Title: Automatic coding of facial expressions displayed during posed and genuine pain

Outline

Figures (4)

Tables (3)

Image and Vision Computing
Volume 27, Issue 12, November 2009, Pages 1797-1803
Automatic coding of facial expressions displayed during posed and genuine pain
Author links open overlay panelGwen C. Littlewort a, Marian Stewart Bartlett a, Kang Lee b

https://doi.org/10.1016/j.imavis.2008.12.010Get rights and content
Abstract
We present initial results from the application of an automated facial expression recognition system to spontaneous facial expressions of pain. In this study, 26 participants were videotaped under three experimental conditions: baseline, posed pain, and real pain. The real pain condition consisted of cold pressor pain induced by submerging the arm in ice water. Our goal was to (1) assess whether the automated measurements were consistent with expression measurements obtained by human experts, and (2) develop a

classifier to automatically differentiate real from faked pain in a subject-independent manner from the automated measurements. We employed a machine learning approach in a two-stage system. In the first stage, a set of 20 detectors for facial actions from the Facial Action Coding System operated on the continuous video stream. These data were then passed to a second machine learning stage, in which a classifier was trained to detect the difference between expressions of real pain and fake pain. Naïve human subjects tested on the same videos were at chance for differentiating faked from real pain, obtaining only 49% accuracy. The automated system was successfully able to differentiate faked from real pain. In an analysis of 26 subjects with faked pain before real pain, the system obtained 88% correct for subject independent discrimination of real versus fake pain on a 2-alternative forced choice. Moreover, the most discriminative facial actions in the automated system were consistent with findings using human expert FACS codes.
 * Previous article in issue
 * Next article in issue
Keywords
Machine learning
Computer vision
Malingering
Facial expression
Spontaneous behavior
Automated FACS
1\. Introduction
The computer vision field has advanced to the point that we are now able to begin to apply automatic facial expression recognition systems to important research questions in behavioral science. This paper is among the first applications of fully automated facial expression measurement to such research questions. It explores the application of a machine learning system for automatic facial expression measurement to the task of differentiating fake from real expressions of pain. This application involves measurement of spontaneous expressions, in which the data conditions are much less constrained than posed expressions of basic emotions, on which most automated systems are developed and tested. Ability to process spontaneous expressions shows that automated expression measurement systems can perform effectively in real applications.
The ability to distinguish real pain from faked pain, (malingering) is an important issue in medicine [10]. Naïve human subjects are near chance for differentiating real from fake pain from observing facial expression (e.g. [13]). In the absence of direct training in facial expressions, clinicians are also poor at assessing pain from the face (e.g. [24], [25], [12]). However, a number of studies using the Facial Action Coding System (FACS) [6] have shown that information exists in the face for differentiating real from posed pain (e.g. [14], [3], [22]). In fact, if subjects receive corrective feedback, their performance improves substantially [15]. Thus it appears that a signal is present, but that most people do not know what to look for.
Recent advances in automated facial expression measurement open up the possibility of automatically differentiating posed from real pain using computer vision systems (e.g. [1], [18], [2], [20]). This paper explores the application of a system for automatically detecting facial actions to this problem. The goal of the paper is to (1) assess whether the automated measurements were consistent with expression measurements obtained by human experts, and (2) develop a classifier to automatically differentiate real from faked pain in a subject-independent manner from the automated measurements.

In this study, 26 participants were videotaped under three experimental conditions: baseline, posed pain, and real pain. The real pain condition consisted of cold pressor pain induced by submerging the arm in ice water. We employed a machine learning approach in a two-stage system. In the first stage, the video was passed through a system for detecting facial actions from the Facial Action Coding System [1]. These data were then passed to a second machine learning stage, in which a classifier was trained to detect the difference between expressions of real pain and fake pain. Naïve human subjects were tested on the same videos to compare their ability to differentiate faked from real pain.

Section 2 describes the human subject methods and experiments. Section 3 describes the computer vision system employed for the first stage of the 2-stage approach, in which automated detectors were developed for 20 facial actions. Section 4.1 analyzes the output of the automated facial action detectors and assesses the degree to which they match previous studies based on human coding. Section 4.2 describes the second machine learning stage, in which a classifier for real versus faked pain is trained on the output of the 20 facial action detectors.

The ultimate goal of this work is not the detection of malingering per se, but rather to demonstrate the ability of the automated systems to detect facial behavior that the untrained eye might fail to interpret, and to differentiate types of neural control of the face. It holds out the prospect of illuminating basic questions pertaining to the behavioral fingerprint of neural control systems, and thus opens many future lines of inquiry.

The contribution of this paper is to show that this system for automated FACS can drive further analysis of behavior by the addition of extra computational layers. We show that the automated system is consistent with earlier research, and brings added value by providing measurement of the face when expert human coding is impractical and enabling measurement of expression dynamics than was infeasible with human coding due to the time required. We hope that the availability of such tools will lead to new experimental designs.

1.1. The facial action coding system

FACS [6] is arguably the most widely used method for coding facial expressions in the behavioral sciences. The system describes facial expressions in terms of 46 component movements, which roughly correspond to the individual facial muscle movements. An example is shown in Fig. 1. FACS provides an objective and comprehensive way to analyze expressions into elementary components, analogous to decomposition of speech into phonemes. Because it is comprehensive, FACS has proven useful for discovering facial movements that are indicative of cognitive and affective states. See [8] for a review of facial expression studies using FACS. The primary limitation to the widespread use of FACS is the time required to code. FACS was developed for coding by hand, using human experts. It takes over a hundred hours of training to become proficient in FACS, and it takes about two hours for human experts to code each minute of video. The authors have been developing methods for fully automating the facial action coding system (e.g. [5], [1]). In this paper, we apply a computer vision system trained to automatically detect FACS to the problem of differentiating posed from real expressions of pain.

1. Download: Download high-res image (84KB)
2. Download: Download full-size image

Fig. 1. Example facial action decomposition from the facial action coding system. A prototypical expression of fear is decomposed into seven component movements. Letters indicate intensity. A fear brow (1 + 2 + 4) is illustrated here.

In previous studies using manual FACS coding by human experts, at least 12 facial actions showed significant relationships with pain across multiple studies and pain modalities. Of these, the ones specifically associated with cold pressor pain were 4, 6, 7, 9, 10, 12, 25, 26 [4], [22]. See Table 1 and Fig. 2 for names and examples of these AU?s. A previous study compared faked to real pain, but in a different pain modality (lower back pain). This study found that when faking, subjects tended to display the following AU?s: 4, 6, 7, 10, 12, 25. When faked pain expressions were compared to real pain expressions, the faked pain expressions contained significantly more brow lower (AU 4), cheek raise (AU 6), and lip corner pull (AU 12) [3]. These studies also reported substantial individual differences in the expressions of both real pain and faked pain, making automated detection of faked pain a challenging problem.

Table 1. AU detection performance on posed and spontaneous facial actions. Values are area under the roc (A?) for generalization to novel subjects.

AU	Name	Posed	Spont
1	Inner brow raise	.90	.88
2	Outer brow raise	.94	.81
4	Brow lower	.98	.73
5	Upper lid raise	.98	.80
6	Cheek raise	.85	.89
7	Lids tight	.96	.77
9	Nose wrinkle	.99	.88
10	Upper lip raise	.98	.78
12	Lip corner pull	.97	.92
14	Dimpler	.90	.77
15	Lip corner depress	.80	.83
17	Chin raise	.92	.80
18	Lip pucker	.87	.70
20	Lip stretch	.98	.60
23	Lip tighten	.89	.63
24	Lip press	.84	.80
25	Lips part	.98	.71
26	Jaw drop	.98	.71
1, 1 + 4	Distress brow	.94	.70
1 + 2 + 4	Fear brow	.95	.63
Mean		.93	.77

1. Download: Download high-res image (40KB)
2. Download: Download full-size image

Fig. 2. Sample behavior: (a) faked pain and (b) real pain.

1.2. Spontaneous expressions

The machine learning system presented here was trained on spontaneous facial expressions. The importance of using spontaneous behavior for developing and testing computer vision systems becomes apparent when we examine the neurological substrate for facial expression. There are two distinct neural pathways that mediate facial expressions, each one originating in a different area of the brain. Volitional facial movements originate in the cortical motor strip, whereas it has been suggested that spontaneous facial expressions originate in the subcortical areas of the brain (see [26], for a review).

These two pathways have different patterns of innervation on the face, with the cortical system tending to give stronger innervation to certain muscles primarily in the lower face, while the subcortical system tends to more strongly innervate certain muscles primarily in the upper face (e.g. [19]).

The facial expressions mediated by these two pathways have differences both in which facial muscles are moved and in their dynamics [7], [8]. Subcortically initiated facial expressions (the spontaneous group) are characterized by synchronized, smooth, symmetrical, consistent, and reflex-like facial muscle movements whereas cortically initiated facial expressions (posed expressions) are subject to volitional real-time control and tend to be less smooth, with more variable dynamics [26], [11], [27], [2]. Given the two different neural pathways for facial expressions, it is reasonable to expect to find differences between genuine and posed expressions of pain. Moreover, it is crucial that the computer vision model for detecting genuine pain is based on machine learning of spontaneous examples of real pain expressions.

2\. Human subject methods

Video data were collected of 26 human subjects during real pain, faked pain, and baseline conditions. Human subjects were university students consisting of 6 men and 20 women. The pain condition consisted of cold pressor pain induced by immersing the arm in cold water at 3 °C. For the baseline and faked pain conditions, the water was 20 °C. Subjects were instructed to immerse their forearm into the water up to the elbow, and hold it there for 60 s in each of the three conditions. The order of the conditions was baseline, faked pain, and then real pain. For the faked pain condition, subjects were asked to manipulate their facial expressions so that an ?expert would be convinced they were in actual pain.? Participants facial expressions were recorded using a digital video camera during each condition.

A second subject group underwent the conditions in the counterbalanced order, with real pain followed by faked pain. This ordering involves immediate memory of the pain just felt, which is a fundamentally different task from imagining unknown pain. The present paper therefore analyzes only the first subject group. The second group will be analyzed separately in a future paper, and compared to the first group.

After the videos were collected, a set of 170 naïve observers were shown the videos and asked to guess whether each video contained faked or real pain. Subjects were undergraduates with no explicit training in facial expression measurement. They were primarily Psychology majors at U. Toronto. Mean accuracy of naïve human subjects for discriminating fake from real pain in these videos was near chance at 49.1% (SD = 13.7). These observers had no specific training in facial expression and were not clinicians. One might suppose that clinicians would be more accurate. However, previous studies suggest that clinicians judgments of pain from the face are similarly unreliable (e.g. [25], [12]). Facial signals to differentiate real from faked pain do appear to exist however [15], [3], [22], [23]. A system based on machine learning from examples of real and faked pain expressions could pick up these signals. This paper investigated whether an automated system could outperform the naïve human subjects on the same set of videos.

3\. Automated facial action detection

The first stage of the computer vision analysis was to employ a system for fully automated facial action coding developed previously by the authors [1], [18]. In this paper, we evaluate this system on the task of measuring facial expressions of pain, and then extend it by developing a second stage classifier for real versus faked pain that operates on the facial action output channels. Here we describe the facial action detection system employed in Stage One. It is a user independent fully automatic system for real time recognition of facial actions from the FACS. The system automatically detects frontal faces in the video stream and codes each frame with respect to 20 action units. It is a machine learning system on image-based features. In previous work, we conducted empirical comparisons of image features, including Gabors, independent components, and flow-based features (e.g. [5]), classifiers such as AdaBoost, support vector machines, and linear discriminant analysis, as well as feature selection techniques [18]. An overview of the system is shown in Fig. 3.

1. Download: Download high-res image (89KB)
2. Download: Download full-size image

Fig. 3. Overview of the automated facial action recognition system.

3.1. Real time face and feature detection

The system employs a real-time face detection system that uses boosting techniques in a generative framework [9] and extends work by Viola and Jones [29]. Enhancements to Viola and Jones include employing Gentleboost instead of AdaBoost, smart feature search, and a novel cascade training procedure, combined in a generative framework. The source code for the face detector is freely available at http://kolmogorov.sourceforge.net. Accuracy on the CMU-MIT dataset, a standard public data set for benchmarking frontal face detection systems [28], is 90% detections and 1/million false alarms, which is state-of-the-art accuracy. The CMU test set has unconstrained lighting and background. With controlled lighting and background, such as the facial expression data employed here, detection accuracy is much higher. All faces in the training datasets, for example, were successfully detected. The system presently operates at 24 frames/s on a 3 GHz Pentium IV for 320 × 240 images. The automatically located faces were rescaled to 96 × 96 pixels. The typical distance between the centers of the eyes was roughly 48 pixels. Previously, automatic eye detection [9] was employed to align the eyes in each image. The current system aligns the face based on four points ? the centers of the eyes, nose and mouth. For this study only, we applied temporal smoothing to these feature positions and linear compensation for motion-related alignment anomalies. There is a choice of Procrustes alignment or unconstrained least square alignment. We used procrustes for this experiment. The images were then passed through a bank of Gabor filters eight orientations and 9 spatial frequencies (2?32 pixels per cycle at 1/2 octave steps). Output magnitudes were then passed to the action unit classifiers.

3.2. Facial action classifiers

The training data for the facial action classifiers came from four data sets, not including the pain data ? three posed datasets and one dataset of spontaneous expressions. The facial expressions in each dataset were FACS coded by certified FACS coders. The first posed datasets was the Cohn?Kanade DFAT-504 dataset [16]. This dataset consists of 100 university students who were instructed by an experimenter to perform a series of 23 facial displays, including expressions of seven basic emotions. The second posed dataset consisted of directed facial actions from 24 subjects collected by Ekman and Hager. Subjects were instructed by a FACS expert on the display of individual facial actions and action combinations, and they practiced with a mirror. The resulting video was verified for AU content by two certified FACS coders. The third posed dataset consisted of a subset of 50 videos from 20 subjects from the MMI database [21]. The spontaneous expression dataset consisted of the FACS-101 dataset collected by Mark Frank [1]. Thirty-three subjects underwent an interview about political opinions on which they felt strongly. Two minutes of each subject were FACS coded. The total training set consisted of 5500 examples, 2500 from posed databases and 3000 from the spontaneous set. Linear Support Vector Machines were trained for each of 20 facial actions. Separate binary classifiers, one for each action, were trained to detect the presence of the action in a ?one versus all? manner. Positive examples consisted of the apex frame for the target AU. Negative examples consisted of all apex frames that did not contain the target AU plus neutral images obtained from the first frame of each sequence. Eighteen of the detectors were for individual action units, and two of the detectors were for specific brow region combinations: fear brow (1 + 2 + 4) and distress brow (1 alone or 1 + 4). All other detectors were trained to detect the presence of the target action regardless of co-occurring actions. A list is shown in Table 1.

The output of the system was a real valued number indicating the distance to the separating hyperplane for each classifier. Previous work showed that the distance to the separating hyperplane (the margin) contained information about action unit intensity (e.g. [1]).

In this paper, area under the ROC (A?) is used to assess performance rather than overall percent correct, since percent correct can be an unreliable measure of performance, as it depends on the proportion of targets to non-targets, and also on the decision threshold. Similarly, other statistics such as true positive and false positive rates depend on decision threshold, which can complicate comparisons across systems. A? is a measure derived from signal detection theory and characterizes the discriminative capacity of the signal, independent of decision threshold. The ROC curve is obtained by plotting true positives against false positives as the decision threshold shifts from 0% to 100% detections. The area under the ROC (A?) ranges from 0.5 (chance) to 1 (perfect discrimination). A? can also be interpreted in terms of percent correct. A? is equivalent to the theoretical maximum percent correct achievable with the information provided by the system when using a 2-Alternative Forced Choice testing paradigm.

Table 1 shows performance for detecting facial actions in posed and spontaneous facial actions. Generalization to novel subjects was tested using threefold cross-validation on the images in the FACS training set (not the pain video). Performance was separated into the posed set, which was 2500 images, and a spontaneous set, which was 1100 images from the FACS-101 database which includes speech. The performance on the spontaneous data is lower than posed data because of less restrictive conditions on head position, lighting, movement and the lower intensity of spontaneous expressions. Despite this performance loss, particularly in the very important brow lowering AU, task labels based on one minute worth of this system output could be learned successfully. In addition to FACS accuracy, the decision is implicitly influenced by combinations and temporal effects.

4. Automated measurement of real and fake pain expressions

Applying this system to the pain video data produced a 20 channel output stream, consisting of one real value for each learned AU, for each frame of the video. We first examine the output of the automated facial action detectors. We assess which AU outputs contained information about genuine pain expressions, faked pain expressions, and show differences between genuine versus faked pain. The results are compared to studies that employed expert human coding. Section 4.2 goes on to develop a second-stage classifier to automatically discriminate real from faked pain from the 20-channel output stream.

4.1. Characterizing the difference between real and fake pain expressions

We first examined which facial action detectors were elevated in real pain compared to the baseline condition. *Z* -scores for each subject and each AU detector were computed as *Z* = (*x* ? *?*)/*?* , where (*?* , *?*) are the mean and variance for the output of frames 100?1100 in the baseline condition (warm water, no faked expressions). The mean difference in *Z* -score between the baseline and pain conditions was computed across the 26 subjects. Table 2 shows the action detectors with the largest difference in _Z_ -scores. We observed that the actions with the largest _Z_ -scores for genuine pain were Mouth opening and jaw drop (25 and 26), lip corner puller by zygomatic (12), nose wrinkle (9), and to a lesser extent, lip raise (10) and cheek raise (6). These facial actions have been previously associated with cold pressor pain (e.g. [22], [4]).

Table 2. _Z_ -score differences of the three pain conditions, averaged across subjects. FB, fear brow 1 + 2 + 4; DB, distress brow (1, 1 + 4).

Action unit	25	12	9	26	10	6
(A) Real pain vs baseline						
_Z_ -score	1.4	1.4	1.3	1.2	0.9	0.9

Action unit	4	DB	1	25	12	6	26	10	FB	9	20	7
(B) Faked pain vs baseline												
_Z_ -score		2.7	2.1	1.7	1.5	1.4	1.4	1.3	1.3	1.2	1.1	1.0
0.9												

Action unit	4	DB	1
(C) Real pain vs faked pain			
_Z_ -score difference	1.8	1.7	1.0

The _Z_ -score analysis was next repeated for faked versus baseline. We observed that in faked pain there was relatively more facial activity than in real pain. The facial action outputs with the highest _Z_ -scores for faked pain relative to baseline were brow lower (4), distress brow (1 or 1 + 4), inner brow raise (1), mouth open and jaw drop (25 and 26), cheek raise (6), lip raise (10), fear brow (1 + 2 + 4), nose wrinkle (9), mouth stretch (20), and lower lid raise (7).

Differences between real and faked pain were examined by computing the difference of the two _Z_ -scores. Differences were observed primarily in the outputs of action unit 4 (brow lower), as well as distress brow (1 or 1 + 4) and inner brow raise (1 in any combination).

Individual subject differences between faked and real pain are shown in Table 3. Difference-of-_Z_ -scores between the genuine and faked pain conditions were computed for each subject and each AU. There was considerable variation among subjects in the difference between their faked and real pain expressions. However, the most consistent finding is that 9 of the 26 subjects showed significantly more brow lowering activity (AU 4) during the faked pain condition, whereas none of the subjects showed significantly more AU 4 activity during the real pain condition. Also seven subjects showed more cheek raise (AU 6), and six subjects showed more inner brow raise (AU 1), and the fear brow combination (1 + 2 + 4). The next most common differences were to show more 12, 15, 17, and distress brow (1 alone or 1 + 4) during faked pain.

Table 3. Individual subject differences between faked and genuine pain. Differences greater than 2 standard deviations are shown. _F_ > _P_ : number of subjects in which the output for the given AU was greater in faked than genuine pain. _P_ > _F_ : number of subjects for which the output was greater in genuine than faked pain. FB: fear brow 1 + 2 + 4. DB: distress brow (1, 1 + 4).

AU	1	2	4	5	6	7	9	10	12	14	15	17	18	20	23	24	25	26	FB	DB
_F_ > _P_	6	4	9	1	7	4	3	6	5	3	5	5	1	4	3	4	4	4	6	5
_P_ > _F_	3	3	0	0	4	0	4	4	4	2	3	1	3	1	1	1	2	4	2	0

Paired _t_ -tests were conducted for each AU to assess whether it was a reliable indicator of genuine versus faked pain in a within-subjects design. Of the 20 actions tested, the difference was statistically significant for three actions. It was highly significant for AU 4 (_p_ < .001), and marginally significant for AU 7 and distress brow (_p_ < .05).

In order to characterize action unit combinations that relate to the difference between fake and real pain expressions, principal component analysis was conducted on the difference-of-_Z_ -scores. The first eigenvector had the largest loading on distress brow and inner brow raise (AU 1). The second eigenvector had the largest loading on lip corner puller (12) and cheek raise (6) and was _lower_ for fake pain expressions. The third eigenvector had the largest loading on brow lower (AU 4). Thus when analyzed singly, the action unit channel with the most information for discriminating fake from real pain was brow lower (AU 4). However, when correlations were assessed through PCA, the largest variance was attributed to two combinations, and AU 4 accounted for the third most variance.

4.1.1. Comparison with previous studies that employed human expert coding

Overall, the outputs of the automated system showed similar patterns to previous studies of real and faked pain using manual FACS coding by human experts. Exaggerated activity of the brow lower (AU 4) during faked pain is consistent with previous studies in which the real pain condition was exacerbated lower back pain [3], [14]. Another study performed a FACS analysis of fake and real pain expressions with cold pressor pain, but with children ages 8?12 [17]. This study observed significant elevation in the following AU?s for fake pain relative to baseline: 1 4 6 7 10 12 20 23 25 26. This closely matches the AU?s with the highest _Z_ -scores in the automated system output of the present study (Table 2B). LaRochette et al. did not measure AU 9 or the brow combinations. When faked pain expressions were compared with real cold pressor pain in children, LaRochette et al found significant differences in AU?s 1 4 7 10. Again the findings of the present study using the automated system are similar, as the AU channels with the highest _Z_ -scores were 1, 4, and 1 + 4 (Table 2C), and the _t_ -tests were significant for 4, 1 + 4 and 7.

4.1.2. Comparison with human expert coding of a subset of the video data

In order to further assess the validity of the automated system findings, we obtained FACS codes for a portion of the video data employed in this study. FACS codes were obtained by an expert coder certified in the Facial Action Coding System. For each subject, the last 500 frames of the fake pain and real pain conditions were FACS coded (about fifteen seconds each). It took 60 man hours to collect the human codes, over the course of more than 3 months, since human coders can only code up to two hours per day before having negative repercussions in accuracy and coder burn-out.

The sum of the frames containing each action unit were collected for each subject condition, as well as a weighted sum, multiplied by the intensity of the action on a 1?5 scale. To investigate whether any action units successfully differentiated real from faked pain, paired _t_ -tests were computed on each individual action unit. (Tests on specific brow region combinations 1 + 2 + 4 and 1, 1 + 4 have not yet been conducted.) The one action unit that significantly differentiated the two conditions was AU 4, brow lower, (_p_ < .01) for both the sum and weighted sum measures. This finding is consistent with the analysis of the automated system, which also found action unit 4 most discriminative.

The above analysis examined the outputs of the automated facial action detectors and compared them to studies that employed expert human coding. We next turned to the problem of automatically discriminating genuine from faked pain expressions from the facial action output stream.

4.2. Automated classification of real vs faked pain

This section describes the second machine learning stage, in which a classifier was trained to discriminate genuine from faked pain from the output of the facial action detectors. The task was to perform subject-independent classification. If the task were to simply detect the presence of a red-flag set of facial actions, then differentiating fake from real pain expressions would be relatively simple. However, it should be noted that subjects display actions such as AU 4, for example, in both real and fake pain, and the distinction is in the magnitude and duration of AU 4. Also, there is inter-subject variation in expressions of both real and fake pain, there may be combinatorial differences in the sets of actions displayed during real and fake pain, and the subjects may cluster. We therefore applied machine learning to the task of discriminating real from faked pain expressions.

A second-layer classifier was trained to discriminate genuine pain from faked pain based on the 20-channel output stream. The system was trained using cross-validation on the 26 subject videos described in Section 2. The input to this second stage consisted of the 20 facial action detector outputs from the full minute of video in each condition. In the cross-validation approach, the system was trained and tested 26 times, each time using data from 25 subjects for parameter estimation and reserving a different subject for testing. This provided an estimate of subject-independent detection performance.

Prior to learning, the system performed an automatic reliability estimate of the face alignment based on the smoothness of the eye positions. Abrupt shifts of two to five pixels tend to occur during eyeblinks. (A more recent version of the eye detector corrects this issue.) Those frames with abrupt shifts of 2 or more pixels in the returned eye positions were automatically detected and the feature positions were recomputed by interpolating from neighboring frames. This eye position filter had a relatively small effect on performance. The analysis of Table 2 was repeated under this criterion, and the _Z_ -scores improved by about 0.1.

The first approach we examined was to employ input features that consisted of window-based statistics. The sixty second from each condition was broken up into six overlapping segments of 500 frames, the windows. For each segment, the following five statistics were measured for each of the 20 AU?s: median, maximum, range, first to third quartile difference and ninetieth to one hundredth percentile difference. Thus the input to the classifier for each segment contained 100 dimensions. Each cross-validation trial contained 300 training samples (25 subjects × 2 conditions × 6 segments).

For this second-layer classification step, we first explored SVM?s, Adaboost, and linear discriminant analysis. Nonlinear SVM?s with radial basis function kernels gave the best performance. Linear classifiers may be inadequate to deal with the combinations from different behavioral clusters for the same task.

A nonlinear SVM trained to discriminate posed from real facial expressions of pain obtained an area under the ROC curve of 0.72 for generalization to novel subjects. This was significantly higher than performance of naïve human subjects, who obtained a mean accuracy of 49% correct for discriminating faked from real pain on the same set of videos.

Our later approach was based on statistics of integrated ?events?. We applied temporal filters at eight different fundamental frequencies to the AU output stream. Whenever these filter outputs were continuously greater than zero for a particular AU, we integrated the intensity into one ?event?. The histograms of these integrated intensities were used as a representation to train a Gaussian SVM. The percent correct 2-alternative forced choice of fake versus

real pain on new subjects was 88%.
This Integrated Event representation, shown in Fig. 4, contains useful dynamic information allowing more accurate behavioral analysis. This suggests that this decision task depends not only on which subset of AU?s are present at which intensity, but also on the duration and number of AU events. More detailed analysis of the dynamics of the action units is currently underway.
1. Download: Download high-res image (66KB)
2. Download: Download full-size image

Fig. 4. Example of integrated event representation, for one subject, showing AU 4 for two of the eight frequencies used.

5\. Discussion

The field of automatic facial expression analysis has advanced to the point that we can begin to apply it to address research questions in behavioral science. Here we describe a pioneering effort to apply fully automated facial action coding to the problem of differentiating fake from real expressions of pain. While naïve human subjects were only at 49% accuracy for distinguishing fake from real pain, the automated system obtained 88% correct on a 2-alternative forced choice. Moreover, the pattern of results in terms of which facial actions may be involved in real pain, fake pain, and differentiating real from fake pain is similar to previous findings in the psychology literature using manual FACS coding.
Here we applied machine learning on a 20-channel output stream of facial action detectors. The machine learning was applied to samples of spontaneous expressions during the subject state in question. Here the state in question was fake versus real pain. The same approach can be applied to learn about other subject states, given a set of spontaneous expression samples. For example, we recently developed a related system to detect driver drowsiness from facial expression [30].
While the accuracy of individual facial action detectors is still below that of human experts, automated systems can be applied to large quantities of video data. Statistical pattern recognition on this large quantity of data can reveal emergent behavioral patterns that previously would have required hundreds of coding hours by human experts, and would be unattainable by the non-expert. Moreover, automated facial expression analysis will enable investigations into facial expression dynamics that were previously intractable by human coding because of the time required to code intensity changes. Future work in automatic discrimination of fake and real pain will include investigations into facial expression dynamics.

Acknowledgments

Portions of the research in this paper use the MMI Facial Expression Database collected by M. Pantic and M.F. Valstar. Support for this work was provided by NSF CNS-0454233 and NSF ADVANCE award 0340851. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. [1]
M.S. Bartlett, G.C. Littlewort, M.G. Frank, C. Lainscsek, I. Fasel, J.R. Movellan
Automatic recognition of facial actions in spontaneous expressions
J. Multimedia, 1 (6) (2006), pp. 22-35
View in ScopusGoogle Scholar
2. [2]

J.F. Cohn, K.L. Schmidt
The timing of facial motion in posed and spontaneous smiles
J. Wavelets Multi-resolution Inform. Process., 2 (2) (2004), pp. 121-132
Google Scholar
3. [3]
K.D. Craig, S. Hyde, C.J. Patrick
Genuine, supressed, and faked facial behaviour during exacerbation of chronic low back pain
Pain, 46 (1991), pp. 161-172
CrossrefView in ScopusGoogle Scholar
4. [4]
K.D. Craig, C.J. Patrick
Facial expression during induced pain
J. Pers. Soc. Psychol., 48 (4) (1985), pp. 1080-1091
CrossrefView in ScopusGoogle Scholar
5. [5]
G. Donato, M.S. Bartlett, J.C. Hager, P. Ekman, T.J. Sejnowski
Classifying facial actions
IEEE Trans. Pattern Anal. Mach. Intell., 21 (10) (1999), pp. 974-989
Google Scholar
6. [6]
P. Ekman, W. Friesen
Facial Action Coding System: A Technique for the Measurement of Facial Movement
Consulting Psychologists Press, Palo Alto, CA (1978)
Google Scholar
7. [7]
P. Ekman
Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage
W.W. Norton, New York (2001)
Google Scholar
8. [8]
P. Ekman, E.L. Rosenberg (Eds.), What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the FACS, Oxford University Press, Oxford (2005)
9. [9]
I. Fasel, B. Fortenberry, J.R. Movellan
A generative framework for real-time object detection and classification
Comput. Vis. Image Understand. (2005), p. 98
Google Scholar
10. [10]
D.A. Fishbain, R. Cutler, H.L. Rosomoff, R.S. Rosomoff
Accuracy of deception judgments
Pers. Soc. Psychol. Rev., 10 (3) (2006), pp. 214-234
Google Scholar
11. [11]
M.G. Frank, P. Ekman, W.V. Friesen
Behavioral markers and recognizability of the smile of enjoyment
J. Pers. Soc. Psychol., 64 (1) (1993), pp. 83-93
View in ScopusGoogle Scholar
12. [12]
S. Grossman, V. Shielder, K. Swedeen, J. Mucenski
Correlation of patient and caregiver ratings of cancer pain
J. Pain Symp. Manage., 6 (2) (1991), pp. 53-57

View PDFView articleView in ScopusGoogle Scholar
13. [13]
H.D. Hadjistavropoulos, K.D. Craig, T. Hadjistavropoulos, G.D. Poole
Subjective judgments of deception in pain expression: accuracy and errors
Pain, 65 (2?3) (1996), pp. 251-258
View PDFView articleCrossrefView in ScopusGoogle Scholar
14. [14]
M.L. Hill, K.D. Craig
Detecting deception in pain expressions: the structure of genuine and deceptive facial displays
Pain, 98 (1?2) (2002), pp. 135-144
View PDFView articleView in ScopusGoogle Scholar
15. [15]
M.L. Hill, K.D. Craig
Detecting deception in facial expressions of pain: accuracy and training
Clin. J. Pain, 20 (6) (2002), pp. 415-422
Google Scholar
16. [16]
T. Kanade, J.F. Cohn, Y. Tian, Comprehensive database for facial expression analysis, in: Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG?00), Grenoble, France, 2000, pp. 46?53.
Google Scholar
17. [17]
A.C. Larochette, C.T. Chambers, K.D. Craig
Genuine, suppressed and faked facial expressions of pain in children
Pain, 15;126 (1?3) (2006), pp. 64-71
Dec
View PDFView articleCrossrefView in ScopusGoogle Scholar
18. [18]
G. Littlewort, M.S. Bartlett, I. Fasel, J. Susskind, J. Movellan
Dynamics of facial expression extracted automatically from video
J. Image Vis. Comput., 24 (6) (2006), pp. 615-625
View PDFView articleView in ScopusGoogle Scholar
19. [19]
R.J. Morecraft, J.L. Louie, J.L. Herrick, K.S. Stilwell-Morecraft
Cortical innervation of the facial nucleus in the non-human primate: a new interpretation of the effects of stroke and related subtotal brain trauma on the muscles of facial expression
Brain, 124 (1) (2001), pp. 176-208
View in ScopusGoogle Scholar
20. [20]
M. Pantic, A. Pentland, A. Nijholt, T. Huang, Human computing and machine understanding of human behaviour: a survey, in: Proceedings of the ACM International Conference on Multimodal Interfaces, 2006, pp. 239?248.
Google Scholar
21. [21]
M. Pantic, M.F. Valstar, R. Rademaker, L. Maat, Web-based Database for facial expression analysis, in: Proceedings of the IEEE International Conference on Multimedia and Expo (ICME?05), Amsterdam, The Netherlands, July 2005.
Google Scholar
22. [22]
KM. Prkachin
Dissociating deliberate and spontaneous expressions of pain

Pain, 51 (3) (1992), pp. 57-65
View PDFView articleView in ScopusGoogle Scholar
23. [23]
KM. Prkachin
Effects of deliberate control on verbal and facial expressions of pain
Pain, 114 (2005), pp. 328-338
View PDFView articleCrossrefView in ScopusGoogle Scholar
24. [24]
K.M. Prkachin, I. Schultz, J. Berkowitz, E. Hughes, D. Hunt
Assessing pain behaviour of low-back pain patients in real time: concurrent validity and examiner sensitivity
Behav. Res. Ther., 40 (5) (2002), pp. 595-607
View PDFView articleView in ScopusGoogle Scholar
25. [25]
K.M. Prkachin, P.A. Solomon, A.J. Ross
The underestimation of pain among health-care providers
Can. J. Nurs. Res., 39 (2007), pp. 88-106
View in ScopusGoogle Scholar
26. [26]
W.E. Rinn
The neurophysiology of facial expression: a review of the neurological and psychological mechanisms for producing facial expression
Psychol. Bull., 95 (1984), pp. 52-77
View in ScopusGoogle Scholar
27. [27]
K.L. Schmidt, J.F. Cohn, Y. Tian
Signal characteristics of spontaneous facial expressions: automatic movement in solitary and social smiles
Biol. Psychol., 65 (1) (2003), pp. 49-66
View PDFView articleView in ScopusGoogle Scholar
28. [28]
H. Schneiderman, T. Kanade, Probabilistic modeling of local appearance and spatial relationships for object recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1998, pp. 45?51.
Google Scholar
29. [29]
P. Viola, M. Jones
Robust real-time face detection
J. Comput. Vis., 57 (2) (2004), pp. 137-154
Google Scholar
30. [30]
E. Vural, A. Ercil, G.C. Littlewort, M.S. Bartlett, J.R. Movellan, Machine learning systems for detecting driver drowsiness, in: Proceedings of the Biennial Conference on Digital Signal Processing for in Vehicle and Mobile Systems, 2007.
Google Scholar

Part of special issue
Visual and multimodal analysis of human spontaneous behaviour:
Edited by
Maja Pantic, Jeffrey F. Cohn
Download full issue
Other articles from this issue
 * ### Social signal processing: Survey of an emerging domain
November 2009
Alessandro Vinciarelli, ?, Hervé Bourlard
View PDF
 * ### Being bored? Recognising natural interest by extensive audiovisual integration for real-life application
November 2009
Björn Schuller, ?, Hitoshi Konosu
View PDF
 * ### Automatic nonverbal analysis of social interaction in small groups: A review
November 2009
Daniel Gatica-Perez
View PDF
View more articles
Recommended articles
No articles found.

digital experience. For more information, see our Cookie Policy and the list of Google Ad-Tech Vendors.

You may choose not to allow some types of cookies. However, blocking some types may impact your experience of our site and the services we are able to offer. See the different category headings below to find out more or change your settings.

Allow all

Manage Consent Preferences

Strictly Necessary Cookies

Always active

These cookies are necessary for the website to function and cannot be switched off in our systems. They are usually only set in response to actions made by you which amount to a request for services, such as setting your privacy preferences, logging in or filling in forms. You can set your browser to block or alert you about these cookies, but some parts of the site will not then work. These cookies do not store any personally identifiable information.

Cookie Details List?

Functional Cookies

Functional Cookies

These cookies enable the website to provide enhanced functionality and personalisation. They may be set by us or by third party providers whose services we have added to our pages. If you do not allow these cookies then some or all of these services may not function properly.

Cookie Details List?

Performance Cookies

Performance Cookies

These cookies allow us to count visits and traffic sources so we can measure and improve the performance of our site. They help us to know which pages are the most and least popular and see how visitors move around the site.

Cookie Details List?

Targeting Cookies

Targeting Cookies

These cookies may be set through our site by our advertising partners. They may be used by those companies to build a profile of your interests and show you relevant adverts on other sites. If you do not allow these cookies, you will experience less targeted advertising.

Cookie Details List?

Back Button

Cookie List

Search Icon

Filter Icon

Clear

checkbox label label

Apply Cancel

Consent Leg.Interest

checkbox label label

checkbox label label

checkbox label label

Confirm my choices