Recognizing Upper Face Action Units for Facial Expression Analysis

Ying-li Tian [1]   Takeo Kanade[1] and Jeffrey F. Cohn[1,2]

[1] Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213
[2] Department of Psychology, University of Pittsburgh, Pittsburgh, PA 15260
Email: {yltian, tk}@cs.cmu.edu   jeffcohn@pitt.edu

Abstract

We develop an automatic system to analyze subtle changes in upper face expressions based on both permanent facial features (brows, eyes, mouth) and transient facial features (deepening of facial furrows) in a nearly frontal image sequence. Our system recognizes fine-grained changes in facial expression based on Facial Action Coding System (FACS) action units (AUs). Multi-state facial component models are proposed for tracking and modeling different facial features, including eyes, brows, cheeks, and furrows. Then we convert the results of tracking to detailed parametric descriptions of the facial features. These feature parameters are fed to a neural network which recognizes 7 upper face action units. A recognition rate of 95% is obtained for the test data that include both single action units and AU combinations.

1. Introduction

Recently facial expression analysis has attracted attention in the computer vision literature [3, 5, 6, 8, 10, 12, 15]. Most automatic expression analysis systems attempt to recognize a small set of prototypic expressions (i.e. joy, surprise, anger, sadness, fear, and disgust) [10, 15]. In everyday life, however, such prototypic expressions occur relatively infrequently. Instead, emotion is communicated by changes in one or two discrete facial features, such as tightening the lips in anger or obliquely lowering the lip corners in sadness [2]. Change in isolated features, especially in the area of the brows or eyelids, is typical of paralinguistic displays; for instance, raising the brows signals greeting. To capture the subtlety of human emotion and paralinguistic communication, automated recognition of fine-grained changes in facial expression is needed.

Ekman and Friesen [4] developed the Facial Action Coding System (FACS) for describing facial expressions by action units (AUs). AUs are anatomically related to contraction of specific facial muscles. They can occur either singly or in combinations. AU combinations may be *additive*, in which case combination does not change the appearance of the constituents, or *non-additive*, in which case the appearance of the constituents changes. Although the number of atomic action units is small, numbering only 44, more than 7,000 combinations of action units have been observed [11]. FACS provides the necessary detail with which to describe facial expression.

Automatic recognition of AUs is a difficult problem. AUs have no quantitative definitions and as noted can appear in complex combinations. Mase [10] and Essa [5] described patterns of optical flow that corresponded to several AUs but did not attempt to recognize them. Several researchers have tried to recognize AUs [1, 3, 8, 14]. The system of Lien *et al.* [8] used dense-flow, feature point tracking and edge extraction to recognize 3 upper face AUs (AU1+2, AU1+4, and AU4) and 6 lower face AUs. A separate hidden Markov model (HMM) was used for each AU or AU combination. However, it is intractable to model more than 7000 AU combinations separately. Bartlett *et al.* [1] recognized 6 individual upper face AUs (AU1, AU2, AU4, AU5, AU6, and AU7) but no combinations. Donato *et al.* [3] compared several techniques for recognizing 6 single upper face AUs and 6 lower face AUs. These techniques include optical flow, principal component analysis, independent component analysis, local feature analysis, and Gabor wavelet representation. The best performances were obtained using a Gabor wavelet representation and independent component analysis. All of these systems [1, 3, 8] used a manual step to align the input images with a standard face image using the center of the eyes and mouth.

We developed a feature-based AU recognition system. This system explicitly analyzes appearance changes in localized facial features. Since each AU is associated with a specific set of facial muscles, we believe that accurate geometrical modeling of facial features will lead to better recognition results. Furthermore, the knowledge of exact facial feature positions could benefit the area-based [15], holistic analysis [1], or optical flow based [8] classifiers. Figure 1 depicts the overview of the analysis system. First, the head orientation and face position are detected. Then, appear-

ance changes in the facial features are measured base on the multi-state facial component models. Motivated by FACS action units, these changes are represented as a collection of mid-level feature parameters. Finally, AUs are classified by feeding these parameters into two neural networks (one for the upper face, one for the lower face) because facial actions in the upper and the lower face are relatively independent [4]. The networks can recognize AUs whether the AUs occur singly or in combinations. We [14] recognized 11 AUs in the lower face and achieved a 96.7% average recognition rate.
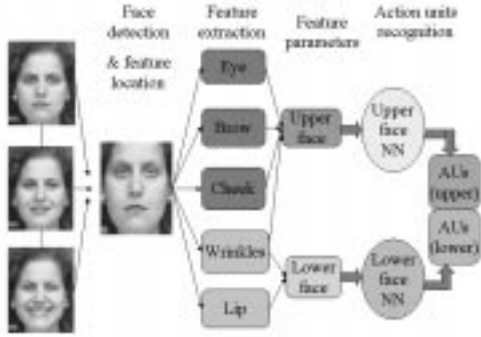
**Figure 1. Feature based action unit recognition system.**

In this paper, we focus on recognizing the upper face AUs and AU combinations. Fifteen parameters describe eye shape, motion, and state, and brow and cheek motion, and upper face furrows. A three-layer neural network is employed to classify the action units using these feature parameters as inputs. Seven basic action units in the upper face are identified whether they occurred singly or in combinations. Our system achieves a 95% average recognition rate. Difficult cases in which AUs occur either individually or in *additive* and *non-additive* combinations are handled. Moreover, the generalizability of our system is evaluated on independent databases recorded under different conditions and in different laboratories.

2. Multi-State Models For Facial Components

2.1. Dual-state Eye Model

Figure 2 shows the dual-state eye model. Using information from the iris of the eye, we distinguish two eye states, open and closed. When the eye is open, part of the iris normally will be visible. When closed, the iris is absent. For the different states, specific eye templates and different algorithms are used to obtain eye features.

For an open eye, we assume the outer contour of the eye is symmetrical about the perpendicular bisector to the line connecting two eye corners. The template, illustrated in Figure 2 (d), is composed of a circle with three param-
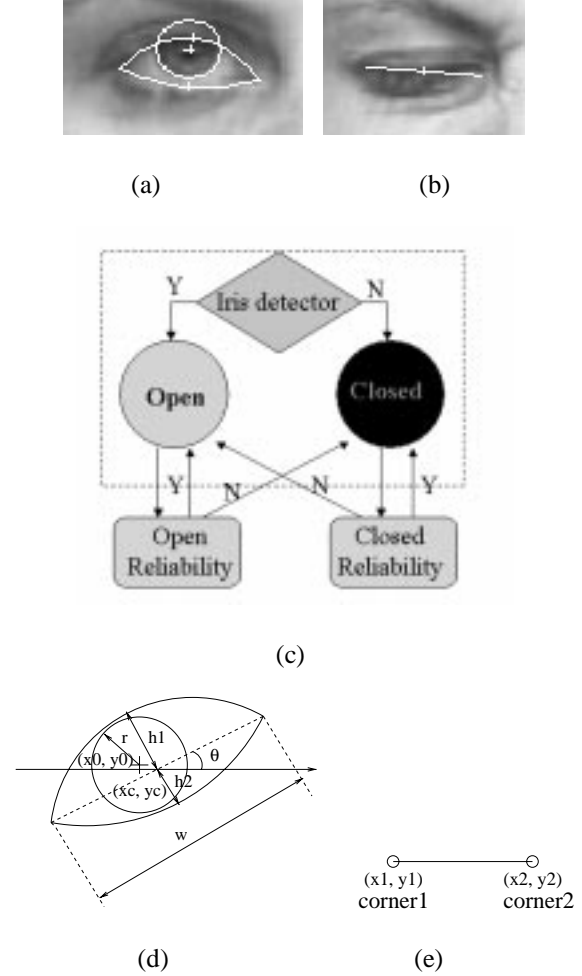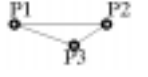

Figure 2. Dual-state eye model. (a) An open eye. (b) A closed eye. (c) The state transition diagram. (d) The open eye parameter model. (e) The closed eye parameter model.

eters (x_0, y_0, r) and two parabolic arcs with six parameters $(x_c, y_c, h_1, h_2, w, \theta)$. This is the same eye template as Yuille's except for two points located at the center of the whites [16]. For a closed eye, the template is reduced to 4 parameters for each of the eye corners (Figure 2 (e)).

2.2. Brow, Cheek and Furrow Models

The models for brow, cheek and furrow are described in Table 1. For the brow and cheek, we use separate single-state models; these are a triangular template with six parameters $P1(x1, y1)$, $P2(x2, y2)$, and $P3(x3, y3)$. Furrows have two states: present and absent.

Table 1. Brow, cheek, and furrow models of a front face

Component	State	Description/Feature
Brow	Present	
Cheek	Present	
Furrow	Present	
	Absent	

3. Upper Face Feature Extraction

Contraction of the facial muscles produces changes in both the direction and magnitude of the motion on the skin surface and in the appearance of permanent and transient facial features. Examples of permanent features are eyes, brow, and any furrows that have become permanent with age. Transient features include facial lines and furrows that are not present at rest. In analyzing a sequence of images, we assume that the first frame is a neutral expression. After initializing the templates of the permanent features in the first frame, both permanent and transient features are automatically extracted in the whole image sequence. Our method robustly tracks facial features even when there is out of plane head rotation. The tracking results can be found at http://www.cs.cmu.edu/~face.

3.1. Eye Features

Most eye trackers developed so far are only for open eyes and simply track the eye locations. However, for facial expression analysis, a more detailed description of the eye is needed. The dual-state eye model is used to detect an open eye or a closed/blinking eye.

The default eye state is open. Locating the open eye template in the first frame, the eye's inner corner is tracked accurately by feature point tracking. We found that the outer corners are hard to track and less stable than the inner corners, so we assume the outer corners are on the line that connects the inner corners. Then, the outer corners can be obtained by the eye width, which is calculated from the first frame.

An iris provides important information about the eye state. Intensity and edge information are used to detect the iris. A half-circle iris mask is used to obtain correct iris edges (Figure 3). If the iris is detected, the eye is open and the iris center is the iris mask center (x_0, y_0). In an image sequence, the eyelid contours are tracked for open eyes by feature point tracking. For a closed eye, tracking of the eyelid contours is omitted. A line connecting the inner and outer corners of the eye is used as the eye boundary. Some eye tracking results for different states are shown in

Figure 4. The detailed eye feature tracking techniques can be found in [13].
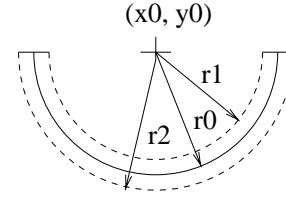
Figure 3. Half circle iris mask. (x_0, y_0) **is the iris center;** r_0 **is the iris radius;** r_1 **is the minimum radius of the mask;** r_2 **is the maximum radius of the mask**

3.2. Brow and Cheek Features

To model the position of brow and cheek, separate triangular templates with six parameters $(x1, y1)$, $(x2, y2)$, and $(x3, y3)$ are used. Brow and cheek are tracked by a modified version of the gradient tracking algorithm [9]. Figure 4 also includes some tracking results of brows for different expressions.

(a)   (b)

Figure 4. Tracking results in (a) narrowly-opened eye and in (b) widely-opened eye with blinking.

3.3. Transient Features

Facial motion produces transient features. Wrinkles and furrows appear perpendicular to the motion direction of the

activated muscle. These transient features provide crucial information for the recognition of action units. Contraction of the corrugator muscle, for instance, produces vertical furrows between the brows, which is coded in FACS as AU 4, while contraction of the medial portion of the frontalis muscle (AU 1) causes horizontal wrinkling in the center of the forehead.

Table 2. Basic upper face action units or AU combinations

AU 1	AU 2	AU 4
Inner portion of the brows is raised.	Outer portion of the brows is raised.	Brows lowered and drawn together
AU 5	AU 6	AU 7
Upper eyelids are raised.	Cheeks are raised.	Lower eyelids are raised.
AU 1+4	AU 4+5	AU 1+2
Medial portion of the brows is raised and pulled together.	Brows lowered and drawn together and upper eyelids are raised.	Inner and outer portions of the brows are raised.
AU 1+2+4	AU1+2+5+6+7	AU0(neutral)
Brows are pulled together and upward.	Brow, eyelids, and cheek are raised.	Eyes, brow, and cheek are relaxed.

Crows-feet wrinkles appearing to the side of the outer eye corners are useful features for recognizing upper face AUs. For example, the lower eyelid is raised for both AU6 and AU7, but the crows-feet wrinkles appear for AU6 only. Compared with the neutral frame, the wrinkle state is present if the wrinkles appear, deepen, or lengthen. Otherwise, it is absent.

After locating the outer corners of the eyes, edge detectors search for crows-feet wrinkles. We compare edge pixel numbers E of the current frame with the edge pixel numbers E_0 of the first frame in the wrinkle areas. If E/E_0 is larger than the threshold T, the crows-feet wrinkles are present. Otherwise, they are absent.

4. Upper Face AUs and Feature Representation

4.1. Upper Face Action Units

Action units can occur either singly or in combinations. The action unit combinations may be additive, in which case the combination does not change the appearance of the constituents (e.g., AU1+5), or nonadditive, in which case the appearance of the constituents does change (e.g., AU1+4). Table 2 shows the definitions of 7 individual upper face AUs and 5 non-additive combinations involving these action units. As an example of a non-additive effect, AU4 appears differently depending on whether it occurs alone or in combination with AU1, as in AU1+4. When AU4 occurs alone, the brows are drawn together and lowered. In AU1+4, the brows are drawn together but are raised by the action of AU 1. As another example, it is difficult to notice any difference between the static images of AU2 and AU1+2 because the action of AU2 pulls the inner brow up, which results in a very similar appearance to AU1+2. In contrast, the action of AU1 alone has little effect on the outer brow.

4.2. Upper Face Feature Representation

To recognize subtle changes of face expression, we represent the upper face features as 15 parameters. Of these, 12 parameters describe the motion and shape of eyes, brows, and cheeks. 2 parameters describe the state of crows feet wrinkles, and 1 parameter describes the distance between brows.

To define these parameters, we first define a coordinate system. Because we found that the inner corners of the eyes are the most stable features in the face and are insensitive to deformation by facial expressions, we define the x-axis as the line connecting two inner corners of eyes and the y-axis as perpendicular to the x-axis. Figure 5 shows the coordinate system and the parameter definitions. The definitions of upper face parameters are listed in Table 3. In order to remove the effects of the different size of face images in different image sequences, all the parameters (except the furrow parameters) are normalized by dividing by the distances between each feature and the line connecting two inner corners of eyes in the neutral frame.

5. Image Databases

Two databases were used to evaluate our system, the CMU-Pittsburgh AU-Coded face expression image database (CMU-Pittsburgh database) [7] and Ekman and Hager's facial action exemplars (Ekman&Hager database). The later was used by Donato and Bartlett [1, 3]. We use the Ekman&Hager database to train the network. During testing both databases were used. Moreover, in part of our evaluation, we trained and tested on completely disjoint databases

Table 3. Upper face feature representation for AU recognition

Permanent features (Left and right)		
Inner brow motion (r_{binner})	Outer brow motion (r_{bouter})	Eye height ($r_{eheight}$)
r_{binner} $=\frac{bi-bi_0}{bi_0}$. If $r_{binner}>0$, Inner brow move up.	r_{bouter} $=\frac{bo-bo_0}{bo_0}$. If $r_{bouter}>0$, Outer brow move up.	$r_{eheight}$ $=\frac{(h1+h2)-(h1_0+h2_0)}{(h1_0+h2_0)}$. If $r_{eheight}>0$, Eye height increases.
Eye top lid motion (r_{top})	Eye bottom lid motion (r_{btm})	Cheek motion (r_{cheek})
r_{top} $=\frac{h1-h1_0}{h1_0}$. If $r_{top}>0$, Eye top lid move up.	r_{btm} $=-\frac{h2-h2_0}{h2_0}$. If $r_{btm}>0$, Eye bottom lid move up.	r_{cheek} $=-\frac{c-c_0}{c_0}$. If $r_{cheek}>0$, Cheek move up.
Other features		
Distance of brows (D_{brow})	Left crows-feet wrinkles (W_{left})	Right crows-feet wrinkles (W_{right})
D_{brow} $=\frac{D-D_0}{D_0}$.	If $W_{left}=1$, Left crows-feet wrinkle present.	If $W_{right}=1$, Right crows-feet wrinkle present.
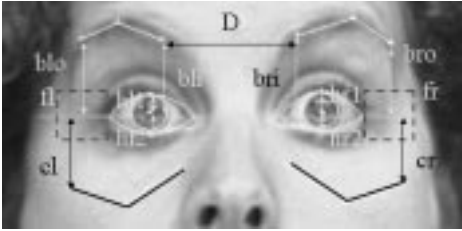
Figure 5. Upper face features. $hl(hl1+hl2)$ and $hr(hr1+hr2)$ are the height of left eye and right eye; D is the distance between brows; cl and cr are the motion of left cheek and right cheek. bli and bri are the motion of the inner part of left brow and right brow. blo and bro are the motion of the outer part of left brow and right brow. fl and fr are the left and right crows-feet wrinkle areas.

that were collected by different research teams under different recording conditions and coded (ground-truth) by separate teams of FACS coders. This is a more rigorous test of generalizability than the more customary method of dividing a single database into test and training sets.

Ekman&Hager database: This image database was obtained from 24 Caucasian subjects, consisting of 12 males and 12 females. Each image sequence consists of 6-8 frames, beginning with a neutral or with very low magnitude facial actions and ending with a high magnitude facial actions. For each sequence, action units were coded by a certified FACS coder.

CMU-Pittsburgh database: This database currently consists of facial behavior recorded in 210 adults between the ages of 18 and 50 years. They were 69% female, 31% male, 81% Euro-American, 13% Afro-American, and 6% other groups. Subjects sat directly in front of the camera and performed a series of facial expressions that included single AUs and AU combinations. To date, 1917 image sequences from 182 subjects have been FACS coded for either target AUs or the entire sequence. Approximately fifteen percent of the 1917 sequences were re-coded by a second certified FACS coder to validate the accuracy of the coding. Each expression sequence began from a neutral face.

6. AU Recognition

We used three-layer neural networks with one hidden layer to recognize AUs. The inputs of the neural networks are the 15 parameters shown in Table 3. The outputs are the upper face AUs. Each output unit gives an estimate of the probability of the input image consisting of the associated AUs. In this section, we concluded 3 experiments. In the first, we compare with other results using the same database. In the second, we study the more difficult case in which AUs occur either individually or in combinations. Furthermore, we investigate the generalizability of our system on independent databases recorded under different conditions and in different laboratories. The optimal number of hidden units to achieve the best average recognition rate was also studied.

6.1. Experiment 1: Comparison With Other Approaches

For comparison with the AU recognition results of Bartlett [1], we trained and tested our system on the same database (the Ekman&Hager database). In this experiment, 99 image sequences containing only individual AUs in upper face were used. Two test sets were selected as in Table 4. In *Familiar faces testset*, some subjects appear in both training and testing sets although they had no common sequences. In *Novel faces testset*, to study the robustness of the system to novel faces, we ensured that the subjects do not appear in both training and test sets. Training and testing were performed on the initial and final two frames in each im-

age sequence. For some of the image sequences with large lighting changes, lighting normalizations were performed.

Table 4. Data distribution of each data set for upper face AU recognition (Ekman&Hager Database).

Datasets	AU0	AU1	AU2	AU4	AU5	AU6	AU7	Total
Trainset	47	14	12	16	22	12	8	141
Familiar face Testset	52	14	12	20	24	14	20	156
Novel face Testset	49	10	10	22	28	4	22	143

In the system of Bartlett and Donato [1, 3], 80 image sequences containing only individual AUs in upper face were used. They manually aligned the faces using three coordinates, rotated the eyes to horizontal, scaled the image, and cropped it to a fix size. Their system was trained and tested using leave-one-out cross-validation and the mean classification accuracy was calculated across all of the test cases.

The comparison is shown in Table 5. For 7 single AUs in the upper face, our system achieves an average recognition rate of 92.3% for familiar faces (new images of the faces used for training) on *Familiar faces testset* and 92.9% when we test the system for novel faces on *Novel faces testset* with zero false alarms. From experiments, we found that 6 hidden units gave the best performance. The performance of Bartlett's [1] feature-based classifier on familiar faces, the rate was 85.3%; on novel faces was 57%. Donato *et al.* did not report the details whether the test images were familiar or novel faces [3]. Our system achieved a 95% average recognition rate as the best performance for recognizing 7 single AUs and more than 10 AU combinations in the upper face. Bartlett *et al.* [1] increased the recognition accuracy to 90.9% correct by combining holistic spatial analysis and optical flow with local features in a hybrid system for 6 single upper face AUs. The best performance of Donato's *et al.* [3] system was obtained using a Gabor wavelet representation and independent component analysis (ICA) and achieved a 95% average recognition rate for 6 single upper face AUs and 6 lower face AUs.

From the comparison, we see that our recognition performance from facial feature measurements is comparable to holistic analysis and Gabor wavelet representation for AU recognition.

Table 5. Comparison with Donato and Batlett's system for AU recognition using Ekman&Hager Database.

Methods	Systems	Recognition rate
Feature-based classifier	Our system	92.3% familiar faces
		92.9% novel faces
	Bartlett's system	85.3% familiar faces
		57% novel faces
Best performance	Our system (Feature-based)	95%
	Bartlett's system (Hybrid)	90.0%
	Donato's system (ICA or Gabor wavelet)	95%

## 6.2. Experiment 2: Results for Combinations of AUs

Because FACS consists of potential combinations numbering in the 1000s, an AU recognition system with ability to recognize both single AUs and AU combinations becomes necessary. All previous AU recognition systems [1, 3, 8] were developed to recognize single AUs only. Although, there were some AU combinations, such as AU1+2, AU1+4 in Lien's system and AU9+25, AU10+25 in Donato and Bartlett's system, each of these combinations was considered as a separate AU. Our system attempts to recognize AUs whether they occur singly or in combinations by a neural network. More than one output units of the network could fire for AU combinations. Moreover, we investigate the effects of the non-additive AU combinations.

A total of 236 image sequences from Ekman&Hager database, 99 image sequences containing individual AUs, and 137 image sequences containing AU combinations were used from 23 subjects for the upper face AU recognition.

AU Combination Recognition When Modeling Only 7 Individual AUs: We restrict the output to be 7 individual AUs (Figure 6 (a)). For the AU combinations, more than one output units could fire and the same value is given for each corresponding individual AU in training data set. For example, for AU1+2+4, the outputs are AU1=1.0, AU2=1.0, and AU4=1.0. From experiments, we found that we need to increase the number of hidden units from 6 to 12 to obtain the best performance.

Table 6 shows the recognition results. A 95% average recognition rate is achieved, with a false alarm rate of 6.4%. The false alarm rate comes from the AU combinations. For example, if we obtained the recognition results with
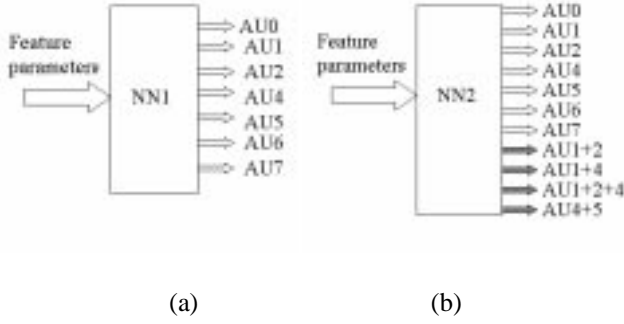
Figure 6. Neural networks of the upper face AU and combination recognition. (a) Model the 7 single AUs only. (b) Separately model the non-additive AU combinations

AU1=0.85 and AU2=0.89 for a human labeled AU2, it was treated as AU1+AU2. This means AU2 is recognized but with AU1 as a false alarm.

Table 6. Upper face AU recognition with AU combinations when modeling 7 single AUs only. The rows correspond to NN outputs, and columns correspond to human labels.

	AU0	AU1	AU2	AU4	AU5	AU6	AU7
AU0	22	0	0	0	0	0	0
AU1	0	18	0	0	0	0	0
AU2	0	0	12	0	0	0	0
AU4	0	0	0	10	0	0	0
AU5	3	0	0	0	7	0	0
AU6	2	0	0	0	0	12	0
AU7	0	0	0	0	0	0	8
Average Recognition rate: 95%							

AU Combination Recognition When Modeling Nonadditive Combinations: In order to study the effects of the non-additive combinations, we separately model the non-additive AU combinations in the network. The 11 outputs consist of 7 individual upper face AUs and 4 non-additive AU combinations (Figure 6 (b)).

The recognition results are shown in Table 7. An average recognition rate of 93.7% is achieved, with a slightly lower false alarm rate of 4.5%. In this case, separately model the non-additive combinations does not improve recognition rate.

6.3. Experiment 3: Generalizability to Other Datasets

To test the generalizability of our system, the independent databases recorded under different conditions and in different laboratories were used. The network was trained on the

Table 7. Upper face AU recognition with AU combinations when separately modeling non-additive AU combinations.

	AU0	AU1	AU2	AU4	AU5	AU6	AU7
AU0	25	0	0	0	0	0	0
AU1	2	20	0	0	0	0	0
AU2	2	0	14	0	0	0	0
AU4	0	0	0	14	0	0	0
AU5	2	0	0	0	8	0	0
AU6	1	0	0	0	0	13	0
AU7	0	0	0	0	0	0	10
Average Recognition rate: 93.7%							

Ekman&Hager database and 72 image sequences from the CMU-Pittsburgh database were used to test the generalizability. The recognition results are shown in Table 8 and a 93.2% recognition rate is achieved. From the results we see that our system is robust and achieves high recognition rate for the new database.

Table 8. Upper face recognition results for test on the Pittsburgh-CMU database when the network is trained on the Ekman&Hager database.

	AU0	AU1	AU2	AU4	AU5	AU6	AU7
AU0	72	0	0	0	0	0	0
AU1	1	51	2	0	0	0	0
AU2	5	2	21	0	0	0	0
AU4	4	0	0	44	0	0	0
AU5	3	0	0	0	31	0	0
AU6	0	0	0	0	0	2	0
AU7	0	0	0	0	0	0	12
Average Recognition rate: 93.2%							

6.4. Major Causes of the Misidentifications

Many of the misidentifications come from AU1/AU2 and AU6/AU7. The confusions between AU1 and AU2 are caused by the strong correlation of AU1 and AU2. The action of AU2 (the outer portion of the brow is raised) pulls the inner brow up (see Table 2). For AU6 and AU7, both of them the lower eyelids raise up, are also confused by human AU coders.

The mistakes of some AUs classified as AU0 are caused by two reasons: (1) Non-additive AU combinations modified the appearances of individual AUs. (2) AUs with lower intensity. For example, in Table 8, the missing of AU1, AU2 and AU4 is cause by the combination AU1+2+4 which modified the appearances of AU1+2 and AU4. If AU4 is in higher intensity, AU1 and AU2 are easy to be ignored.

7. Conclusion

In this paper, we developed a multi-state feature-based facial expression recognition system to recognize both individual AUs and AU combinations. All the facial features were represented in a group of feature parameters. The network was able to learn the correlations between facial feature parameter patterns and specific action units. It has high sensitivity and specificity for subtle differences in facial expressions.

Our system was tested in image sequences for large number of subjects, which included people of African and Asian in addition to European, thus providing a sufficient test of how well the initial training analyses generalized to new image sequences and new databases. From the experimental results, we have the following observations:

1. The recognition performance from facial feature measurements is comparable to holistic analysis and Gabor wavelet representation for AU recognition.

2. 5 to 7 hidden units are sufficient to code 7 individual the upper face AUs. 10 to 16 hidden units are needed when AUs may occur either singly or in complex combinations.

3. For the upper face AU recognition, separately modeling nonadditive AU combinations affords no increase in the recognition accuracy.

4. After using sufficient data to train the NN, our system is robustness to recognize AUs and AU combinations for new faces and new databases.

Unlike a previous method [8] which built a separate model for each AU and AU combination, we developed a single model that recognized AUs whether they occur singly or in combinations. This is an important capability since the number of possible AU combinations is too large (over 7000) for each combination to be modeled separately. An average recognition rate of 95% was achieved for 7 upper face AUs and more than 10 AU combinations. Our system was robust across independent databases recorded under different conditions and in different laboratories.

Acknowledgements

References

[1] M. Bartlett, J. Hager, P.Ekman, and T. Sejnowski. Measuring facial expressions by computer image analysis. *Psychophysiology*, 36:253–264, 1999.

[2] J. M. Carroll and J. Russell. Facial expression in hollywood's portrayal of emotion. *Journal of Personality and Social Psychology.*, 72:164–176, 1997.

[3] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski. Classifying facial actions. *International Journal of Pattern Analysis and Machine Intelligence*, 21(10):974–989, October 1999.

[4] P. Ekman and W. V. Friesen. *The Facial Action Coding System: A Technique For The Measurement of Facial Movement.* Consulting Psychologists Press Inc., San Francisco, CA, 1978.

[5] I. A. Essa and A. P. Pentland. Coding, analysis, interpretation, and recognition of facial expressions. *IEEE Transc. On Pattern Analysis and Machine Intelligence*, 19(7):757–763, JULY 1997.

[6] K. Fukui and O. Yamaguchi. Facial feature point extraction method based on combination of shape extraction and pattern matching. *Systems and Computers in Japan*, 29(6):49–58, 1998.

[7] T. Kanade, J. Cohn, and Y. Tian. Comprehensive database for facial expression analysis. In *Proceedings of International Conference on Face and Gesture Recognition*, March, 2000.

[8] J.-J. J. Lien, T. Kanade, J. F. Chon, and C. C. Li. Detection, tracking, and classification of action units in facial expression. *Journal of Robotics and Autonomous System*, in press.

[9] B. Lucas and T. Kanade. An interative image registration technique with an application in stereo vision. In *The 7th International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.

[10] K. Mase. Recognition of facial expression from optical flow. *IEICE Transactions*, E. 74(10):3474–3483, October 1991.

[11] K. Scherer and P. Ekman. *Handbook of methods in nonverbal behavior research*. Cambridge University Press, Cambridge, UK, 1982.

[12] D. Terzopoulos and K. Waters. Analysis of facial images using physical and anatomical models. In *IEEE International Conference on Computer Vision*, pages 727–732, 1990.

[13] Y. Tian, T. Kanade, and J. Cohn. Dual-state parametric eye tracking. In *Proceedings of International Conference on Face and Gesture Recognition*, March, 2000.

[14] Y. Tian, T. Kanade, and J. Cohn. Recognizing lower face actions for facial expression analysis. In *Proceedings of International Conference on Face and Gesture Recognition*, March, 2000.

[15] Y. Yacoob and L. S. Davis. Recognizing human facial expression from long image sequences using optical flow. *IEEE Transactions On Pattern Analysis and machine Intelligence*, 18(6):636–642, June 1996.

[16] A. Yuille, P. Haallinan, and D. S. Cohen. Feature extraction from faces using deformable templates. *International Journal of Computer Vision,*, 8(2):99–111, 1992.