Chapter 4: Expressing pain: The classification and interpretation of facial pain signals

4.1 Introduction

Pain, assumed to be the fifth vital sign, is an important symptom that needs to be adequately assessed in health care. The visual changes reflected on the face of a person in pain reveals the intensity level. It is often noticed that experts seldom make use of the tools available to diagnose the patient's pain, mainly due to the time constraint or they do not give satisfactory results, rather they blindly believe in the patient's self-report. This technique too suffers from several intricacies; therefore, an attempt is made to analyze and correlate the cognitive aspect of the patient, observer and the expert by a well-designed methodology that will streamline the process of patient assessment, increase its accessibility to physicians and improve the quality of care. Based on the nomenclature of muscle movements of the face, observers were asked to rate the pain score. Addressing pain intensity in the treatment of patients suffering from acute or chronic pain is very much important for proper diagnosis and treatment.

Pain is an unpleasant yet necessary signal that notifies us of actual or impending bodily damage and allows an individual to take action [Huguet *et al.*, 2010]. In clinical settings, this action could translate to patient diagnosis, medications or even a surgical procedure. Thus, measurement of pain is imperative for effective treatment. Signaling pain in others is highly salient [Simon, *et al.* 2008; Rotteveel & Phaf 2004; Gonzalez-Roldan *et al.*, 2011] and elicits empathic responses in addition to changes in observers facial expression and vicarious pain [Gonzalez-Roldan, *et al.*, 2013]. The field of facial expression analysis has recently seen significant progress due to advancements in computer vision. Several studies have shown that facial behavior can be used as a modality for prediction of internal states such as mood and confusion [Bartlett, *et al.*, 2005; Littlewort, *et al*., 2011]. Estimates of pain intensity are commonly obtained in clinical settings via self-report and behavioral measures [Von Baeyer & Spagrud, L.J. 2007; Stinson, *et al*., 2006; Tomlinson, *et al*., 2010].

The self-report measure allows an individual to verbally communicate the amount of experienced pain and suffers from several drawbacks such as subjective bias and patient idiosyncrasies. Moreover, it cannot be employed by verbally impaired patients. On the other hand, observational measures are based on inspecting non-verbal clues e.g., body, face or voice of an individual in pain for reporting pain intensity. Such measures are disrupted by the presence of observer's bias, considerable demands on clinician's time, and the influence of factors such as likeability of patient, underestimation of pain [Prkachin & Mercer, 1994]. Since pain is inherently a subjective and internal experience, mostly preferred measures in comparison over others is the self-report which is considered the gold-standard for conveying pain intensity [Zhou & Horgan, 2008; Craig, 1992]. Facial expression specifically related to pain has been studied using the Facial Action Coding System (FACS) [Craig and Patrick, 1985; Patrick, *et al*.1986; Craig et al. 1991; Ekman & Friesen 1978]. Several facial actions correlated with pain (Figure 4.1) that have been identified includes raised cheeks, lowered brows, closed eyes, tightened eyelids and a raised upper lip or opened mouth [Craig *et al*., 1991, 2001]. Expressions of acute pain-related facial activity have been identified in the general population, i.e. brow lowering, narrowing of the eye aperture from below, raising the upper lip, and blinking [Craig & Patrick, 1985; Patrick *et al*., 1986]. Using FACS, facial activity associated with exposure to the painful stimulus in healthy adults was recorded and six action units (AU) categories that occurred more usually during exposure to the stimulus than during a baseline experience were identified.

In summary, it has been shown that the FACS provides an objective assessment of facial reactions that are the most reflexive and automatic non-verbal indices of pain. Even when facial expressions have been reliably identified in children, infants, adults and the elderly using FACS, there is much less evidence of its utility in critically ill patients that requires more research to fill in the existing gap. When assessing pain (acute and chronic) there are multiple dimensions that should be considered. These dimensions include (a) sensory (e.g., intensity, word descriptors, duration, location, and frequency), (b) affective/cognitive (pain unpleasantness), and (c) the impact of pain on aspects of everyday life e.g. physical, emotional, role functioning and social). While it’s vital to assess each of these domains, the most commonly used parameter in clinical practices is the measurement of the pain intensity or how much it really hurts [Fordyce & Louis, 1976].

The aim is to determine which domains and measures should be used in clinical trials for pain. The two systematic reviews aimed to identify self-report pain intensity [Stinson, *et al*., 2006] and observational measures of pain intensity [Von Baeyer & Spagrud, 2007] with well-established psychometric properties that could be recommended for use in clinical trials. Moreover, use of self-report cannot be used in significant populations, such as young children, patients who cannot communicate, the mentally retarded, and patients who have the need for supportive breathing. In these circumstances, an observer rating is required, and the Faces Pain Scale is commonly used where the observer chooses a face on the scale which best resembles the facial expression of the patient [Prkachin & Craig, 1985].

Our work was focused towards predicting pain intensity using facial expressions. We propose to use the facial expression information (AU) to objectify the process of both detecting and measuring pain intensity in clinical settings. Since pain is a complex signal, such a system should be able to capture both the temporal dynamics and appearance variation of pain expression. This system was evaluated on the basis of predicting both postoperative ongoing pain and experimentally induced acute pain. Since appropriate datasets were required to validate computer vision algorithms for predicting pain, we prepared the database of students while inducing pain on which we tested our hypothesis. We are also in the process of collecting an in-the-wild clinical dataset that could be used to validate the posed research question in the coming future. Such automated methods for measuring pain intensity could be used to aid clinical staff in monitoring patient for long-term. Moreover, observers keeping a watch on the machine could alert clinicians to instances of pain and thus free up resources for more efficient allocation of clinical attention. Such systems are also useful in cases where verbal self-report ratings are not available.

4.2 Methods

4.2.1 Participants

The sample consisted of 20 adults (12 women and 8 men) university students between 18 and 35 years (mean = 24, SD = 3.3). Prospective participants who reported visual impairments (e.g., uncorrected vision), having psychiatric problems or undergoing medical, psychological treatment and taking medication were not included. Participants were given course credits for their participation.

4.2.2 Procedures

Self-report measures were taken after the composed assent structure was marked, members finished the Interpersonal Reactivity Index (IRI) [Davis, 1980]. This survey (questionnaire) was initially intended to evaluate empathy from a multidimensional viewpoint and comprises of 28 things conveyed in four subscales measuring the affective and cognitive perspectives of empathy: perspective taking (PT), surveying the propensity to adopt the psychological viewpoint of others; empathic concern (EC), evaluating the inclination to experience feelings of warmth and sympathy toward others; fantasy (FS), measuring the tendency to recognize oneself candidly with characters in fictional circumstances; and personal distress (PD), measuring self-oriented feelings as a consequence of seeing another's passionate trouble. Participants completed the Self-Assessment Manikin (SAM) [Bradley & Lang, 1994] to rate normal, mild, moderate and severe pain conveyed by affective facial expressions. This consisted of four sets of humanoid figures representing the dimensions of normal to severe pain. Each rating scale included ten levels of intensity, ranging from a smiling to a frowning figure for normal, slight muscle movements for mild followed by a little more for moderate pain and high for severe pain.

Participants were asked to assess their ratings on a 0-3 scale (0-no pain, 1-mild, 2-moderate, 3-severe pain) which they analyzed while viewing each facial expression displayed on the computer screen.

4.2.3 Experimental task

The overall idea was to record and analyze the opinion of observers and experts while they viewed the video frames. The experimental work was similar in nature as used by [Roelofs *et al*., 2010] to examine behavioural responses to affective facial expressions (happy, neutral, and angry) along with a slight modification in which the participants and the practitioners (experts) were asked to rate the pain intensity level for each video frame separately. Their observations included identifying the number of the AUs involved (Figure 4.1) and then predict the pain intensity level, for which sufficient training was given to them.

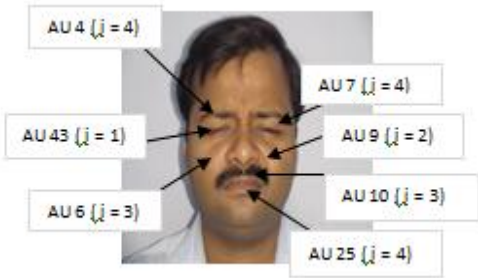

Figure 4.1: An example of a painful face from the self-prepared database with the corresponding Action Units and their intensities. (i = intensity of each AU).

We decided to use dynamic rather than static facial expressions and to include expressions of pain ranging from normal (no pain) followed by mild, moderate and severe pain. From a set of video clips, the stimuli were taken of a self-prepared database of patients undergoing pain treatment in Institute of medical science, SSL Hospital, BHU, Varanasi (India). During this work, twelve faces of males and eight females displaying neutral, mild, moderate and severe painful facial expressions were used as affective stimuli. Pain was induced in the clinical setting using pressure algometer applied to the hand of individuals and increasing the temperature. Individual expressions displayed in original video clips started with a neutral face and ending at the maximum peak of each expression for 1 second. For obtaining a similar presentation time of the facial expressions as used in [Roelofs, *et al.*, 2010], original video clips were slowed down to 2.5 seconds length and presented consecutively in blocks of 20 video clips using the same facial expressions (75 seconds) (Figure 4.2). Four blocks were designed with a white fixation cross and were timely separated by a 20-seconds black screen. The presentation order of blocks was pseudo-randomized across subjects by using a Latin square design. In order to record participant's response they were asked to press numbers on the keyboard e.g., 1 for normal face (no pain), 2 for mild pain, 3 for moderate and 4 for severe pain expressions. The task always started with a 60-seconds black screen followed by a white fixation cross to stabilize subject's concentration.

The coding was done in Super Lab 4.0 software on a 380 screen located at a distance of 200 cm using view angles of 23u (horizontal) and 17u (vertical). The responses were automatically generated on an Ms-excel sheet displaying the following attributes: Participant no., Day, Date and Time of filling, Trial name (Frame no.), Response (1-4 of keys pressed), Error Code (answer correct or wrong) and Reaction Time (time taken to identify the expressions in ms). While viewing the frames, participants sat in a comfortable chair positioned in front of the computer. It was ensured that during experiment there was proper lighting and only participant was present in the room. Participants were also given a short training session prior to starting the experiment.

4.2.4 Video recording and data acquisition

Facial expressions were recorded with a standard webcam (Logitech) at 30 frames per second located in front of the subject and connected to a laptop. Using a simple Pressure Algometer (PA), the recording was done, in which the pressure was increased to the limit such that the subject could resist it. Video recordings were converted into frames. Not all the frames were included in the study, except for those that showed major changes in the facial expressions during the various pain conditions.

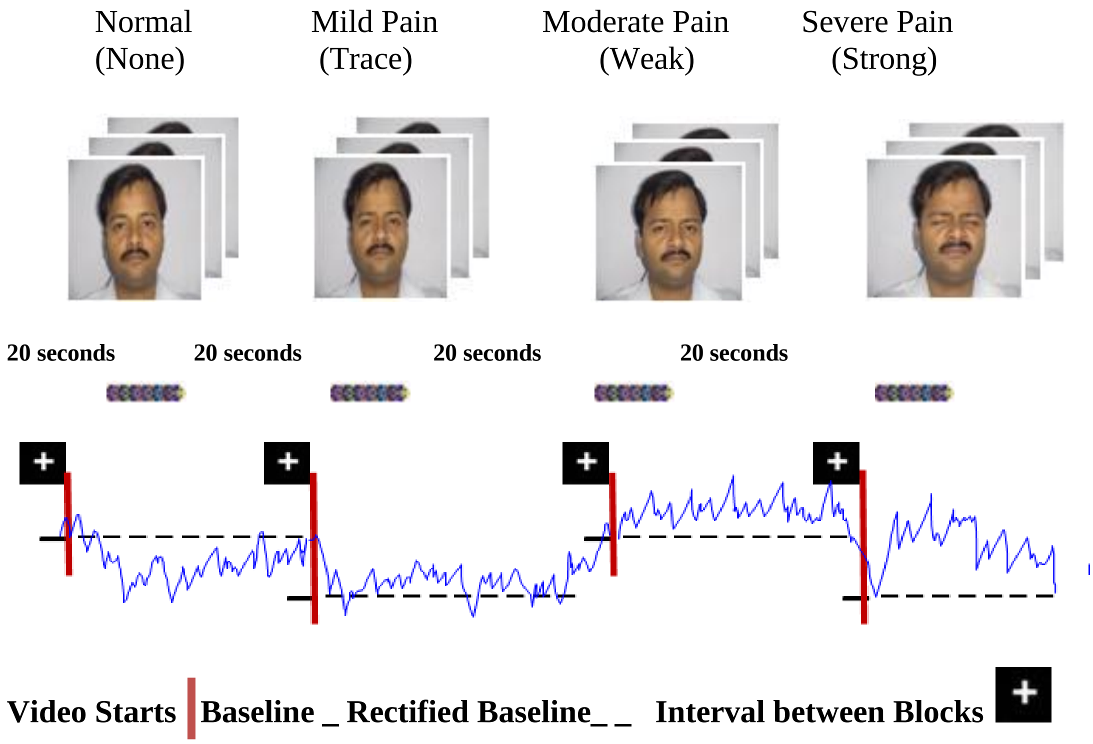

Figure 4.2: Description of experimental task and pain intensity signals elicited when viewing different pain related facial expressions. Data present above the baseline represents the amplitude of facial muscle movements occurring due to pain and data below the baseline correspond to the amplitude of no muscle movements.

4.3 Statistical analysis

Offline, independent observers rated pain intensity (OPI) observing the recorded video frames concentrating on the AUs contributing to pain. Considerable training in the identification of pain expression was provided to observers. The ratings of observer were performed on a 4-point Likert-type scale that ranged from 0 (no pain) to 3 (strong pain). To assess the inter-observer reliability of the OPI pain ratings, the trials were independently rated by the second rater (medical practitioner). The Pearson correlation between the observers and experts OPI was 0.60, ($p < 0.001$), which signifies low inter-observer reliability. Correlation between the subjects self-reported pain on the VAS and the observers rating was 0.71, ($p < 0.001$) for the trials used in the current experiment. A value of 0.70 is viewed as a large effect and is commonly taken as showing high concurrent validity. The inter-method correlation found here signifies moderate to high concurrent validity for pain intensity.

Subjective data were checked initially for normal distribution using the Shapiro-Wilk test. As our datasets significantly deviated from a normal distribution, the effects of the facial expression (normal, mild, moderate and severe pain) on dependent variables were evaluated using the Friedman test. The coding was done in such a manner that the accuracy and the response time in recognizing and classifying the painful facial expressions by both types of observers were automatically generated on an Ms-Excel sheet. SPSS 19.0 statistical package was used to perform all analyzes. For all statistical analyzes, a significance level of p = .05 was used.

4.4 Results

The result obtained clearly specifies that the mean accuracy rate and response time taken by the expert across all the frames were 83.75 and 348.5 (ms) for expert whereas for the observer it came to be 75.5 and 292 (ms). This clearly stated that there was a significant difference in accuracy rate and the time taken in recognizing and categorizing the pain intensity by the two observers. Experts provided more accurate result while taking more time to analyze and classify the frames in comparison to the observer. However, it was observed that the expert often underestimated the pain intensity in comparison to the observer (Table 4.1). The Friedman test yielded a significant effect of facial expressions on normal [χ^2 (2) = 64.51, p, .001], mild [χ^2 (2) = 43.32, p, .001], moderate [χ^2 (2) = 24.21, p, .001] and severe pain ratings [χ^2 (2) = 11.06, p, .01]. Table 4.1 shows the mean and standard deviation of normal to severe pain ratings for each facial expression.

Post-hoc pair-wise mean comparisons revealed that severe pain faces were more unpleasant than normal (no pain) [Z = 4.61, p, .001], mild pain [Z=5.10, p,001] and moderate pain faces [Z = 5.41, p, .001], and that normal faces were also more pleasant than pain faces [Z = 4.87, p, .001]. In addition, pain faces were more arousing than normal faces [Z = 3.19, p, .01].

The Friedman test also revealed significant effects of facial expression on the amplitude of muscle movements (AUs) [χ^2 (2) = 6.03, p< .5]. While doing the post-hoc pairwise mean comparisons revealed that severe pain faces elicited greater muscle movements in comparison to normal faces [Z =2.45,p<.05], whereas there was no difference between mild and moderate pain faces [Z =1.35, NS].

Table 4.1: Mean (and standard deviation) of pleasantness, arousal, accuracy rate and response time.

Facial Expression	Pleasantness	Arousal	Pain Rating Expert	Pain Rating Observer	Accuracy (%) Expert	Accuracy (%) Observer	Reaction time (ms) Expert	Reaction time (ms) Observer
Normal	7.96 (1.5)	3.01 (2.1)	0.5	1.1	90	86	240	202
Mild pain	7.42 (1.4)	3.78 (2.2)	3.3	4.1	81	74	398	308
Moderate pain	3.60 (1.3)	4.08 (2.4)	5.2	7.4	79	62	406	346
Severe pain	2.80 (1.1)	5.27 (2.6)	6.3	9.2	85	80	350	312

Pleasantness ratings ranged from 1 (very unpleasant) to 9 (very pleasant), arousal ratings range from 1 (very calm) to 9 (very excited), Pain ratings 0 (no pain) to 10 (severe pain).

4.5 Discussion

The aim of conducting this work was to evaluate the pain intensity elicited by the observation of facial pain expressions. We compared pain intensity scores provided by the three categories of respondents (observer, physician (expert) and the patient) by observing the facial expressions of pain on a frame by frame basis. The third category of the respondent here were the patients who gave their self-report on the VAS (Visual analogue scale). As it was expected, participants rated normal and severe pain faces with good accuracy, however, there was the difference of opinions with respect to the mild and moderate pain faces. On the other hand, we observed that both normal and severe pain faces took lesser response time in comparison to mild and moderate pain faces. The scores obtained were recorded on a frame by frame basis for both the participant and the expert (practitioner) separately, undergoing the experiment.

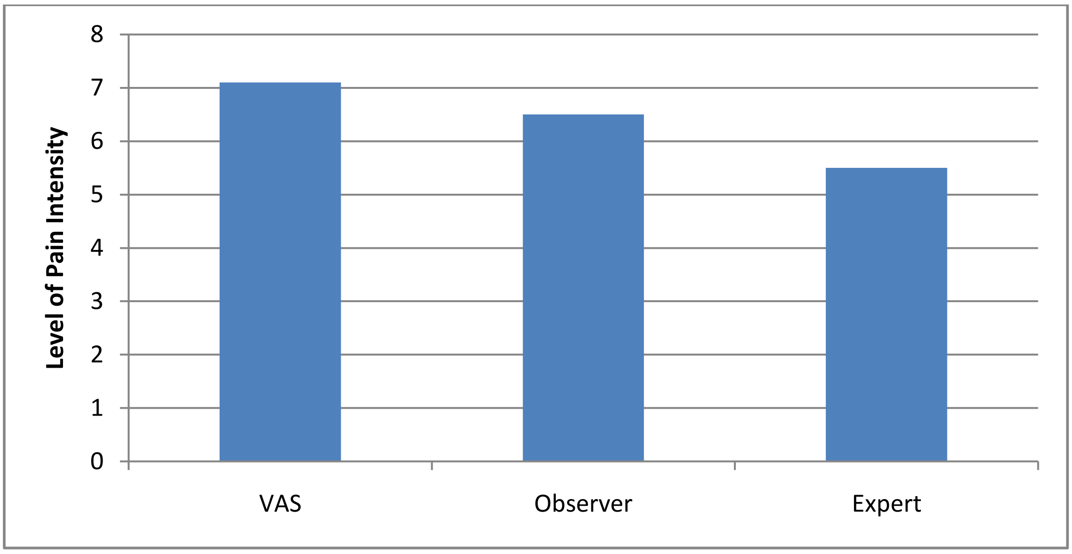

Figure 4.3: The mean of the pain intensity scores received by the three categories of observers i.e., self-report provided by the patient (VAS), observer and expert.

While taking the mean of the scores of all the subjects given by the three types of observers, it was found that the patient using self-report gave a average pain intensity score of 7.1 followed by the observers of 6.42 and experts of 5.52 (Figure 4.3) on a scale of 0-10. With the result obtained, it was clear that the medical practitioner generally underestimated the pain intensity in comparison to the observer. While looking at the other aspect, one can predict that the score of the observer and patient self-report were approximately near to each other. Finally, the unpleasantness was significantly correlated with increased amplitude of facial muscle movements elicited by the intensity level of pain faces. It was also observed that there was more movement of facial muscle (AU) in the case of severe pain which got lesser for the moderate and then for mild pain. The fact that normal faces elicited increased amplitudes below the baseline (Figure 4.2) was entirely different from the pain related facial expressions which were above the baseline. These data indicated that participants needed less time to initiate a normal face than pain faces. In our study, it was noted that the participants displayed greater concentration when they were viewing dynamic facial expressions of pain in comparison with static ones.

Indeed, faces of pain are usually considered as unpleasant and activating (Simon *et al.*, 2008).Thus, taking into account the accurate perception of some other person in pain may be considered as a pertinent cue for social support and delivering effective care when in pain. These methods play a significant role in shedding light on commonly observed underestimation bias in judgments related to pain: the fact that observers, including health care providers, generally underestimated others pain, when the criterion is related to the pain reports of others.

However, some limitations are there of our work that needs further consideration. A first shortcoming of the study was that our sample size was limited. Future research should investigate if gender may also play a significant role in observing the facial expression of pain in others. Therefore, these findings should be taken with caution and considered as exploratory.

At last, our video recording and analysis methods represent novel approaches that have never been used before for quantitative measurement of changes in pain expressions. This method should be used to compare with standardized methods in order to check the reliability of this technique. The current work opens several additional directions for future investigations. To our knowledge, no one has yet compared direct versus indirect measurement of the intensity of pain or other constructs.

Second, following previous work, we measured pain at the frame-by-frame level. However, pain expression is not static but results from the progressive deformation of facial features over time. A next investigation would be to include dynamics when measuring pain intensity.

And third, previous work in both pain and AU detection primarily regards head pose variation as a source of registration error. However, the head pose is itself a potentially informative signal. In particular, head pose changes may themselves be a good indicator of pain and pain intensity. We are currently in the process of exploring the dynamic characteristics of head orientation such as (but not limited to) the speed, velocity, and acceleration of pain indicators. We believe explicit attention to dynamics is an exciting direction for further research.