Intelligent Decision Technologies 16 (2022) 631–641
DOI 10.3233/IDT-220070
IOS Press

V-GRAFFER, a system for Visual GRoup AFFect Recognition, Part I: Foundations

Andreas M. Triantafyllou* and George A. Tsihrintzis
Department of Informatics, University of Piraeus, Piraeus Greece

Abstract. In this paper, we present our research towards building V-GRAFFER, a system for Visual GRoup AFFect Recognition. Specifically, V-GRAFFER aims at the development and provision of services with regard to the recognition of the emotional state of groups of people. At this stage of the V-GRAFFER development, implemented services are oriented towards detecting and drawing conclusions about students who attend educational events such as lectures, question and answer (Q&A) sessions, or lab participation. Specific functionalities of the current version of V-GRAFFER include processes for data collection, lecture experiments under real conditions, algorithms for sample auto-extraction, and optimized approaches. These functionalities allowed the collection of data of various educational events under real conditions and the creation of flexible databases of appropriate samples for drawing conclusions with regard to emotion detection of groups of people. Furthermore, we devised and implemented innovative algorithms to identify and classify group samples from recorded educational events. These algorithms have been evaluated and improved via continuous cycles of development-testing-evaluation in a variety of experiments. Finally, we constructed completed databases of group samples which are correlated with each other based on time, depth of time and educational settings.

Keywords: Visual group affect recognition, group affect recognition, emotion detection, group emotion detection, sentiment detection

1. Introduction

V-GRAFFER is a research work about the smart solutions for emotion detection in several fields in which groups of people are involved. One research path of our works concerns the emotion detection of groups of people who attend educational events. Our goal is to research and develop services which will connect with tutoring systems in order to achieve better results in computer-assisted learning and education. Specifically, our connected services with tutoring systems aim at the customization of e-course/e-lecture flows according to the emotional state of a group of participants. For example, a lecture could appear difficult or boring to students in conjunction with other parameters and in relation to previous lectures or real life variables. For this reason, tutoring systems need to be able to draw conclusions about the emotional state of each group and to adapt their lessons in order to achieve successful course delivery.

Other fields that group emotion detection could be used include the automatic evaluation of music performances, shows or movies through tracking the attendants' reactions. Additionally, emotion recognition of groups may be useful in the selection of music tracks in music festivals according to the audience's sentimental state. Finally, lecture adaptation in tutoring systems constitutes another significant field in which the group emotion recognition will be useful.

Emotion detection can combine data from many human signals to achieve better results when drawing conclusions. An important factor which affects the overall emotional state of a team in some event is the team itself. For this reason, we understand that group emotion detection should include combined rules and special sample types for algorithm research. In our research, we have devised and implemented some important common parts which constitute the foundation for several research paths to be followed in group emotion detection.

*Corresponding author: Andreas M. Triantafyllou, Department of Informatics, University of Piraeus, Piraeus 185 34, Greece. E-mail: aandreasam@gmail.com.

ISSN 1872-4981/$35.00

Specifically, in this paper, we present the definition of the group samples which are useful for many approaches in group emotion detection. We also describe the functionalities of data collection and processing, the pattern recognition and information extraction, and flexible database schemes which are used in our research and implementations. Furthermore, we present algorithms that we have devised for collecting, identifying and classifying the group samples, the training sets and the recovery flow of these data in order to use them in our services. We have also performed evaluations, tests and improvements via several experimental activities. Through these research steps, we have implemented separated software into services in order to use it in flexible procedures.

Accordingly, the above-mentioned research steps form the foundations for devising, implementing and evaluating algorithms and approaches for group emotion detection, as the group sample definition constitutes the first significant step. Defining such group samples, allows sample collection, their auto-classification and training set creation. With regard to emotion detection via facial samples specifically, the group samples are connected facial images which have been taken at the same time in the same educational event and they are also linked to the previous and the subsequent group samples based on time depth as analyzed in the following sections.

Next, we present the algorithms and approaches we have implemented with regard to the automatic exporting and classification of group samples in order to create flexible databases of training sets and testing data. These algorithms have been implemented and comparatively evaluated, which allowed us to come up with better/optimal solutions. All the related experiments and required development of auxiliary software is presented in the following sections.

The rest of the paper is organized as follows: Section 2 is a review of previous related work in the paper area. Section 3, is devoted to the description of data collection, data processing and pattern recognition implemented in V-GRAFFER. Section 4 describes the database schemas we implemented, while Section 5 discusses software implementations as in the current version of V-GRAFFER. Section 6 illustrates experimental results from use of V-GRAFFER in real scenarios. Finally, Section 7 is devoted to summarization of the paper, conclusion drawing and indications of future related work in this area.

2. Previous related work

In the previous works, various papers have dealt with the emotion detection of a person via diverse approaches. The detection and recognition of the emotional state of individuals or groups of people has been an area of intensiver research interest over the past two decades and the relevant literature is extensive [4,6,10,11,13,15–23,25,27–30,33]. Some kinds of these works include the emotion detection of individuals using face recognition. The detection and analysis of facial expressions can help us in various fields. For example the [9] have worked on facial expressions in depression, they have experimented in groups of combined healthy and depressed people in order to research it for individuals. Furthermore, we have seen in [1], algorithms for facial comparison have been investigated in real time responses per person with real-time results. This work has dealt with emotion detection algorithms using different training sets with the same or unknown faces. Facial expressions have been investigated for stability over the time as well. The researchers of [24] have investigated differences in facial expressions for 2 people between 4 and 12 months coming to the conclusion that facial expressions were stable. This fact makes the research on facial expressions important. Furthermore, Facial Expressions have been researched as to the context for classification [26] and standardization in order to be in useful forms for detection algorithms. Emotion detection through facial expression requires an automated way for face detection. For this need, there is an interesting approach to [12] that they have implemented a solution for face detection through artificial neural networks.

In addition to detection of facial expressions, there are works that have dealt with a combination of modalities. Specifically, the [7,11,17,20] have researched and presented modalities for Visual - Facial recognition, Audio-Lingual and Keyboard - Stroke Pattern measures and also they have presented improvements via combinations of these modalities for individuals. Keyboard stroke patterns have been researched to [14,16–18] as well, which is useful for using that knowledge for our Group Emotion detection via combination of modalities. Apart from these types of emotion recognition, there are also useful informations for emotional states through linguistic & paralinguistic audio signals [7,15], meaningful information extraction from body posture or gait [32,33], detection from written texts and Bio-signal measurements [31].

Concerning the group emotion detection, researchers [8] have examined detection of spontaneous emotion of groups at specific times through mesh analysis.

These research fields are about the emotion detection of individuals or for group emotions without correlation based on time flow and context conditions. Over the years, we have been researching the emotion and sentiment detection of groups of people via computers, that people are correlated with each other in the same conditions. Specifically, we have begun the V-GRAFFER series in which special emphasis has been placed on educational events and learning applications. Our research so far has aimed at groups of students in various class settings. However, our research results and our system implementations may be used in several fields in which groups of people are involved.

In the present paper, the main steps, which are about the foundation for further research on emotional state detection of groups of people, have been included. Specifically, we present the definitions for involved components which are invented by ourselves. Furthermore, we analyze and present our approaches to the useful data recognition, collection & processing; the database schemas, our experimental activities, our software implementations and the evaluations. Moreover, the improvements we have implemented during the research are mentioned.

3. Recognition, collection and data processing

There is no doubt that the most significant components in the pattern recognition approaches are the samples. The samples are important for both machine learning and algorithm development sides. For this reason, the first component which should be defined is the appropriate sample. An appropriate sample may consist of datasets which are capable of giving us the requested information.

As regards the emotion detection of groups of people through extracting information from facial images, we define the 'Group Samples' which have been invented through our research [3]. The Group Sample consists of a set of faces, which are correlated to each other based on time and the educational settings. Specifically, a group sample should consist of the facial images from students who attend the same educational event. Furthermore, these images should be taken at the same time during the event.

For Example:

If 10 students attend a 60-minute lecture, then we have:

The first group sample consists of 10 faces at $t = 0$, The second group sample consists of 10 faces at $t = 1$, etc

Bytes arrays/vectors	Time frame (t)
10	0
10	1
10	$..n$

These Group Samples include the information about the time linking with each other during the lecture time.

Moreover, information about the educational event and specifics tags per face should be kept in Group Samples data.

According to the above definition, the data which can be collected from an educational event are about the number of students, the duration of the event and the group sample collecting period.

For instance, if we want to collect 1 group sample per second in a 60-minute lecture which is attended by 10 students, we have:

60 minutes * 60 second * 1 Group Sample = 3600 Group Samples
3600 Group Samples * 10 students = 36000 face records

Clearly, such a number of samples is almost impossible to be collected manually. For this reason, we have researched, implemented and improved the processes of automated group samples extraction.

The automated procedures about the group samples extraction have been achieved through the face detection on frames of recorded videos of educational events. We have implemented procedures about the automatic classification of extracted faces which are associated with each event and the time, thus creating the group samples.

4. Database schemas

In this step of the work, flexible database schemes will be presented which are useful for smart algorithm development as well as for system training. First of all, we mention the core aspects about the entities' creation as we have presented in [6]. Specifically, the Educational Event, the Duration (of an Educational Event), and the Group of students (Total of individuals) who attend an educational event are the main aspects.

According to these aspects, we should keep data about the:

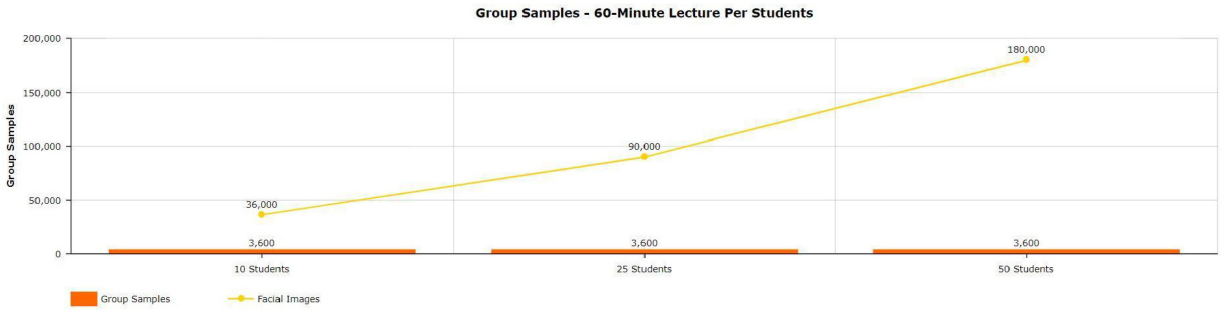


Fig. 1. Group Samples Per Students. 60-minute Lecture.

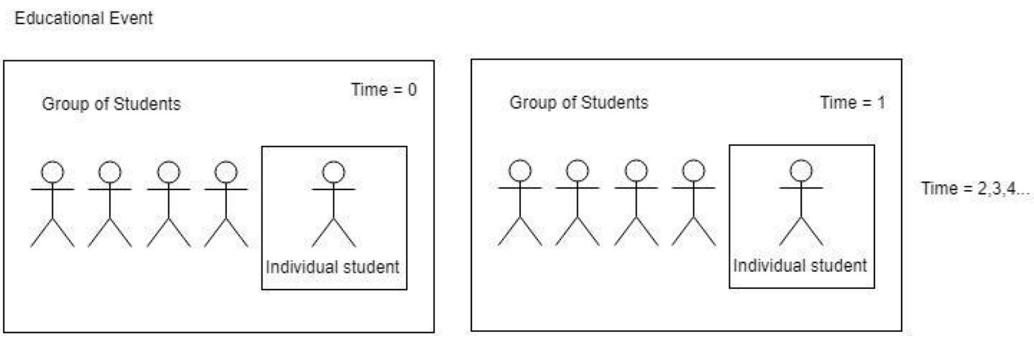


Fig. 2. Initial aspects.

- Identification of Educational Event: That is an ID for each educational event.
- Data – Vectors of student faces: The detected data for faces.
- Time: Time frame/Date time for each sample.
- Data about the face association with time, depth of time, and educational event (Linked Data).

Through this research, we have led to the core database scheme which has been used on our next database expansions in order to collect, keep and use the group samples. This scheme is sufficiently flexible as we can easily extend it.

- EventInfo (Id, EventName, Datetime, OtherDetails)
- Sample (Id, EventId, FrameCount, TypeId, Tag, SampleBytes)
- SampleType (Id, Type, OtherDetails)

We have presented 2 scheme extensions of this database scheme [2]. The first one extension includes a percentage for attending pointer. The second one refers to a database extension with more tables which contain sentimental data collected by various modalities (Face expressions, Audio-Lingual analysis, etc.). These schemes are used by our algorithm approaches on Group Emotion Detection. Furthermore, these schemes could be used for further research and algorithm approaches in this research field.

Having mentioned the Database schemas, we continue in the Software Implementations section about the foundations of V-GRAFFER.

5. Software implementations

The issue about the collection and processing of enormous numbers of group samples needs smart algorithm solutions. For this reason, we present our software implementations in order to succeed in collecting and classifying such samples. First of all, we will mention the 4 milestones about our software solutions and improvements as we have presented in [3–6]. The first important aspect that we have implemented concerns the core of our software for automatic Group Samples collection, extraction, processing and storage. Furthermore, we implemented another very important functionality for Samples and Algorithm evaluation. Through algorithm evaluation, 2 new classification algorithms have been created and we have managed to reduce the error rate. Finally, we developed software extensions to supporting other modules – services and multiple data schemes via intermediate micro service.

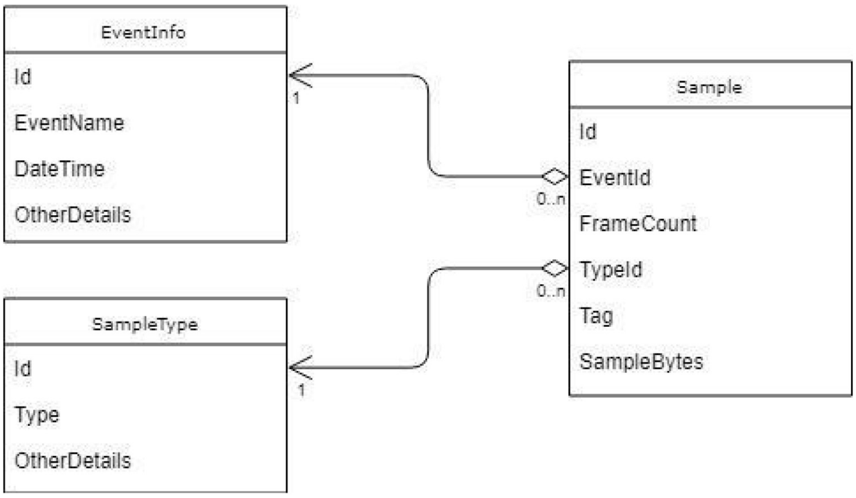


Fig. 3. Core database scheme.

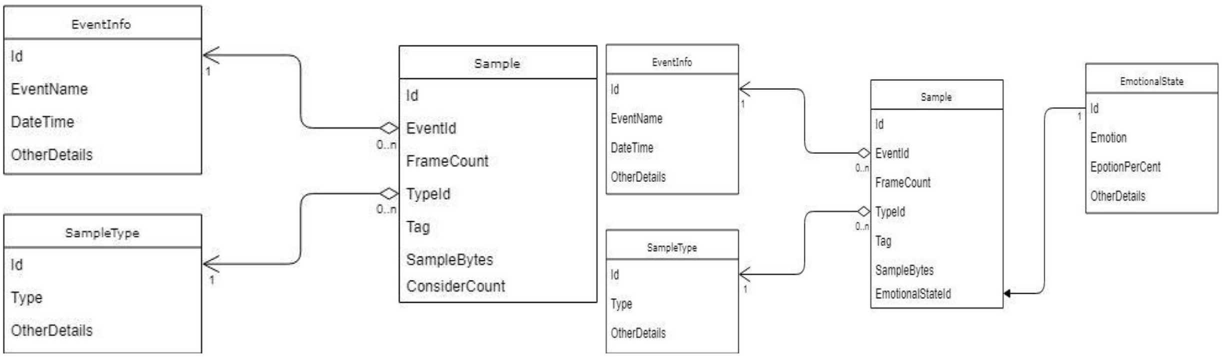


Fig. 4. Database schema extensions.

At this step of the work, we mention the technologies that we have used in order to develop the below features. Specifically, we have used MySql for database solutions. We also had used Microsoft SQL Server before we changed our databases to MySql. For Back-End development the main programming language is C# but we also have used python only for test algorithms. For visual solutions, we have used WPF Application (C# and xaml). Furthermore, our micro-services have been developed with ASP .NET Core with C# which is a cross-origin technology by Microsoft. That means that we can deploy these services to both Linux and Windows Servers as well. We have used Docker images on Linux Servers for microservices deployments. Finally, .Net Core API solutions have been used to separate functionalities in order to transfer data among our software.

First and foremost, we implemented the core of our software. Specifically, we created procedures for reading frames of recorded videos of educational events, face extraction and saving them with the appropriate relations according to the time, the depth of time and each educational event. The first window application (the UI is contained) includes the below options – functions:

Recorded video selection	Insertion of period definition (e.g Each period is 2 seconds of the videos)
Insertion of Educational Event Title and Details	Option about the UI Frames View
Connect with existing Educational Event (in order to continue the samples processing)	Open, Start, Pause, Stop and Next Frame options
Insertion of collections samples per period (e.g. Collect 2 samples per period)	Displayed information related to the videos (Frame and Bit Rate, Width / Height, Is Opened, Details, etc.)
Options in regard to the storage media selection.	Window with the details of the saved Group Samples

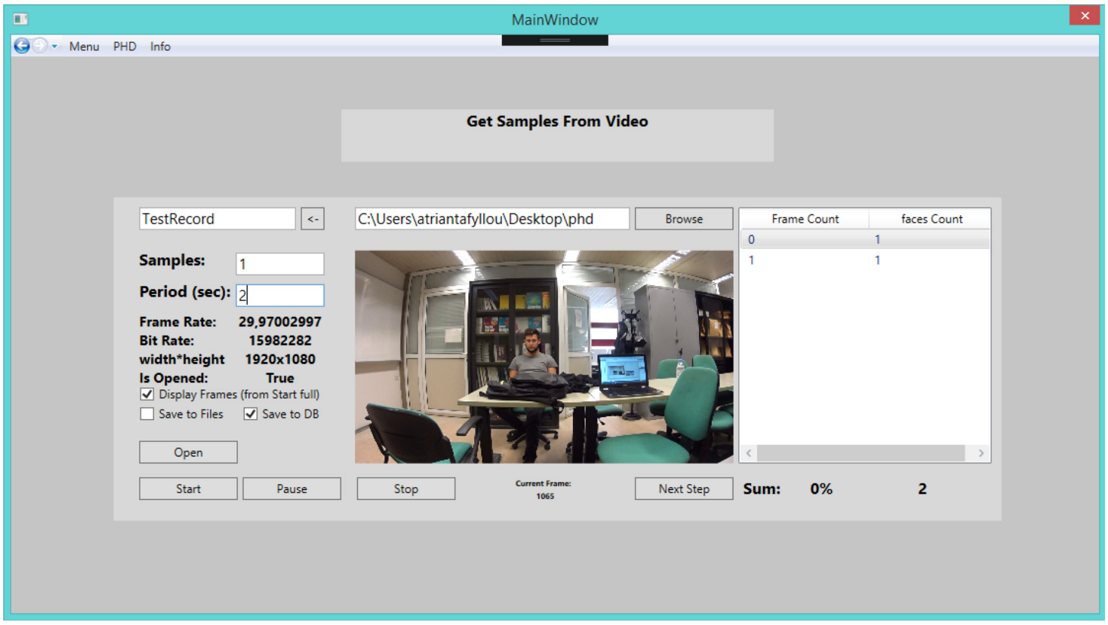


Fig. 5. Main page of the program.

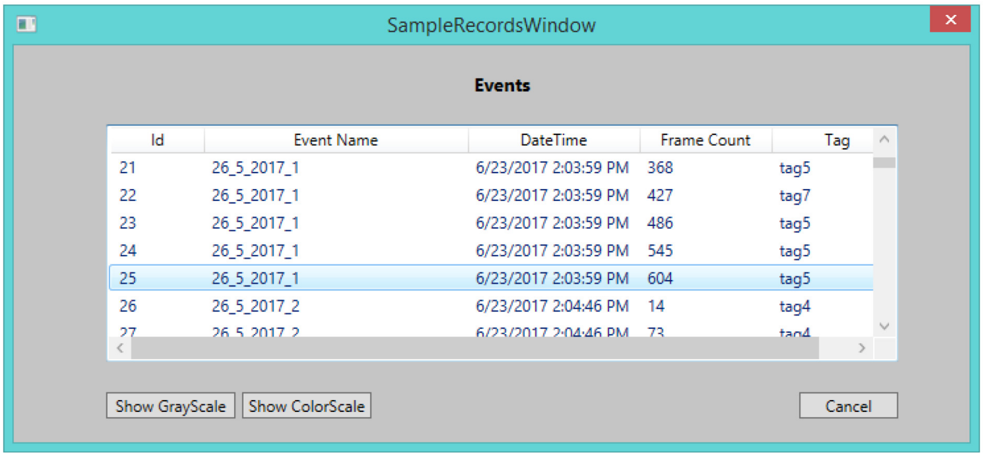


Fig. 6. Window with the group sample details.

The evaluation of the initial classification algorithm was achieved by the next software extensions. More specifically, we have created features about the displaying of group faces in UI images with their tags and details with options about the selection of each educational event and the classified tags in order to observe and fix the wrong classifications. Furthermore, we have implemented the procedures for classification errors fixing, error rate extraction and calculation of statistics about the algorithm efficiency.

These additions to a list:

We are continuing with the 2 new algorithms which

Page – Window with the Grouped Samples	Total faces in educational event
Functionality about the view of the faces grouped by the educational event and the tag	Total faces on each tag
Functionality about the statistics	Functionality for the database updating about the tags fixing

are implemented in order to reduce the classification error rate. This achievement was significant because we can build reliable completed databases in a better time with a lower error rate.

The first one new algorithm contains the procedure

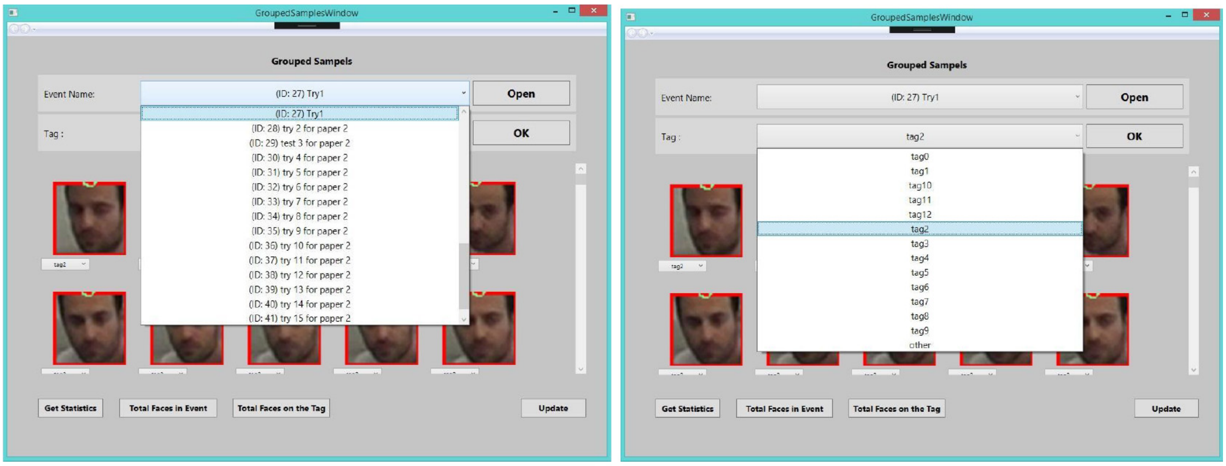


Fig. 7. Grouped facial images and selection of the educational event and classified tags.

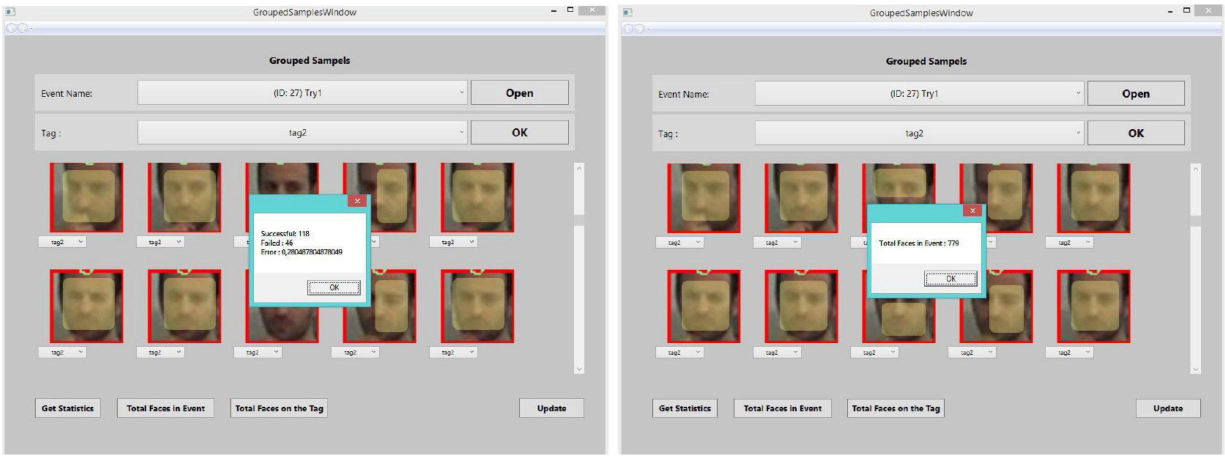


Fig. 8. Statistics.

for manual initial tagging for the students who will attend each educational event. That requires the detection of the distinct students in each event. The new window, in which we can insert the student tag is required as well.

The functionality of the second new algorithm contains the creation of wide data sets with the audience faces with their tags by manual tagging and the use of them as training sets in our automatic classification. That needs the detection of the distinct students in each event as well. Furthermore, it needs parameters about the total range tha we want for each training set and finally, it needs the mentioned-above window for the manual student tagging.

We have implemented and added all the above functionalities in our systems. We can see some photos about the UI of these in use.

In the end, we have extended our software services in order to support the connection of other modules – micro services of our research work. We have separated some of our algorithms in micro services. This separation is useful and flexible enough in order to combine services. Moreover, we have implemented a service which is responsible for database handling. With this service we can easily adapt the new Database schemas and the different database services such as the Microsoft SQL Server and the MySQL.

6. Experimental activities

In this section of our work, we present our experimental activities which were carried out in University of Piraeus in collaboration with the department of in-

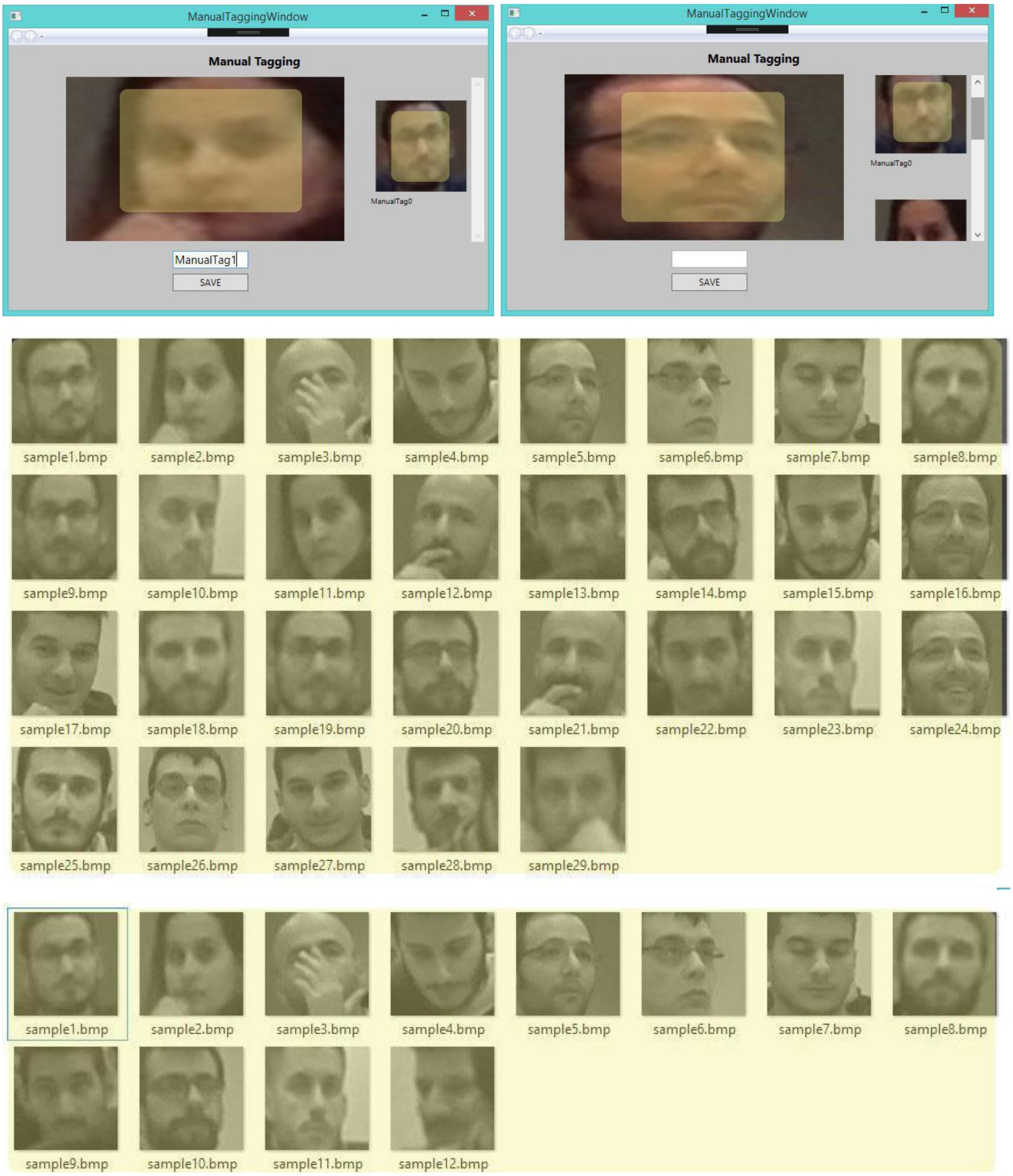


Fig. 9. Software extensions about the 2 improved algorithms.

formatics in order to collect samples which are useful for developing and improving our algorithms. Furthermore, such samples are useful for creating training sets. The samples that we want to collect should be correlated with each other based on time, depth of time, and educational events. For this reason, a lot of videos of students who were attending lectures, workshops and Q&A have been recorded. Specifically, we have recorded over 10 hours of educational events through which we are capable of exporting the vectors – bytes arrays from students' faces. We used a Camera Sony HDR – AS50 with a wide range of video capture (170

Fig. 10. Frames of video lectures in real conditions.

degrees). Through these records we can export a plenty of group samples. For instance, for a 60-minute lecture with 30 frames per second we can have up to 30 samples per second. If we choose to get a group sample per second, we have:

60 minutes * 60 seconds * 1 group sample = 3600 group samples

So, if the lecture was attended by 10 students, we have

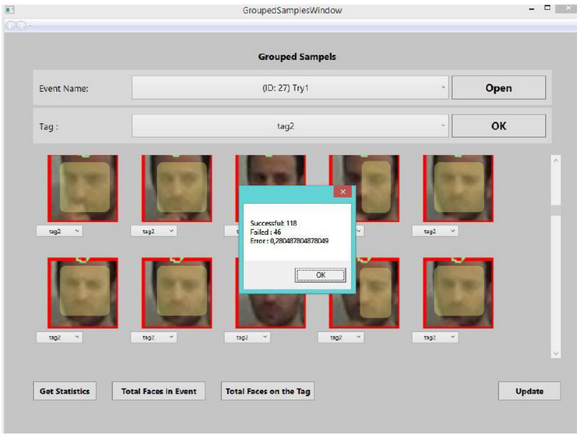
3600 group samples * 10 students' faces = 36000 facial vectors

As we presented above, we have created our algorithms for automatic face detection, exporting and classification. Through this we managed to collect useful data from such records in flexible databases.

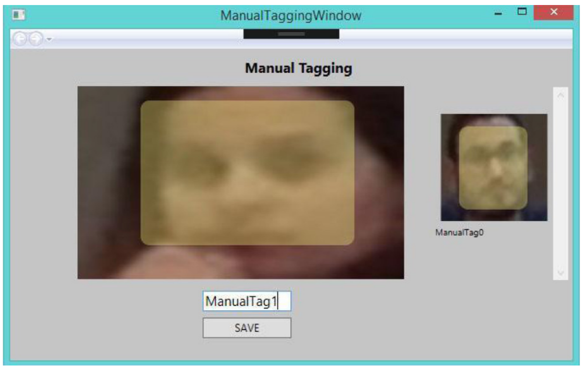
7. Evaluations

As we presented above, we have implemented some improvements in our algorithms in order to achieve better results. This was achieved through our evaluations of all our algorithms which have been presented in our previous papers [3–5]. We have implemented functionalities for helping us in evaluation and also manually error fixing.

Specifically, we have implemented a visual window in our window application which displays the grouped faces per tag. In this window we can mark the wrong faces per tag in order to export the error rate per algorithm. With this functionality we managed to evaluate our final solutions. Performing every test we managed to improve our software with smarter algorithms.



Furthermore, we developed a new window with functionality for manual tagging. That means that we can fix the wrong samples manually in order to get the correct completed training sets.



Through the above evaluation functionality, we export the below error rates after a lot of algorithm runs:

Algorithm approach	Error rate
Initial automatic classification	$0.21 \sim 0.25$
First Improvement (Method 1 [5])	$0.14 \sim 0.18$
Second Improvement (Method 2 [5])	$0.05 \sim 0.09$

We have evaluated the algorithms with ranges because the educational events were carried out in real conditions which were frequently altered.

More specifically, the steps for the evaluation started from the real lectures video recordings. We recorded plenty of lectures at University of Piraeus by a camera with wide frames. We put the Video records in our software and we set the preferences about the samples period and every specific setting for each algorithm. After the Processing step, Group Samples have been exported in order to process them for evaluation. We were marking the wrong samples and we were noting the results in order to calculate the average of the error rates. Subsequently, we repeated the procedure with each of our algorithms.

In the second improved algorithm [5] the error rate has been reduced to $0.05 \sim 0.09$ which is low enough for a totally new algorithm for the above-defined Group Samples.

8. Synopsis and future work

In summary, in this paper we presented the foundations of V-GRAFFER, a software system for Visual GRoup AFFEct Recognition. Specifically, we presented the scope of this work and its ultimate goal. We also defined the group samples which are useful for researching algorithms for emotion detection of groups of people. Additionally, we presented our algorithms about sample collection, information extraction, recognition, and classification. For these functionalities, we implemented corresponding software modules. Furthermore, we designed and developed database schemas in order to support additional research approaches on emotion detection of groups of people. Finally, we presented the evaluation of our algorithms and we mentioned our implemented improvements. Specifically, we started with an error rate of about $0.21 \sim 0.25$ and we managed to reduce it to $0.05 \sim 0.09$.

Our long-term goal is to create and integrate our services, improving the quality of several ways of education. For instance, one of the significant achievements will be the creation of e-tutoring/e-learning systems with the ability to adapt e-lectures in relation to the emotional state of the students who attend them. This functionality will be similar to the adaptation that human teachers make under real conditions when the lessons appear difficult or boring to the students.

Towards this long-term goal, our future research includes the connection of the current version of V-GRAFFER to our improved facial feature extraction. We will also investigate and present further aspects of our emotion detection algorithms of facial features and we will compare them with other approaches. Furthermore, we will combine the facial feature-based algorithms with algorithms from other modalities in order to achieve better results. Finally, we will integrate these additional functionalities in a more comprehensive software oriented towards computer-assisted education. This and other related research is currently in progress and its results will be announced elsewhere in the near future.

References

[1] Triantafyllou A. Detecting human emotion/mood from a computer. MSc Thesis. 2017.

[2] Triantafyllou AM, Tsihrintzis GA, Virvou M, Alepis E. A Bimodal System for Emotion Recognition via Computer of Known or Unknown Persons in Normal or Fatigue Situations, in: G.A. Tsihrintzis and M. Virvou, eds, ADVANCES IN CORE COMPUTER SCIENCE-BASED TECHNOLOGIES, Learning and Analytics in Intelligent Systems. Springer; 2020. vol. 14, pp. 9-35.

[3] Triantafyllou AM, Tsihrintzis GA. Group affect Recognition: Visual – Facial collection, IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI). 2017.

[4] Triantafyllou AM, Tsihrintzis GA. Group affect Recognition: Evaluation of basic automated sorting, IEEE 9th International Conference on Information, Intelligence, Systems and Applications (IISA); 2018.

[5] Triantafyllou AM, Tsihrintzis GA. Group affect ãĂĂRecognition: Optimization of automatic classification, Springer 12th Joint Conference on Knowledge-Based Software Engineering (JCKBSE), 2018.

[6] Triantafyllou AM, Tsihrintzis GA. Group affect Recognition: Completed Databases & Smart Uses, ACM 3th International Conference on E-Education, E-Business and E-Technology (ICEBT), 2019.

[7] Virvou M, Tsihrintzis GA, Alepis E, Stathopoulou I-O, Kabassi K. Toward affect recognition by audio-lingual and visual-facial modalities: empirical studies and multi-attribute decision making. International Journal on Artificial Intelligence Tools. 2012; 21(2): 1240001-1-1240001-28.

[8] Veenendaal A, Daly E, Jones E, Gang Z, Vartak S, Patwardhan RS. Group Emotion Detection using Edge Detection Mesh Analysis. 2014.

[9] Suslow T, Junghanns K, Arolt V. Detection of Facial Expressions of Emotions in Depression, 2001.

[10] Graesser A, McDaniel B, Chipman P, Witherspoon A, D'Mello S, Gholson B. Detection of Emotions during Learning with AutoTutor, 2006.

[11] Stathopoulou I-O, Alepis E, Tsihrintzis GA, Virvou M. On Assisting a Visual-Facial Affect Recognition System with Keyboard-Stroke Pattern Information. Knowledge-based Systems, Elsevier; 2010; 23: 350-356.

[12] Stathopoulou I-O, Tsihrintzis GA. Appearance-based face detection with artificial neural networks. Intelligent Decision Technologies. 2011; 5(2): 101-111.

[13] Stathopoulou I-O, Tsihrintzis GA. Visual Affect Recognition. vol. 214 of Frontiers in Artificial Intelligence and Applications, IOS Press, 2010.

[14] Stathopoulou I-O, Alepis E, George A. Tsihrintzis. Virvou M. On assisting a visual-facial affect recognition system with keyboard-stroke pattern information. Knowl-Based Syst. 2010; 23(4): 350-356.

[15] Alepis E, Stathopoulou I-O, Virvou M, Tsihrintzis GA, Kabassi K. Audio-lingual and Visual-facial Emotion Recognition: Towards a Bi-modal Interaction System. ICTAI (2), 2010, pp. 274-281.

[16] Stathopoulou I-O, Alepis E, Tsihrintzis GA, Virvou M. On Assisting a Visual-Facial Affect Recognition System with Keyboard-Stroke Pattern Information. SGAI Conf. 2009, pp. 451-463.

[17] Tsihrintzis GA, Virvou M, Alepis E, Stathopoulou I-O. Towards Improving Visual-Facial Emotion Recognition through Use of Complementary Keyboard-Stroke Pattern Information. ITNG, 2008, pp. 32-37.

[18] Tsihrintzis GA, Virvou M, Stathopoulou I-O, Alepis E. On Improving Visual-Facial Emotion Recognition with Audio-lingual and Keyboard Stroke Pattern Information. Web Intelligence, 2008, pp. 810-816.

[19] Stathopoulou I-O, Tsihrintzis GA. NEU-FACES: A Neural Network-Based Face Image Analysis System. ICANNGA (2), 2007, pp. 449-456.

[20] Stathopoulou I-O, Tsihrintzis GA. Towards Automated Inferencing of Emotional State from Face Images. ICSOFT (PL/DPS/KE/MUSE), 2007, pp. 206-211.

[21] Stathopoulou I-O, Tsihrintzis GA. A Neural Network-based System for Face Detection in Low Quality Web Camera Images. SIGMAP, 2007, pp. 53-58.

[22] Alepis E, Virvou M, Kabassi K. Affective Reasoning Based on Bi-modal Interaction and User Stereotypes. New Directions in Intelligent Interactive Multimedia, 2008, pp. 523-532.

[23] Matsumoto D, Ekman P. Facial expression analysis. Scholarpedia. 2008; 3(5): 4237.

[24] Cohn JF, Schmidt KL, Gross R, Ekman P. Individual Differences in Facial Expression: Stability over Time, Relation to Self-Reported Emotion, and Ability to Inform Person Identification. ICMI, 2002, pp. 491-498.

[25] Donato G, Bartlett MS, Hager JC, Ekman P, Sejnowski TJ. Classifying Facial Actions. IEEE Trans. Pattern Anal. Mach. Intell, 1999; 21(10): 974-989.

[26] Bartlett MS, Donato G, Movellan JR, Hager JC, Ekman P, Sejnowski TJ. Image Representations for Facial Expression Coding. NIPS, 1999, pp. 886-892.

[27] Bartlett MS, Viola PA, Sejnowski TJ, Golomb BA, Larsen J, Hager JC, Ekman P. Classifying Facial Action. NIPS, 1995, pp. 823-829.

[28] Kapoor A, Qi Y(A), Picard RW. Fully Automatic Upper Facial Action Recognition. AMFG, 2003, pp. 195-202.

[29] Kapoor A, Picard RW. Real-Time, Fully Automatic Upper Facial Feature Tracking. FGR, 2002, pp. 10-15.

[30] Picard RW. Toward computers that recognize and respond to user emotion. IBM Systems Journal. 2000; 39(3&4): 705-719.

[31] Haag A, Goronzy S, Schaich P, Williams J. Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System, 2004, pp. 36-48.

[32] Masuda M, Kato S, Itoh H. Emotion Detection from Body Motion of Human Form Robot Based on Laban Movement Analysis, 2009, pp. 322-334.

[33] Garber-Barron M, Si M. Using body movement and posture for Emotion Detection in non-acted Scenarios: IEEE International Conference on Fuzzy Systems, 2012.