

고유영상에 기반을 둔 동물표현의 수화인식

조용현, 홍성준*
대구가톨릭대학교 컴퓨터정보통신공학부
e-mail:yhcho@cu.ac.kr

Sign Language Recognition for Representing Animals Based on Eigenimages

Yong-Hyun Cho, Seong-Jun Hong*
School of Computer and Information Communications. Eng., Catholic Univ.
of Daegu

요 약

본 논문에서는 고유영상 추출에 기반을 둔 동물표현의 수화인식 기법을 제안하였다. 여기서는 영상으로부터 구조적 특징을 효과적으로 추출하기 위하여 PCA를 이용한 방법을 제안한다. 이는 선형투영 및 차원의 감소에 따른 계산부하의 감소로 인식시간을 줄이기 위함이다. 제안된 기법을 10종류의 240*215 픽셀 동물표현의 수화(160개 : 1명*10종류*16동작)영상을 대상으로 City-block의 분류척도를 이용하여 실험한 결과 우수한 인식성능이 있음을 확인하였다.

1. 서론

인간과 컴퓨터의 상호작용으로 이용되는 한 가지 수단인 몸짓인식은 대부분 손동작으로 이루어진다. 이렇게 몸짓을 컴퓨터의 입력으로 인식하는 것은 장애인들로 하여금 더 쉽게 컴퓨터 등에 접근할 수 있도록 한다[1-3].

일반적으로 인간과 컴퓨터를 연결하기 위한 시각 인터페이스 기술은 시각정보 센싱(입력) 기술과 시각정보 표현(출력) 기술로 나눌 수 있다. 입력 측에서는 정밀한 센싱 기술에 대한 연구가 진행 중이며 주로 얼굴 및 제스처 인식에 대한 연구가 활발히 진행되고 있다[1,2]. 이러한 얼굴 및 제스처 인식을 위한 영상인식기술은 MIT, CMU, 메릴랜드 등의 대학과 마이크로소프트 등 기업체에서 다양한 센싱 기법과 알고리즘들을 개발하고 있다. 얼굴인식 관련기술로 MIT Media Lab.에서는 얼굴특징을 추출하여 정규화한 후 Eigenvalue의 변화를 통해 얼굴을 인식하는 엔진을 개발하였다[4,5]. 이외에 eTrue나 Visonics 등 업체에서 동영상에서의 얼굴영역 분리 및 인증기술, 국부특징 분석에 의한 인식 알고리즘을 개발하였다. 또한 미국 국방성 DARPA에서는 20년 이상 Image Understanding Program을 진행하고 있으며 영상인식 및 이동물체 추적기술은 상당부분

실용화 단계에 이르고 있다[2].

국내에서도 연구소와 대학을 중심으로 시각인터페이스를 위한 입체영상 및 신호처리 기술에 대한 많은 연구들을 진행하고 있다. KIST 등의 연구소에서 입체영상 기술에 대한 연구를 주로 수행하고 있으며, 서울대 등의 대학에서는 양안 및 다안식 입체영상 신호처리 관련연구 및 홀로그램에 관한 기술들을 연구하고 있다.

한편 수화는 청각 장애인이 의사전달을 위해 사용하는 언어이다. 수화는 손의 움직임과 형태로 표현 가능하며, 어휘를 표현하기 위한 수화단어와 자음, 모음, 숫자, 알파벳 등을 표현하기 위한 지화로 구성되어 있다. 이러한 수화의 인식을 위해서는 수화동작 중 의미 없는 손동작과 의미 있는 수화단어를 구분해 인식하는 것이 필요하다. 일반적으로 수화는 손의 동작과 형태 정보를 가지고 정의되기 때문에 손의 모양과 움직임을 동시에 고려해 특징으로 사용한다. 또한 수화인식에서 가장 중요한 문제는 의미 있는 동작과 의미 없는 동작을 구분하는 것이다. 의미 없는 손동작을 모델링하기 위해서는 무한대의 학습데이터가 필요하기 때문에 일반적인 방법으로 해결이 불가능하다[1,2].

한편 표정이나 얼굴 등의 인식을 위한 특징추출에는 holistic analysis에 바탕을 둔 연구들이 주로 이

용되며, 주성분분석(principal component analysis : PCA)[4-6]과 이들에 바탕을 둔 기법들이 연구되고 있다[6]. 이들 방법들은 인식을 위한 영상 속에 포함된 진폭위상만을 나타내는 2차원 통계성을 이용한다. 특히 PCA는 영상을 표현하는 고유벡터를 추출하여 벡터 공간으로 투영시킨 가중치를 이용하는 방식으로, 전체의 영상들을 적은 수의 벡터를 이용하여 표현할 수 있으며 빠르고, 계산이 간단하며, 점진적 변화에 덜 민감하다.

본 논문에서는 수화인식을 위한 영상으로부터 구조적 특징을 추출하기 위하여 PCA를 이용한 방법을 제안한다. 이는 인식에 필요한 계산량을 줄여 특징추출 시간을 줄이기 위함이다. 제안된 기법을 10 종류의 240*215 픽셀 동물표현의 수화(160개 : 1명*10종류*16동작)영상을 대상으로 City-block의 분류 척도를 이용하여 실험하고, 그 성능을 고찰하였다.

2. PCA에 의한 고유영상 추출

PCA는 입력 데이터의 공분산에 의해 결정되는 정규직교기저로 표현되며, 가우스 변수의 모든 정보는 공분산행렬 내에 포함되기 때문에 가우스 분포의 입력 데이터에 대해서는 적당한 것이다. 일반적으로 영상 내 각 픽셀사이의 2차원 기저영상은 주요한 특징이며, 이의 추출에 PCA를 이용하였다.

자기상관행렬(autocorrelation matrix) $R_{xx} = \langle xx^T \rangle$ 를 가진 영 평균(zero-mean)된 입력벡터 $x = [x_1, x_2, \dots, x_n]^T$ 에 대해서 생각해 보자. 여기서 T는 전치를 나타내며, $\langle \cdot \rangle$ 는 기대치(expected value)를 나타낸다. 또한 $\hat{w}_1, \hat{w}_2, \dots, \hat{w}_n$ 이 R_{xx} 의 고유벡터(eigen vector)와 직교되는 연결가중치 벡터라 할 때, $\hat{w}_1 = [\hat{w}_{11}, \hat{w}_{12}, \dots, \hat{w}_{1n}]^T$ 는 가장 큰 고유값(eigen value) λ_1 과 일치하며, $\hat{w}_2 = [\hat{w}_{21}, \hat{w}_{22}, \dots, \hat{w}_{2n}]^T$ 는 두 번째로 큰 고유값 λ_2 , 그리고 $\hat{w}_n = [\hat{w}_{n1}, \hat{w}_{n2}, \dots, \hat{w}_{nm}]^T$ 는 가장 작은 고유값 λ_n 과 각각 일치한다. 이상의 관계를 행렬방정식으로 나타내면 식 (1)과 같다.

$$R_{xx}\hat{w}_j = \lambda_j\hat{w}_j, \quad (j=1,2,\dots, n) \quad (1)$$

여기서 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ 이다. 주어진 입력벡터 x 를 위한 첫 번째 k개의 주요 특징을 나타내는 고유벡터 y 는 다음의 선형변환식 (2)로 나타

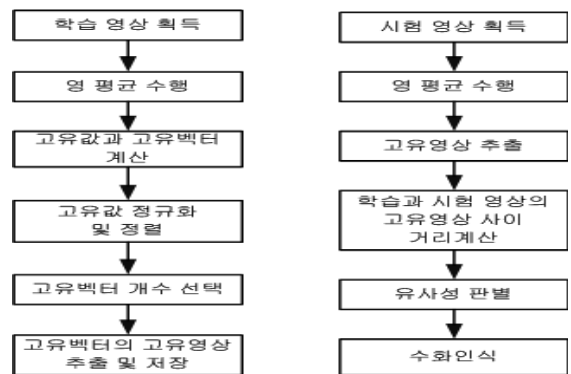
낼 수 있다.

$$y = \hat{W}x \quad (2)$$

여기서 $\hat{W} = [\hat{w}_1, \hat{w}_2, \dots, \hat{w}_k]^T \in \mathbb{R}_{k \times n}$ 이며, 이 식에서 연결가중치행렬 \hat{W} 의 행은 가장 큰 고유값과 일치하는 상관행렬 R_{xx} 의 고유벡터임을 의미한다. 다시 말하면, 입력데이터 공간에서 k차원의 주요특징을 나타내는 부공간은 R_{xx} 의 k개 주요 고유벡터에 의해 구성된 서브공간으로 정의된다. 결국 PCA는 $\|\hat{w}_j^T x\|^2 >$ 가 최대인 고유벡터 $\hat{w}_1, \hat{w}_2, \dots, \hat{w}_k$ 의 방향을 찾는 것이다. 일반적으로 얻어지는 고유값은 크기에 따라 정렬하고 고유벡터도 해당 고유값의 위치대로 정렬한다. PCA에서 순서대로 정렬된 고유값의 뒤쪽은 0에 가까운 값을 가지게 되어 이를 삭제할 수 있다. 이는 고유벡터의 작은 값들을 고려하지 않음으로써 입력 데이터의 차원을 줄이기 위함이다.

한편 PCA를 좀 더 효과적으로 수행하기 위해 신호의 영 평균(zero-mean)을 수행한다[4-6]. 이는 신호의 1차적 통계성을 고려한 정규화로 PCA를 간략화 하기 위함이다. 신호의 영 평균은 신호벡터 x 에서 평균값 x^* 를 뺀 차로 $x = x - x^*$ 이다.

따라서 수화영상을 대상으로 PCA를 수행하면 2차의 고유영상을 추출할 수 있다. 이는 선형투영 및 차원의 감소에 따른 계산부하의 감소로 인식시간을 줄이기 위함이다. PCA를 이용한 제안된 얼굴인식 과정을 요약하면 그림 1과 같다. 여기서 고유벡터의 개수 설정은 차원의 감소를 가져오며, 경험에 의해서 설정할 수 있다.



(a) 학습 (b) 인식
그림 1. 제안된 수화영상 인식과정

학습과정에서 추출된 고유벡터는 수화의 특징벡터로 2차원의 속성을 가지는 기저영상이며 시험영상의 판별을 위해 저장된다. 한편 식별과정에서 유사성 판별을 위한 거리척도로는 city-block, Euclidean, negative angle들이 이용된다. 본 논문에서는 City-block의 척도를 이용하였다.

3. 실험 및 결과분석

PCA에 의한 고유영상을 이용한 제안된 동물표현의 수화인식 방법의 성능을 평가하기 위해 240*215 픽셀의 160개(1명*10종류*16수화) 영상을 대상으로 실험하였다. 여기서 곰, 소, 사슴, 오리, 코끼리, 염소, 돼지, 토끼, 호랑이, 늑대의 10가지 동물을 나타내기 위한 수화동작으로 16개 연속동작 프레임을 이용하였다. 각 동물의 수화동작 프레임 16중에서 홀수 프레임(1, 3, 5, ..., 15) 영상은 학습에, 나머지 짝수 프레임(2, 3, 6, ..., 16) 영상은 시험에 이용하였다. 인식성능을 확인하기 위하여 City-block의 판별척도를 이용하였다. 실험환경은 펜티엄IV-3.0G 컴퓨터에서 Matlab 7.1로 구현하였다. 그림 2와 그림 3은 각각 10가지 동물에 대한 학습과 시험영상을 나타낸 것이다. 여기서 상단에서 하단으로 곰, 소, ..., 늑대를 나타내고, 좌측에서 우측으로 각각 프레임 1부터 15, 또는 프레임 2부터 16까지를 나타낸다.



그림 2. 10가지 동물표현의 수화 학습영상

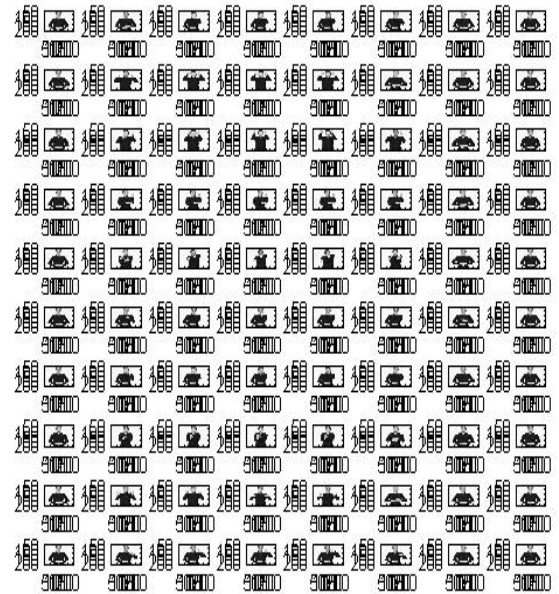


그림 3. 10가지 동물표현의 수화 시험영상

한편 그림 4는 곰을 나타내는 8개의 수화 학습영상을 대상으로 주요성분분석을 이용하여 추출된 8개의 고유영상과 이들로부터 복원된 영상을 나타낸 것이다. 이때 복원된 영상과 원 학습영상과의 전체오차는 약 0.12정도로 아주 유사함을 알 수 있다.

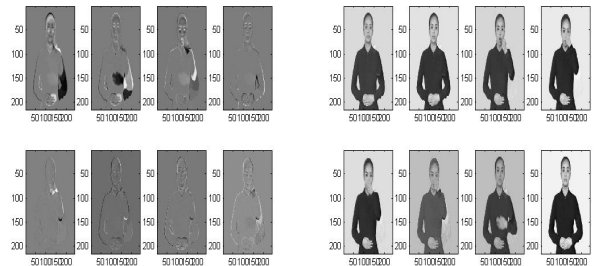


그림 5. 곰 학습영상의 고유영상 및 복원영상

<표 1>은 고유영상을 8개로 하였을 때, 학습영상 8개와 시험영상 8개를 대상으로 City-block 거리척도에 따른 인식결과를 각각 나타낸 것이다.

<표 1> 10개 동물의 8개 학습과 시험영상의 인식결과

거리척도	입력영상	PCA	
		인식률(%)	오차평균
City-block	학습영상	100	0
	시험영상	97.5	0.201714

참고문헌

여기서 오차평균은 올바르게 인식된 영상만을 대상으로 학습의 기저영상과 시험의 기저영상 사이 오차의 절대치 평균이다. 표에서 보면 학습된 수화영상은 모두 100%의 인식률을 보이나, 시험 수화영상은 약 97.5%의 인식률을 보인다. 시험 영상의 경우 돼지의 6번 프레임 영상이 소로, 호랑이의 3번 프레임 영상은 돼지로 각각 잘못 인식되었다. 이는 이들 프레임 영상 사이가 육안으로도 서로 유사성이 있음을 알 수 있다.

따라서 PCA를 이용한 수화동작 인식은 빠르면서도 정확한 인식이 가능함을 확인할 수 있다.

4. 결론

본 논문에서는 고유영상 추출에 기반을 둔 동물표현의 수화인식 기법을 제안하였다. 여기서는 영상으로부터 구조적 특징을 효과적으로 추출하기 위하여 PCA를 이용한 방법을 제안한다. 이는 선형투영 및 차원의 감소에 따른 계산부하의 감소로 인식시간을 줄이기 위함이다.

제안된 기법을 10종류의 240*215 픽셀 동물 표현 수화(160개 : 1명*10종류*16동작) 영상을 대상으로 City-block의 분류척도를 이용하여 실험한 결과, 학습 수화영상에서는 100%, 시험 수화영상에서는 97.5%의 인식결과를 각각 가져 우수한 인식성능이 있음을 확인하였다.

향후 제안된 방법을 좀 더 큰 규모의 인식문제와 다양한 분야에 적용하는 연구가 지속적으로 이루어져야 할 것이다.

- [1] 박호식, 배철수, “수화인식을 위한 얼굴과 손 추적 알고리즘,” 한국통신학회논문지, 제 31권 11C호, 2006년 11월
- [2] 김희곤, 오형철, 양윤모, “수화인식을 위한 특징량 추출하드웨어 설계,” 대한전자공학회 추계학술대회 논문집, 제 18권 2호, 1995년 12월
- [3] 조재현, 엄성용, “다중마커를 이용한 수화 인식 기법,” 서울여자대학교대학원 논문집, 제 1998권 6호, 1998년
- [4] J. T. Tou and R. C. Gonzalez, ‘Pattern Recognition Principles,’ Addison-Wesley Pub., London, 1974
- [5] Y. L. Tian, T. Kanada, and J. F. Cohn, “Recognizing Action Units for Facial Expression Analysis,” *EEE. Trans. on Pattern Analysis and Machine Intelligence*, Vol.23, No.2. pp. 1-19, Feb. 2001
- [6] K. I. Diamantaras and S. Y. Kung, *Principal Component Neural Networks : Theory and Applications, Adaptive and learning Systems for Signal Processing, Communications, and Control*, John Wiley & Sons, Inc., 1996