Drive Safe: An Intelligent System for Monitoring Stress and Pain from Drivers' Facial Expressions

[1]R. Manoharan, [2]S. Chandrakala and [3]W. Khan
[1]*Rajalakhsmi Engineering College, Chennai, Tamil Nadu, India*
[2]*Sri Krishna College of Engineering & Technology, Coimbatore, Tamil Nadu, India*
[3]*Bournemouth University, United Kingdom*

***Abstract* — Stress and abnormal pain experienced by drivers during driving is one of the major causes of road accidents. Most of the existing systems focus on drivers being drowsy and monitoring fatigue. In this paper, an effective intelligent system for monitoring drivers' stress and pain from facial expressions is proposed. A novel method of detecting stress as well as pain from facial expressions is proposed by combining a CK data set and Pain dataset. Initially, AAM (Active Appearance Models) features are tracked from the face; using these features, the Euclidian distance between the normal face and the emotional face are calculated and normalized. From the normalized values, the facial expression is detected via trained models. It has been observed from the results of the experiment that the developed system works very well on simulated data. The proposed system will be implemented on a mobile platform soon and will be proposed for android automobiles.**

***Keywords* — emotion; stress; pain; detection; driver monitoring.**

I. INTRODUCTION

Road safety is one of the main objectives in designing driver assistance systems. Due to car crashes, occurring on average every 30 seconds, one person dies somewhere in the world. The cost of accidents in the United States is estimated to be about $300 billion annually [1], i.e. about 2% of its gross domestic product. Conservative estimates suggest that a high proportion of fatalities and injuries due to traffic accidents involve impaired drivers. It is predicted that these figures could be increased by 65% in the next 20 years unless novel driving risk reduction methods are leveraged [2]. Among all fatal traffic accidents, 95% are caused by human error [3]. The three major causes of these human errors, which are often referred to as the "Big Three," are alcohol, drowsiness, and inattention [4]. Statistics show that 25% of fatal accidents in Europe [5], 32% in the United States [6], and 38% in Canada [7] are caused by drunk drivers.

Applying modern computer vision technologies for enhancing the safety of vehicle driving has been investigated for several decades. Most of the research has focused on detecting the drowsiness of the driver, which according to [8], causes a large percentage of the car accidents. Recently, reports [9] also show that the emotional status (e.g. stress, impatience) of the driver may also endanger safety. From the viewpoint of behaviour scientists, a high level of stress may damage self-confidence, narrow attention and eventually disrupt concentration. This often leads to aggressive driving and makes the driver pay less attention to the traffic situation. To reduce riskiness from a stressed state, it is necessary to detect such emotions and take certain actions to relax the driver. Abnormal pain during driving, for example a heart attack or any spinal pain, may also lead to an accident, so the detection of pain from facial expressions along with other emotions is also vital. According to [10], registration errors, head-pose variations, illumination variations, identity bias and occlusion are the main challenges in automatic affect recognition. Before measuring facial expressions, head-pose variations have to be modeled. Due to head movements, even under constant illumination, illumination variations can be problematic. Registration errors are mainly caused by registration techniques and occlusions may occur due to camera movement or head movement.

In this report, an effective intelligent system for monitoring drivers' stress and pain from facial expressions is proposed to avoid accidents due to driving with stress and pain. A heart attack or any sort of body pain during driving may completely distract the driver and may lead to an accident. A novel method of detecting stress as well as pain from facial expressions is proposed by combining a CK data set and a Pain dataset. Initially, AAM (Active Appearance Model) features are tracked from the face in each frame. Using these features, the Euclidian distance between the normal face and the emotional face are calculated and normalized. From the normalized values, the facial expression is detected via trained models. The CK and Pain datasets are used for modeling the classifiers. It was observed from the experimental results that the developed system worked very well on simulated data. The proposed system will be implemented on a mobile platform soon and will be proposed for android automobiles in the near future.

II. RELATED WORK

There have been some significant separate previous studies about emotion and pain monitoring from facial expressions. Many computer vision-based schemes have been developed for non-intrusive, real-time detection of drivers' stressed states with the help of various visual cues and observed facial features. An observed pattern of the movement of eyes and head and changes in facial expressions are known to reflect the person's stress and pain levels. Eye closure, head movement, jaw drop, eyebrow shape and eyelid movement are examples of some features typical of high stress and expressed pain of a person. To make use of these visual cues, a remote camera is usually mounted on the dashboard of the vehicle which, with the help of various extracted facial features, analyses the driver's physical conditions and classifies the current state as stress/non-stress and pain/non-pain.

It has been concluded that computer vision techniques are non-intrusive, practically acceptable and hence are the most promising for determining the driver's physical conditions and monitoring driver fatigue [9]. Most of the previous work on stress detection applied to physiological features (such as electromyogram, electrocardiogram, respiration, and skin conductance) [12]. It was found that in real-world driving tasks, skin conductivity and heart rate metrics were most closely correlated with the driver's stress levels [12]. However, those measurements were intrusive, so were less comfortable in real applications. A non-intrusive stress detection system was developed in [13], in which a physiological measure based on skin temperature was used. In [9], acoustic signals were used for measuring the stress level. However, performance might be affected by the noisy in-car driving environment. A system in [14] fused several physiological signals and visual features (eye closure, head movement) to monitor driver drowsiness and stress in a driving simulator. Liao *et al*. [15] applied facial expression, head motion and eye gaze as the visual cues for stress inference and evidence of the different signal modalities are combined with Dynamic Bayesian Networks (DBN). [16] classified the driver's facial emotion from thermal images, which provided a natural combination of visual evidence and skin temperature.

Affect recognition systems usually recognize the appearance of facial actions or the emotions conveyed by the actions. Facial Action Coding System (FACS) is where the former set of systems usually relies. [23]. Facial configurations are described by the facial Action Units (AUs), e.g. AU 43 is eye closure. In interpreting emotional displays the production of a facial action plays an important role in the temporal evolution of facial action unit [24], [25]. It is typically modeled with four temporal segments [26]: neutral, onset, apex and offset. No signs of expressionless phase (no muscular activity) are called Neutral. Onset denotes the period when muscular contraction begins and increases in intensity. Apex is a plateau where the intensity usually reaches a stable level, whereas Offset is a phase of muscular action relaxation. Although the order of these phases is usually neutral-onset-apex-offset, alternative combinations such as multiple-apex actions are also possible [27]. Automatic detection of pain from the face through facial action units is proposed by Pucey *et al*. in [28]; based on the action units extracted from the series of frames, the pain is detected from the facial expression.

In [29], Discriminative Shared Gaussian Processes for Multi-View and View-Invariant Facial Expression Recognition was proposed. First, a discriminative manifold shared by multiple views of a facial expression was learned. Subsequently, facial expression classification in the expression manifold was performed. Finally, classification of an observed facial expression was carried out, either in the view-invariant manner (using only a single view of the expression), or in the multi-view manner (using multiple views of the expression). In [30] a novel framework for expression recognition was proposed by using the appearance features of selected facial patches. A few prominent facial patches, depending on the position of facial landmarks, were extracted which are active during emotional elicitation.

The face is a reliable source of information for judging the pain experienced by others. Until now, most of the studies investigating the facial expression of pain have used a descriptive method (i.e. Facial Action Coding System). However, the facial features that are relevant for the observer in the identification of the expression of pain remain largely unknown despite the strong medical impact that misjudging pain can have on a patient's well-being. A non-action unit based recognition system is proposed using the bubbles method in [31] with high speed and moderate accuracy. But when it comes to driver monitoring, whatever developed should have the potential to run faster and be more efficient. In our previous work [17], an effective driver fatigue and distraction alert system was implemented using the android openCV. However, stress related emotions were left for future work. As a result, in this proposed work an effective driver stress and pain monitoring system is proposed.

III. OVERVIEW OF THE PROPOSED SYSTEM

In this section the overview of the proposed system is explained in detail. Figure 1 shows the block diagram of the proposed driver stress and pain monitoring system. At first, the AAM features are tracked from the face. The initial frame is considered to be the normal face (neutral emotion) and the frame with the highest change compared to the first frame is considered to be the peak emotional face. The Euclidian distances between these two frames are calculated. The calculated value is normalized and from the normalized value, the emotion is recognized using trained models (Support Vector Machine).

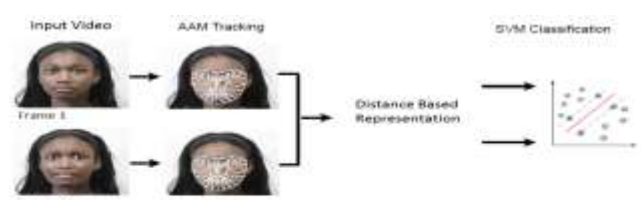


Fig. 1. Block diagram of the proposed stress and pain monitoring system

Active Appearance Models (AAMs) have been shown to be a good method of aligning a pre-defined linear shape model that also has linear appearance variation, to a previously unseen source image containing the object of interest. In general, AAMs fit their shape and appearance components through a gradient-descent search, although other optimization methods have been employed with similar results [18].

The distance between two points in the Euclidean space of a straight line can be stated as the Euclidean distance or Euclidean metric. With this distance, Euclidean space becomes a metric space. The associated norm is called the Euclidean norm. Here the Euclidean distance between the two sets of points (normal face and expression face features) are calculated. The calculated values are normalized to avoid glitches in the data. The normalized values are then analyzed for emotions. This is undertaken using trained classifiers. The Support Vector Machine classifier is trained with the sample emotions and is used for classifying the new emotions.

IV. Stress And Pain Detection

This section describes the implementation details of the individual modules. This includes the definition of the detection target and the methodologies. Different people may behave or express emotions differently under stress; it is hard to find a universal pattern to define the stress emotion. Moreover, it is not easy to collect data for training offline models. To ease the problem, we define stress and pain in terms of basic emotions which have corresponding facial expressions that are well defined and universal [19]. The expression of stress is defined as anger, disgust, or a combination of these two fundamental facial expressions [20]. The expression of pain is defined using the pain dataset which will be explained later in this paper. Figure 3 shows the complete workflow of the proposed driver stress and pain recognition system.

A. *Major steps:*

1. Face Tracking

The camera placed in front of the driver records a video of the driver's face and sends it to the face tracking module. The facial points are tracked using AAM methodology. Figure 2 shows the real time tracking of the driver's face inside the vehicle.

Fig. 2. Sample frame marked with 66 facial AAM landmarks

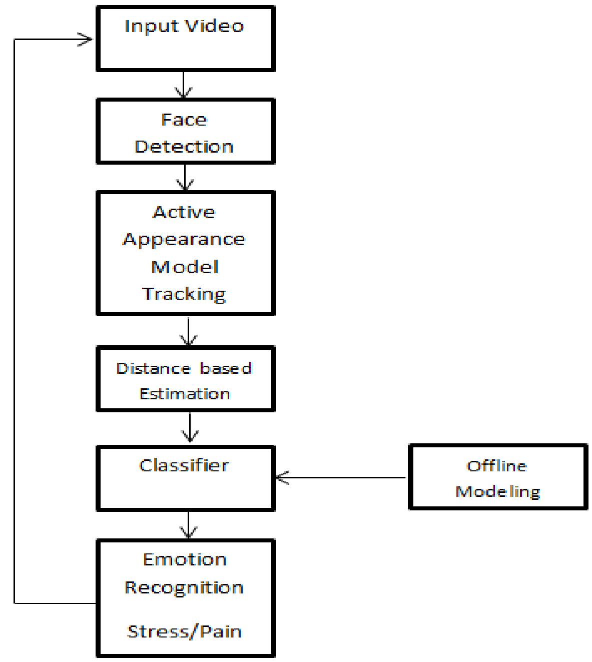


Fig. 3. Workflow of the proposed stress and pain monitoring system

2. Distance Based Estimation

During the training phase of the classifier, the first frame is considered to be the normal frame and the last frame is considered to be the peak frame. The Euclidian distance between these two frames is calculated to get the two column vector of facial landmark co-ordinates. These co-ordinates are normalized to avoid glitches.

3. Pain and Stress Detection

The classifier is trained basically using the CK [21] dataset and the Pain dataset [22]. The CK database contains 486 sequences across 97 subjects. Each of the sequences contains images from the onset (neutral frame) to peak expression (last frame). This project uses the AAM features of the 486 sequences. The AAM feature comprises 68 facial landmarks which are used to calculate the facial expressions. Figure 4 shows the example frames in the CK and CK+ dataset.

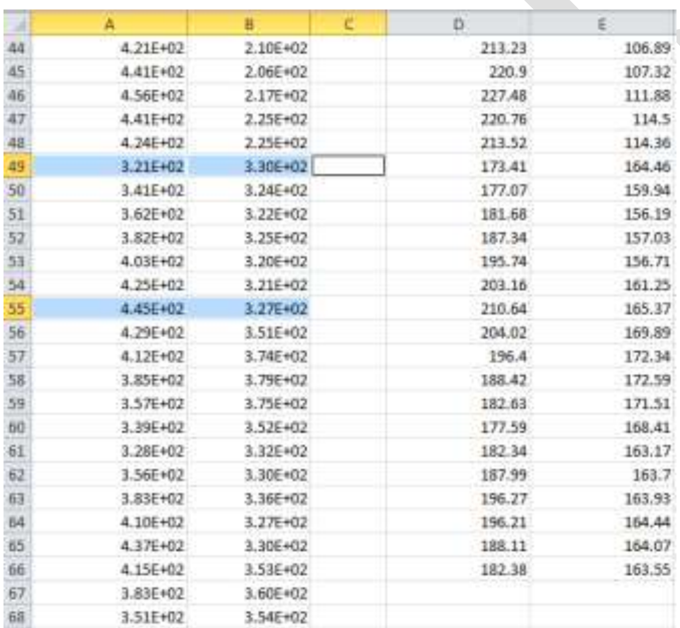
Fig. 4. Example of CK & CK+ Dataset

In the McMaster UNBC Pain dataset there are 200 sequences across 25 subjects, which totals 48,398 images. Spontaneous expressions of pain from patients with shoulder problems are shown in the image sequences. All frames have their own AAM features which comprise 66 facial landmarks from neutral to peak pain expressions. Figure 5 shows the example frames in the Pain dataset.

Fig. 5. Example McMaster UNBC Pain Dataset

4. Combining the CK and the Pain Datasets

Combining the CK dataset and the Pain dataset is itself a challenging task in which the AAM features have to be made common to make sense of these two datasets. The 68 points in the CK dataset are reduced to 66 points by removing the landmarks near lip corners. In the actual pain dataset the values of the AAM differs. As a result, in order to make it common, the AAM features of the CK data set are multiplied by 100 and divided by 2, so that the features of both datasets share a common format and are ready for processing.



	A	B	C	D	E
44	4.21E+02	2.10E+02		213.23	106.89
45	4.41E+02	2.06E+02		220.9	107.32
46	4.56E+02	2.17E+02		227.48	111.88
47	4.41E+02	2.25E+02		220.76	114.5
48	4.24E+02	2.25E+02		213.52	114.36
49	3.21E+02	3.30E+02		173.41	164.46
50	3.41E+02	3.24E+02		177.07	159.94
51	3.62E+02	3.22E+02		181.68	156.19
52	3.82E+02	3.25E+02		187.34	157.03
53	4.03E+02	3.20E+02		195.74	156.71
54	4.25E+02	3.21E+02		203.16	161.25
55	4.45E+02	3.27E+02		210.64	165.37
56	4.29E+02	3.51E+02		204.02	169.89
57	4.12E+02	3.74E+02		196.4	172.34
58	3.85E+02	3.79E+02		188.42	172.59
59	3.57E+02	3.75E+02		182.63	171.51
60	3.39E+02	3.52E+02		177.59	168.41
61	3.28E+02	3.32E+02		182.34	163.17
62	3.56E+02	3.30E+02		187.99	163.7
63	3.83E+02	3.36E+02		196.27	163.93
64	4.10E+02	3.27E+02		196.21	164.44
65	4.37E+02	3.30E+02		188.11	164.07
66	4.15E+02	3.53E+02		182.38	163.55
67	3.83E+02	3.60E+02			
68	3.51E+02	3.54E+02			

Fig. 6. Removal of two points from CK AAM features

Figure 6 shows the numerical representation of the facial points. Columns A and B represent the AAM features of the CK dataset and columns D and E represent the AAM features of the pain dataset. In the A and B columns, the 49th and 55th rows are removed (which are lip corners) so that the 68 points are reduced to 66 points. Figure 7 shows the graphical plot of the 66 point facial landmarks. After the CK+ dataset has been reduced from 68 to 66 points, all the features will look close to Figure 7 when it is plotted.

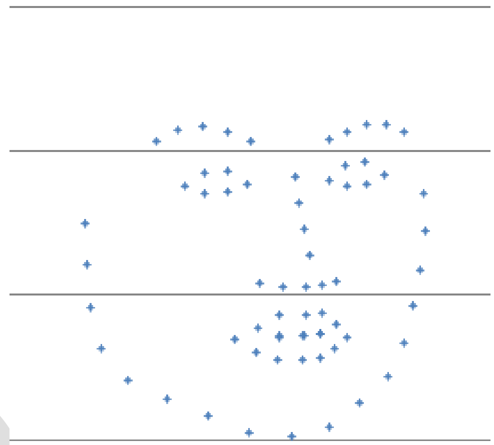


Fig. 7. 66 point graphical plot of pain dataset

V. EXPERIMENT AND ANALYSIS

The detection of stress and pain is based on reading basic facial expressions. The Ekman's six basic emotions and neutral expressions are considered to be the standard emotions in most works on facial expression recognition in the literature, e.g. the six basic emotions are anger, disgust, fear, happiness, sadness and surprise. Until now there has been no work based on combining facial emotion detection with pain expression. This is the first attempt to detect pain along with other expressions for driver monitoring.

To train classifiers for stress and pain recognition, images from two facial expression databases, i.e. the CK+ database and the Pain database, are used. The images in both databases are in frontal and upright pose, and evenly illuminated with posed facial emotions. Multi-class classifiers are trained using the extracted features. The classifiers are implemented with support vector machines (SVM) in a one-vs-all manner. Automatic sigma estimation (sigest) for RBF or Laplase kernel is used. The partition ratio for the training and testing of data is 3:1; i.e. 75% of the total data is used for training the models and 25% of the remaining data is used for testing.

Implementation is undertaken using the R language coded in R studio IDE. Experiment results show that the proposed method is comparable with state of the art methods available to date. Table I shows the performance comparison of various classifiers in which the SVM with sigma estimation provided the maximum result percentage of 98.3%.

Table II shows the confusion matrix of the pain and stress recognition system in which the SVM classifier using automatic sigma estimation for Laplace kernel attained a detection rate of 98.3%. In the confusion matrix, labels 1, 2 and 3 represent the class names, normal stress and pain respectively.

TABLE I. PERFORMANCE ANALYSIS OF VARIOUS CLASSIFIERS

Classifier Type	**Detection Percentage**
Support Vector Machine (SVM)	98.3%
SVM without sigma estimation	97%
Decision Tree	83%

TABLE II. CONFUSION MATRIX OF SVM CLASSIFIER

	Normal	***Stress***	***Pain***
Normal	*30*	*0*	*0*
Stress	*0*	*29*	*0*
Pain	*0*	*1*	*30*

VI. CONCLUSION AND FUTURE WORK

To reduce the number of accidents due to driver stress and abnormal pain during driving, an effective non-intrusive driver stress and pain monitoring system is proposed in this paper. A novel method of detecting stress as well as pain from facial expressions is proposed by combining the CK dataset and the Pain dataset. Experiment results show that the developed system operates very well under simulated conditions. The proposed system will be implemented in a mobile platform soon and will be proposed for android automobiles. The proposed system will be integrated with a driver fatigue monitoring system and implemented in a mobile platform as a future work.

REFERENCES

[1]. "Crashes vs. congestion, what is the cost to society?" American Automobile Association (AAA), Heathrow, FL, USA, 2011.
[2]. Doshi and M.M. Trivedi, "On the roles of eye gaze and head dynamics in predicting driver's intent to change lanes," IEEE Trans. Intell. Transp.Syst., vol. 10, no. 3, pp. 453–465, Sep. 2009.
[3]. Amditis, M. Bimpas, G. Thomaidis, M. Tsogas, M. Netto, S.Mammar, A. Beutner, N. Möhler, T.Wirthgen, S. Zipser, A. Etemad, M. Da Lio, and R. Cicilloni, "A situation-adaptive lane-keeping support system: Overview of the SAFELANE approach," IEEE Trans. Intell. Transp. Syst., vol. 11, no. 3, pp. 617–629, Sep. 2010.
[4]. K. Murata, E. Fujita, S. Kojima, S. Maeda, Y. Ogura, T. Kamei, T. Tsuji, S. Kaneko, M. Yoshizumi, and N. Suzuki, "Noninvasive biological sensor system for detection of drunk driving," IEEE Trans. Inf. Technol. Biomed., vol. 15, no. 1, pp. 19–25, Jan. 2011.
[5]. L.-W. Zhu, Z.-Y. Zhang, Z.-J. Bao, and Y. Sun, "Study on the drink driving behavior of drivers in Beijing based on the theory of plan behavior," in Proc. LEITS, 2010, pp. 1–5.
[6]. D. Jiangpeng, T. Jin, B. Xiaole, S. Zhaohui, and X. Dong, "Mobile phone based drunk driving detection," in 4th Int. Conf. Pervasive Health, Munich, Germany, 2010, pp. 1–8.
[7]. Y.C. Liu and C.H. Ho, "The effects of different breath alcohol concentration and post alcohol upon drivers driving performance," in Proc. IEEE IEEM, 2007, pp. 505–509.
[8]. Y. Dong, Z. Hu, K. Uchimura and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," IEEE Transactions on Intelligent Transportation Systems, vol. 12, no. 2, pp. 596–614, 2011.
[9]. Nass, I.-M. Jonsson, H. Harris, B. Reaves, J. Endo, S. Brave and L. Takayama, "Improving Automotive Safety by Pairing Driver Emotion and Car Voice Emotion," in Intl. Conf. on HCI, 2005.
[10]. E. Sariyanidi, H. Gunes, and A. Cavallaro, "Automatic analysis of facial affect: A survey of registration, representation and recognition", IEEE Transactions On Pattern Analysis And Machine Intelligence, August 2013.
[11]. Conf. Ocular Measures of Driver Alertness, Washington, DC, Apr.26–27,1999.
[12]. J. Healy and R. Picard, "Detecting Stress During Real-World Driving Tasks Using Physiological Sensors," IEEE Transactions on Intelegent Transportation Systems, vol. 6, no. 3, 2005.
[13]. H. Yoshida, H. Kataoka, M. Yasuda, A. Saijo and M. Osumi, "Development of a skin temperature measuring system for non-contact stress evaluation," in the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1998, vol. 2, pp. 940–943.
[14]. Manstetten, M. Rimini-Doering, U. Landsatter, T. Altmueller and M. Mahler, "Monitoring driver drowsiness and stress in a driving simulator," in First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, 2001, pp. 58–63.
[15]. W. Liao, W. Zhang, Z. Zhu and Q. Ji, "A real-time human stress monitoring system using dynamic Bayesian network," in IEEE Conference on Computer Vision and Pattern Recognition - Workshops, 2005, pp. 70–70.
[16]. Kolli, A. Fasih, F. Al Machot and K. Kyamakya, "Non-intrusive car driver's emotion recognition using thermal camera," in Joint 3rd Int'l Workshop on Nonlinear Dynamics and Synchronization (INDS), 2011, pp. 1–5.
[17]. R. Manoharan and S. Chandrakala, "Android OpenCV based effective driver fatigue and distraction monitoring system", IEEE ICCCRT, 2015.
[18]. T. Cootes, G. Edwards and C. Taylor. "Active Appearance Models". IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6):681–685, 2001.
[19]. P. Ekman, "Universal and Cultural Differences in Facial Expressions of Emotion," J. Cole ed, Nebraska Symposium on Motivation, vol. 19, pp. 207–282, 1972.
[20]. H. Gao, A. Y¨uce and J.-P. Thiran, "Detecting Emotional Stress From Facial Expressions For Driving Safety", ICIP, 2012.
[21]. J.F. Cohn, J. Saragih, T. Kanade, Z. Ambadar, I. Matthews and P. Lucey, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression", Proc. of IEEE workshop on CVPR, San Francisco, USA, 2010.
[22]. P. Lucey, J.F. Cohn, K.M. Prkachin, P.E. Solomon and I. Matthews, "PAINFUL DATA: The UNBC-McMaster Shoulder Pain Expression Archive Database", submitted to the IEEE International Conference on Automatic Face and Gesture Recognition (FG2011), Santa Barbara, USA, 2011.
[23]. P. Ekman, W. Friesen and J. Hager, The Facial Action Coding System, 2nd ed., London: Weidenfeld and Nicolson, 2002.
[24]. Z. Ambadar, J.W. Schooler and J. Cohn, "Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions," Psychological Science, vol. 16, no. 5, pp. 403–410, 2005.
[25]. R. Rosenthal and N. Ambady, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta–analysis," Psychological Bulletin, vol. 11, no. 2, pp. 256–274, 1992.

[26]. P. Ekman, W. Friesen and J. Hager, “The Facial Action Coding System”, 2nd ed., London: Weidenfeld and Nicolson, 2002.
[27]. J. Cohn and K.L. Schmidt, “The timing of facial motion in posed and spontaneous smiles,” Int’l J. of Wavelets, Multiresolution and Information Processing, vol. 02, no. 02, pp. 121–132, 2004.
[28]. P. Lucey, J. Cohn, I. Matthews, S. Lucey, S. Sridharan, J. Howlett and K. Prkachin, “Automatically detecting pain in video through facial action units,” IEEE Trans. Systems, Man and Cybernetics – Part B, vol. 41, no. 3, pp. 664–674, 2011.
[29]. S. Eleftheriadis, O. Rudovic and M. Pantic, "Discriminative Shared Gaussian Processes for Multiview and View-Invariant Facial Expression Recognition", IEEE Transactions On Image Processing, Vol. 24, No. 1, January 2015.
[30]. S.L. Happy and A. Routray, "Automatic Facial Expression Recognition Using Features of Salient Facial Patches", IEEE Transactions On Affective Computing, March 2015.
[31]. F.Gosselin, R.P. Rainville, C. Blais and D. Fiset, "Efficient information for recognizing pain in facial expressions", Europian Journal of Pain, 2015.