

아이겐포인트를 이용한 표정 인식

홍성희^o 변혜란

연세대학교 컴퓨터과학과

{lovehite^o, hbyun}@cs.yonsei.ac.kr

Facial expression recognition using eigen-points

Sunghee Hong^o Hyeran Byun

Dept. of Computer Science, Yonsei University

요 약

본 논문에서는 사람의 얼굴표정을 구분하기 위해서 무표정 영상으로부터 18개의 특징점을 찾고, 그 특징점 간의 거리를 템플릿으로 이용하는 방법을 연구하였다. 얼굴표정인식을 위해 정의된 기본 템플릿과 입력 표정 영상에서의 특징점 간의 상대적인 거리의 차이와 특징점의 좌표변위 차이를 이용하여 표정을 구분하도록 하였다. 각 테스트 표정영상의 특징점은 주요 얼굴요소로부터 아이겐포인트(eigen-point)를 자동으로 추출하였다. 표정 인식은 신경망 학습을 통해서 기쁨, 경멸, 놀람, 공포, 슬픔 등 5가지로 나누어 실험하였고, 신경망의 인식 결과와 사람의 인식 결과를 통해서 비교한 결과, 72%의 인식성능을 보여주었다.

1. 서론

얼굴은 사람의 감정을 표현할 수 있는 가장 유용하고 자연스러운 수단이며, 즉흥적인 감정의 반응을 내포하고 있으므로 이를 이용하면 얼굴표정을 통해 인간의 감정을 자동으로 인식할 수 있다.

초기의 얼굴표정의 연구는 주로 연구자의 주관적 판단에 의해서 이루어졌다. 그러나 1970년대 말에 이르러 체계적으로 얼굴근육의 움직임을 부호화하여 각 정서 상태를 표현하는 연구가 시작되었다. Ekman과 Friesen[1]은 Facial Action Coding System (FACS)를 개발하여 얼굴 근육의 움직임을 얼굴 표정으로 구분하였다. 90년대 이후, 얼굴의 특징 벡터를 이용한 얼굴 표정에 관한 연구들을 살펴보면, 다음과 같다.

(1)정지영상에서의 표정인식 : Zhang[2]은 34개의 수동으로 특징을 추출하여 무표정영상과 표정영상간의 좌표변이를 비교 하여 6개의 표정을 인식하였고, Huang and Huang[3]은 90개의 point distribution model(PDM)을 수동으로 생성하여 표정을 인식하였다.

(2)동영상에서의 표정인식 : Black 과 Yacoob[4]는 얼굴의 특징이 되는 부분 (Rigid face) 은 planar model을 이용하고, 나머지 얼굴 영역 (Nonrigid face) 은 affine-plus-curvature 를 이용하여 표정을 인식하였고, Wang[5]은 19개의 수동 point를 이용하여 Labeled graph model 을 그리고 유사도를 이용하여 표정을 분류하였다. Cohn [6]은 Optical flow를 이용하여 특징점을 추적하며 High gradient edge를 추출하여 유사도를 비교하여 표정을 인식하였고, Tian[7]은 Geometric and motion parameters 사용하고 Canny edge 를 이용하여 주름 찾고, 신경망으로 학습시켰다.

이들 논문에서는 기쁨, 경멸, 놀람, 슬픔 등 네 가지 혹은 공포, 혐오를 포함한 여섯 가지의 표정에 대해 좋은 인식 성능을 보여주고 있지만 얼굴의 특징점들의 위치를 수동으로 추출하므로 불편할 뿐만 아니라 현실적으로 적용하기 어려움이 있다. 본 논문에서

는 아이겐포인트를 이용하여 얼굴의 특징점을 자동으로 찾아주고, 이를 기반으로 무표정영상과 표정영상의 특징점간의 상대적인 거리와 좌표변위차를 이용하여 얼굴표정을 구분하는 방법을 제안한다.

2. 전처리

본 논문에서는 무배경인 그레이 영상을 사용함으로써 색상에 대한 보정은 하지 않고, 밝기 변화를 보정하기 위하여 히스토그램 평활화를 사용하였다. 히스토그램 평활화는 명암값 분포를 재분배함으로써 영상의 명암의 분포를 일정한 분포로 만들어 주게 된다.

3. 제안하는 방법

3.1.아이겐포인트 (eigen-points)

얼굴의 특징점에 관한 상세한 정보를 얻기 위해서 무표정영상을 템플릿으로 사용한다. 본 논문에서 사용하는 템플릿은 [그림1]의 (2)에서 볼 수 있는 것처럼 40개의 특징점을 수동으로 표시하며, 각 특징점이 얼굴의 특정부위의 윤곽을 나타내도록 하였다. 본 논문에서는 왼쪽 눈의 안쪽위치를 고정시키고 이를 기준으로 하여 얼굴 전체에서 특징점을 찾았다. 140 * 160 픽셀크기의 얼굴영상으로부터 66 * 60 픽셀크기로 얼굴 영역만 추출하여 40개의 특징점을 수동으로 찾았다. 40개의 특징점을 사용한 이유는 적절한 특징점 개수를 찾아주어야 얼굴의 윤곽을 나타낼 수 있는 아이겐포인트[8][9]를 추출할 수 있기 때문이다.

아이겐포인트는 PCA 의 기본 원리를 토대로 새로운 영상에서 특징점을 추출할 수 있다. 우선 n개의 얼굴영상을 I_i 라고 하자.

\bar{I} 는 영상의 평균이라면 훈련영상들의 공분산 행렬은 아래와 같다.

$$S = AA^T \quad (A = [I_1 - \bar{I}, I_2 - \bar{I}, \dots, I_n - \bar{I}])$$

$A^T A$ 의 고유벡터를 V 이라 한다면, 행렬(U)는 다음과 같이



(2) 40개의 수동 특징점 위치

(1) 영상 (3) Eigen point 를 이용한 특징점 위치
[그림 1] 얼굴 형태

나타낸다.

$$U = [U_1, U_2, \dots, U_n] = AV \quad (1)$$

식(1)은 식(2)와 같이 다시 표현될 수 있다. P 는 영상의 각 특징점의 평균값이며, Ep_i 는 특징점의 위치를 나타내는 벡터값(아이겐포인트)이다.

$$[Ep_1, Ep_2, \dots, Ep_n] = [P_1 - P, P_2 - P, \dots, P_n - P] V \quad (2)$$

새로운 영상에서 특징점의 벡터값(P_{new})을 얻기 위해서 우선, 고유벡터의 선형결합으로써 새로운 영상(I_{new})을 표현한다.(식3)

$$I_{new} = \sum (I_{new} - I) U^T * U_i + \bar{I} \quad (3)$$

추출된 아이겐포인트(Ep)와 전체영상(I_{new})은 같은 선형결합을 사용함으로써 표현한다. 새로운 영상의 특징점(P_{new})은 식 (3)과 같은 선형결합을 통해서 식 (4)과 같은 방법으로 구할 수 있다.

$$P_{new} = \sum (I_{new} - I) U^T * Ep + P \quad (4)$$

식(4)를 통해서 새로운 영상의 특징점(P_{new})를 추출할 수 있다.

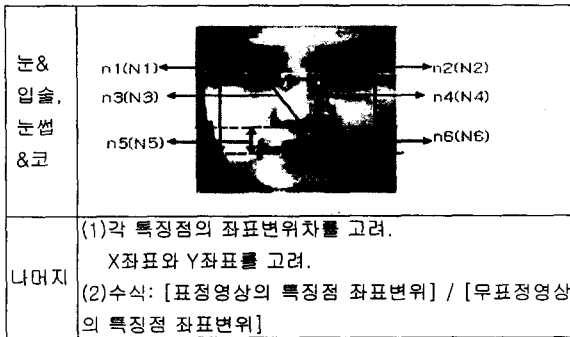
실험을 위해서 104개(연세대 인지과학연구소 데이터베이스 12개 + 연세대 얼굴 데이터베이스 (<http://cvpr.yonsei.ac.kr/facetracking>) 92개)의 각각 다른 사람, 다른 표정의 얼굴 영상을 가지고 40개의 특징점을 수동으로 찍어서 훈련을 하였다. 40개의 특징점은 순서는 가지고 있으며, 신경망학습을 위해 표정변화에 가장 민감한 18개의 특징점을 추출하여 실험하였다. ([그림 1]의 (3)은 18개의 특징점의 위치)

3.2 표정 학습

18개의 특징점을 가지고 무표정영상에서 특징점간의 가변적인 거리차와 표정영상에서 특징점의 좌표변위의차를 이용하여 표정을 구분하게 된다.

요소	학습방법
변	(1)눈의 크기 차이 이용 (2)무표정영상에서 눈의 안쪽과 눈의 가장 높은 부분의

변	<p>거리차: h_1, 눈의 안쪽과 눈의 가장 낮은 부분의 거리차: h_2, 눈의 안쪽과 눈의 가장 낮은 부분의 거리차: h</p> <p>(3)표정영상에서 각 거리의 차는 무표정영상에서의 거리차와 같다. 이를 H_1, H_2, H로 표현한다.</p> <p>(4)수식: $(H_1 - h_1) / h_1, (H_2 - h_2) / h_2, (H - h) / h$</p>
눈썹 & 눈	<p>(1)눈썹의 움직임과 눈썹간의 간격차이 이용</p> <p>(2)무표정영상에서 왼쪽눈썹가장자리와 왼쪽눈사이의 간격차: b_1, 왼쪽눈썹안쪽과 왼쪽눈사이의 간격차: b_2, 오른쪽눈썹가장자리와 오른쪽눈사이의 간격차: b_3, 오른쪽눈썹안쪽과 오른쪽눈사이의 간격차 : b_4, 양쪽눈썹간의 간격차: b</p> <p>(3)표정영상에서 각 거리의 차는 무표정영상에서의 거리의 차와 같다. 이를 B_1, B_2, B_3, B_4, B로 표현한다.</p> <p>(4)수식: $(B_1 - b_1) / b_1, (B_2 - b_2) / b_2, (B_3 - b_3) / b_3, (B_4 - b_4) / b_4, (B - b) / b$</p>
입술	<p>(1)입술의 크기변화 이용</p> <p>(2)무표정영상에서 입술의 가장 아래부분과 입술중앙의 거리차: g_1, 입술의 가장 윗부분과 입술중앙의 거리차: g_2, 입술의 가장 윗부분과 아래부분의 거리차: g, 입술의 넓이: w</p> <p>(3)표정영상에서 각 거리의 차는 무표정영상에서의 거리의 차와 같다. 이를 G_1, G_2, G, W로 표현한다.</p> <p>(4)수식: $(G_1 - g_1) / g_1, (G_2 - g_2) / g_2, (G - g) / g, (W - w) / w$</p>
눈 & 입술, 눈썹 & 코	<p>(1)얼굴전체의 움직임의 변화 이용</p> <p>(2)무표정영상에서 왼쪽눈썹끝과 입술위쪽의 중앙까지 거리차: n_1, 왼쪽눈썹중앙과 코 중앙까지 거리차: n_3, 왼쪽코끝과 입술중앙까지 거리차: n_5, 오른쪽눈썹끝과 입술의 중앙까지 거리차: n_2, 오른쪽눈썹중앙과 코 중앙까지 거리차: n_4, 오른쪽코끝과 입술중앙까지 거리차: n_6</p> <p>(3)표정영상에서 각 거리의 차는 무표정영상에서의 거리의 차와 같다. 이를 $N_1, N_2, N_3, N_4, N_5, N_6$로 표현한다.</p> <p>(4)수식: $(N_1 - n_1) / n_1, (N_2 - n_2) / n_2, (N_3 - n_3) / n_3, (N_4 - n_4) / n_4, (N_5 - n_5) / n_5, (N_6 - n_6) / n_6,$</p>



[표 1] 얼굴모델

3.3 신경망 학습

신경망 학습을 위해 MLP를 사용하였고, 54개의 입력층 ([표1] 참조), 10개의 은닉층을 사용하였으며, 90% 이상의 성능이 보일 때, 학습을 멈추도록 하였다. 입력층의 입력값은 0 에서 1사이로 정규화하여 실험하였다.

4. 실험

제한하는 알고리즘을 수행하기 위해서 140 * 160 의 해상도를 가지고 있는 12명의 12개의 무표정영상과 101개의 표정영상을 가지고 얼굴 부분만 추출하여 66 * 60 해상도로 정규화하여 사용하였다. 신경망 학습을 위해서 연세대 인지과학연구소의 표정데이터베이스로부터 5명의 56개의 표정영상을 이용하였으며, 각 영상은 수동으로 특징점을 두 번씩 표시하였으며, 총 112개의 영상을 훈련하였다. 연세대 데이터베이스로부터 7명의 55개의 표정영상을 2번씩 특징점을 찾아서 테스트를 하였다. 수동으로 특징점을 찾은 것과 아이겐포인트를 이용한 자동으로 특징점을 찾은 것과 결과를 비교하였다. 얼굴 표정의 구분은 기쁨, 경멸, 놀람, 공포, 슬픔 등 5가지로 나누었다.

표정	기쁨	경멸	놀람	공포	슬픔	전체
수동 특징점을 이용한 검출률 (%)	100	87	80	80	50	87
자동 특징점을 이용한 검출률 (%)	100	97	80	50	40	72

[표 2] 실험결과

[표2] 에서 보는 것과 같이 수동을 이용한 전체 검출률은 87%였지만, 아이겐포인트를 이용한 전체 검출률은 72%로 나왔다.

5. 결론

차후 연구에서는 보다 정확한 표정 분류 연구에 초점을 맞출 것이다. 또한 "기쁨" 검출률에 비해 "공포"나 "슬픔"의 검출률이 상대적으로 낮는데, 이것은 "공포"나 "슬픔" 사이에 특징벡터의 차이가 크지 않기 때문으로 해석된다. 따라서 추가적인 특징벡터를 추출하여 검출 성능을 향상시키도록 할 것이다.

감사의 글 : 본 연구는 학술진흥재단(KRF-2002-005-H20002)의

자원을 받아 수행하였습니다.

6. 참고문헌

[1] P. Ekman and W. Friesen, "The Facial Action Coding System: A Technique for The Measurement of Facial Movement", San Francisco: Consulting Psychologists Press, 1978.

[2] Z. Zang, M. Lyons, M. Schuster and S. Akamatsu, "Comparison between Geometry-Based and Gabor Wavelets-Based Facial Expression Recognition Using Multi-Layer Perceptron", *Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp.454-459, 1998.

[3] C.L. Huang and Y.M. Huang, "Facial Expression Recognition Using Model-Based Feature Extraction and Action Parameters Classification", *J. Visual Comm. and Image Representation*, Vol. 8, no. 3, pp. 278-290, 1997.

[4] M.J. Black and Y. Yacoob, "Recognizing Facial Expressions in Image Sequences Using Local Parameterized Models of Image Motion", *International Journal on Computer Vision*, Vol.1, 1997.

[5] M. Wang, Y. Iwai and M. Yachida, "Automatic Expression recognition from time-sequential facial images by use of expression change model", *Proceedings of Third IEEE International Conference on Face and Gesture Recognition*, pp.324-329, 1998.

[6] J.F Cohn, A.J. Zlochower, J.J Lien and T. Kanade, "Feature point tracking by optical flow discriminates subtle differences in facial expression", *Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp.396-401, 1998.

[7] Y. Tian, T. Kanade and J.F. Cohn, "Recognizing action units for facial expression analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.97-115, 2001.

[8] M. Covell and C. Bregler, "Eigen_points[image_matching]", *Proceedings of International Conference on Image Processing*, Vol.3, pp. 471 -474, 1996.

[9] A. Kapoor and R.W. Picard, "Real-time, fully automatic upper facial feature tracking", *Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, pp.8-13, 2002.

[10] Y.-I. Tian, T. Kanade and J.F Cohn, "Recognizing action units for facial expression analysis", *IEEE Transactions on Pattern Analysis and Machin Intelligence*, Vol.23, pp.97-115, 2001.

[11] F. Bourel, C.C Chibelushi and A.A Low, "Robust facial expression recognition using a state-based model of spatially localised facial dynamics", *Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, pp.106-111, 2002.