He is now a Full Professor at ESUP Tecnocampus Mataro-Maresme since 2010 and heads the Signal Processing Group. His research interests lie in the fields of biometrics and speech coding. He is author of more than 50 papers indexed in ISI Journal citation report, more than 100 conference papers, ten books, and responsible for ten national and European research projects.

Stéphan Clémençon received the Ph.D. degree in applied mathematics from the University Denis Diderot, Paris 7, France, in 2000. In October 2001, he joined the faculty of the University Paris X as Associate Professor and successfully defended his habilitation thesis in 2006. Since October 2007, he has been professor and researcher with Telecom ParisTech, the leading school in the field of information technologies in France. His research interests include machine-learning, Markov processes and nonparametric statistics.

Gennaro Cordasco is Assistant Professor in the Department of Psychology at the Second University of Naples, and affiliate of the International Institute for Advanced Scientific Studies (IIASS).