

매니폴드 학습을 이용한 자연스런 표정인식

천영재^o 신중주 성재원 김대진

포항공과대학교 컴퓨터공학과 지능형 미디어 연구실

{ yjcheon^o, jjshin, jwsing, dkim } @ postech.ac.kr

Natural facial expression recognition using manifold learning

Yeongjae Cheon^o Jong-Ju Shin Jaewon Sung Daijin Kim

Intelligent Media Lab, Dept. of CSE, POSTECH

얼굴표정은 사람의 중요한 표현 수단 가운데 하나이다. 따라서 사람과 컴퓨터간의 상호작용을 위한 방법으로 얼굴표정인식은 활발하게 연구되어지고 있다. 얼굴표정인식의 연구는 크게 정적인(statistic) 방법과 동적인(dynamic) 방법으로 나뉜다. 정적인 분류기는 한 프레임의 이미지만을 가지고 어떤 표정에 속하는지를 분류해 낸다.[1, 2] 정적인 분류기는 연속된 여러 프레임의 이미지로부터의 정보를 이용하여 최종적으로 표정을 분류해 낸다. Cohen[3]은 고차원의 HMM을 이용하여 temporal spotting까지 할 수 있는 2계층의 HMM모델을 제안하였다. Chang[4,5,6]은 매니폴드를 이용한 modified Lipschitz Embedding에서의 표정의 확률모델로 표정을 인식하고 있다.

본 논문에서 제안하고 있는 사람에 독립적인 자연스런 표정인식 시스템은 표정의 인식 성능을 높이기 위한 방법으로 주요한 세 가지 모듈로 구성된다. 첫 번째는 서로 다른 사람들 사이에 생기는 차이로 인하여 특징점이 overlap되는 문제를 극복하기 위하여 differential feature 를 정의하고 사용하였다는 것이다. 두 번째는 시간상의 사람의 얼굴 표정 이미지 사이의 관련성을 모델링하기 위하여 manifold learning을 통하여 표정 스페이스를 학습하였다. 마지막으로 노이즈한 특징점에도 강인하게 표정을 인식하기 위하여 매니폴드 공간상에서의 시퀀스의 trajectory를 k-NN majority voting 방법을 이용하여 표정을 인식하였다. 다음은 앞에서 설명한 세 가지 모듈에 대한 자세한 설명이다.

Differential feature : 표정인식을 위한 특징점 추출에서 얼굴 표정의 변화에 따른 특징이 아닌 같은 표정임에도 사람간의 차이에서 오는 특징 등은 정규화 되어야 함이 마땅하다. 본 논문은 이를 위한 방법으로 'Differential feature'를 제안한다. 시간 t에서의 Differential feature는 시간 t에서의 얼굴 이미지의 AAM coefficient feature[7] 와 현재 사람의 reference AAM coefficient feature 의 차로 정의한다. 여기서 reference AAM coefficient feature는 각 사람의 feature domain의 offset의 의미를 가진다.

얼굴 이미지로부터 좋은 성능의 표정인식을 위한 특징점(Feature)을 추출하고자 할 때 사람에 따른 차이를 제거하고 정규화하는 과정은 인식기의 성능에 중요한 영향을 미친다. 본 논문은 사람마다 서로 다른 표정 도메인을 가지고 있지만 그 형태가 유사한 것에 착안하여 무표정(Neutral Facial Expression)을 표정 도메인의 reference로 하고 현재 이미지의 특징점과의 차이를 differential feature로 하여 정규화한다. 이를 위하여 reference로 사용하기 위한 무표정한 이미지를 찾는 방법으로 얼굴표정 feature의 motion를 관찰한다. 시간 t에서의 Differential feature는 시간 t에서의 얼굴 이미지의 AAM coefficient feature 와 현재 사람의 reference AAM coefficient feature 의 차로 정의한다. 여기서 reference AAM coefficient feature는 각 사람의 feature domain의 offset의 의미를 가지게 된다.

Manifold learning / k-NN majority voting : 무표정으로부터 각 표정으로의 differential feature를

이용하여 manifold를 학습하였다. 학습된 manifold 공간상에 표정 시퀀스를 embedding하여 gallery를 만든다. 그리고 실시간의 입력 시퀀스에 대하여 manifold 공간상에서 k-NN majority voting한다. distance measure로는 Modified Hausdorff Distance[8] 의 일부인 directed distance를 사용하였다.

앞에서 정의하고 설명된 'Differential feature', Manifold learning 그리고 k-NN majority voting 방법으로 구현된 시스템은 사람에 무관하게 실시간으로 4가지 표정을 인식한다. 특별히 본 논문에서 제안하고 있는 'Differential feature'를 통하여 특징점을 검출하는 방법은 각 사람에 대하여 자연스런 표정까지도 정확히 분류 할 수 있음을 실험을 통하여 확인 할 수 있었다. 제안된 시스템은 22명의 표정 database에 대하여 94.44%의 높은 성능을 보임을 확인 할 수 있었다.

참고문헌

- [1] Y. Tian, T. Kanade, and J.F. Cohn, "Recognizing action units for facial expression analysis," IEEE Trans. On Pattern Analysis and Machine Intelligence 23 (2), 2001.
- [2] M.J. Lyons, S. Akamatsu, M. Kamachi, and J. Gyob, "Coding facial expressions with gabor wavelets," Proceedings of Int. Conf. on Automatic Face and Gesture Recognition, 1998.
- [3] I. Cohen, N. Sebe, L. Chen, A. Garg, and T. Huang, "Facial expression recognition from video sequences: temporal and static modeling," Computer Vision and Image Understanding, 91 (1), pp. 160-187, 2003.
- [4] Y. Chang, C. Hu, and M. Turk. "Mainfold of facial expression." IEEE Int. Workshop on Analysis and Modeling of Faces and Gestures, 2003.
- [5] Y. Chang, C. Hu, and M. Turk. "Probabilistic expression analysis on manifolds," IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2004.
- [6] C. Hu, Y. Chang, R. Feris, and M. Turk. "Manifold based analysis of facial expression," CVPR Workshop on Face Processing in Video, 2004.
- [7] J. Xiao, S. Baker, I. Matthew, and T. Kanade, "Real-time combined 2D+3D active appearance models," IEEE Conf. on Computer Vision and Pattern Recognition, pp. 535-542, 2004.
- [8] M. Dubuisson and A.K. Jain, "A Modified Hausdorff Distance for Object Matching," Proceedings of Int. Conf. on Pattern Recognition, pp 566-568, 1994.