Fake Emotions

Ikechukwu Ofodile, *Student Member, IEEE,* Kaustubh Kulkarni, *Member, IEEE,* Ciprian Adrian Corneanu, *Student Member, IEEE,* Sergio Escarela, *Member, IEEE,* Sylwia Hyniewska, *Member, IEEE,* Jüri Allik, and Gholamreza Anbarjafari, *Senior Member, IEEE*

Abstract—The abstract goes here.

Index Terms—Fake emotion recognition.

1 Introduction

Humans are skilled at simulating and observers tend to perform barely above chance level when asked to detect such insincere behaviours [1] [2], particularly when relying on visual cues only [3]. An emotional expression can be considered insincere or fake when not accompanying a relevant genuine emotional state. It has been argued that deceivers experiencing emotions would be betrayed by the leakage of their genuine emotional states through their non-verbal behaviour [4] [5]. Numerous studies involving videos of genuine and deceitful behaviours focused on cues of deceit in directed interviews, eg. [2]. Rare are the studies that have studied cues of deceit while controlling for the emotional state of expressers. These studies mostly originated based on Duchenne's work, who was the first to differentiate facial actions observed in genuine and in false displays of emotions [6]. The most notorious example of a genuine expression is the involvement of the orbicularis oculi muscle (causing lifting of the cheeks and crow's feet around the eyes) together with the zygomaticus major muscle (pulling of the lips upwards, producing a smile). This combination is since known as the Duchenne smile and considered the genuine expression of enjoyment. It has been argued that certain facial muscles cannot be activated voluntarily, while appearing in spontaneous expressions [4]. Although some authors argued that orbicularis oculi is absent from fake or voluntary expressions of enjoyment, one study showed that when looking at 105 posed smiles, 67% were accompanied by the orbicularis oculi activation [7]

Social experiment conducted [8] shows that social rejection improves detection of real and fake smiles. Participants were shown faces exhibiting either Duchenne (considered a "true" smile) or non-Duchenne smile and had to decide whether each was fake or real having been surveyed through an essay to determine a degree to which they feel socially rejected.

In this paper, we present our approach to identifying insincere expressions of emotions. The paper is organised as follows: Section 2 describes Related Work; Section 3 explains the Proposed Method...

2 Related Work

[8]–[17]

Several literature works have discussed and tried to overview a growing understanding into emotion recognition. An overview of various facial and emotional states is given in [10] with parameters on speech and different emotions based on psychological structure and boundaries. An automatic emotion recognition system that employs neural network as a learning algorithm for input streams was developed in [9] based on a hybrid system Emotionally Rich Man-machine Intelligent System (ERMIS) capable of recognising peoples emotion based on speech and facial expression.

Major approaches based on the computer vision perspective relating to MultiModal Human Computer Interaction (MMHCI) and its applications had been reviewed in [11] with emphasis on body, gesture, gaze and affective interaction dealing with facial expression recognition and emotion recognition in audio. The author describes the process of large scale body movements, gestures and gaze under vision techniques and proposed a dynamic Markov network that uses mean field Monte Carlo algorithm as well as obtained mathematical model that consider both spatial and temporal characteristics of hand gestures. However, this may not be efficient with real time systems due to complexity in analysis and considering that only specific commands based on hand posture are recognised. Two common scales were used to describe emotions with respect to the affective interaction: Valence which describes a variation in pleasant feelings, with pleasant (happiness) on one end, unpleasant (disgust)

- *I. Ofodile and G. Anbarjafari are with the the iCV Research Group, Institute of Technology, University of Tartu, Tartu, Estonia. E-mail: shb@icv.tuit.ut.ee*
- *S. Escarela is with Computer Vision Center and University of Barcelona, Barcelona, Spain. E-mail: sergio@maia.ub.es*
- *S. Hyniewska is with University of Bath, Bath, UK. E-mail: s.hyniewska@bath.ac.uk*
- *J. Allik is with Department of Psychology, University of Tartu and The Estonian Center of Behavioral and Health Sciences, Estonia. E-mail: juri.allik@ut.ee*
- *G. Anbarjafari is also with Department of Electrical and Electronic Engineering, Hasan Kalyoncu University, Gaziantep, Turkey.*

Manuscript received April 19, 2005; revised August 26, 2015.

on the other. The other scale is arousal whereby sadness has low arousal and surprise has high arousal.

The stages of identification of human emotions from facial expression as indicated in [11] and [12] involve segmentation and then classification based on either feature or region of interest. The dynamic rather than static form of analyses using several video frames to perform the classification by analysing temporal patterns of the region of interest has proven to be more effective. The authors in [12] employ fuzzy logic approach to analyse facial expressions by segmenting and localizing individual frames into regions of interest. They also provided a fuzzy scheme for controlling the transition of emotion dynamics towards a desired state and uses fuzzy-c-means (FCM) clustering for segmentation of the facial images into important regions of the face such as mouth, eye and eyebrows, then obtained the fuzzy emotions using a fuzzy reasoning algorithm.

The use of an ensemble of features consisting of hierarchical Gaussianization (HG), scale-invariant feature transform (SIFT) , and some coarse motion features (MF) was introduced in [13] with a form of preprocessing involving de-interlacing, key-point or area detection and alignment where active shape model (ASM) is performed to detect facial feature points. The ensemble of features are extracted from the faces detected which includes the SIFT at selected key points, HG features which is robust to occlusion and alignment problems and then MFs for a better validation on accuracy of the training data. The method uses person-dependent and person-independent classifier method for training with the former having a better performance.

Based on the general implementation of automated facial expression and recognition analysis (FERA), [14] proposes a multi-kernel (MK) learning system by initially combining different features to detect facial action units (AU) in images of faces. The system uses one MK support vector machine (SVM) for training of each AU. The training samples used in the MK SVMs are computed using local Gabor binary pettern (LGBP) histograms and active appearance model (AAM) coefficient vectors and then an average filtering is done on the output of each SVM classifier for temporal information.

The idea on ”fake emotion” is also discussed in [15] with analysis to suggest that cheaters have the ability to disguise negative emotional expressions that signal their uncooperative attitude. It involved evaluating emotional intensity and trustworthiness for facial photographs of cheater and cooperators. Based on the high intensity shown by cheaters on fake smile, cheater detection based on processing of negative facial expression may be inaccurate. The study concludes that cheaters were rated to be more emotionally expressive, where in the case of angry expression, they were judged to be more aggressive. A study [16] aimed at investigating the age differences in recognising facial and vocal emotion expression revealed that the largest age differences were found for negative emotions. The study noted that asides age, education, sex, personality traits also have their own independent contribution to emotion recognition.

A classifier method was proposed and results presented in [17] with the aim of distinguishing between posed and spontaneous smiles. This method is based on the hypothesis that the eyelid movements can identify smiles mainly by the activity of the orbicularis oculi muscle. The movement is modeled by continuous hidden Markov models (CHMM). Based on their result of classification rates with different feature groups, the eyelid movement provided the highest classification rates. The result also noted that simple classifiers such as k-nearest neighbours and naive Bayes could be used instead of CHMM which is time consuming during the training phase.

2.1 Temporal

[18]–[22]

Using temporal image sequence, a facial recognition method was developed in [18] by applying Mutual Subspace Method (MSM) to define similarities between subspaces. First the face is identified by forming a subspace with the image sequence where two subspaces are obtained for distinguishing between facial and non facial features. The MSM defines similarities by the angle between two subspaces based on several input images. Results show that the MSM is stable and has high similiarity in comparison with the conventional subspace method which represents each feature by its own subspace based on ”single shot”. This method was influenced by observation that the recognition system will be unreliable due to failure in real time image processing or a noisy input image based on the single shot image recognition input. The temporal image sequence was also used with HMM classifier computation in [19] [20]. Cohen *et al.* proposed a multi-level HMM classifier for auto segmenting and FER while comparing different Bayesian classifiers. The Bayesian network classifier is used to compute the probability of set labels of observed features and classify the features with most probable level. The authors in [20] proposed an algorithm to adapt HMM in order to enhance modelling over time which according to their results had a better error rate in comparison with temporal HMM or the baseline image-based algorithm which they used based on individual principal component analysis (PCA).

A method for AU detection [21] and its temporal dynamics is presented based on changes in position of facial points tracked from a face profile view image sequence. The facial points are tracked using particle filtering. In the paper, Pantic *et al.* provided an overview of the procedure of automatic detection of AUs.

A 3D aging model for face recognition is developed in [22] by learning aging pattern based on PCA coeffiecients in seperated facial textures. The model developed in 3D domain for the aging simulation from 2D database aids better compensation for pose and lightening variations. The performance of the method is evaluated by comparing face recognition before and after aging simulation. Although the method had failure in small percentage of cases in matching after the simulation as discussed, it is capable of handling growth and development and adult face aging.

3 The Proposed Method

4 Dataset

The SASE-FE database consist of XX:total number of video different samples videos which have been recorded by

GoPro-Hero camera with XXX spatial resolution and 100 frames per second. XXX people of age 19-36 were recorded under the following set up. The people were asked to act the six basic emotions, namely, Happy, Sad, Anger, Disgust, Contempt and Surprise. In order for the candidate to have the correct feeling, they were shown some videos in line with the asked emotions so they will act the emotion more naturally. In all the set up the candidate was asked to start from a neutral emotion. The length of each acted emotion is about 3-4 seconds. After each acted emotion, the candidate was asked to be neutral and then they were asked to act an emotion which was an opposite of the recently acted emotion. This was achieved by asking people to act Happy after they were Sad, Surprise after they were Sad, Disgust after they were Happy, Sad after they were Happy, Anger after they were Happy, Contempt after they were Happy. None of the candidate were aware of that they will be asked to act an opposite emotion at all. The candidate's first two seconds of reaction that they tried to fake the opposite emotion has been recorded by the same device with the same configuration. As a result for each candidate we have collected 12 different videos which 6 of them are acted emotions and other 6 are fake emotions. For the purpose of labelling of the dataset, the sequence of returning to a natural state was introduced to show the variance from the acted emotion being displayed to the fake emotion to be used in the training and learning process. The process has been closely supervised by experimental psychologists so that the set up will results into realistic recording of fake emotions. In SASE-FE there are XX female and XX male samples.

5 Experimental Results and Discussions

5.1 Set up

6 Experimental Results

6.1 Discussions

We had decided to increase the number of samples for the database to 60+, in order to have about 1000 data of fake and real (acting) emotions for the learning process. The temporal learning will be done mostly using Convolutional Neural Network(CNN), however we will conduct further study on still images to show the importance of temporal information. An overview of recent research done in temporal analysis should be included [8]–[22]. The initial set up will be in the way that the system will learn two classes, namely real and fake and given the input data it decide which class we belong to. We will use leave-on-out strategy.

two setup

- fake vs not fake;
- effect of temporal information;
- change of spatial and temporal resolution and see their effect (25%, 50%, 75%);
- Leave-one-subject-out strategy.

7 Conclusion

The conclusion goes here.

Structure of the paper

8 Introduction

9 Related Work

10 Problem Formulation and Notation

- Let the image sequence be represented as a time-series, or sequence of vectors X, Y, Z
- $X_{1:T_X} = (x_1, \ldots, x_t, \ldots, x_{T_X})$where$x \in^k$

$X_{t_i:t_j}$
$1 \leq t_i \leq t_j \leq T_X$

11 Method

11.1 Dynamic Time warping Kernel

11.2 Fisher Kernel

In this section, we formulate the problem of incorporating the emotions as defined in the introduction into a Fisher vector framework. We first introduce the general concept of Fisher vectors and the formulas when Fisher vector is constructed from an Gaussian Mixture Model(GMM) or a Hidden Markow model. Following which we introduce the novelty of incorporating the actemes in the Fisher vector. Let $X = \{x_m\}_{m=1}^{M}$ be a set of M observations where $x_m \in \mathcal{X}$. Here,x_m is a d dimensional vector. We assume that the elements in $\mathcal{X}$ are drawn out of a probability density function (pdf) $P(X|\theta)$ where, θ is a vector that parametrizes the pdf and $\theta \in R^N$. *Fisher Score* measures how sensitive is the probability density function to the change in θ. It is the Jacobian of the log-likehood w.r.t to the model parameters θ

$$J_\theta^X = \nabla_\theta \log p(X|\theta) \tag{1}$$

11.2.1 Gaussian Mixture Models

In this case we assume that the observations in X are drawn from a Gaussian mixture model. Each GMM is learned in a supervised manner using the observations of the corresponding class. Let the per class GMM have K components. Assuming that the observations in X are statistically independent the log-likelihood of X given θ_c is:

$$\log P(X|\theta_c) = \sum_{m=1}^{M} \log \sum_{k=1}^{K} w_k^c \mathcal{N}(x_m; \mu_k^c, (\sigma_k^c)^2) \tag{2}$$

where, $\sum_{k=1}^{K} w_k^c = 1$ and $\theta_c = \{w_k^c, \mu_k^c, (\sigma_k^c)^2\}$. We assume diagonal covariance matrices. The parameters of the per-class GMMs are estimated with the Expectation maximization (EM) algorithm to optimize the maximum likelihood (ML) criterion. To keep the magnitude of the Fisher vector independent of the number of observations in X we normalize it by M. Following the formulation of [?] we can now write the formulas for the gradients of the log-likelihood $P(X|\theta_c)$ w.r.t to the individual parameters of the GMM as:

$$\mathcal{J}_{w_k^c}^X = \frac{1}{M\sqrt{w_k^c}} \sum_{m=1}^{M} \gamma_k(m) - w_k^c \tag{3}$$

$$\mathcal{J}_{\mu_k^c}^X = \frac{1}{M\sqrt{w_k^c}} \sum_{m=1}^{M} \gamma_k(m) \left(\frac{x_m - \mu_k^c}{(\sigma_k^c)^2} \right) \tag{4}$$

$$\mathcal{J}^{X}_{(\sigma^c_k)^2} = \frac{1}{M\sqrt{2w^c_k}} \sum_{m=1}^{M} \gamma_k(m) \left[\frac{(x_m - \mu^c_k)^2}{(\sigma^c_k)^2} - 1 \right] \quad (5)$$

where, $\gamma_k(m)$ is the posterior probability or the responsibility of assigning the observation x_m to component k. It is given as:

$$\gamma_k(m) = \frac{w^c_k \mathcal{N}(x_m; \mu^c_k, (\sigma^c_k)^2)}{\sum_{i=1}^{K} w^c_i \mathcal{N}(x_m; \mu^c_i, (\sigma^c_i)^2)} \quad (6)$$

Note that the normalizing factor other than M comes from the FIM. The Fisher vector is constructed by concatenating the gradients of the log-likelihood w.r.t. to the parameters of the class conditioned GMMs as shown in (**??**)

11.2.2 Hidden Markov Models

In the case when the observations are a time series $X = (x_1, x_2, \ldots, x_t, \ldots, x_{T_X})$ where $x_t \in R^d$. Let each class c be modelled by an HMM having $S \in \{s_1, s_2, \ldots, S\}$ states and each state's output distribution modelled by a GMM with K components, $\mathcal{N}(\mu^c_{s,k}, \Sigma^c_{s,k})$ where $k \in 1, \ldots, K$. Assuming the covariance matrix to be diagonal.

Gradient Formulas are given by:

$$\mathcal{J}^{X}_{\mu^c_{s,k}} = \gamma_{s,k}(m) \left(\frac{x_m - \mu^c_{s,k}}{(\sigma^c_{s,k})^2} \right) \quad (7)$$

$$\mathcal{J}^{X}_{(\sigma^c_{s,k})^2} = \gamma^c_{s,k}(m) \left[\frac{(x_m - \mu^c_{s,k})^2}{(\sigma^c_{s,k})^3} - \frac{1}{\sigma^c_{s,k}} \right] \quad (8)$$

Computing the posteriors with the forward backward algorithm and approximately using the Viterbi algorithm [**?**].

11.3 Recursive Neural Networks

11.4 Differential-RNN

12 Feature Descriptors and Encoding

12.1 Local feature descritpors

12.1.1 SIFT

12.1.2 Improved Dense trajectories

12.2 Convolutional Neural Network (Appearance)

12.3 Trajectory pooled features

12.4 Geometric features

In order to perform landmark localisation we have used the Continuous Supervised Descent Method (CSDM) [23] a cascaded regression approach exploiting the second partial derivative wrt. the main modes of variation of the features. It is a natural extension to the Supervised Descent Method (SDM) [24] and Global Supervised Descent Method (GSDM) [25] which adapts the regressors for individual instances to rotation and illumination changes.

Lets consider $\underline{X}^i \in {}^{n\times m}$ the m targets for each of n samples at a given cascade step i, $\Delta\underline{\Phi}^i \in {}^{n\times(k+1)} = \underline{\Phi}^i - \overline{\underline{\Phi}^i}$ the difference of the feature vectors of length k from the mean, with a column vector of ones added in order to account for the bias, and $\underline{R}^i \in {}^{(k+1)\times m}$ the linear regressor for each of the m parameters. Then the update formula for SDM can be expressed as follows:

$$\underline{X}^{i+1} = \underline{X}^i + (\underline{\Phi}^i - \overline{\underline{\Phi}^i})\underline{R}^i = \underline{X}^i + \Delta\underline{\Phi}^i\underline{R}^i \quad (9)$$

This can be seen as learning a linear approximation of the first-order partial derivatives for each parameter. To make this approximation, the slope is considered homogeneous for any point of the feature space. This assumption does not hold for most problems, where the gradient direction suffers from large variations on different locations of that space. On Global SDM [25] these variations are handled by partitioning the space into different regions and learning a linear regressor for each one.

In the moethod used [23] a continuous formulation is introduced, where a set of bases are learnt for the regressors, effectively learning a linear approximation of the second derivative. To do so, first a set of main modes of variation are learnt from either $\Delta\underline{X}^*$ or $\Delta\underline{\Phi}^i$ using Principal Component Analysis (PCA):

$$\Delta\widetilde{\underline{\Phi}^i} = \left[\Delta\underline{\Phi}^i\underline{P}_{1:l}, \underline{1}_n \right], \quad (10)$$

Where l represents the number of bases to learn and $\underline{P}_{1:l}$ is the projection matrix. $\underline{1}_n \in {}^{n\times 1}$ denotes an all-ones vector. Given that the total number of learnable parameters for one of m targets equals $p = (k+1)(l+1)$, learning the second derivative for all parameters ($l = k$) would drastically increase the problem dimensionality. Estimating the second derivative on the l main variation modes is a more treatable problem. Given one of the targets $\Delta\underline{X}^i_j \in {}^{n\times 1}$, its associated second order regressor is expressed as the solution to the following minimisation problem:

$$\arg\min_{\underline{R}^i_j} ||(\Delta\underline{\Phi}^i \circ (\Delta\widetilde{\underline{\Phi}^i}\underline{R}^i_j))\underline{1}_{(k+1)} - \Delta\underline{X}^i_j||^2_2 \quad (11)$$

At test time, the parameters are updated with the following equation:

$$\underline{X}^{i+1}_j = \underline{X}^i_j + (\Delta\underline{\Phi}^i \circ (\Delta\widetilde{\underline{\Phi}^i}\underline{R}^i_j))\underline{1}_{(k+1)} \quad (12)$$

This formula estimates the regressor weights and bias for the current value of the principal components $\widetilde{\underline{\Phi}^i}$, and applies it to the features.

12.5 Early Fusion of geometric features and Appearance features

12.6 Encoding

12.6.1 Bag-of-Words

12.6.2 Fisher Vectors

13 Experiments

13.1 Datasets

Acknowledgments

This work is supported Estonian Research Council Grant (PUT638), European Regional Development Fund and the Spanish Projects TIN2013-43478-P and TIN2016-74946-P.

References

[1] M. Hartwig and C. F. Bond Jr, "Why do lie-catchers fail? a lens model meta-analysis of human lie judgments," 2011.

[2] L. Ten Brinke, P. Khambatta, and D. R. Carney, "Physically scarce (vs. enriched) environments decrease the ability to tell lies successfully." *Journal of experimental psychology: general*, vol. 144, no. 5, p. 982, 2015.

[3] C. F. Bond and B. M. DePaulo, "Accuracy of deception judgments," *Personality and social psychology Review*, vol. 10, no. 3, pp. 214–234, 2006.

[4] C. Darwin, *The expression of the emotions in man and animals, New York: D*, 1872.

[5] M. G. Frank and P. Ekman, "The ability to detect deceit generalizes across different types of high-stake lies." *Journal of personality and social psychology*, vol. 72, no. 6, p. 1429, 1997.

[6] G. Duchenne de Bologne, "The mechanism of human facial expression (ra cuthbertson, trans.)," *Paris: Jules Renard*, 1862.

[7] K. L. Schmidt and J. F. Cohn, "Human facial expressions as adaptations: Evolutionary questions in facial expression research," *American journal of physical anthropology*, vol. 116, no. S33, pp. 3–24, 2001.

[8] M. J. Bernstein, S. G. Young, C. M. Brown, D. F. Sacco, and H. M. Claypool, "Adaptive responses to social exclusion social rejection improves detection of real and fake smiles," *Psychological Science*, vol. 19, no. 10, pp. 981–983, 2008.

[9] N. Fragopanagos and J. G. Taylor, "Emotion recognition in human–computer interaction," *Neural Networks*, vol. 18, no. 4, pp. 389–405, 2005.

[10] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal processing magazine*, vol. 18, no. 1, pp. 32–80, 2001.

[11] A. Jaimes and N. Sebe, "Multimodal human–computer interaction: A survey," *Computer vision and image understanding*, vol. 108, no. 1, pp. 116–134, 2007.

[12] A. Chakraborty, A. Konar, U. K. Chakraborty, and A. Chatterjee, "Emotion recognition from facial expressions and its control using fuzzy logic," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 39, no. 4, pp. 726–743, 2009.

[13] U. Tariq, K.-H. Lin, Z. Li, X. Zhou, Z. Wang, V. Le, T. S. Huang, X. Lv, and T. X. Han, "Recognizing emotions from an ensemble of features," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 4, pp. 1017–1026, 2012.

[14] T. Senechal, V. Rapp, H. Salam, R. Seguier, K. Bailly, and L. Prevost, "Facial action recognition combining heterogeneous features via multikernel learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 4, pp. 993–1005, 2012.

[15] M. Okubo, A. Kobayashi, and K. Ishikawa, "A fake smile thwarts cheater detection," *Journal of Nonverbal Behavior*, vol. 36, no. 3, pp. 217–225, 2012.

[16] A. Mill, J. Allik, A. Realo, and R. Valk, "Age-related differences in emotion recognition ability: a cross-sectional study." *Emotion*, vol. 9, no. 5, p. 619, 2009.

[17] H. Dibeklioglu, R. Valenti, A. A. Salah, and T. Gevers, "Eyes do not lie: spontaneous versus posed smiles," in *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 2010, pp. 703–706.

[18] O. Yamaguchi, K. Fukui, and K.-i. Maeda, "Face recognition using temporal image sequence," in *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*. IEEE, 1998, pp. 318–323.

[19] I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang, "Facial expression recognition from video sequences: temporal and static modeling," *Computer Vision and image understanding*, vol. 91, no. 1, pp. 160–187, 2003.

[20] X. Liu and T. Cheng, "Video-based face recognition using adaptive hidden markov models," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 1. IEEE, 2003, pp. I–340.

[21] M. Pantic and I. Patras, "Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 36, no. 2, pp. 433–449, 2006.

[22] U. Park, Y. Tong, and A. K. Jain, "Face recognition with temporal invariance: A 3d aging model," in *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on*. IEEE, 2008, pp. 1–7.

[23] M. Oliu, C. Corneanu, L. A. Jeni, J. F. Cohn, T. Kanade, and S. Escalera, "Continuous supervised descent method for facial landmark localisation," in *ACCV*, 2016.

[24] X. Xiong and F. De la Torre, "Supervised descent method and its applications to face alignment," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 532–539.

[25] ——, "Global supervised descent method," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 2664–2673.

PLACE PHOTO HERE

Ikechukwu Ofodile Biography text here.

Sergio Escarela Biography text here.

Jüri Allik Biography text here.

Gholamreza Anbarjafari Biography text here.