Pain Intensity Evaluation Through Facial Action Units

Zuhair Zafar
Dept. of Electrical Engineering, SBASSE,
Lahore University of Management Sciences,
Lahore, Pakistan
11060007@lums.edu.pk

Nadeem Ahmad Khan
Dept. of Electrical Engineering, SBASSE,
Lahore University of Management Sciences,
Lahore, Pakistan
nkhan@lums.edu.pk

***Abstract*—In this work we present a system that enables automatic estimation of Pain from image sequences with frontal views of faces. The system uses facial characteristic points to characterize different Action Units (AU) of pain and is able to operate in cluttered and dynamic scenes. Geometric features are computed using 22 facial characteristic points. We use k-NN classifier for classifying AU. Only action units relevant to pain are classified. Validation studies are done on UNBC McMaster Shoulder Pain Archive Database [8]. We also classify action unit intensities for evaluating pain intensity on a 16 point scale. Our system is simpler in design compared to the already reported works in literature. Our system reports AU intensities on a standard scale and also reports pain intensity to assess pain. We have achieved more than 84% accuracy for AU intensity levels and 87.4% *area under ROC curve* for pain assessment as compared to 84% of state-of-the-art scheme.**

Keywords—Pain Expression; Geometric Features; K-NN Classifier; Action Units; ROC curve

I. INTRODUCTION

Faces in our daily life convey a lot of information. Especially in face to face communication scenarios, facial expressions play a vital role in better conveying the intended meaning of message. Recognizing pain expression using image processing and computer vision techniques is not a straightforward job. Pain is assessed through self-report of a patient or by observer report. In case of young children self-report cannot be used, similarly for patients in transient states of consciousness, and those that require assisted breathing [17]. Observer report also has its own limitation. It is highly inefficient and impractical if the observer is required for a longer period of time which could be the case for a patient in an intensive care unit.

In order to address these shortcomings, automatic recognition of pain expression becomes an essential need especially for elderly care and by using computer vision and machine learning algorithms it is potentially possible. There are two main approaches in order to recognize facial expression: emotion detection on the base of the exhibited facial expression [4] while the other approach fundamentally focused on activation and deactivation of facial action units or simply action units (AUs) [5]-[7]. The latter methodology detects the activation of AU irrespective of the overall emotion. Later these AUs are deciphered into an expression. AUs are primarily the actions of human face muscles that correspond to a certain expression. Upward motion of an eyebrow is an AU1 and closing of eyes is AU43. Ekman and Friesen build up facial action coding system (FACS) for manual labeling of actions units. Out of 46 facial action units, any facial expression can automatically be recognized by the combination of 32 action units [9]. The rationale behind recognizing AUs is that this approach is independent of formulating higher order assessment. This higher order assessment is done by FACS system. In this paper we present a scheme for assessment of spontaneous pain by using geometric based features. AUs are then classified using K nearest neighbor (K-NN) classifier.

The outline of this paper is as follows: Section II describes literature survey of state-of-the-art schemes; Section III explains the proposed methodology in detail; In Section IV we assess the performance of the proposed methodology and discuss results; we summarized the paper in Section V.

II. LITERATURE SURVEY

Limited efforts have been made so far to recognize pain expression [2,3,10,11]. Recently, Kaltwang et al. [2] used appearance based features, discrete cosine transform (DCT) and linear binary patterns (LBPs), extracted from facial images of subjects displaying different intensities of pain. They have shown that combination of appearance features outperforms separately trained classifiers on different datasets. The work by Lucey et al. [3] also addresses AU and pain detection based on SVMs. They detect pain either directly using image features or by applying a two-step approach, where first AUs are detected and then this output is fused by Logistical Linear Regression (LLR) in order to detect pain. Littlewort et al. [10] proposed a two-layer SVM-based approach for the classification of image sequences in terms of real pain and posed pain. In their approach, the presence of Facial Action Units (AUs) per frame is detected with a set of AU-specific SVM classifiers based on Gabor features. The outputs of AU-specific SVMs are then temporally filtered and used as an input to the SVM classifier. Brahnam et al. [11] used Principal Component Analysis, Linear Discriminant Analysis and Support Vector Machines (SVMs) for binary classification of pain images (i.e., pain vs. no pain).

The current state-of-the-art scheme for recognition of expressions, specifically using AUs, are schemes reported by Lucey et al. [3] and Valstar & Pantic [1]. The former scheme builds upon active appearance model (AAMs) technique where

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.803

AAM model fit shape and appearance components through a gradient descent search method. Support vector machine (SVM) is used for classification purpose. The critical part in AAMs approach lies in model fitting. Fitting a model perfectly on a human face is challenging task and require highly efficient algorithms. As compared to AAM approach, [1] exploits geometric features as mid-level features to achieve a large feature set. Features are then selected using GentleBoost. SVM is applied for classification. These geometric features hold the information of facial components and their deformations.

It can be seen from the literature review that the proposed approaches on pain recognition uses multiple features sets and classifiers for detection of pain which are then fused to yield high performance. Moreover, these approaches seldom present their results on action unit intensity and pain intensity on standardized scales. The proposed approach in this paper not only remove the above shortcomings but also supersede in performance.

III. Proposed Methodology

The main contribution of this work is the improvement in recognition rate that is reported in the literature work. Instead of appearance features we have used carefully selected pertinent geometric feature derived from just 22 facial characteristic points (FCPs). We show that these features are more powerful as only a single classifier (K-NN classifier) yields promising results. The pain model is thus simpler than the other approaches used. The scores generated are directly on standard scale. We use multistage classification scheme such that we use K-NN classifier for recognizing the action units of pain and estimated their intensities in the first stage. For evaluating pain we use pain-scale method given by Prkachin and Solomon [12] in the second stage.

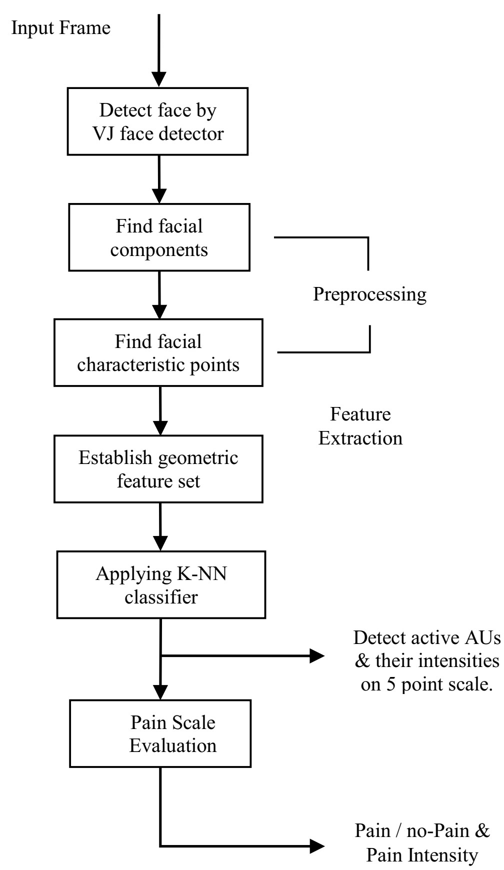


Fig.1. Schematic Flow of Proposed approach.

Fig. 1 shows the schematic flow of our approach. Our approach works best for images/videos with near-frontal view of face. Any skew or rotation in a face can be normalized first during face registration. Another common assumption used is that each sequence starts with a neutral expression on a face. Firstly, face is detected in a neutral frame, the region of face is segment out from the whole frame and facial characteristic points are extracted. For this we have used the detector provided by Valstar and Pantic [1]. The outcome of facial point detector is 22 characteristic points. From these 22 facial characteristic points we establish a feature vector based on geometric features. After classification of action units we use pain evaluation scale provided by Prkachin and Solomon [12] to establish pain condition and its intensity. The schematic flow of the approach is shown in Fig. 1. In the following lines various modules of the proposed approach is discussed in detail.

A. Database

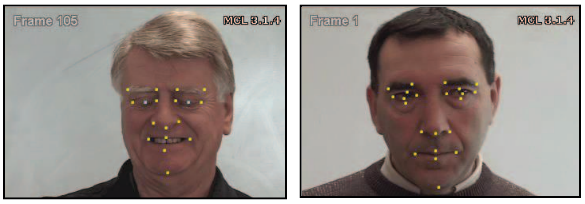
We have trained the samples on the standard database, UNBC McMaster Shoulder Pain Archive database [8]. Each frame of the database is AU coded by FACS coders, self-reports and observer measures. Database contains 200 video sequences containing spontaneous expressions. In this database there are 48398 FACS coded frames. Pain score of every frame is also available on 16 point scale. We have used more than 20000 samples in our approach. The sub-dataset of these samples is divided into two parts one for training and other for testing. 14670 samples of different subjects expressing pain or no-pain are used for training while approximately 6830 samples are used for testing.

B. Face Detection

This is the fundamental and essential step in the analysis of face. There are several approaches for face detection [13] and almost all of them detect near-frontal or near-profile face. Viola-Jones face detector is perhaps the robust real-time face detector. This face detector consists of cascade of classifiers trained by AdaBoost. Haar-like features are used on the integral image which can be computed very fast at any location of the integral image [14]. The performance of the face detector depends upon number of training images. However, Viola-Jones face detector does not cater high rotations.

C. Facial Characteristics Points

For extraction of information from face, most researchers [1] find characteristic points on face. These characteristic points hold the information of facial components and their transformations. For this Holden and Owen [15] used log Gabor features to localize feature points, Cristinacce and



(a) (b)

Fig.2. Extended Version of Facial Characteristic Point Detector. **(a)** indicates eye closure by white color as compared to yellow color **(b)**

Cootes [16] used shape and template models to find the best matching shape of face.

We employed the facial point detector established by Valstar and Pantic [1], because of its higher accuracy and reliability of the detector. Their facial point detector algorithm localizes 22 points on human face accurately with more than 90% success rate [1]. The detector performs best with near frontal view. Their detector behaves poorly with close eyes. Hence we modified the facial point detector such that if eyes were closed the marked vertical points over the eyes vertically were collapsed to a single point in the mid of the iris. Fig. 2 shows 22 facial characteristic points with and without eyes open.

The summary of facial point detection is as follows: firstly face is detected using Viola-Jones face detector, divides the face region in three parts i.e. two eyes parts and a mouth part and then facial components are localized by analyzing the histogram in these regions. After that a fine search is performed in a search window along these facial components, Gabor responses are calculated in this search window and fed into GentleBoost based point detectors. The position with maximum output corresponds to a point. The system is reliable for near upright faces.

D. Feature Extraction

Feature extraction is the key phase of our approach. We use geometric features for feature extraction task. The motive behind using geometric features is that it based on the shape of facial components which gives us the information of the expression. Also pain expression consists of AUs that considerably transform the shape of facial components as discussed earlier. Before extraction of features we explain the AUs that are active or play a role in pain expression.

Our choice of geometric features is inspired by the one used by Valstar and Pantic [1]. We, however, use a smaller set of geometric features which helps to keep the dimensionality of the feature vector lower. Based on the established 22 facial characteristics points, we define the following set of geometric features:

Value of x- coordinates of these 22 FCPs, where i, j = {1 to 22}

$$f_1 = p(x_i) \tag{1}$$

Value of y- coordinates of each FCPs:

$$f_2 = p(y_i) \tag{2}$$

Distance between the pair of FCPs, where i and j are not equal.

$$f_3 = \left\| p(x_i, y_i) - p(x_j, y_j) \right\| \tag{3}$$

The angle between all pair of with respect to horizontal:

$$f_4 = \tan^{-1}\left(\frac{p(y_i) - p(y_j)}{p(x_i) - p(x_j)}\right) \tag{4}$$

The displacement of all the above feature values from the neutral frame

$$\langle f_5 \dots f_8 \rangle = \langle C(f_1) \dots C(f_4) \rangle \tag{5}$$

where, $C(x(t)) = x(t) - x(0)$

Combining all these feature vectors, this results in a feature vector of size 1012 for each frame of input image sequence.

E. Pain-Scale Expression

Since the goal is to assess pain, we use facial action units in order to estimate pain and its intensity. According to Prkachin and Solomon [12], there are 4 core AUs (or pair of AUs) that are active during pain. These action units are AU4 (Brow Lowering), AU6/AU7 (Orbital Tightening/Cheek Raising), AU9/AU10 (Wrinkling of Nose/Raising of Upper Lip Raiser) and AU43 (Eye Close) as shown in fig. 3. These

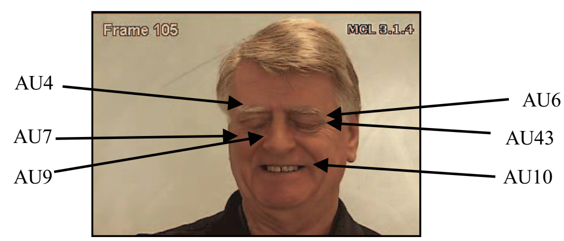


Fig.3. Active AUs of Pain.

set of action units play a vital role in identifying pain.

Prkachin and Solomon [12] defined pain as the sum of AUs of brow lowering, orbital tightening, levator contraction and eye closure. The pain scale expression is given by:

$$\text{Pain} = \text{AU4} + (\text{AU6}|\text{AU7}) + (\text{AU9}|10) + \text{AU43} \tag{6}$$

The sum of AU4, (max. of AU6, AU7), (max. of AU9, AU10) and AU43 give pain intensity value. The scale employed for AU intensity is 1-5 which provides the pain scale range to be 0-16. On this pain scale, we determine pain/no-pain condition based on the selected threshold value. Value higher than threshold shows pain expression

F. K-NN Classifier

K-nearest neighbor (k-NN) algorithm is a classification method based on statistical theory. In this algorithm the Euclidean distance is usually chosen as the similarity measure, which relates to all attributes. We use k-NN classifier in classification stage for the classification of AUs. The basic reason behind it is that k-NN is a statistical classifier and is totally based on the inter-relationship of attributes. Because of the fact that we use geometric features which are highly correlated to each other (attribute-wise) strengthening the use of this classifier. One important thing in k-NN classifier is the careful selection of k. This is done by testing several values of k with different features each time and select the best k out of it. We have fixed the value of k to 3 after testing several times.

We have used multi-stage classification scheme. In the first stage we have used K-NN for detection of active action units and their intensities. In the second stage we have used pain scale evaluation method given by Prkachin and Solomon [12] (eq-6) to detect pain on a frame by frame level using AU intensities. The pain-scale method returns pain intensities on a 16 point scale (0-16), a similar method used by Lucey et al. [3].

IV. Performance Assessment

We have shown two set of experimentations. In the first set of experiment, we evaluated the performance of AUs and their intensities of pain. In the second set, we estimated the performance of pain-scale method for validating pain assessment.

In our testing we have recognized six AUs of pain expression as described in Section 3. Some of them are recognized mutually as they are closely interlinked i.e. AU4, AU6/AU7, AU9/AU10 and AU43. We have recognize six AUs in which AU6 and AU7 are logically pooled because both these action units describe the same muscle movement that is tightening of orbital or eyes and upper movement of cheek. The rationale behind this conclusion is because of the fact that mostly when cheeks are raised, the orbital or eyes tightened automatically. Similarly for AU9 and AU10, both these action units are analogous to each other i.e. raising of upper lip and wrinkling of nose. The reason for this is again to the fact that mostly wrinkling of nose constitutes upper lip for upward movement.

A. Action Unit Estimation

Fig. 4 shows four different subjects exhibiting different AUs of pain. For fig. 4(a) our approach, using k-NN, classified AU4, AU6/AU7 and AU43 to be active while using ground truth available for each frame of the dataset shows same AUs plus AU10 is also active. Since, there are not many points on nose thus making it hard to recognize AU10. For fig. 4(b) our system report AU6/AU7 as active action units while the ground truth for this frame shows AU4 and AU6 to be active. For fig. 4(c) and (d) both our classifier and ground truth are exactly same. From the experimental results we have shown that our system can achieve more than 84% accuracy for AU of pain.

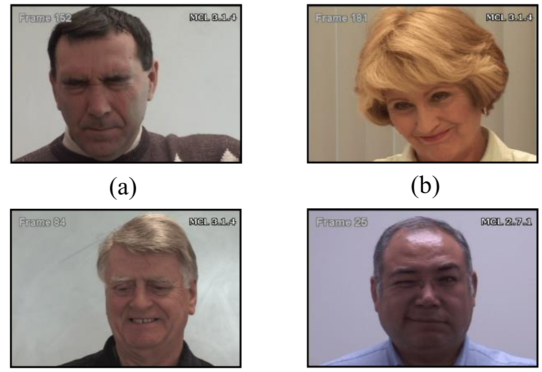


Fig.4. (a) Ground Truth: AU4, AU6, AU10, AU43; Our Classifier: AU4, AU6/AU7, AU43 **(b)** Ground Truth: AU4, AU6; Our Classifier: AU6/AU7 **(c)** Ground Truth: AU6, AU7; Our Classifier: AU6/AU7 **(d)** Ground Truth: AU6, AU43; Our Classifier: AU6/AU7, AU43

For AU intensity level we use K-NN classifier to classify six classes. First class shows inactive action unit and the remaining six classes show different levels of action unit intensities. Owners of UNBC McMaster shoulder pain database has provided ground truths of active action unit intensity for each frame. Making this a multi-class problem with 6 classes we classified each intensity level of AUs of pain. Hence separate classifier approach is used for each AU of pain. These AU intensities are computed on a 5 point scale (0-5) with '0' showing inactive AU, '1' shows active AU with low intensity, '2' shows active AU with medium-low intensity, '3' shows active AU with medium intensity, '4' shows AU with medium-high and '5' shows high intensity for active action unit.

TABLE 1. Validation results of different AU/AU pairs for AU intensity levels (0-5). 0 level shows inactive action unit.

	AU6/7	**AU4**	**AU9/10**	**AU43**	**Avg.**
Level 0	91.43	94.89	93.48	96.17	94.00
Level 1	82.10	84.38	81.56	86.89	83.73
Level 2	80.67	79.68	80.71	-	80.35
Level 3	84.39	81.36	77.63	-	81.12
Level 4	77.54	83.45	68.36	-	76.54
Level 5	70.63	-	-	-	70.63

In Table 1, we have shown AU intensity level results. We have represented the intensity levels with low, medium-low, medium, medium-high and high with low being the lowest intensity level and high for the maximum intensity of an action unit. Inactive intensity level shows inactive action unit. Also we can see, high intensity levels have low recognition rate as compared to low intensity levels. The reason for this is that, few samples are available in the database with high intensity level for these AUs. AU43 is by definition closing of eyes and is a binary action unit value.

As seen in Table 1, accuracy for AU9/AU10 is low as compared to other action units. AU9 and AU10 are wrinkling of nose and upper lip raising. The reason behind this low recognition is due to the fact that there are not enough FCPs on nose to track the movement. One solution is to locate more

FCPs on face but increasing points also enhance the feature set.

B. Pain Estimation

We have achieved 86.21% of accuracy for pain assessment using the pain scale equation (eq-6) on UNBC McMaster shoulder pain archive database [8]. We have fixed the threshold value to ‘2’ and above that value we termed the sample as pain. In this way we not only assess the pain/no-pain label for each frame but also show pain intensities on 16 point scale along with the action units.

Frame-level hit rate for pain is 86% and false acceptance rate is 25% as compared to [17], their hit rate is 82% and false acceptance rate is 30% which shows much improvement in the overall results when compared with state-of-the-art approaches.

C. Comparison with State-of-the-art

The most recent work in the context of pain is done by Kaltwang et al. [2] and Lucey et al. [3]. Lucey et al. [3] evaluate their performance using ROC curves. Table 2 shows the comparison of area under ROC curve of K-NN approach with Lucey et al.

TABLE 2. Comparison of Area under ROC curve of AUs classified by Lucey et al. versus K-NN approach.

	AU6/7	AU4	AU9/10	AU43	Pain
AUC (K-NN)	91.2	78.77	92.10	96.53	87.34
AUC Lucey et al.	86.2	53.7	79.8	90.9	84.7

For AU9/10 and AU4, area under ROC curve for Lucey et al. [3] is considerably low as compared to of K-NN approach. On the other hand Kaltwang et al. [2] evaluate their performance using MSE. Table 3 shows the comparison of MSE of K-NN approach with Kaltwang et al. [2].

TABLE 3. Comparison of Mean Square Error of AUs classified by Kaltwang et al. versus K-NN approach.

	MSE (K-NN)	MSE (Kaltwang et al.)
AU6/7	0.2430	0.480
AU4	0.1200	0.242
AU9/10	0.0699	0.071
AU43	0.0168	0.179
Pain	0.5069	1.633

For AU6/AU7 and AU4, MSE for Kaltwang et al. [2] is twice of MSE of K-NN approach and for AU9/AU10 it is almost the same. The reason behind reduced MSE and higher area under ROC curve (compared to Lucey et al. [3]) is the precise recognition of AUs and accurate estimation of their intensities. We have formulate the intensity estimation problem into a classification problem that enables us to estimate intensities accurately.

We have achieved 84.02% recognition rate for action unit recognition which is higher than state-of-the-art schemes. Lucey et al. [3] approach using AAM approach achieved 78% accuracy for action unit classification and 82% area under ROC curve for pain assessment [17]. We have used the same dataset i.e. used by Lucey et al. [3] and we have achieved 87.34% area under ROC curve for pain assessment, higher than the state-of-the-art scheme.

V. Conclusion

Automatic recognition of pain expression has many applications. Using facial characteristic points we extract geometric (shape) features. We have shown that by using only 22 facial points, our system can assess pain with more than 86% reliability. This reflects the fact that the proposed representation and features successfully capture the dynamics of pain expression. This selection has yielded good results with K-NN as classifier even with training lower than used in literature for recognizing action units and their intensity levels. Six AUs of pain are than used to generate the score for pain intensity. Our approach is also distinct in presenting results on the standard intensity scales both for AUs and pain. Validation studies are done on UNBC McMaster Shoulder Pain Archive Database. Our system reports 84.02% classification rate for action units of pain and 87.4% area under ROC curve for Pain Expression.

References

[1] Valstar, M.F.; Pantic, M., "Fully Automatic Recognition of the Temporal Phases of Facial Actions," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, vol.42, no.1, pp.28,43, Feb. 2012.

[2] Kaltwang, S. and Rudovic, O. and Pantic, M. (2012) Continuous pain intensity estimation from facial expressions. In: Proceedings of the 8th International Symposium on Advances in Visual Computing, ISVC 2012, 16-18 July 2012, Crete, Greece. pp. 368-377.

[3] Lucey, P.; Cohn, J.F.; Matthews, I.; Lucey, S.; Sridharan, S.; Howlett, J.; Prkachin, K.M., "Automatically Detecting Pain in Video Through Facial Action Units," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on , vol.41, no.3, pp.664,674, June 2011

[4] M. Pantic and M. Bartlett, “Machine analysis of facial expressions,” in Face Recognition, K. Delac and M. Grgic, Eds. Vienna, Austria: I-Tech Educ. Publishing, 2007, pp. 377–416.

[5] M. S. Bartlett, P. Viola, T. Sejnowski, B. Golomb, J. Larsen, J. Hager, and P. Ekman, “Classifying facial actions,” in Proc. Adv. Neural Inf. Process. Syst., 1996, vol. 8, pp. 823–829.

[6] J. Lien, T. Kanade, J. Cohn, and C. Li, “Subtly different facial expression recognition and expression intensity estimation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recog., 1998, pp. 853–859.

[7] M. Pantic, L. J. M. Rothkrantz, and H. Koppelaar, “Automation of nonverbal communication of facial expressions,” in Proc. Conf. Euromedia, 1998, pp. 86–93.

[8] Lucey, P.; Cohn, J.F.; Prkachin, K.M.; Solomon, P.E.; Matthews, I., "Painful data: The UNBC-McMaster shoulder pain expression archive database," Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on , vol., no., pp.57,64, 21-25 March 2011

[9] P. Ekman and W. Friesen, “The repertoire of nonverbal behavioral categories—Origins, usage and coding,” Semiotica, vol. 1, pp. 49–98, 1969.

[10] Littlewort, G.C., Bartlett, M.S., Lee, K.: Automatic coding of facial expressions displayed during posed and genuine pain. Image and Vision Computing 27, 1797–1803 (2009)

[11] Brahnam, S., Chuang, C.F., Shih, F.Y., Slack, M.R.: Machine recognition and representation of neonatal facial displays of acute pain. Artificial Intelligence in Medicine 36, 211–222 (2006)

[12] K. Prkachin and P. Solomon, "The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain," Pain, vol. 139, no. 2, pp. 267–274, Oct. 2008.

[13] E. Hjelmås and B. Low, "Face detection: A survey," Comput. Vis. Image Understand., vol. 83, no. 3, pp. 236–274, 2001.

[14] P. Viola and M. Jones, "Robust real-time object detection," Int. J. Comput. Vis.., vol. 57, no. 2, pp. 137–154, May 2004.

[15] E. Holden and R. Owens, "Automatic facial point detection," in Proc. Asian Conf. Comput. Vis., 2002, pp. 731–736.

[16] D. Cristinacce and T. Cootes, "Facial feature detection and tracking with automatic template selection," in Proc. IEEE Int. Conf. Autom. Face Gesture Recog., 2006, pp. 429–434.

[17] A. Ashraf, S. Lucey, J. Cohn, K. M. Prkachin, and P. Solomon, "The painful face II—Pain expression recognition using active appearance models," Image Vis. Comput., vol. 27, no. 12, pp. 1788–1796, Nov. 2009.