This event is endorsed and organized by

1st EAI International Conference on Design, Learning & Innovation

MAY 2–3, 2016 | ESBJERG, DENMARK

Program: ArtsIT and DLI 2016

EACH DELEGATE PRESENTATION MUST TAKE 25 MINUTES MAX – 20 MINUTES + 5 Q & A

TIME	MONDAY	
	ARTSIT	**DLI**
09:00-09:15	Welcome: Tony & Eva **C1.117**	Welcome: Eva & Tony **C1.117**
09:15-10:00 Session chairs Tony Brooks Eva Brooks	**KEYNOTE**: Antoni Jaume-i-Capó **C1.117**	**KEYNOTE**: Antoni Jaume-i-Capó **C1.117**
10:00 – 10:15	**COFFEE/TEA BREAK**	**COFFEE/TEA BREAK**
10:15-10:40 Session chair	Paper: **The Farm Game: A Game Designed to Follow Children's Playing**	Paper: **The Farm Game: A Game Designed to Follow Children's Playing**

Important dates

Workshops proposal deadline
24 January 2016

Full Paper Submission deadline
14 March 2016

C117 Uriel Martinez Hernandez	Authors: *Zidianakis, Emmanouil, Stratigi, Kalliopi, Ioannidi, Danae, Partarakis, Nikolaos, Antona, Margherita, Stephanidis, Constantine **C1.117**	Authors: *Zidianakis, Emmanouil, Stratigi, Kalliopi, Ioannidi, Danae, Partarakis, Nikolaos, Antona, Margherita, Stephanidis, Constantine **C1.117**
10:40-11:05	Paper: **A Tangible Augmented Reality Toy Kit: Interactive Solution** Authors: Zhu, Yujie, *Wang, Stephen, Jia **C1.117**	Paper: **A Tangible Augmented Reality Toy Kit: Interactive Solution** Authors: Zhu, Yujie, *Wang, Stephen, Jia **C1.117**
11:05-11:30	Paper: **SceneMaker: Creative Technology for Digital StoryTelling** Authors: Paul McKevitt et al. **C1.117**	Paper: **SceneMaker: Creative Technology for Digital StoryTelling** Authors: Paul McKevitt et al. **C1.117**
11:30-11:55	Paper: **Structuring Design and Evaluation of an Interactive Installation** Authors: *Erkut, Cumhur **C1.117**	Paper: **Structuring Design and Evaluation of an Interactive Installation** Authors: *Erkut, Cumhur **C1.117**
11:55-12:30	**LUNCH BREAK**	**LUNCH BREAK**

Notification and Registration opens
25 March 2016

Camera-ready deadline
18 April 2016

Start of Conference
2 May 2016

End of Conference
3 May 2016

Sponsors

Technical sponsors

Tweets by @EAIchannel

12:30-12:55 Session chair C117 Erkut, Cumhur C119 Triantafyllidis, Georgios	Paper: **Gamify HCI: Device's Human Resolution for Dragging on Tou...** Authors: Christensen, Allan, Pedersen, Simon, Andre, *Knoche, Hendrik **C1.117**	Paper: **Maze and Mirror Game Design for Increasing Motivation in ...** Authors: *Hojjat, Sara, Ikemoto, Chiaki, Sowa, Tomoyuki **C1.119**
12:55-13:20	Paper: **Towards a Wearable Interface for Immersive Telepresence** Authors: *Martinez-Hernandez, Uriel, Szollosy, Michael, Boorman, Luke, W., Kerdegari, Hamideh & Prescott, Tony, J. **C1.117**	Paper: **Designing Digital Tools for Physiotherapy** Authors: Barbu Postolache, Gabriela, Oliveira, Raul, *Postolache, Octavian **C1.119**
13:20-13:45	Paper: **Enriching Location-based Games with Navigational Game Act** Authors: Nadarajah, Stephanie, Overgaard, Benjamin, Pedersen, Peder, Schnatterbeck, Camilla, *Rehm, Matthias **C1.117**	Paper: **Pairing Craft-making with Mandarin eBooks: An Investigati...** Authors: *Tan, Wil-Kie, Wang, Stephen, Jia, Janet, Jeff

EAI Innovation
@EAIchannel

On a quest to protect industrial control systems from cyber attacks: bit.ly/29VGVwe

8h

EAI Innovation Retweeted

Urban Experience
@UrbExperience

The latest Urban Experience! paper.li/UrbExperience/… Thanks to @EAIchannel @LaraMastra @aborgnino #agrogeneration #roma

Sadiq: London IS still 'open for…
www.standard.co.uk Leading fig…
paper.li

9h

EAI Innovation
@EAIchannel

Widespread adoption of semi-autonomous cars shows significant

		C1.119
13:45-14:10	Paper: **Toward a Decolonizing Approach to Game Studies: Philosoph...** Authors: *Cho, Hyunkyoung, Yoon, Joonsung **C1.117**	Paper: **CollaTrEx – Collaborative Context-Aware Mobile Training a...** Authors: *Botev, Jean, Marschall, Ralph, Rothkugel, Steffen **C1.119**
14:10-14:35	Paper title: **Increasing the Perceived Camera Velocity in 3D Racing Game...** Authors: *Holm, Kristoffer Lind, Skovhus, Nicolai, Kraus, Martin **C1.117**	Paper: **Assessment of Stand-Alone Displays for Time Management in...** Author: *Frimodt-Møller, Søren **C1.119**
14:35-14:50	**COFFEE/TEA BREAK**	**COFFEE/TEA BREAK**
14:50-15:15 Session chair C117AntoniJaume-i-Capó	Paper: **Sandtime- A Tangible Interaction Featured Gaming Installa...** Authors: *Yang, Chulin, Wang, Stephen, Jia **C1.117**	Paper: **Sandtime- A Tangible Interaction Featured Gaming Installa...** Authors: *Yang, Chulin, Wang, Stephen, Jia **C1.117**

societal and economic benefits:bit.ly/2a4KJuX

10h

15:15-15:40	Paper: **The Imitation Game to Cultural Heritage: A Human-like Int...** Authors: *Marulli, Fiammetta, Vallifuoco, Luca **C1.117**	Paper: **The Imitation Game to Cultural Heritage: A Human-like Int...** Authors: *Marulli, Fiammetta, Vallifuoco, Luca **C1.117**
15:40-16:05	Paper: **Enhancing the Multisensory Environment with Adaptive Game...** Authors: *Challis, Ben, Kang, Angela, Rimmer, Rachel, Hildred, Mark **C1.117**	Paper: **Enhancing the Multisensory Environment with Adaptive Game...** Authors: *Challis, Ben, Kang, Angela, Rimmer, Rachel, Hildred, Mark C1.117
16:05-16:30	Paper: **Investigating the Effect of Scaffolding in Modern Game De...** Authors: *Jensen, Kasper, Halkjær, Kraus, Martin **C1.117**	Paper: **Investigating the Effect of Scaffolding in Modern Game De...** Authors: *Jensen, Kasper, Halkjær, Kraus, Martin **C1.117**

19:00-21:00 (TO BE CONFIRMED)	**WELCOME RECEPTION: FISKERI- OG SØFARTSMUSEET (FISH & SHIPPING MUSEUM)** **SALTVANDSAKVARIET** **(SALT WATER AQUARIUM)** TARPHAGEVEJ 2 6710 ESBJERG PHONE: 7612 2000	**WELCOME RECEPTION: FISKERI- OG SØFARTSMUSEET (FISH & SHIPPING MUSEUM)** **SALTVANDSAKVARIET** **(SALT WATER AQUARIUM)** TARPHAGEVEJ 2 6710 ESBJERG PHONE: 7612 2000

TIME	**TUESDAY**	
	ARTSIT	**DLI**
09:00-09:15	Welcome: Tony & Eva **C1.119**	Welcome: Eva & Tony **C1.119**
09:15-10:00 Session chairs Tony Brooks Eva Brooks	**KEYNOTE**: Sudarshan Khanna&Surabhi Khanna Title: **Toys & Tales with Everyday Materials:** The Relevance of Ingenious, Playful Ideas for Design, Learning and Innovation **C1.119**	**KEYNOTE**: Sudarshan Khanna&Surabhi Khanna Title: **Toys & Tales with Everyday Materials:** The Relevance of Ingenious, Playful Ideas for Design, Learning and Innovation **C1.119**
10:00-10:15	**COFFEE/TEA BREAK**	**COFFEE/TEA BREAK**
10:15-10:40 Session chair	Paper: **Multi-Kinect Skeleton Fusion for Enactive Games**	**INVITED TALK**: Hildegunn Juulsgaard Johannesen Title: **Playful Learning and**

	C117 Paul McKevitt C119 Eva Brooks + Frimodt-Møller, Søren	Authors: Støvring, Nikolaj, Marimo, Kaspersen, Esbern, Torgard, Korsholm, Jeppe, Milling, Najim, Yousif, Ali Hassan, Makhlouf, Soraya, Khani, Alireza, &*Erkut, Cumhur **C1.117**	**Digital Technology:** **C1.119**
	10:40-11:05	Paper: **Analysing Emotional Sentiment in People's YouTube Channel...** Authors: *Mulholland, Eleanor, McKevitt, Paul, Lunney, Tom Schneider, Karl-Michael **C1.117**	Paper: **Mobile Device Applications for Head Start Experience in Music** Authors: *Chung, Szu-Ming, Wu Chun-Tsai **C1.119**
	11:05-11:30	Paper: **The Effect of Interacting with Two Devices when Creating...** Authors: Bork, Andreas, Heldbjerg, Bech, Christoffer, *Memborg, Jakob, Birch, Rosenlund, Lasse, Schøne, Kraus, Martin **C1.117**	Paper: **A Multimodal Interaction Framework for Blended Learning** Authors: *Triantafyllidis, Georgios, Vidakis, Nikolaos, Kalafatis, Kostas **C1.119**

11:30-11:55	Paper:**Widening the Experience of Artistic Sketchbooks** Authors: *Christiansen, Henning, Laursen, Bjørn **C1.117**	Paper: **Considerations and Methods for Usability Testing with School Children** Authors: Hjortboe Andersen, Malene, *Khalid, Md. Saifuddin, Brooks, Eva Irene **C1.119**
11:55-12:30	**LUNCH BREAK**	**LUNCH BREAK**
12:30-12:55 Session chair C117 Ben Challis C119 Eva Brooks	Paper: **An Adaptation Framework for Turning Real-Life Events into**... Authors: Therkildsen, Sacha, Kjærhus, Bunkenborg, Nanna, Cassøe, *Juel Larsen, Lasse **C1.117**	**12:30-13:20** **Workshop:** Sudarshan Khanna, Surabhi Khanna **Toys & Tales Design** Hands-on to experience the points discussed in the keynote presentation **C1.119**
12:55-13:20	Paper: **Emotion Index of Cover Song Music Video Clips based on** Authors: *Triantafyllidis, Georgios, Kavalakis, Georgios, Vidakis, Nikolaos	**Workshop:** Sudarshan Khanna, Surabhi Khanna **C1.119**

	C1.117	
13:20-13:45	Paper: **The Opportunities of Applying the 360° Video** Authors: *Stolyarov, Denis, Borisov, Nikolay, Smolin, Artem, Shcherbakov, Pavel, Trushin, Vasiliy **C1.117**	**14:50-16:05** Panel track: **Design of interactive environments for inclusion** **Chair: Eva Brooks** Paper: **Learning Together Apart – the Impact on Participation whe...** Authors: *Sorensen, Elsebeth, Korsgaard, Andersen, Hanne, Voldborg Paper: **Learning by Designing Interview Methods in Special Education** Author: *Jönsson, Lise Paper: **Powerlessness or Omnipotence – the Impact of Structuring ...** Authors: Andersen, Hanne, Voldborg, *Sorensen, Elsebeth, Korsgaard **C1.119**

13:45-14:10	Paper: **Pyramid Algorithm Framework for Real-Time Image Effects** Authors: ArbuésSangüesa, Adrià, Ene, Andreea-Daniela, Jørgensen, Nicolai, Krogh, Larsen, Christian, Aagaard, *Michelsanti, Daniel, Kraus, Martin **C1.117**	Panel track **Design of interactive environments for inclusion** **C1.119**
14:10-14:35	Paper: **Multimodal Detection of Music Performance for** Authors: Oxholm, Esben, Hansen, Ellen, K., *Triantafyllidis, Georgios **C1.117**	Panel track **Design of interactive environments for inclusion** **C1.119**
14:35-14:50	**COFFEE/TEA BREAK**	**COFFEE/TEA BREAK**
14:50-15:15 Session chair C117 Michael, Boelstofte C119	Paper: **Engaging with the Intangible Cultural Heritage of the City** Authors: *Rehm, Matthias, Rodil, Kasper **C1.117**	**14:50-16:05** **ROUND TABLE DISCUSSION:** **CO-CREATION: DESIGNING FOR INCLUSIVE LEARNING – EMPOWERING YOUNG INNOVATORS** **C1.119**

Eva Brooks		
15:15-15:40	Paper: **Aesthetic Computing for Representation of the Computing** Author: *Jeon, Myounghoon **C1.117**	**ROUND TABLE DISCUSSION** **C1.119**
15:40-16:05	Paper: **AcuTable: A Touch-enabled, Actuated Tangible User Interface** Authors: Dibbern, Simon, Rasmussen, Kasper, Vestergaard, Ortiz-Arroyo, Daniel, Holte, Michael, Boelstofte **C1.117**	**ROUND TABLE DISCUSSION** **C1.119**
16:15-16:30	**CONCLUSIONS AND OUTLOOK** Tony & Eva **C1.119**	**CONCLUSIONS AND OUTLOOK** Eva & Tony **C1.119**
19:00-21:00	**GALA DINNER: DRONNING LOUISE** Torvet 19 (main square)	**GALA DINNER: DRONNING LOUISE** Torvet 19 (main square)

	6700 Esbjerg Phone: 7513 1344	6700 Esbjerg Phone: 7513 1344

DLIROUND TABLE DISCUSSION:

CO-CREATION: DESIGNING FOR INCLUSIVE LEARNING – EMPOWERING YOUNG INNOVATORS

Creativity, innovation capacity, problem-solving skills, being critical and reflective, participation, skills related to teamwork, and sharing information and knowledge, might all be key competence areas starting from young children. From this perspective, educational settings and non-formal ways of learning should empower children with the skills they need from early on in life to become young innovators. Empowering the young through skills and competencies for creativity, innovation, and participation is important to build more inclusive societies, offering opportunities to all in order to address inequalities. Technologies, processes, resources, and open environments can support and encourage young people' experimentation and participation through inclusion and interdisciplinary approaches. The DLI round table will discuss these challenges by addressing questions of how learning and teaching technologies and activities can be designed to foster creativity and innovation related skills.

Speakers: please attend auditorium prior to session start to ensure all is good with your presentation and media files on the presentation computer or if using your own to ensure a hot swap for minimal time used in change-over – thanks in advance

Session chairs: Please attend the auditorium prior to your session to coordinate with speakers that they are all there and set up. You will have a 5 minute and 1 minute sign to show to speakers. It is imperative that each time slot is strictly adhered to – thanks in advance.

The welcome reception (2/5) at FISKERI- OG SØFARTSMUSEET (FISH & SHIPPING MUSEUM) : The museum is located outside the city along the coast – about a five minute taxi ride from downtown.

The Gala Dinner (3/5) at the DRONNING LOUISE (Queen Louise Restaurant): The restaurant is located at the main square downtown.

Session chairs confirmed:ArtsIT

CumhurErkut

GeorgiosTriantafyllidis

AntoniJaume-i-Capó

Paul McKevitt

Uriel Martinez Hernandez

Ben Challis

Session chairs confirmed: DLI

SørenFrimodt-Møller

EAI Institutional Members

View all EAI Members

Unione Europea
F E S R

Ministero dello Sviluppo
Economico

Provincia Autonoma
di Trento

Investiamo nel vostro futuro

Programma Operativo 2007-2013
Fondo Europeo di Sviluppo Regionale (FESR)

Emotion Index of Cover Song Music Video Clips based on Facial Expression Recognition

G. Kavalakis [(1)], N. Vidakis [(1)] and G. Triantafyllidis [(2)]

(1) Informatics and Multimedia Dept, TEI of Crete, Greece
(2) Medialogy Section, ADMT, Aalborg University Copenhagen, Denmark

gkavalakis@gmail.com, nv@ie.teicrete.gr & gt@create.aau.dk

Abstract. This paper presents a scheme of creating an emotion index of cover song music video clips by recognizing and classifying facial expressions of the artist in the video. More specifically, it fuses effective and robust algorithms which are employed for expression recognition, along with the use of a neural network system using the features extracted by the SIFT algorithm. Also we support the need of this fusion of different expression recognition algorithms, because of the way that emotions are linked to facial expressions in music video clips.

Keywords: Facial Expression Recognition, SIFT, Cover song video clips

1 Introduction

Human expression recognition is a fundamental problem in computer vision attracting great interest from the research community over the last years. Human expressions involve facial expressions, gestures, voices, etc. In this work, we will focus on the facial expression recognition which is the task of automatically identifying and classifying expressions in an image or video sequence. This is still a difficult task for computer vision to perform, although humans recognize facial expressions without effort or delay.

It is also known that music is a basic way of expressing human emotions. But there is somehow a rather different way of face expressing emotions in music performances and in music video clips compared to the same emotions in everyday life. When someone is singing or playing music, he/she is usually more expressive and also may employ different or more intense or even new expressions for specific emotions. For example, a singer looking down usually indicates a sad expression, while this not always true in everyday life expressions. In this context, this paper considers a specific and interesting case of facial expressions recognition: recognition in music video clips of cover songs, aiming at producing an emotion index, labeling the video clip.

The proposed scheme employs already known and used methods of classifying emotions such as the Logistic Regression (LogReg) [1], the Classification and Regression

Trees (CR-tree) [2], Linear discriminant analysis (LDA) [3], k-Nearest neighbor (k-NN) [4] and Quadratic discriminant analysis (QDA) [5]. The paper also suggests the use of SIFT feature extraction algorithm [6] for emotion classification. SIFT is an algorithm in computer vision that detects and describes local features in images. In this context, for any face in an image, interesting points can be extracted to provide a "feature description" of the face. This description can then be used to recognize the facial emotions in the image with a use of a neural network.

Before processing the facial expressions recognition, an algorithm for face detection should be employed, in order to detect faces that will be evaluated accordingly. However, detecting faces in music video clips is also a challenging task. Face position, lighting, occlusions, video quality are factors that affect the face detection performance.

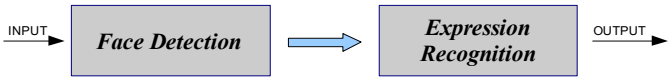


Fig. 1. General scheme

For the facial expression recognition task, we employ a novel scheme with a fusion decision system for each detected face in the music performance. More specifically, it fuses the aforementioned expression recognition techniques for getting a more reliable final decision regarding the emotion index. This fusion of such techniques improves the efficiency of the system, since the facial expressions during performing music may be different or even more difficult to be classified compared to the everyday expressions, because of the way that emotions are linked to facial expressions while performing music.

After this short introduction the rest of the paper is organized as follows: In the following chapter, we present the methodology and the algorithms of the proposed system, next the experimental results for the case of cover song music video clips are presented and finally the conclusions and future work.

2 Methodology

2.1 General scheme

The program's input is a music video clip of a cover song. The general scheme is illustrated in Figure 1. A group of frames is selected (eg 5 frames per second) and processed with the face detection algorithm. If the detected face is considered as acceptable (see next section for more details) for the expression analysis, this face image is stored

for further processing. This includes the expression analysis which fuses different methods and produces a decision for that face image. This decision is an emotion index. If the detected face is not acceptable or there is not any detected face, the algorithm checks the next group of frames until it finds an acceptable face image for processing.

2.2 Face Detection

The first step of the proposed scheme is the detection of faces within the video frames of a musical video clips. A critical issue is the question if the detected faces are suitable for further (expression) analysis. So it is essential that the basic features on the faces (i.e. two eyes and mouth) should be detected, helping in extracting information. In this context, the face detection algorithm of [7] was employed, which uses a cascaded classifier. In this technique of face detection, we also added the constraint of detecting the two eyes and the mouth, since these are the main features which export the information for the facial expression. Once we detect the two eyes and the mouth, this face image is accepted and we proceed to the next step of the expression recognition.

2.3 Expression Recognition

The facial expression recognition has been a subject of research in the computer science for a long time and there is a lot of research on this area. However, the classification rules are somehow different in recognizing expressions of an artist in a music video clip, since an artist playing music or singing a song, presents specific facial expressions, according to his/her emotions, which may differ in a way from the everyday expressions. This fact makes the expression recognition even harder and proves the correctness of the fusion approach which is suggested in this paper. Some more details about the algorithms we used for expression recognition:

Logistic Regression (LogReg) [1]**:** The logistic regression analysis offers an elegant possibility of examining the influence of several (of quantitative or qualitative) arguments or "factors of risk". The idea of the logistic regression is based on the conception that the probability of an event with an involution model can be functionally described. The influence of the arguments can be modeled directly. A further advantage is that these variables can be usually transferred in their original form to the model. Besides, only the involution coefficients must become estimated, which reduces the number of necessary statistic tests.

Classification and Regression Trees (CR-Tree or CART) [2]: The CR-Tree decision tree is a binary recursive partitioning procedure capable of processing continuous and nominal attributes as targets and predictors. Data are handled in their raw form; no binning is required or recommended. Beginning in the root node, the data are split into two children, and each of the children is in turn split into grandchildren. Trees are grown to a maximal size without the use of a stopping rule; essentially the tree-growing process stops when no further splits are possible due to lack of data. The maximal-sized tree is then pruned back to the root (essentially split by split) via the novel method of cost-complexity pruning. The next split to be pruned is the one contributing least to the

overall performance of the tree on training data (and more than one split may be removed at a time). The CR-Tree mechanism is intended to produce not one tree, but a sequence of nested pruned trees, each of which is a candidate to be the optimal tree.

Linear discriminant analysis (LDA) [3]**:** Linear Discriminant Analysis (LDA) is a method of finding a linear combination of variables which best separates two or more classes. In itself LDA is not a classification algorithm, although it makes use of class labels. However, the LDA result is mostly used as part of a linear classifier. The other alternative use is making a dimension reduction before using nonlinear classification algorithms.

k-Nearest neighbor (k-NN) [4]: is a method for classifying objects based on closest training examples in the feature space. k-NN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The k-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbor.

Quadratic discriminant analysis (QDA) [5]**:** is one of the most commonly used nonlinear techniques for pattern classification. In the QDA framework, the class conditional distribution is assumed to be Gaussian, however, with an allowance for different covariance matrices. In such cases, a more complex quadratic boundary can be formed. It is therefore reasonable to believe that QDA better fits the real data structure. However, due to the fact that more free parameters are to be estimated (C covariance matrices, where C denotes the number of classes) compared to those in an LDA-based solution (1 covariance matrix), QDA is more susceptible to the so-called small sample size (SSS) problem where the number of training samples is smaller or comparable to the dimensionality of the sample space.

Scale-invariant feature transform (SIFT) [6]**:** is an algorithm in computer vision to detect and describe local features in images. The algorithm was published by David Lowe in 1999 and computes scale-space extrema of the space Laplacian, and then samples for each one of these extrema a square image patch. SIFT method is actually proposes descriptors that are invariant to image translations and rotations, to scale changes (blur), and robust to illumination changes. It is also surprisingly robust to large enough orientation changes of the viewpoint (up to 60 degrees). The initial goal of the SIFT method is to compare two images (or two image parts) that can be deduced from each other (or from a common one) by a rotation, a translation, and a zoom [8]. The method turned out to be also robust to large enough changes in view point angle, which explains its success also in object recognition [9].

So, the proposed scheme employs the SIFT method for expression recognition since the SIFT feature keypoints are highly distinctive, in the sense that a single feature can be correctly matched with high probability against a large database of features from many images. So, we actually build a neural network and train it by using as input the SIFT feature keypoints of several face images and as output the respective emotion.

Regarding this SIFT-based neural network, the problem that we face is that there are many feature keypoints which are not associated to the expressions, but only with the

face characteristics. Therefore, the efficient choice of the face regions is critical. Moreover, and to solve this problem, we may perform a different SIFT-based neural network to the eyes region (SIFT top) and to the mouth region (SIFT down), since these regions are greatly affected by facial expressions.

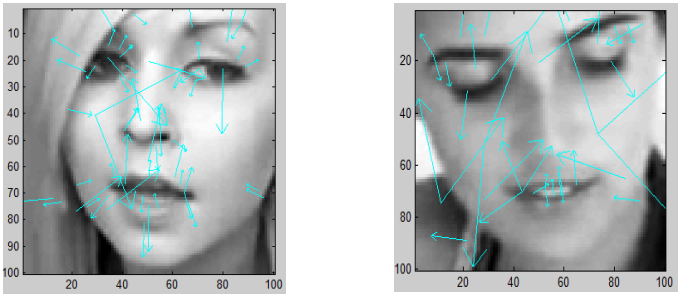


Fig. 2. SIFT features keypoints

The facial expression flowchart is depicted in Figure 3. Each one of the faces which were detected in the face detection part of the proposed scheme is analyzed by using all six algorithms presented above. These outcomes are fused to result the final decision for the expression recognition of the specific face image which was analyzed. Next, we apply the same algorithm to other detected faces and we finally produce the emotion index of the whole video.

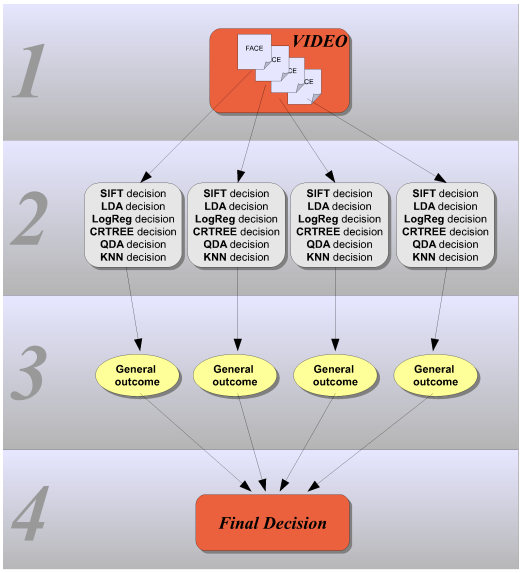


Fig. 3. Emotion Recognition Analysis

3 Experimental Results

Figure 4 depicts a screenshot of the interface (which is still in a beta version). The user may select the music video clip that will be processed for the facial expression recognition. The interface has been designed in such a way where there is a window playing the video clip, while the faces detected to be suitable for the expression analysis are shown below this video window. The right panel of the interface presents the results (in this case we use the happy or sad emotions) obtained from each face image according to each algorithm.

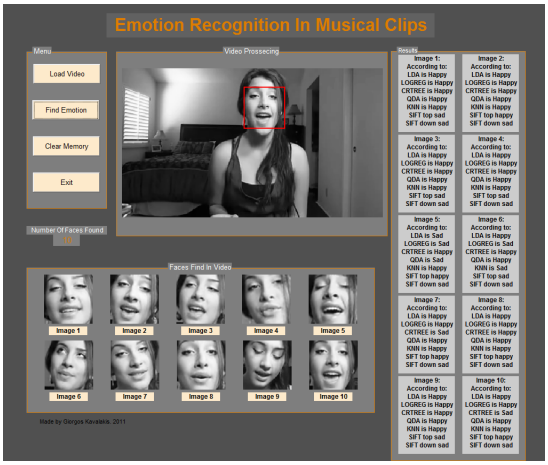
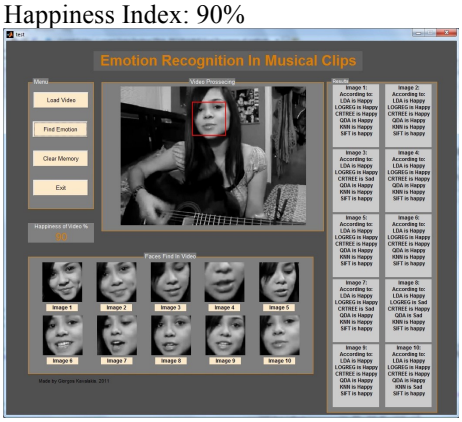


Fig. 4. Program's Interface

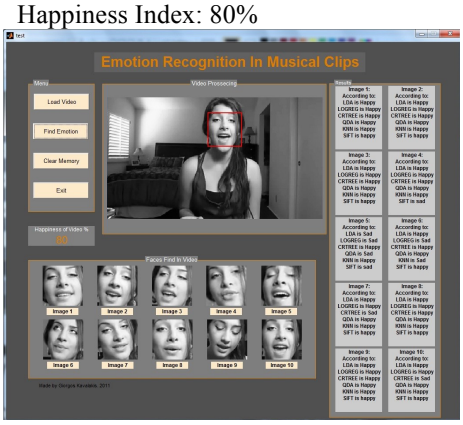
So taking into consideration the errors and the difficulty of each algorithm in obtaining results, we can reach to a final emotion expression decision for each face image (see Figure 3). Then, considering these decisions, we can draw a degree (index) of the emotion (e.g. happiness or sadness) for each video.

In our experiments we used 4 randomly selected cover song video clips from YouTube. We applied the suggested scheme and concluded to an overall index regarding the emotion of happiness. Results are shown in Table 1. Results greater than 70% indicate a happy song, lower than 30% indicate a sad song, while results around 50% indicate a rather neutral song.

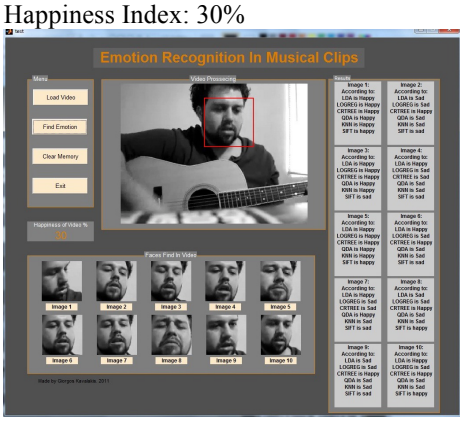
Cover Song 1: Song Title: I am yours
www.youtube.com/watch?v=dQz0U6LV-ME
Happiness Index: 90%



Cover Song 2: Song Title: Just the way you are
www.youtube.com/watch?v=dW6VGm3I8jY
Happiness Index: 80%



Cover Song 3: Song Title: Creep
www.youtube.com/watch?v=Ph1GaDwdecI
Happiness Index: 30%



Cover Song 4: Song Title: The writer
www.youtube.com/watch?v=92dxTXL46h4
Happiness Index: 60%

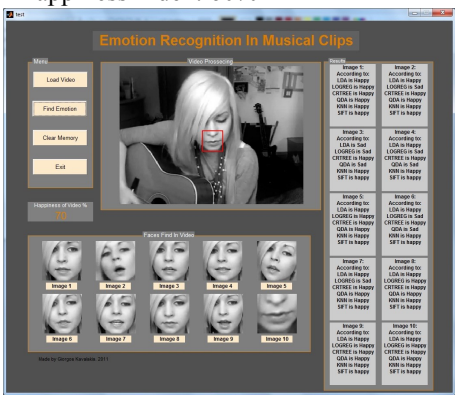


Table 1. Emotion (happiness) index on 4 Youtube cover songs videos

4 Conclusions and Future work

Automatic emotion labeling of music video clips based on the facial expression recognition is a challenging task, but also a very useful tool. There are internet services offering huge collections of music video clips that may use such methods for classifying the videos in a data base accordingly. Also users may use the proposed scheme for organizing and sorting their personal collections of music video clips. In this way and

depending on the user's mood a music player may automatically choose the right music video clip for playing.

In this context, this paper presents a preliminary study of creating an emotion index of cover song video clips by recognizing facial expressions. Improvements, such as more effective segmentation of areas such as the mouth and the eyes and the selection of proper feature keypoints may produce more accurate SIFT-based neural network's decisions. Also, the creation of a more effective data base is one of our priorities, aiming to facial expression recognition methods to produce better results. Finally, the range of the emotions that can be recognized from the system is about to grow and cover the real range of emotions of a music video clip.

5 References

[1]. Agresti, Alan: Categorical DATA analysis. New York: Wiley, 1990.

[2]. D. Steinberg, P. Colla - San Diego, CA: Salford Systems, 1997

[3]. S. Balakrishnama and A. Ganapathiraju, "Linear Discriminant Analysis—A Brief Tutorial,"Mississippi State Univ., Starkville, MS, 1998.

[4]. Hall P, Park BU, Samworth RJ, "Choice of neighbor order in nearest-neighbor classification". Annals of Statistics, 2008.

[5]. Jie Wang, Konstantinos N. Plataniotis, Juwei Lu, Anastasios N. Venetsanopoulos: Kernel quadratic discriminant analysis for small sample size problem. Pattern Recognition 41(5), pp 1528-1538, 2008.

[6]. David G. Lowe, "Distinctive image features from scale-invariant keypoints" International Journal of Computer Vision, 60, pp. 91-110, 2004.

[7]. W. Kienzle, G. Bakir, M. Franz and B. Scholkopf: Face Detection - Efficient and Rank Deficient. In: Advances in Neural Information Processing Systems 17, pp. 673-680, 2005

[8]. J.J. Foo and R. Sinha. Pruning SIFT for scalable near-duplicate image matching. Proceedings of the eighteenth conference on Australasian database-Volume 63, pages 63–71, 2007.

[9]. P. Moreels and P. Perona. Common-frame model for object recognition. Proc. NIPS, 2004.