Thermal Analysis
of Facial Muscles Contractions

Sophie Jarlier, Didier Grandjean, Sylvain Delplanque, Karim N'Diaye, Isabelle Cayeux,
Maria Inés Velazco, David Sander, Patrik Vuilleumier, and Klaus R. Scherer

Abstract—Facial expressions can be systematically coded using the Facial Action Coding System (FACS) that describes the specific action unit (AU) or combination of AUs elicited during different kinds of expressions. This study investigated the thermal patterns concomitant to specific action units performance. As thermal imaging can track dynamic patterns in facial temperature at any distance (> 0.4 m), with high temporal (< 20 m) and thermal (< 20 mK@300 K) resolutions, this noninvasive technique was tested as a method to assess fluctuations of facial heat patterns induced by facial muscles contractions. Four FACS-trained coders produced nine different AUs or combination of AUs at various speeds and intensities. Using a spatial pattern approach based on PCA decomposition of the thermal signal, we showed that thermal fluctuations are specific to the activated AUs and are sensitive to the kinetics and intensities of AU production. These results open new avenues for studying patterns of facial muscle activity related to emotion or other cognitively induced activities, in a noninvasive manner, avoiding potential lighting issues.

Index Terms—Facial expression, FACS, muscle contraction, thermography.

✦

---

1 INTRODUCTION

FACIAL expressions are complex muscular patterns that carry complex social signals. A facial expression results from one or more motions or positions of the muscles of the face. In order to study and analyze facial muscles contractions, researchers use techniques such as Facial Action Coding System (FACS) coding, video-taped images recording, and electromyography (EMG).

Ekman et al. [1] have developed the most popular standard system to classify the physical aspects of facial expressions: the FACS. This system is based on the anatomy of the facial muscles and is composed of action units (AUs) that describe all visible facial movements at different intensities. Since 1976, when the FACS system was first developed, 7,000 different combinations of AUs have been identified in a large number of facial expressions (44 AUs for changes in facial expression and 12 AUs for changes in gaze direction and head orientation [2]). Nowadays, a standardized procedure exists (FACS Final Test) to train researchers who are interested in becoming FACS coders not only to decode the precise AUs expressed by others, but also to produce these AUs on demand. Although very informative, this process of coding each action unit in a facial expression is very time-consuming.

Several attempts have been made to automatize the coding though. For instance, Lien et al. [3] used video-taped images under visible spectrum lighting to automatically detect, track, and classify the AUs implied in the expressions. Unfortunately, the influence of lighting on image quality (contrast fluctuations or low light) limits this visible-spectrum imagery technique.

To circumvent this problem, researchers can directly record the electrical activity of muscles that subtend the AUs by means of facial EMG, which measures muscle contraction (even the visually imperceptible). This technique is particularly sensitive to measure the kinetics and intensity of that muscular contraction [4]. However, EMG recording is not without drawbacks: 1) It can be difficult to record the precise activity of a specific muscle involved in a given AU because of the diffusion of electrical activity from one muscle to another (i.e., cross-talk phenomenon); 2) electrodes must be fixed on many areas of the face, a constraint that could hamper natural muscular contraction; and 3) theoretically, there should be as many electrodes as there are different muscles related to the AUs. This last point constitutes a severe limitation for the use of EMG as a noninvasive method. However, to date no technique has been developed that would allow the simultaneous recording of all facial muscle activity, being sensitive to the intensity and the temporal dynamics of the contractions, and without hindering natural AU production or facing light problems.

• S. Jarlier, S. Delplanque, and K.R. Scherer are with the Swiss Center for Affective Sciences (CISA), University of Geneva, 7 rue des Battoirs, 1205 Geneva, Switzerland.
E-mail: {sophie.jarlier, sylvain.delplanque, klaus.scherer}@unige.ch.
• D. Grandjean and D. Sander are with the Department of Psychology, University of Geneva, 40 bd du Pont d'Arve, 1205 Geneva, Switzerland.
E-mail: {didier.grandjean, david.sander}@unige.ch.
• K. N'Diaye is with the Equipe Cogimage (ex-LENA CNRS UPR 640), CRICM—UPMC/INSERM UMR-S975/CNRS UMR7225, 47 bd de l'Hôpital, F-75651 Paris Cedex 13. E-mail: karim.ndiaye@upmc.fr.
• I. Cayeux and M.I. Velazco are with the Firmenich SA—PO Box 239, Route des Jeunes 1, CH-1211 Geneva 8, Switzerland.
E-mail: isabelle.cayeux@firmenich.com.
• P. Vuilleumier is with the Laboratory of Behavioral Neurology and Imaging of Cognition (LabNIC), Department of Neuroscience/Physiology, University Medical Center (CMU), Bat. A, 7th floor, room 7008, 1 rue Michel-Servet, CH-1211 Geneva, Switzerland.
E-mail: patrik.vuilleumier@unige.ch.
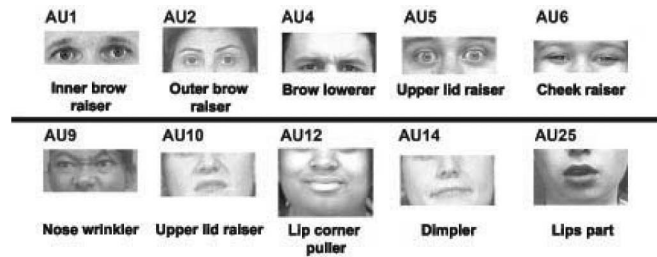
Fig. 1. Experimental settings.

Fig. 2. Action units produced by FACS coders [9].

procedure to extract the facial features from thermal images, i.e., the representative heat maps for each requested action unit. In particular, a spatial pattern detection procedure (using a principal component analysis) was undertaken, allowing the detection of coherent heat changes in the face without any a priori assumption about the particular facial area that would be activated.

2 METHOD

2.1 Participants (Coders)

Four trained and certified FACS coders (three women, all right handed, 28-51 years old) participated in our experiment as AUs coders.

2.2 Procedure

The coders were seated on a comfortable chair in a dimly lit room at a temperature between 20 and 23° C. Their heads were immobilized with a noninvasive head fixation system made for this purpose (head frame, Fig. 1). Participants were asked not to use any makeup or facial products the day of the experiment. In addition, they were asked not to eat or drink hot substances and not to smoke during the hour preceding the experiment. Their facial skin was washed with 70 percent alcohol to remove any interfering substances. The experiment was then described and the participants rested for 15 min to acclimatize with the surrounding temperature.

The participants were requested to perform nine different AUs or combinations of AUs (i.e., AUs numbers 4, 5, 12, 14, 25, 6 + 12, 12 + 25, 9 + 10, and 1 + 2; see Fig. 2) at two different speeds (fast and slow) and with three different intensities (just perceptible, normal, and high).[1] Those AUs integrated muscles from lower (e.g., 12, 25, 9 + 10) and upper (e.g., 4, 5, 1 + 2) parts of the face, but also several closed smiles (12, 14, 6 + 12, 12 + 25). The latter were chosen to study the sensitivity of the thermal imaging technique to discriminate close muscles contraction. The coders had to produce each of the requested AUs in five different intensity-speed combinations: just perceptible and fast, normal and fast, normal and slow, high and fast, and high and slow. For each given AU, these five combinations always followed the same order to ensure better intensity control. This sequence was performed twice by AU in randomized order. Thus, during the entire experiment, the coders performed 90 trials (9 AUs × 5 intensity-speed combinations × 2 repetitions). A computer screen placed in front of the participants indicated the requested combination of AU(s), speed, and intensity. Before

Thermal imaging, which has recently been used in domains such as public services (e.g., security, firefighters, and military) and medical diagnosis, may be a promising alternative for the investigation of AU production. A facial AU represents the contraction of a specific muscle or a combination of muscles, and research has demonstrated that such muscle contraction induces an increase in skin temperature (e.g., [5]). For this reason, thermal imaging analyses might be well suited to detect AU production. Moreover, in contrast to EMG recordings, this technique is noninvasive (no electrodes on the face) and can record the whole face activity at once. Furthermore, as demonstrated by studies on human breath tracking [6] and on vessel blood flow [7], thermal imaging techniques provide very high temporal resolution, given appropriate image sampling rates. This latter characteristic is of particular relevance where the dynamic of muscle contraction is concerned. In sum, thermography could be used as a noninvasive method, and without the visible-spectrum illumination limitations, to detect the location, the intensity, and the speed of thermal changes related to muscle contractions. However, to date, no study has attempted to systematically validate this technique to reveal facial heat modulations related to AUs performance.

This study constitutes a first attempt to investigate the suitability and the sensitivity of the thermal imaging technique to detect specific facial muscles' heat patterns. The objective of this work was to examine whether specific AUs are associated with a specific activated heat pattern. To test the specificity of the heat pattern produced, four trained FACS coders were recorded while they voluntarily produced different AUs. Moreover, to test whether thermal imaging is sensitive to both the intensity and the speed of muscle contraction, FACS coders were asked to activate the different AUs at different intensities and speeds. Analyses of thermal images followed the standard process [8] used in facial expression recognition: 1) location of the face in the images and faces normalization, 2) facial features extraction, and 3) expression classification based on facial features motion and/or appearance. Here, we used an analytical

1. The AU selection, speed, and intensities, as well as the timing, have been elaborated in collaboration with the FACS coders.

Fig. 3. Experimental protocol.
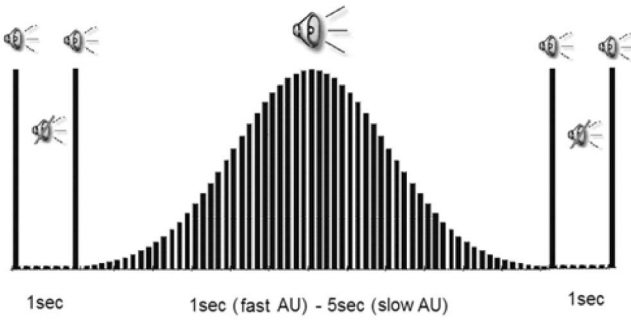

Fig. 4. Image registration method: (a) Intratrial alignment, (b) intertrials, and (c) normalization with the facial template.

each trial, the coders trained as long as necessary (with the help of a mirror) to maximize the accuracy of producing the requested muscle contraction during the recording session. Every trial (see Fig. 3) began with a beeping sound, indicating to participants that they had to stay calm with a relaxed face and be ready to produce the AU. One second later, another beeping sound signaled that they had to start producing the AU at the requested speed and intensity. To help the coder to produce the AU, a sound was presented that mimicked the requested contraction. During the whole expression production, the pitch of the sound represented the requested intensity of the contraction (the stronger the contraction, the higher the frequency), and the duration of the sound represented the requested speed of the contraction (long lasting sounds = 5 s for slow contractions; short lasting sounds = 1 s for fast contractions). A final series of two more beeps separated by 1 s were then presented to indicate a return to baseline, during which time the coders were requested to stay calm with a relaxed face and to wait for the next trial.

2.3 Thermal Image Acquisition

We used a thermal camera (FLIR ThermaCAM SC3000 Quantum Well Infrared Photodetector) that provides a high thermal sensitivity of less than $0.02°$ C for temperatures between -20 and $80°$ C. The camera was set for human skin emissivity ($e = 0.98$). Using this emissivity, temperature fluctuations brought on by illumination and other ambient changes will not affect the system. The temperature data were recorded with FLIR ThermaCAM Researcher Professional software. The image acquisition rate was fixed to 60 Hz (one image was recorded every 17 ms). Simultaneously with the thermal recording, we used a visible-spectrum camera to control for the accuracy of AU production.

Each trial was time locked to the beginning of the muscles contraction (determined by visual inspection of the motion), and it was defined as a 1,700 ms (100 thermal images) baseline period followed by 5 seconds of thermal signals.

2.4 Data Analysis

2.4.1 Image Preprocessing

All recorded images were reduced to the face area and rescaled to a particular size (210 pixels high $\times$ 150 pixels wide) by means of a bilinear interpolation algorithm to optimize speed calculation and disk storage, using Matlab (Matlab, Release 14, The Mathworks, Inc.). First, a rigid 2D translation and rotation procedure was applied to each trial
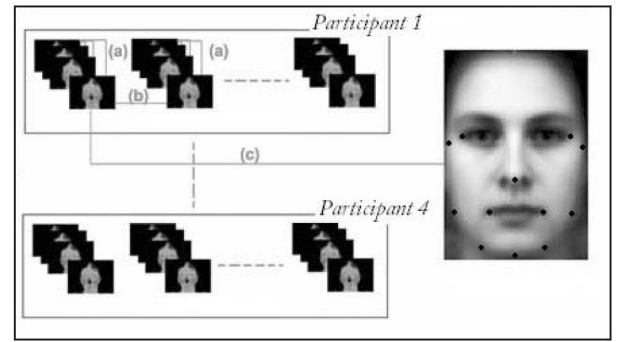
(baseline included) to align the images composing it (Fig. 4a). Images were realigned with each other using a Matlab's basic optimization routine (*fminsearch*) to find the transformation that restores the original image shape. The same procedure was then applied subsequently to align the trials within individual participants (Fig. 4b). For each participant, the algorithm determines an affine transformation that matched 12 control points placed on the individual's face (first image of participant's first trial) with those placed on the average face in the Karolinska Directed Emotional Faces [10] database (Fig. 4c). All facial images of a given participant were then spatially normalized to the average face according to this affine transformation (Fig. 4c).

Examples of the results obtained with this thermal image transformation technique are presented in Fig. 5. Finally, to eliminate the temperature changes not related to the AU production, we subtracted the mean facial temperatures during baseline from each image and for each trial.

2.4.2 Topography, Latency, and Amplitude of Thermal Changes

Spatial patterning approach. The rationale for assessing temperature change patterns in the spatial domain was to find the areas in the face where temperature values reliably covary, rather than focusing on a priori determined facial areas. Thus, we performed a spatial PCA on each trial, treating the pixels as variables and the temperature value at each time as observations. Without any a priori assumption on our part about the shape or number of pertinent areas in the data set, the PCA determines the complex relationships between the many temperature values measured for each pixel. These
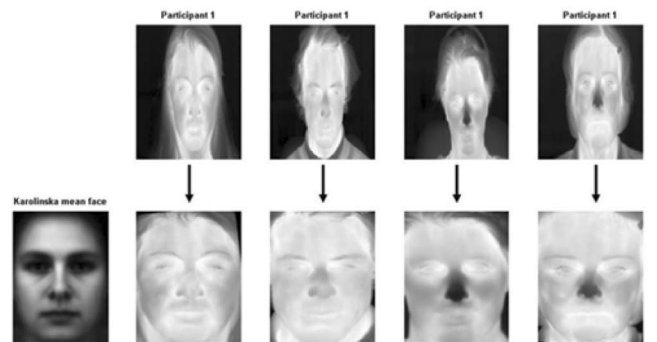

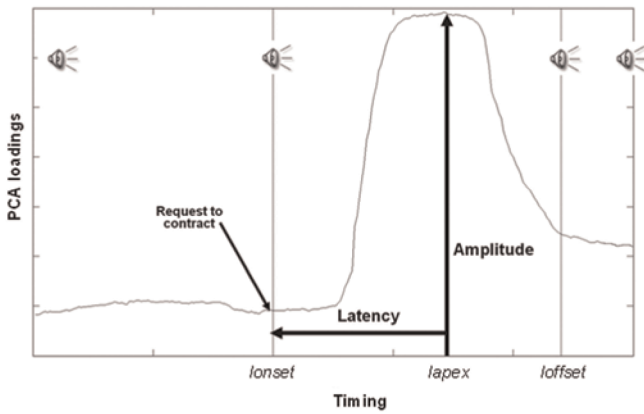Fig. 5. Example of the results of image normalization.

Fig. 6. Procedure to determine the apex of the trials (highest thermal amplitude after baseline).

relationships are summarized as latent-dependent variables corresponding to the spatial components. The result of such a PCA is a set of factor loadings that correspond to the contribution of the spatial components to the original temperature values (i.e., how much the spatial component accounts for the temperatures recorded at each time). The PCA also provides a set of factor scores that correspond to the contribution of each observation to each new spatial component. On average, the first two spatial components explained 86.28 percent of the variance in the data set ($SD = 5.91$) and, for each trial, the variance explained by the two first components was greater than 66.05 percent. The factor loadings of the first two components were retained and averaged, providing the areas in the face where temperature values reliably covary during AUs performance.

Apex determination. The temperature peak (apex) is the moment when the facial temperature is maximal. For each trial, a distribution is calculated averaging the two first component loadings convoluted with the trial thermal values (time points $\times$ pixels) in time. We calculated the apex when the maximum amplitude of this temporal distribution is reached after the baseline period. For each trial, the apex was calculated by performing an automatic baseline-to-peak analysis to find the maximum temperature variations. The latency of the apex gave us information about the speed of temperature changes related to AU production, and the amplitude of the signal at the apex informed us about the intensity of temperature change related to the AU(s) (Fig. 6).

Statistics on speed and intensities. To test whether thermal recordings are sensitive to different speeds of contraction, we used a permutation procedure performed on the apex latency values. Specifically, from the two groups of apex latency values (slow and fast), we calculated the means observed difference between the two samples (T(obs)). Then, differences in means samples are calculated after randomly dividing all the pooled latency values into two groups, for 1,000 times. From the distribution of these 1,000 differences, we could finally test the significance of the differences between the two distributions (slow and fast speeds). Differences are significant (at $p < 0.05$) when the observed difference (T(obs)) is not included in the 95 percent of the distribution of the randomly calculated differences. The same procedure was applied with the amplitude values

at the apex to test the significance of the differences between the levels of thermal intensities.

Topography of thermal changes (features extraction). We characterized the AUs produced by examining the localization of temperature changes related to muscle contractions of the face. We calculated the representative temperature maps (Map) for each trial by subtracting each temperature value of the image obtained at the apex (Iapex) from the spatially corresponding temperature image obtained at the end of the baseline (Ionset) ($\mathrm{Map} = \mathrm{Iapex} - \mathrm{Ionset}$). This procedure is intended to facilitate the interpretation of the results by avoiding the use of the loadings and the scores resulting from the PCA decomposition, parameters that should be interpreted as temperature variations around the mean temperature of the face. Finally, we calculated a mean representative topographic map for each requested AU by averaging (across time/subject) all individual maps obtained in each condition.

2.4.3 Classification of AUs

In order to quantify the robustness of the thermal imaging technique to discriminate the muscles contractions, we used a classifier based on the leave-one-out cross-validation method [11]. Several classes of classification algorithms exist, such as Neural Networks and Decision Trees or Linear Discriminant Analysis [12]; however, the K-nearest neighbor model (with $K = 1$) is the simplest algorithm to implement [13], [14]. This classifier could automatically split classes of objects (here the AUs) by using a predefined parameter; in our case, the Pearson's linear correlations between a given image (the image at the apex minus the image at the onset; see above) with the specific AU representative maps (obtained with the $n - 1$ participants). More precisely, for each trial, we recalculated the representative maps with the remaining three participants who had not performed the current trial. And we calculated the correlations between the image for this trial (Iapex-Ionset) with the representative map of all requested AUs. The highest of these correlations should then be an indicator for the AU that was requested on this trial. This procedure was then repeated for each trial.

3 RESULTS

3.1 Control of AU Production Accuracy

The visible-spectrum videos were coded by an external trained and certified FACS coder from another laboratory. He systematically coded the AUs of each facial expression produced. By comparing the coding done by the outside FACS coder with our experimental AUs, we found that the requested AUs were produced in 74.26 percent of cases (with a very low performance for AU9+10 at 6.06 percent). This comparison also led to the finding that 66.48 percent of the fast trials were judged as fast contractions and 73.04 percent of the slow trials were judged as slow contractions. Finally, 50.81 percent of the strong contractions, 58.62 percent of the normal contractions, and 57.14 percent of the just-perceptible contractions were judged as belonging to the intended category.
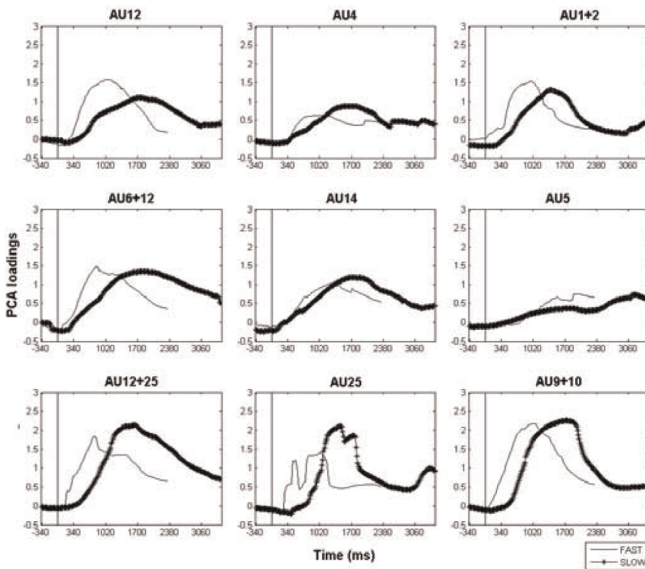
Fig. 7. Investigation of kinetics: For each AU, temperature mean distribution over all participants at normal intensity in the corresponding area at slow and fast speed.

3.2 Sensitivity to Speed

The apex of each facial muscle contraction was reliably detected. A representative sample of the temperature as it unfolded for the slow and fast speed conditions is represented in Fig. 7. On average, the apex (i.e., the temperature peak) was detected at 1,282 ms ($SD = 439$) after the end of the baseline for fast AUs, and at 1,976 ms ($SD = 550$) for slow AUs. Interestingly, when the AUs are considered together, the latency values obtained were significantly different between the slow and fast conditions (permutation test; $p < 0.001$). The same permutation procedure performed for each AU separately revealed that the two different speeds were significantly discriminated for all requested AU(s) (Table 1; at least $p < 0.02$).

3.3 Sensitivity to Intensity

A representative sample of the temporal unfolding of the temperature for the three intensities is presented in Fig. 8. The permutation procedure applied to the apex amplitude values for all the AUs together revealed a significant difference between the just perceptible and the high intensity conditions (permutation test; $p < 0.001$). More precisely, the approach led to a significant difference between just perceptible and high intensity contraction for all AUs except AUs 1 + 2 and 5. Just perceptible and normal intensities were hardly dissociated, and normal and high intensities were differentiated in rare cases (Table 1).

3.4 Specificity of the Facial Heat Patterns

Using a spatial patterning approach, we obtain, for each requested AU, a coherent representative thermal map (Fig. 9). Globally, in these map differences, we detected important temperature variations in the facial areas that corresponded to the location of the contracted muscle(s). For instance, we observed an increase in temperature in the zygomaticus region during a smile (i.e., when AU12 was requested), or a decrease in temperature in the frontalis region during the raising of the brows (AU1+2 requested).

TABLE 1
Results of the Permutation Tests Performed
on Apex Amplitudes and Latencies Derived

Speed	SLOW	FAST	P_{S-F}		
AU12	1900(433)	1245(328)	p<.001		
AU4	1961(600)	1299(429)	p<.01		
AU1+2	1811(322)	1274(450)	p<.01		
AU6+12	2053(521)	1182(316)	p<.001		
AU14	2074(382)	1332(400)	p<.001		
AU5	2389(880)	1690(486)	p<.02		
AU12+25	1869(348)	1111(374)	p<.001		
AU25	1925(539)	1324(570)	p<.01		
AU9+10	1790(600)	1052(305)	p<.001		
Intensity	JP	N	H	P_{JP-N}	P_{N-H}	P_{JP-H}
AU12	0.89(.49)	1.53(.86)	2.2(1.15)	n.s.	n.s.	p<.05
AU4	0.76(.45)	1.07(.38)	1.7(.76)	n.s.	p<.05	p<.02
AU1+2	1.26(.77)	1.51(1.04)	2.06(.99)	n.s.	n.s.	n.s.
AU6+12	1.2(.49)	1.61(.59)	2.36(1.05)	n.s.	p<.05	p<.02
AU14	0.99(.54)	1.37(.35)	1.8(.45)	n.s.	p<.02	p<.01
AU5	0.73(.34)	1.01(.63)	1.03(.8)	n.s.	n.s.	n.s.
AU12+25	1.46(.35)	2.3(.93)	2.92(.89)	n.s.	n.s.	p<.02
AU25	1.08(.68)	2.51(1.46)	2.99(1.22)	p<.05	n.s.	p<.01
AU9+10	2.03(.68)	2.38(.65)	2.9(.8)	n.s.	n.s.	p<.05

The permutation tests compared for each requested AU, the two speeds (SLOW and FAST in ms after the end of the baseline with the standard deviation in parentheses) and the three intensities (JP, N, and H in Kelvins with the standard deviation in parentheses). Slow-Fast = permutations between slow and fast contractions; JP-H = permutations between just perceptible and high intensity contractions; JP-N = permutations between just perceptible and normal intensity contractions; N-H = permutations between normal and high intensity contractions.

3.5 Classification of AUs

The classification rates of AUs prodution are displayed in Table 2. The highest classification rate found was 83.9 percent for the detection of the corrugator contraction (AU4, brow lower), the lowest being for the detection of the Duchenne smile (AU6+12, 15.2 percent). Whereas low classification rates were found for AU5 (30.3 percent), we obtained better results for AUs 9 + 10 (75 percent), 12 + 25 (72.4 percent), 14 (69.7 percent), 1 + 2 (67.7 percent), and 25 (60 percent). The
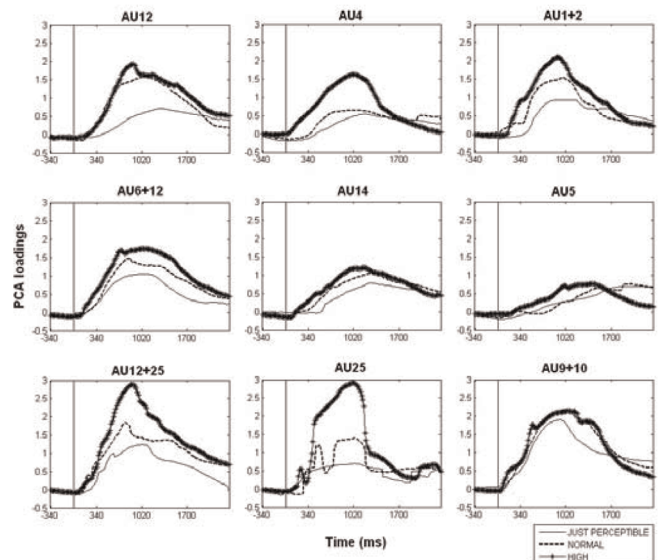

Fig. 8. Investigation of intensities: For each AU, mean temperature distribution over all participants at fast speed in the corresponding area at just perceptible, normal, and high intensities.
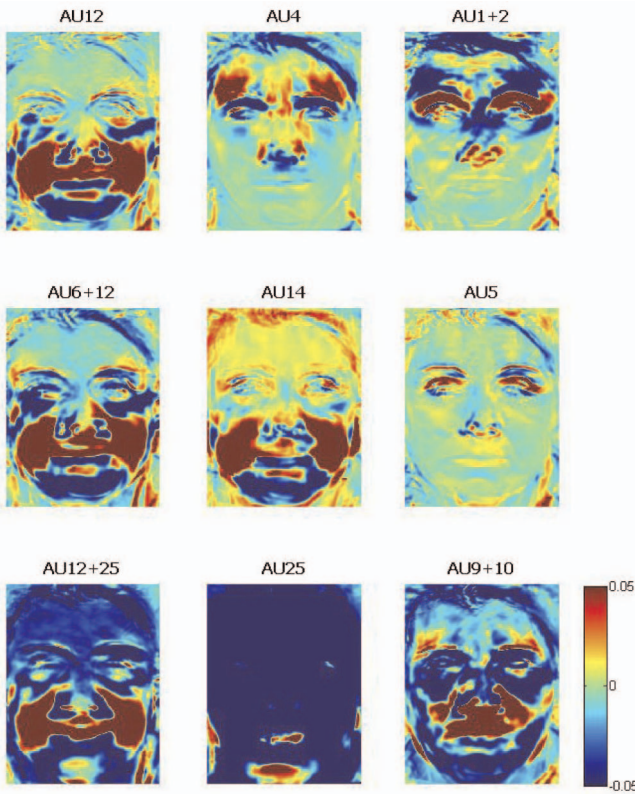
Fig. 9. Representative heat patterns changes (Iapex-Ibaseline) for each requested AU (averaged across FACS coders).

TABLE 2
Confusion Matrix

Requested AUs	Classified AUs								
	12	6+12	25	12+25	14	9+10	4	1+2	5
12	33.3	27.3	0.0	0.0	27.3	0.0	9.1	3.0	0.0
6+12	51.5	15.2	0.0	15.2	15.2	0.0	0.0	3.0	0.0
25	0.0	0.0	60.0	0.0	3.3	20.0	16.7	0.0	0.0
12+25	20.7	0.0	3.4	72.4	3.4	0.0	0.0	0.0	0.0
14	12.1	0.0	0.0	6.1	69.7	3.0	0.0	6.1	3.0
9+10	0.0	12.5	0.0	9.4	0.0	75.0	3.1	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	6.5	83.9	6.5	3.2
1+2	0.0	3.2	3.2	3.2	3.2	6.5	6.5	67.7	6.5
5	0.0	3.0	6.1	3.0	0.0	15.2	27.3	15.2	30.3

The results of correct classifications are expressed as percentages.

to define the specific facial heat patterns of AUs production, which was accomplished using PCA (Fig. 9). Our analyses strongly suggest that AU production induces a temperature *increase* in certain areas of the face and a *decrease* in others. For instance, the zygomaticus region seemed to increase in temperature when contracting, whereas the temperature of the frontalis region tended to decrease when activated. Our main interpretation is that cold skin structures such as eyebrows were sliding over the examined muscle region (e.g., frontalis) and thus decreasing in temperature, whereas such a skin structure dislocation does not exist for other muscle locations (e.g., zygomaticus). Another possible explanation for a decrease in temperature could be the crumpling of the skin in the muscle region during the contraction. Indeed, the temperature at the skin surface also depends on the underlying blood flows, which could be modified and/or more difficult to track during the crumpling of the skin. However, on the basis of this study, we are not able to dissociate the potential origins of the observed thermal variations; actually, these fluctuations of temperature might be related to the changes in blood flow occurring within the muscles during the contractions, or they might be more related to the movement of the skin. It could also be possible that both occur simultaneously, with a skin displacement more or less intense according to the position of the contracted muscle. Finally, a third objective of this study was to quantify the robustness of this thermal imaging technique to classify the AU productions. The results of the classifier showed that thermal imaging can differentiate the wrinkling of the nose (AU9+10, 75.0 percent) from the brow lowerer (AU4, 83.9 percent) and from the brow raiser (AU1+2, 67.7 percent). However, the approach led to weak classification rates in the discrimination of the upper lid raiser expression (AU5, 30.3 percent) from all other requested AUs. A plausible interpretation could be the small size of the muscle that underlies the production of AU5 (levator palpebrae superioris muscle), which is a muscle deeply hidden behind the eye. In addition, the limitation of this classifier arose in discriminating different closed smiles, such as a simple smile (AU12) and the Duchenne smile (AU6+12). Indeed, the classifier predicted AU12 in 33.3 percent of the cases but also AUs 6 + 12 or 14 both in 27.3 percent of the cases when the requested AU was AU12. The low classification rates for AUs 12 or 6 + 12 can be explained by the fact that the method included the production of both AU12 and AU6+12 in the same category. Another problem in discriminating AU6+12 is the quality of the stimuli: The FACS coders reported that this combination was difficult to perform. Indeed, from the external FACS coder report, AU6 was sometimes coded

close smiles seemed difficult to differentiate with this classifier. For instance, only 33.3 percent of AU12 were recognized; most of the time, they were confounded with close smiles such as AUs 6 + 12 (27.3 percent) and 14 (27.3 percent).

4 DISCUSSION

The main objective of this study was to test whether thermal imaging can be used as a tool to investigate specific facial heat patterns associated with the production of facial AUs. We tested whether the analyses based on the thermal images could specifically discriminate not only the contraction of a particular muscle related to the production of AUs or combinations of AUs, but also their intensities and speeds of contraction. We used a spatial patterning approach by using PCA on temperature values to extract specific facial heat maps associated with the requested AUs. The power of the approach for discriminating the speed and the intensity of contraction was tested with a permutation procedure and the capacity of thermal imaging to discriminate the different AUs was tested with a classification procedure.

Globally, this spatial patterning approach led to good results. One objective of the study was to assess whether we could provide reliable information on the strength and the speed of AU production. The permutation procedures performed on the amplitudes and latencies of the thermal responses measured at the apex led to a significant speed discrimination for all AUs, and to a significant discrimination between weak and strong contractions for all AUs except AUs 1 + 2 and 5 (Table 1). A second objective of this study was

when AU12 was requested, and AU6 was not always coded in AU6+12 trials. Even when the classifier did not allow for the precise disentanglement of the different kinds of smiles, we could correctly indicate whether the temperature related to muscle contraction changed in the lower or upper part of the face. For example, the production of AU12+25 was well recognized at 72.4 percent but was confounded with AU12 in 20.7 percent of the cases, which corresponds to contraction of muscles from the lower part of the face, whereas it was not confounded with AUs corresponding to contraction of muscles from the upper part of the face (0 percent for AU4, 0 percent for AU1+2, and 0 percent for AU5).

In sum, the thermal imaging technique seems to constitute a promising approach to detecting and evaluating changes in facial muscles contraction in relation with the production of AUs. Indeed, we demonstrated that it led to reasonable results in terms of detection of contraction locations, their kinetics, and their strength. Moreover, it avoids the problems of lighting encountered when using traditional cameras and the use of hampering electrodes on the face when using electromyography. An experimental limitation of the procedure mentioned here was that we used a system to immobilize the head of the participants. Further head tracking techniques would be necessary to allow the characterization of facial muscles contractions in more natural scenarios, for example in response to emotionally relevant situations. In addition, in this study, the participants were requested to voluntarily produce specific facial muscles contractions. In the future, it would be interesting to differentiate among the various spontaneous muscular activities relative to AU productions, as well as among temperature changes related to other physiological activity like vasoconstriction and vasodilatation. Indeed, thermal imaging was recently used to investigate stress and bodily temperature responses in emotional reactions in rats [15], monkeys [16], and humans ([17], [18]). Tanaka et al. [19] showed that the temperature of the nose increases with anger, joy, and sadness, but decreases with fear. These studies constitute the first valuable attempts to reveal emotional reactions by thermal imaging of the face. Future research should continue investigating facial temperature changes during laboratory induced emotions.

5   CONCLUSION

To our knowledge, this study is the first to use thermal imaging to discriminate specific facial temperature patterns related to muscle contractions corresponding to facial action units. Given the promising results, we suggest that thermography is an advantageous method for the investigation of AU discrimination. We used a spatial pattern approach to classify nine different AUs or combinations of AUs and to differentiate their speed and strength of contraction. Using this technique, we also found specific facial heat patterns associated to the different muscle contractions. Finally, thermography may prove to be a useful tool to unobtrusively analyze fine-grained elements of facial expressions. It remains to be determined whether it can be used to detect and characterize more spontaneous expressions in other situations.

REFERENCES

[1] P. Ekman, W.V. Friesen, and J.C. Hager, *Facial Action Coding System.* Consulting Psychologist Press, 1978.
[2] P. Ekman, W.V. Friesen, and P. Ellsworth, "Conceptual Ambiguities," *Emotion in the Human Face,* Cambridge Univ. Press, pp. 98-110, 1982.
[3] J.J.J. Lien, T. Kanade, C.C. Li, and J.F. Cohn, "Detection, Tracking and Classification of Action Units in Facial Expression," *IEEE J. Robotics and Autonomous Systems,* special issue: face expression in human-robot interaction systems, vol. 31, pp. 131-146, 2000.
[4] S. Delplanque, D. Grandjean, C. Chrea, L. Aymard, I. Cayeux, C. Margot, M.I. Velazco, D. Sander, and K.R. Scherer, "Sequential Unfolding of Novelty and Pleasantness Appraisals of Odors: Evidence from Facial Electromyography and Automatic Reactions," *Emotion,* vol. 9, no. 3, pp. 316-328, 2009.
[5] J. González-Alonso, B. Quistorff, P. Krustrup, J. Bangsbo, and B. Saltin, "Heat Production in Human Skeletal Muscle at the Onset of Intense Dynamic Exercise," *J. Physiology,* vol. 524, pp. 603-615, 2000.
[6] Z. Zhu, J. Fei, and I. Pavlidis, "Tracking Human Breath in Infrared Imaging," *Proc. Fifth IEEE Symp. Bioinformatics and Bioeng.,* pp. 227-231, 2005.
[7] N. Sun, I. Pavlidis, M. Garbey, and J. Fei, "Harvesting the Thermal Cardiac Pulse Signal," *Medical Image Computing and Computer Assisted Intervention Soc.,* vol. 4191, pp. 569-576, 2006.
[8] M. Pantic, "Facial Expression Recognition," *Encyclopedia of Biometrics,* Springer, pp. 400-406, 2009.
[9] T. Kanade, J.F. Cohn, and Y. Tian, "Comprehensive Database for Facial Expression Analysis," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition,* 2000, http://vasc.ri.cmu.edu/idb/html/face/facial_expression/index.html.
[10] D. Lundqvist and J.E. Litton, *The Averaged Karolinska Directed Emotional Faces (AKDEF).* Dept. of Clinical Neuroscience, Psychology Section, Karolinska Inst., 1998.
[11] A. Webb, *Statistical Pattern Recognition.* Wiley, pp. 254-255, 2002.
[12] M.Y. Kiang, "A Comparative Assessment of Classification Methods," *Decision Support Systems,* vol. 35, pp. 441-454, 2003.
[13] R.O. Duda, *Pattern Classification,* second ed. Wiley, 2001.
[14] A. Webb, *Statistical Pattern Recognition.* Wiley, 2002.
[15] D. Vianna and P. Carrive, "Changes in Cutaneous and Body Temperature during and after Conditioned Fear to Context in the Rat," *European J. Neuroscience,* vol. 21, pp. 2505-2512, 2005.
[16] K. Nakayama, S. Goto, K. Kuraoka, and K. Nakamura, "Decrease in Nasal Temperature of Rhesus Monkeys (Macaca Mulatta) in Negative Emotional State," *J. Physiology and Behavior,* vol. 84, pp. 783-790, 2005.
[17] I. Pavlidis, J. Levine, and P. Baukol, "Thermal Image Analysis for Anxiety Detection," *Proc. Int'l Conf. Image Processing,* vol. 2, pp. 315-318, 2001.
[18] C. Puri, L. Olson, I. Pavlidis, J. Levine, and J. Starren, "StressCam: Non-Contact Measurement of Users' Emotional States through Thermal Imaging," *Proc. Conf. Human Factors in Computing Systems,* pp. 1725-1728, 2005.
[19] H. Tanaka, H. Ide, and Y. Nagashima, "Emotion Analysis Using FST," *J. Robotics and Mechatronics,* vol. 10, pp. 266-271, 1998.

Sophie Jarlier received the master's degree in computer science imaging from the Universities of Lyon and Geneva, working on different projects (such as 3D sound, facial animation, hip joint modelization). She is currently working toward the PhD degree under the supervision of Patrik Vuilleumier and Didier Grandjean on emotions elicited by olfaction using a new technique: thermography. The aim of her work is to better understand the relationship between olfactory-elicited emotion processing, facial vasodilatation, and vaso-constrictions through thermal patterning of the face and cerebral activities. She is with the Swiss Center for Affective Science and the Laboratory for Neurology and Imaging of Cognition at the University of Geneva.

Didier Grandjean is an assistant professor and director of the Neuroscience of Emotion and Affective Dynamics laboratory in the Department of Psychology and Educational Sciences and the Swiss Center for Affective Sciences at the University of Geneva. In 2005, he finished his thesis about the dynamics of appraisal processes using electroencephalographic methods. He has published more than 30 peer-reviewed articles in international scientific journals in psychology and neuroscience about emotional processes related to emotional prosody perception and production, appraisal processes, the emergence of feelings, music and emotion, olfaction and emotion, and emotional facial expression perception and production.

Sylvain Delplanque received the postgraduate degree in neuropsychology and cognitive neuroscience and the PhD degree in cognitive sciences from the University Pierre et Marie Curie, Paris, France, in 2004. He studied human physiology, neurophysiology, and psychophysiology at the University of Lille I, France. He is now a senior researcher in the Swiss Centre for Affective Sciences, University of Geneva, Switzerland. His empirical work focuses on the cortical and autonomic concomitants to the appraisal processes underlying the genesis and the differentiation of odor-elicited emotions. The techniques he uses are mainly psychophysiological recordings, EEG-based indicators, and subjective reports through questionnaires and scales.

Karim N'Diaye received the PhD degree in cognitive neuroscience from the Laboratoire de Neurosciences Cognitive et Imagerie Cérébrale in Paris, France, in 2006, under the supervision of V. Pouthas and L. Garnero. He is currently a postdoctoral researcher in the COGIMAGE group at the Brain and Spine Institute and the Centre d'Economie de la Sorbonne in Paris, France, after a first postdoctoral fellowship at the Laboratory of Neurology and Imaging of Cognition in Geneva, Switzerland. He is mainly interested in metacognition and the neuroeconomy of the feeling of control, and in the evolutionary cognitive neuroscience of social processes and emotions working with various experimental psychology and brain imaging methods in humans.

Isabelle Cayeux is a doctor in food sciences responsible for sensory analysis research at Firmenich.

Maria Inés Velazco received the PhD degree in neuroscience from the University of Geneva for her work on the nicotinic Acetylcholine Receptor and its implication in the autoimmune disease *Myasthenia gravis*. She pursued a postdoctoral position at the Geneva University Hospital in the field of infectious diseases. She is a senior vice president of the Corporate R&D Division at Firmenich, SA, the largest privately owned company in the perfume and flavor business. She is responsible for scientific research in the fields of chemistry and human perception of flavors and fragrances, sensory analysis, psychophysics, emotion, analytical chemistry, and discovery of new active ingredients from cature. Her research interests include the physiology and psychology of taste and odor perception mechanisms in humans, cognitive neuroscience, polymers, and physical chemistry.

David Sander received the postgraduate degree in psychology and cognitive neuroscience and the PhD degree in cognitive sciences from the University Louis Lumière, Lyon, France. He studied mathematics and psychology at the University René Descartes, Paris, France. In 2002, he joined the Geneva Emotion Research Group in the Department of Psychology at the University of Geneva, Switzerland. He has also been the scientific coordinator of the Swiss Center for Affective Sciences since its creation in 2005. He is now an associate professor in the Department of Psychology at the University of Geneva, where he directs the Laboratory for the study of Emotion Elicitation and Expression (E3 Lab).

Patrik Vuilleumier is a neurologist who is using brain imaging techniques (such as functional resonance magnetic imaging, fMRI) to study the cerebral mechanisms of emotion and awareness. After first training in neuropsychology at Geneva and Lausanne, he pursued his research in cognitive neurosciences at the University of California, Davis (1997-1999), and then at University College London (1999-2002). He now works at the Laboratory for Neurology and Imaging of Cognition at the Geneva University Medical Center and Hospital. His current research investigates neural circuits in the human brain enabling emotion signals to influence perception and guide behavior, as well as the effect of brain lesions on emotion processes and consciousness.

Klaus R. Scherer holds the chair for emotion psychology at the University of Geneva and directs the Geneva Emotion Research Group. His major research interest is the further theoretical development and empirical test of his Component Process Model of Emotion (CPM), specifically the modeling of appraisal-driven processes of motor expression and physiological reaction patterns, as well as the reflection of these processes in subjective experience. Other major research foci are the study of the expression of affect in voice and speech, and applied emotion research.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.