creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

조선대학교
CHOSUN UNIVERSITY

[UCI]I804:24011-200000266444

2017년 8월

석사학위논문

키넥트 센서 데이터를 이용한 K-pop 포인트 안무 분석에 관한 연구

조 선 대 학 교 대 학 원

제어계측공학과

김 동 현

조선대학교
CHOSUN UNIVERSITY

키넥트 센서 데이터를 이용한 K-pop 포인트 안무 분석에 관한 연구

K-pop Point Dance Analysis Based on Using Kinect Sensor Data

2017년 8월 25일

조 선 대 학 교 대 학 원

제어계측공학과

김 동 현

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

키넥트 센서 데이터를 이용한 K-pop 포인트 안무 분석에 관한 연구

지도교수 곽 근 창

이 논문을 공학석사학위신청 논문으로 제출함.

2017년 4월

조 선 대 학 교 대 학 원

제어계측공학과

김 동 현

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

김동현의 공학석사학위논문을 인준함

위원장 조선대학교 교수 장순석 인

위 원 조선대학교 교수 반성범 인

위 원 조선대학교 교수 곽근창 인

2017년 5월

조선대학교 대학원

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

목 차

제1장 서론 ····· 1

제2장 ELM(Extreme Learning Machine) ····· 5

제1절 ELM의 구조 ····· 5

제2절 ELM 분류기 ····· 6

제3절 R-ELM 분류기 ····· 10

제3장 K-pop 포인트 댄스 분류 방법 ····· 13

제1절 K-pop 포인트 댄스 특징 추출 ····· 13

제2절 PCA + LDA 차원축소 알고리즘 ····· 19

제3절 K-pop 포인트 댄스 분류 ····· 23

제4장 실험 및 결과 ····· 26

제1절 K-POP 포인트 댄스 DB 구축 ····· 26

제2절 실험결과 ····· 34

제5장 결론 ····· 42

참고문헌 ····· 44

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 목 차

표 2-1 활성함수의 종류 ····· 8
표 3-1 접점을 가지는 사이 각 정보 ····· 15
표 4-1 Kinect 성능 비교 ····· 26
표 4-2 Kinect V2 Joint 정보 ····· 27
표 4-3 K-POP 곡 정보 ····· 29
표 4-4 컴퓨터 사양 ····· 36
표 4-5 통계적 각도 특징을 사용한 분류 성능 ····· 38
표 4-6 PCA를 통해 차원 축소된 특징을 사용한 분류 성능 ····· 38
표 4-7 PCA+LDA를 통해 차원 축소된 특징을 사용한 분류 성능 ····· 39

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

그 림 목 차

그림 2.1 일반적인 ELM 구조 ····· 5
그림 3.1 K-pop 포인트 댄스 분류 과정 ····· 13
그림 3.2 3차원 좌표 A,B,C ····· 14
그림 3.3 몸의 기울기를 구하기 위한 좌표 A,B,C,D ····· 15
그림 3.4 좌 팔꿈치 사이 각도 ····· 16
그림 3.5 우 어깨 사이 각도 ····· 17
그림 3.6 좌 팔꿈치 통계적 각도 ····· 18
그림 3.7 우 어깨 통계적 각도 ····· 18
그림 3.8 K-pop 포인트 댄스 특징 데이터 ····· 19
그림 3.9 임의의 3차원 데이터의 주성분 벡터 ····· 20
그림 3.10 클래스 내 분산과 클래스 간 분산 ····· 20
그림 3.11 PCA+LDA와 R-ELM 분류기를 이용한 K-pop 포인트 댄스 분류 시스템 ·· 25
그림 4.1 키넥트 V2 Joint ····· 27
그림 4.2 데이터 구축 환경 ····· 28
그림 4.3 포인트 댄스 촬영영상 ····· 33
그림 4.4 포인트 댄스 실루엣 ····· 33
그림 4.5 포인트 댄스 프레임 별 변화도 ····· 34
그림 4.6 포인트 댄스 데이터의 PCA 고유 값 ····· 35
그림 4.7 판별 벡터 수에 따른 유클리디안 거리 분류 정확도 ····· 36
그림 4.8 K-pop 포인트 댄스의 피셔 댄스 ····· 36
그림 4.9 PCA+LDA 특징 벡터 사용 분류기 분류 속도 ····· 39
그림 4.10 은닉 노드 수 변화에 따른 sigmoid 함수 기반의 R-ELM 분류 성능 ····· 40
그림 4.11 은닉 노드 수 변화에 따른 sin 함수 기반의 R-ELM 분류 성능 ····· 41
그림 4.12 은닉 노드 수 변화에 따른 radbas 함수 기반의 R-ELM 분류 성능 ····· 41

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

ABSTRACT

K-pop Point Dance Analysis Based on Using Kinect Sensor Data

Kim, Dong-Hyeon
Advisor : Prof. Kwak, Keun Chang, Ph. D.
Dept. of Control and Instrumentation Eng.,
Graduate School of Chosun University

In this paper, we suggests a method of classifying Korean pop (K-pop) dances. In order to accomplish this, we construct a K-pop dance database, including 200 dance movements produced by four professional dancers and forty amateur dancers from skeletal joint data obtained by a Kinect sensor. First, we obtain thirteen core angles representing important motion features from 25 markers in each frame. The calculated angles are normalized with a min-max normalization technique and four statistical values (minimum value, average value, maximum value, variance) are calculated every 30 frames. The 13 statistical angles calculated through the above procedure are used as feature vectors by connecting them. Then, a dimensionality reduction is performed with a combination of principal component analysis and Fisher's linear discriminant analysis. Finally, we design a Regularized-Extreme Learning Machine (R-ELM) classifier using ridge regression analysis. Experimental results show that classification performance is better than existing methods such as k-NN (K-Nearest Neighbor), SVM (Support Vector Machine) and ELM (Extreme Learning Machine).

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제1장 서론

컴퓨터 비전 기술과 그래픽 기술, 다양한 센서들의 발달로 인간과 컴퓨터와의 상호작용을 위한 중요기술 중 하나인 모션 인식 기술이 활발하게 연구 및 발전되고 있다[1-6]. 모션 인식 기술은 사용자의 움직임 및 위치 등을 인식하는 각종 센서를 이용하는 기술을 의미한다. 현재 대한민국대중음악(K-pop)은 2000년도 후반부터 아시아를 시작으로 유럽, 남미, 중동지역, 세계 각국에서 사랑 받고 있다. K-pop은 대한민국의 모든 대중음악의 장르를 포함하는 말이지만 1990년대 이후 댄스, 힙합, R&B, 발라드, 록, 일렉트로닉 음악을 일컫는 말로 불러진다. 온라인 커뮤니티의 활성화로 K-pop은 전 세계의 젊은이들에게 급속도로 전파되었다. K-pop이 세계적으로 사랑을 받는 이유 중 하나로 커버댄스(cover dance)가 있다. 커버댄스는 팬 코스프레의 일종으로, 좋아하는 가수의 노래, 댄스를 모방하는 행위를 가리킨다. 하지만 위와 같이 개인이 아닌 상업적인 이유로 댄스를 모방하는 경우 저작권 표절 문제가 생긴다. 2016년 중국 걸 그룹 'AOS'가 국내 걸 그룹 '여자친구'의 곡의 콘셉트와 댄스를 표절했다는 의혹이 제기된 바 있다. 이와 반대로 대한민국의 가수 '싸이'는 '젠틀맨' 컴백 당시, 안무 저작권을 정식 구입하여 주목을 받았다. 과거에는 댄스의 저작권료를 따로 지불한 사례가 없었으나 그는 안무 팀 측에 저작권료를 지급하는 선례를 남겼다. '젠틀맨'의 포인트 댄스 '시건방 춤'이라 불리는 포인트 댄스를 사용하면서 '시건방 춤' 제작 댄스 팀에 정식 리메이크 절차를 밟은 것이다. 저작권 비용 지급 사례가 화재가 된 이유는 법적으로 안무는 연극 저작물로 분류되어 보호를 명시하고 있지만, 관행적으로 사용자가 안무 저작권을 구입할 의무에서 비켜서 있던 까닭이다. 현재까지는 누군가 만든 안무들을 변화시켜 사용하거나, 학원과 같은 교육기관에서 타인이 만든 안무를 사용해도 논란만 일으킬 뿐 이를 규제할 수 없다. 이러한 관행적인 문제 해결 목적으로 카메라 기반의 모션 인식 기술을 이용하여 K-pop 포인트 댄스를 분류 한다.

모션 인식 기술은 크게 '카메라 기반의 모션 인식 기술', '추가적인 하드웨어 기반의 모션 인식 기술', '주파수 기반의 모션 인식 기술'로 구분한다[7]. 카메라 기반의 모션 인식 기술은 사용자가 장비를 착용하지 않고 카메라를 통해 얻어진 영상 정보를 이용한 영상처리 기술을 기반으로 한다[8]. 비용, 인지범위

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

등의 단점이 있지만 카메라만 있으면 접근이 쉬운 기술이며 카메라 범위에서 높은 정확도로 실시간 인식이 가능하다는 장점이 있다[9]. Microsoft사의 키넥트(Kinect)[10], Intel사의 리얼센스(Real-Sense)[11], 립 모션(Leap Motion)[12]이 대표적인 카메라 기반의 모션 인식 기술이다. 추가적인 하드웨어 기반의 모션 인식 기술은 직접 센서를 부착 및 기기를 사용하는 기술이며 크게 광학식, 기계식, 자기식 시스템이 있다. 광학식은 가장 정확도가 높으며 실시간 인식이 가능하다. 그러나 많은 비용과 넓은 공간이 필요하다는 단점이 있다. Vicon 사의 Vicon System이 대표적 광학식 시스템이다[13]. 기계식은 잡음에 강하며 노이즈가 적은 데이터 획득이 가능하다[14]. 그러나 사용자가 무거운 기계 장치를 부착 하므로 자연스러운 동작 표현의 어려움이 있다. 대표적인 기계식 시스템으로 Analogues사의 Gypsy가 있다[15]. 자기식은 사용자의 신체에 자기장을 발생 시키는 센서를 부착하여 데이터를 획득하는 기술이며 유지비가 적으며 준비 과정이 단순한 장점이 있다. 그러나 주변 환경의 영향을 받아 정확도가 떨어지는 단점이 있다[16]. 대표적 자기식 시스템으로 Ascension사의 Motion Star Wireless가 있다[17]. 주파수 기반의 모션 인식 기술은 상용화가 되지 않은 기술이지만 연구를 통해 발전하고 있으며 카메라 기반, 추가 하드웨어 기반의 기술의 단점을 보완하는 기술이다. 그러나 다른 주파수에 의해 데이터가 영향을 받는 단점이 있으며 SoundWave[18], Wisee[19]가 대표적인 기술이다. 본 논문에서는 카메라 기반의 모션 인식 기술 중 여러 분야에서 연구되고 있는 Microsoft사에서 개발된 Kincet를 사용하여 얻어진 3차원 관절 데이터를 사용한다[20-29].

기계학습 혹은 머신러닝(Machine Learning)은 인공지능(AI: Artificial Intelligence) 기술의 한 분야로 컴퓨터를 사용해 스스로 학습할 수 있도록 하는 기술을 말한다. 컴퓨터 스스로 어떠한 작업(task)에 대해 데이터로부터 얻어지는 경험(experience)을 통해 그 작업에 대한 성능(performance)을 향상시키는 것으로 정의된다. 기계학습은 경험에 대한 좋은 품질의 데이터를 많이 가지고 있다면 보다 높은 성능을 낼 수 있다. 기계학습은 데이터에 대한 목표 값(targeted value)의 유무에 따라 지도 학습(supervised learning)과 비지도 학습(unsupervised Learning) 크게 두 가지로 나뉜다. 지도 학습은 학습데이터와 학습데이터 대한 목표 값을 입력하면 학습데이터에서 목표 값에 대한 규칙을 찾아 매개변수나 가중치의 수치를 수정하고 마지막으로 검증데이터에 대해 규칙이 적용되는 비율이나 수준(support level)과 신뢰수준(confidence)을 계산하여 일정한 수준을 보장하는

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

규칙만을 남기는 과정을 반복함으로써 높은 정확도를 가지는 결과를 도출하는 학습 방법이다. 지도학습은 목표 값에 따라 분류(classification), 회기(regression)등에 적용된다. 대표적 기법으로 k-NN(k-Nearest Neighbor), SVM(Support Vector Machine), 결정 트리(decision tree)가 있다. 비지도 학습 혹은 자율 학습은 입력데이터에 대한 특성을 발견하는 방법으로 군집화(clustering), 분포 추정(underlying probability density estimation), 차원축소(dimensionality) 등의 기술 모델링(descriptive modeling)에 활용된다. 그 밖에 지도 학습과 비지도 학습의 특성을 결합한 반지도 학습(semi supervised Learning)은 현재 상태(state)에 대해 어떤 행동(action)을 취하는 것이 최적의 방법인지를 학습하는 것이다. 행동에 대해 외부 환경에서 보상(reward)이 주어지는데 이를 최대화 하는 방향으로 학습이 진행되고 위 방법들과는 다르게 실시간 학습을 진행하는 것이 일반적이다. 기계학습의 문제 해결과정은 적용을 위해 문제에 대한 목표를 정의하고, 적절한 방법을 선택한다. 문제에 대해 필요한 데이터를 구축하고 선택된 기계학습 방법을 사용하여 데이터에 대해 모델링 과정을 수행한다. 실제 문제에 구축된 모델을 적용하여 성능을 평가하고 결과를 도출한다[30]. 기계 학습의 지도학습 방법을 사용하는 ELM(Extreme Learning Machine)[31] 모델은 기존 신경회로망의 역 전파(back-propagation) 알고리즘의 단점을 해결하기 위해 제안되었다. 입력 층과 은닉 층을 연결하는 입력 가중치에 임의의 값을 할당하여 반복 학습을 하지 않고 일반화된 역행렬과 최소자승법을 사용해 출력 가중치를 계산함으로서 빠른 학습 속도와 높은 정확도를 가진다. ELM은 이미지 초 해상도[32], 객체 인식[33], 풍력 발전소 측정도, 전기 요금 예측[34-35], 모션 인식[36], 3D 모양분할 및 라벨링(labeling)[37]등 다양한 예측 및 분류 문제[38-43]에 대해 연구 되고 있다. 본 논문에서는 일반적인 ELM 분류기의 학습 과정에서 규제화(regularize) 방법이 적용된 R-ELM(Regularized-Extreme Learning Machine)을 사용한다. ELM은 하나의 은닉 층을 가지는 신경회로망과 같은 SLFNs(Single Hidden Layer Feed-forward Network) 구조를 가진다. ELM은 신경회로망의 과적합(over fitting)과 같은 문제 해결하기 위해 제안된 방법이다. 과적합 문제란, 학습 데이터에 대해 과하게 학습 되어 학습데이터에 대해서는 오차가 감소하지만 실제 데이터에 대해서 오차가 증가하는 현상이며 가중치의 수, 가중치의 자유도가 높을수록 발생 가능성이 높다. 신경회로망과 같은 구조를 가지는 알고리즘에서 가중치의 자유도는 표현 능력과 직결된다. 가중치의 자유도를 낮추는 것은 과다한

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

자유도를 가지는 경우를 제외하면 바람직하지 않다. 따라서 학습 과정에서 가중치의 자유도를 제약하는 규제화 방법을 적용하여 위의 문제를 완화한다[44-45]. 본 논문에서는 R-ELM 분류기를 사용하여 K-pop 포인트 댄스 모션을 분류한다.

Kinect를 사용한 댄스 모션 인식 연구는 일부 발레 댄스[46-48], 살사 댄스[49],인도네시아 전통춤[50]에 관해 연구된바 있다. 본 논문에서는 전문 댄서와 아마추어 댄서를 통해 K-pop 포인트 댄스 데이터베이스를 구축 한다. 포인트 댄스 분류는 특징 추출, 차원 축소, 분류 알고리즘 적용 과정을 가진다. 특징 추출 단계에서 Kinect를 사용하여 얻어진 3차원 관절 데이터를 사용하여 각각의 댄스 모션이 나타내는 13개의 특징 각도를 구한다. 구해진 각도는 각각 30프레임마다 최솟값, 평균, 최댓값, 분산의 통계적 수치를 계산한다. 계산된 13개각도의 통계적 수치를 벡터의 형태로 하나로 이어 붙여 포인트 댄스 모션의 특징 데이터로 사용한다. 차원 축소 과정에서는 주성분 분석(PCA : Principal Component Analysis)[51]과 선형 판별 분석(LDA : Linear Discriminant Analysis)[52]을 사용하여 K-pop 포인트 댄스의 피셔 댄스(Fisher dance)를 구한다. 마지막으로 R-ELM분류기의 학습 데이터로 전문 댄서의 피셔댄스, 검증 데이터로 아마추어 댄서의 피셔댄스를 입력 값으로 사용하여 포인트 댄스를 분류 한다. 제안된 방법은 별도의 장비 없이 Kinect만을 사용하여 쉽게 데이터 취득이 가능하며 K-pop 포인트 댄스 모션을 잘 나타내는 13개의 각도의 통계적 수치와 R-ELM 분류기를 사용하여 빠르고 정확한 분류가 가능하다. R-ELM 분류기의 성능 검증을 위해 기존의 기계학습 알고리즘인 k-NN(k-nearest neighbors), SVM(Support Vector Machine), 일반적인 ELM을 사용하여 비교 분석하고 기존의 분류기 보다 높은 성능을 보임을 증명했다.

2장은 ELM 분류기와 R-ELM 분류기에 대하여 설명하고, 3장은 K-pop 포인트 댄스 데이터 취득과 특징 추출, 분류과정에 대한 설명, 4장은 실험 결과 및 분석에 대해 다룬다. 그리고 5장으로는 결론으로 마무리 한다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제2장 ELM(Extreme Learning Machine)

제1절 ELM의 구조

인공지능의 한 분야인 신경회로망은 최급 경사 법(gradient descent method)에 기반을 두고 좋은 성능을 내기 위해 많은 반복 학습, 즉 역 전파 알고리즘(Back propagation learning algorithm)을 사용한다. 하지만 반복 학습은 적절하지 못한 학습에 의하여 낮은 성능을 보이거나 국부 최소해(local minimum)에 수렴하여 시간이 오래 걸리는 단점을 가진다. Huang[31]에 의하여 제안된 ELM(Extreme Learning Machine) 모델은 신경회로망의 위와 같은 단점을 해결하기 위해 제안된 기계 학습 알고리즘이다.

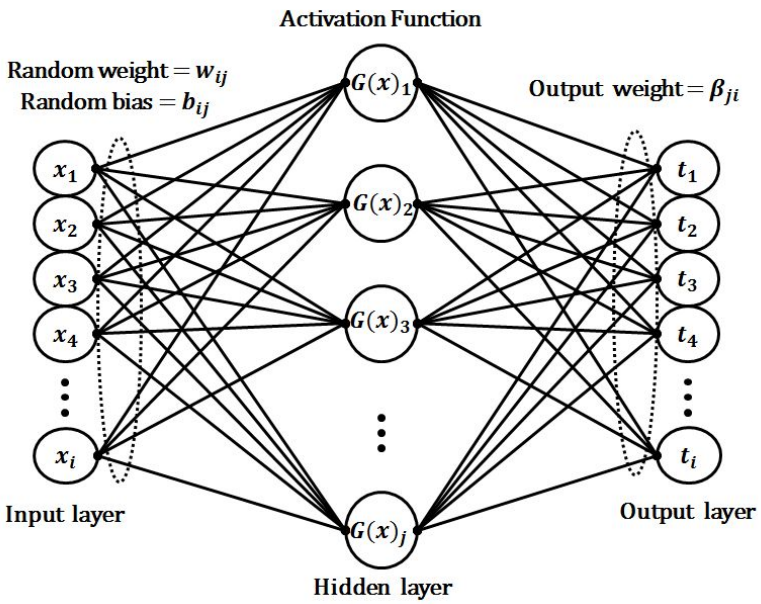


그림 2.1 일반적인 ELM 구조

그림 2.1은 다양한 예측 및 분류 문제에 적용하여 연구 되고 있는 일반적인 ELM의 구조를 나타낸다. 입력 층(input layer), 은닉 층(hidden layer), 출력 층(output layer) 형태로 구성되며 ELM의 학습 과정은 목표 값을 가지는 지도 학습

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

과정을 가진다. 출력 가중치와 바이어스 항에 임의의 값을 할당하고 최소 자승법(least square estimation)과 모어-페로스의 일반화된 역행렬(moore-penrose generalized inverse)을 사용하여 분석적으로 계산되며 반복 학습이 이루어지지 않는다. 반복 학습을 하지 않고 전 방향(feed forward)단계에서 학습이 완료됨으로써 신경 회로망의 계산 속도에 비해 월등히 좋은 성능을 가지며 과잉 적합(over fitting)으로 인한 성능 문제가 나타나지 않는다. ELM의 파라미터는 입력 층과 은닉 층을 연결하는 입력 가중치 w, 입력 바이어스(bias) b를 가지고 은닉 층과 출력 층을 연결하는 출력 가중치 β 가 존재하며 사용자에 의해 결정되는 은닉 층의 노드 개수, 은닉 층에서 사용되는 활성 함수(active function)가 있다.

제 2절 ELM 분류기

ELM 분류기를 사용함에 따라 학습 시간을 단축하여 빠르고 정확한 분류기를 구현할 수 있게 되고 높은 정확도를 얻을 수 있다. 일반적인 ELM 분류기의 분류과정을 설명하기 위해 입력 학습 데이터 $x_i = [x_1, x_2, x_3, ..., x_i]$, 학습 데이터 목표 출력 값 $t_i = [t_1, t_2, t_3, ..., t_i]$, 검증 데이터, $y_i = [y_1, y_2, y_3, ..., y_k]$, 검증 데이터 목표 출력 값 $z_k = [z_1, z_2, z_3, ..., z_k]$, 은닉 층 노드 j개, 활성 함수 $G(x)$가 주어진다고 가정 할 때 학습 데이터의 출력은 식 (2.1)로 표현된다.

$$t_i = \sum_{j=1}^{n} \beta_j \cdot G_j(x_i) = \sum_{j=1}^{n} \beta_j \cdot G_j(w_{ij} \cdot x_i + b_{ij}) \qquad (2.1)$$

$G_j(x_i)$는 은닉 층의 j번째 노드의 출력이다. w_{ij}는 i번째 입력 층의 노드와 j번째 은닉 층의 노드 사이의 연결 가중치를 의미한다. b_j는 은닉 층의 j번째 노드의 바이어스 값이며 β_j는 출력 가중치를 나타낸다. 학습 데이터의 은닉 층 출력이 식(2.2)와 같은 때 식 (2.3)과 같이 간단히 나타 낼 수 있다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

$$\mathrm{H} = \begin{bmatrix} G(w_{11} \cdot x_1 + b_{11}) & \cdots & G(w_{ij} \cdot x_1 + b_{ij}) \\ \vdots & \cdots & \vdots \\ G(w_{1j} \cdot x_1 + b_{1j}) & \cdots & G(w_{ij} \cdot x_i + b_{ij}) \end{bmatrix} \tag{2.2}$$

$$\mathrm{H}\beta = t \tag{2.3}$$

ELM의 입력 가중치 w와 바이어스 값 b는 확률 분포 함수를 사용하여 임의의(random)값을 할당한다. 최적의 성능을 얻을 수 있는 은닉 층과 출력 층 사이의 출력 가중치 β를 구하기 위해 식 (2.4) 최소 자승법을 사용한다.

$$J = \min_{\beta} \| H\beta - t \| \tag{2.4}$$

식 (2.5) 을 사용하여 출력 가중치 β를 계산한다. H^{+}는 H의 무어-페로스의 일반화된 역행렬이다.

$$\beta = H^{+} T \tag{2.5}$$

학습 과정에서 사용된 입력 가중치 w와 바이어스 값 b와 계산된 출력 가중치 β를 사용하여 식 (2.6 ~ 2.7)과 같이 검증 데이터 y_k의 출력을 계산한다.

$$s_k = \sum_{j=1}^{n} \beta_j \cdot G_j(y_k) = \sum_{j=1}^{n} \beta_j \cdot G_j(w_{ij} \cdot y_k + b_{ij}) \tag{2.6}$$

$$D = \begin{bmatrix} G(w_{11} \cdot y_1 + b_{11}) & \cdots & G(w_{ij} \cdot y_1 + b_{ij}) \\ \vdots & \cdots & \vdots \\ G(w_{1j} \cdot y_1 + b_{1j}) & \cdots & G(w_{ij} \cdot y_k + b_{ij}) \end{bmatrix} \tag{2.7}$$

s_k는 검증 데이터의 출력을 나타내며 각 클래스에 대한 확률 값이 생성되며, D는 검증 데이터의 은닉 층 출력을 나타낸다. 분류 오차 MCR의 초기 값을 0으로 설정하고 각 항의 검증 데이터 목표 출력 값 z_k와 검증 데이터 출력 s_k의 최댓값이 식(2.8)을 만족하지 못할 때 식 (2.9)가 반복되어 분류 오차 값은 1씩 증

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

가한다. 마지막으로 계산된 분류 오차와 식 (2.10)을 통해 분류 정확도를 계산 한다.

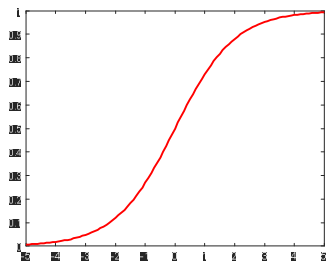
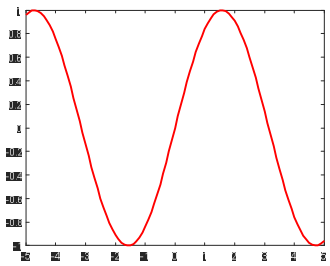
$$z_k = \max(s_k) \tag{2.8}$$

$$MCR = MCR + 1 \tag{2.9}$$

$$Classification\ \ accuracy = 1 - \frac{MCR}{k} \tag{2.10}$$

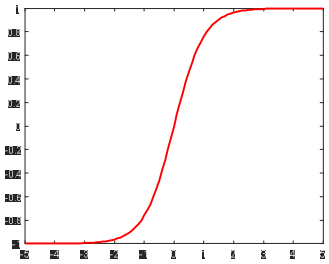
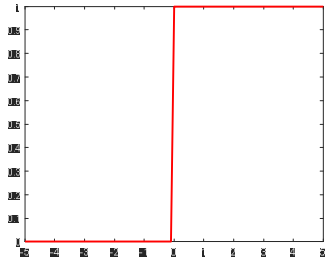
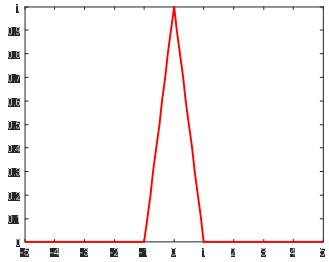
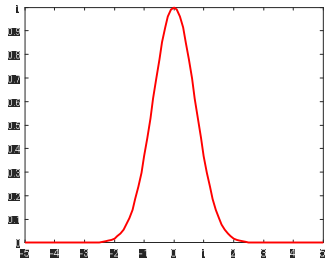
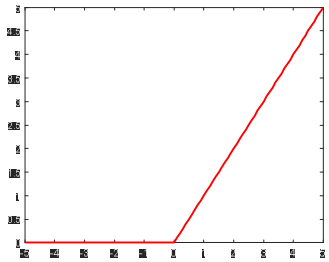
ELM의 은닉 층의 노드에서 사용되는 활성 함수 $G(x)$의 역할은 입력된 데이터와 가중치, 바이어스를 사용해 계산된 값이 일정 임계값을 넘었을 때 은닉 층의 출력 값을 조절하는 역할을 한다. 다양한 종류의 활성 함수가 존재 하며 사용자의 선택에 따라 결정된 활성 함수에 따라 학습 속도 및 정확도가 달라진다. 표 2-1은 ELM에서 사용가능한 활성 함수의 종류를 나타낸다.

표 2-1 활성 함수의 종류

Active function	Expression	
Sigmoid	$G(x) = \frac{1}{1+e^{-x}}$	
Sin	$G(x) = \sin(x)$	

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

Tanh	$G(x) = \tanh(x)$	
Hard-limit	$G(x) = \begin{cases} 0, & x < 0 \\ 1, & x > = 1 \end{cases}$	
Triangular basis	$G(x) = \begin{cases} 1 - abs(x), & \text{if } -1 \le n \le 1 \\ 0, & otherwise \end{cases}$	
Radial basis	$G(x) = e^{-x^2}$	
ReLU	$G(x) = \max(0, x)$	

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제3절 R-ELM

본 논문에서 사용한 R-ELM 분류기는 학습 과정에서 출력 가중치의 자유도를 제약하는 규제화 방법을 사용하여 학습 모델의 설명도를 유지하면서 복잡도를 줄이는 방법이다. 규제화 방법 중 많이 사용하는 방법으로 학습 모델에 패널티(penalty)를 주는 방법 중 하나인 Ridge 회귀 분석법이 사용된다. Ridge 회귀 분석법의 해를 유도하기 위해 식 (2.10)은 잔차 제곱식(RSS)를 나타낸다.

$$RSS = \sum_{i=1}^{n}(y_i - (\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad (2.10)$$

RSS를 최소화하는 β를 찾음으로서 회귀 분석의 계수 값을 구할 수 있다. 잔차 행렬식으로 나타기 위해 식(2.11 ~2. 12)와 같이 가정한다.

$$X = \begin{bmatrix} \underline{x_1^T} \\ \underline{x_2^T} \\ \vdots \\ \underline{x_n^T} \end{bmatrix} = \begin{bmatrix} (x_{11}\, x_{12} \cdots x_{1p}) \\ (x_{21}\, x_{22} \cdots x_{2p}) \\ \vdots \\ (x_{n1}\, x_{n2} \cdots x_{np}) \end{bmatrix} \qquad (2.11)$$

$$\underline{y} = \begin{bmatrix} y_1 - \overline{y} \\ y_2 - \overline{y} \\ \vdots \\ y_n - \overline{y} \end{bmatrix}, \underline{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \qquad (2.12)$$

β_0 의 값은 $\overline{y}$로 추정되며 목표 값 y는 $n \times 1$의 크기를 갖는 벡터이며, 찾고자 하는 해 β는 $p \times 1$ 의 크기를 갖는 벡터이다. 위에서 정의된 값을 이용해 식 (2.13)과 같은 β에 대한 손실함수(loss function) 형태로 나타낼 수 있다.

$$\begin{aligned} L(\beta) &= \frac{1}{2}\sum_{i=1}^{n}((y_i - \overline{y}) - (\beta_1 x_{i1} + \cdots + \beta_p x_{ip})^2 \\ &= \frac{1}{2}(y - X\beta)^T (y - X\beta) \end{aligned} \qquad (2.13)$$

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

손실함수의 값을 최소화 하는 벡터 β를 찾아 회귀분석의 계수로 사용하는 방법이 일반적인 회귀분석 방법이다. 여기서 Ridge 회귀 분석방법은 식 (2.14)과 같이 L2 놈 즉 유클리디안 놈(euclidean norm)을 사용하여 패널티를 준다.

$$\begin{aligned} L(\beta) &= \frac{1}{2}(y-X\beta)^T(y-X\beta)+\frac{\lambda}{2}\sum_{i=1}^{p}\beta_i^2 \\ &= \frac{1}{2}(y-X\beta)^T(y-X\beta)+\frac{\lambda}{2}\|\beta\|_2^2 \\ &= \frac{1}{2}(y-X\beta)^T(y-X\beta)+\frac{\lambda}{2}\beta^T\beta \end{aligned} \tag{2.14}$$

유클리디안 놈은 다차원 공간에서 벡터의 크기를 나타내는 한 가지 단위이며, $\|\beta\|_2=\sqrt{\beta_1^2+\beta_2^2+\cdots\beta_p^2}$와 같이 계산된다. 식 (2.15)는 패널티가 포함된 손실함수 식을 β에 대해 한번 미분된 식을 나타낸다.

$$\frac{\partial L(\beta)}{\partial\beta}=-X^T(y-X\beta)+\lambda\beta \tag{2.15}$$

위의 식의 값을 0으로 설정하고 대응되는 β 값을 찾으면 식 (2.16)과 같은 Ridge 회귀 분석법의 해를 찾을 수 있다. λ는 패널티를 조절 하는 역할을 하고 0보다 항상 큰 값을 가진다.

$$\begin{aligned} -X^T(y-X\beta)+\lambda\beta &= 0 \\ \Rightarrow (X^TX+\lambda I)\beta &= X^Ty \\ \Rightarrow \beta^{Ridge} &= (X^TX+\lambda I)^{-1}X^Ty \end{aligned} \tag{2.15}$$

R-ELM 분류기의 구조는 ELM 분류기의 구조와 같으며 학습 과정을 설명하기 위해 학습 데이터 $x_i=[x_1,x_2,x_3,\ldots,x_i]$, 학습 데이터 목표 출력 값 $y_i=[y_1,y_2,y_3,\ldots,y_i]$ 은닉 층 노드 j개, 활성 함수 $G(x)$가 주어질 때 출력은 식 (2.16)로 표현된다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

$$y_i = \sum_{j=1}^{n} \beta_j \cdot G_j(x_i) = \sum_{j=1}^{n} \beta_j \cdot G_j(w_{ij} \cdot x_i + b_j) \tag{2.16}$$

$G(x)$는 은닉 층의 출력, w는 입력 층과 은닉 층을 연결하는 입력 가중치, b는 은닉 층의 바이어스 항이며 입력가중치와 바이어스 항은 임의의 값을 할당한다. β는 출력 가중치를 나타내며 은닉 층 출력은 식 (2.17)와 같이 나타낼 수 있다.

$$\mathrm{H} = \begin{bmatrix} G(w_{11} \cdot x_1 + b_{11}) & \cdots & G(w_{ij} \cdot x_1 + b_{ij}) \\ \vdots & \cdots & \vdots \\ G(w_{1j} \cdot x_1 + b_{1j}) & \cdots & G(w_{ij} \cdot x_i + b_{ij}) \end{bmatrix} \tag{2.17}$$

최적의 성능을 얻을 수 있는 출력 가중치 β를 구하는 것이 일반적인 ELM의 학습 목표이다. R-ELM은 일반적인 ELM에서 사용하는 최소 자승법과 무어페로스의 일반화된 역행렬을 사용하지 않고 식 (2.18)의 Ridge 회귀 분석 방법을 적용하여 출력 가중치를 계산한다[44-45].

$$\beta^{ridge} = (HH^T + \lambda I)^{-1} H^T y \tag{2.18}$$

I는 1의 값을 갖는 대각 원소 행렬이며, λ는 임의의 값이 입력되며 β^{ridge}은 R-ELM 학습 과정의 최종 목표인 출력 가중치를 나타낸다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제3장 K-pop 포인트 댄스 분류 방법

본 논문에서 제안된 K-pop 포인트 댄스 분류 과정은 그림 3.1과 같다. Kinect를 통해 얻어진 3차원 관절 좌표 데이터를 사용하여 포인트 댄스의 특징 벡터를 생성하고 PCA와 LDA 알고리즘을 사용하여 차원 축소 과정을 통해 얻어진 데이터를 R-ELM 분류기의 입력 데이터로 사용해 분류한다.

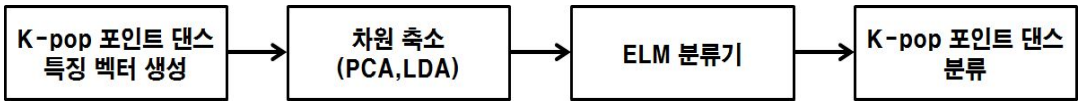


그림 3.1 K-pop 포인트 댄스 분류 과정

제1절 K-pop 포인트 댄스 특징 추출

3차원 관절 좌표를 사용하여 프레임마다 K-pop 포인트 댄스의 특징을 나타내는 13개(좌/우 팔꿈치 각도, 좌/우 어깨각도, 좌/우 무릎 각도, 좌/우 손 벌어짐, 좌/우 발 벌어짐, 양발 각도, 양손 각도, 몸의 기울기)각도를 계산한다. 포인트 댄스 데이터의 첫 프레임부터 10번째까지의 프레임을 제외하고 이후의 프레임에서 계산된 각도 값은 min-max 정규화 과정을 수행한다. 포인트 댄스 데이터는 곡에 따라 크기가 다르기 때문에 가장 긴 프레임을 기준으로 제로 패딩(zero padding) 기법을 사용하여 데이터 크기를 동일하게 한다. 30프레임마다 각각의 13개각도의 통계적 수치(최솟값, 평균, 최댓값, 분산)를 계산한다. 계산 된 값은 하나의 벡터로 이어 붙여 K-pop 포인트 댄스의 특징 벡터를 생성 한다.

1. K-pop 포인트 댄스 특징 각도 계산

몸의 기울기 각도를 제외하고 나머지 12개의 각도는 하나의 접을 가지는 사이각을 계산하게 된다. 각도를 구하는 방법을 설명하기 위해 임의의 3차원 좌표 A , B , C 가 그림 3.2와 같다고 가정하자. C를 접점으로 벡터 $\overrightarrow{AC}$, $\overrightarrow{BC}$의 사이각 θ는 식 (3.1 ~ 3.6)을 통해 계산된다.

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

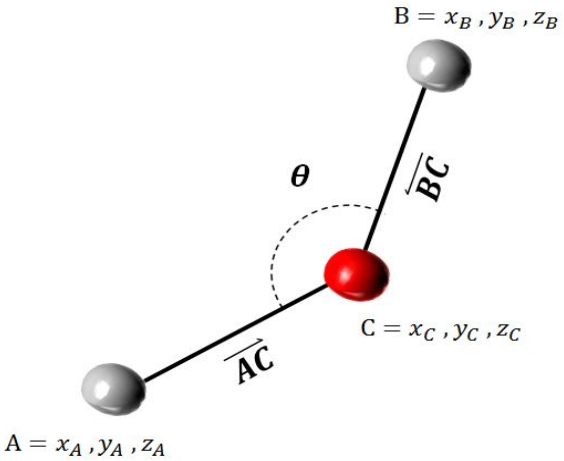


그림 3.2 3차원 좌표 A,B,C

$$A = x_A, y_A, z_A \tag{3.1}$$

$$B = x_B, y_B, z_B \tag{3.2}$$

$$C = x_C, y_C, z_C \tag{3.3}$$

$$\overrightarrow{AC} = (x_A - x_C, y_A - y_C, z_A - z_C) \tag{3.4}$$

$$\overrightarrow{BC} = (x_B - x_C, y_B - y_C, z_B - z_C) \tag{3.5}$$

$$\Theta = \cos^{-1} \frac{\overrightarrow{AC} \cdot \overrightarrow{BC}}{|\overrightarrow{AC}||\overrightarrow{BC}|} \times \frac{180}{\Pi} \tag{3.6}$$

표 3-1은 접점을 가지는 사이 각도를 구하기 위해 사용된 관절 좌표를 나타낸다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 3-1 접점을 가지는 사이 각 정보

사이 각도	A	B	C(접점)
좌 팔꿈치	좌 손목	좌 어깨	좌 팔꿈치
우 팔꿈치	우 손목	우 어깨	우 팔꿈치
좌 어깨	좌 팔꿈치	어깨 척추	좌 어깨
우 어깨	우 팔꿈치	어깨 척추	우 어깨
좌 무릎	좌 발목	좌 엉덩이	좌 무릎
우 무릎	우 발목	우 엉덩이	우 무릎
좌 손 벌어짐	좌 손	골반 척추	좌 어깨
우 손 벌어짐	우 손	골반 척추	우 어깨
좌 발 벌어짐	좌 무릎	골반 척추	좌 엉덩이
우 발 벌어짐	우 무릎	골반 척추	우 엉덩이
양발	좌 발	우 발	골반 척추
양손	좌 손	우 손	어깨 척추

몸의 기울기는 접점을 사용하지 않고 4개의 관절 좌표를 사용하여 계산하게 된다. 그림 3.3은 몸의 기울기를 구하기 위해 사용되는 관절 좌표를 나타낸다. 식 (3.7 ~ 3.13)과 같이 3차원 관절 좌표 A(머리), B(골반척추)를 지나는 벡터 $\overrightarrow{AB}$, 좌표 C(우 발)와 D(좌 발)를 지나는 $\overrightarrow{CD}$ 를 사용해 몸의 기울기각 Θ를 계산한다.

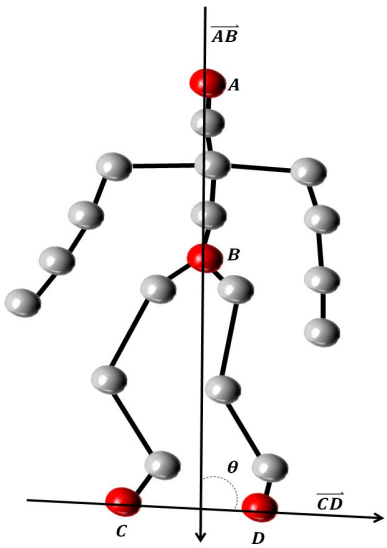


그림 3.3 몸의 기울기를 구하기 위한 좌표 A,B,C,D

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

$$A = x_A, y_A, z_A \tag{3.7}$$

$$B = x_B, y_B, z_B \tag{3.8}$$

$$C = x_C, y_C, z_C \tag{3.9}$$

$$D = x_D, y_D, z_D \tag{3.10}$$

$$\overrightarrow{AB} = (x_A - x_B, y_A - y_B, z_A - z_B) \tag{3.11}$$

$$\overrightarrow{CD} = (x_C - x_D, y_C - y_D, z_C - z_D) \tag{3.12}$$

$$\Theta = \cos^{-1} \frac{\overrightarrow{AB} . \overrightarrow{CD}}{|\overrightarrow{AB}||\overrightarrow{CD}|} \times \frac{180}{\Pi} \tag{3.13}$$

그림 3.4 ~ 3.5는 위의 과정을 통해 계산된 13개의 각도 중 좌 팔꿈치와 우 어깨의 각도 값을 min-max 정규화과정을 통해 계산된 값을 나타낸다.

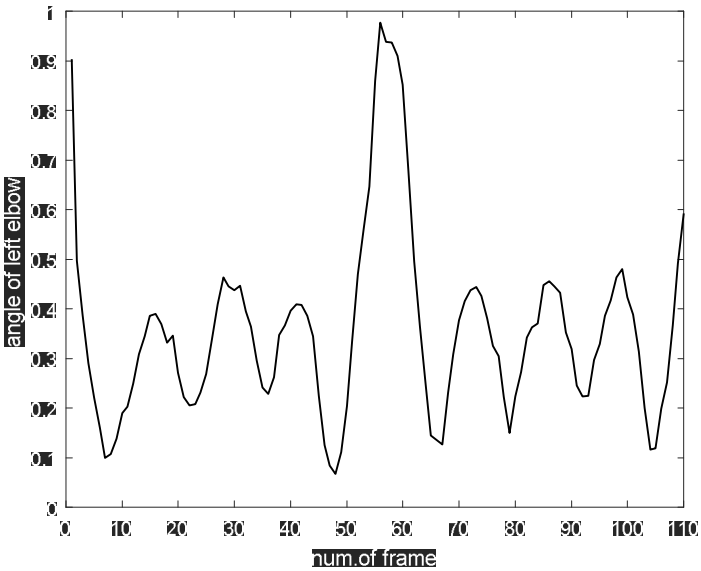


그림 3.4 좌 팔꿈치 사이 각도

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

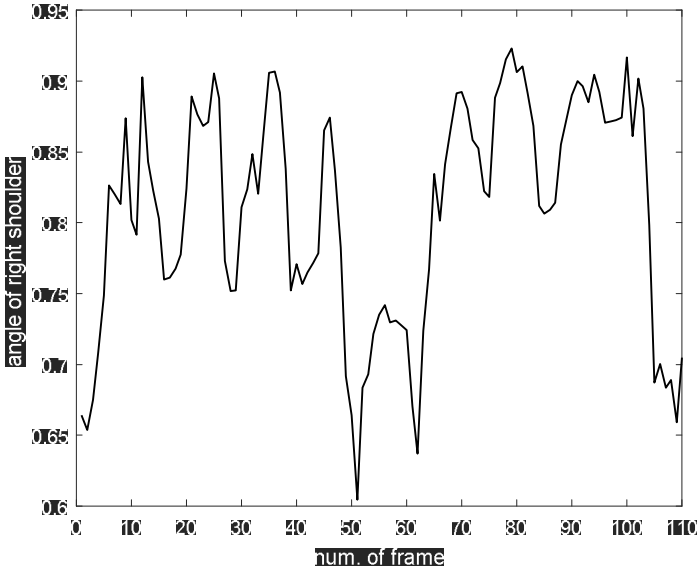


그림 3.5 우 어깨 사이 각도

첫 번째 부터 10번째까지의 프레임을 제외하고 각각의 프레임에서 얻어진 각도 값을 30프레임마다 통계적 수치(최솟값, 평균, 최댓값, 분산)를 계산한다. min-max 정규화 과정을 거치면 데이터 값은 0과 1사이의 값을 갖게 되며 K-pop 포인트 댄스는 각각의 곡에 따라 프레임 크기가 다르기 때문에 가장 긴 곡 276프레임을 기준으로 계산을 위해 280프레임으로 제로 패딩 기법을 사용해 0을 대입하여 데이터 크기를 정규화 한다. 계산된 13개의 통계적 각도 값을 하나의 벡터로 이어 붙여 특징 벡터를 생성 한다. 그림 3.6 ~ 3.7은 좌 팔꿈치와 우 어깨의 통계적 수치 값을 나타내며 그림 3.8은 각각의 각도에서 계산된 통계적 수치를 하나로 이어 붙인 K-pop 포인트 댄스의 특징 데이터를 나타낸다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

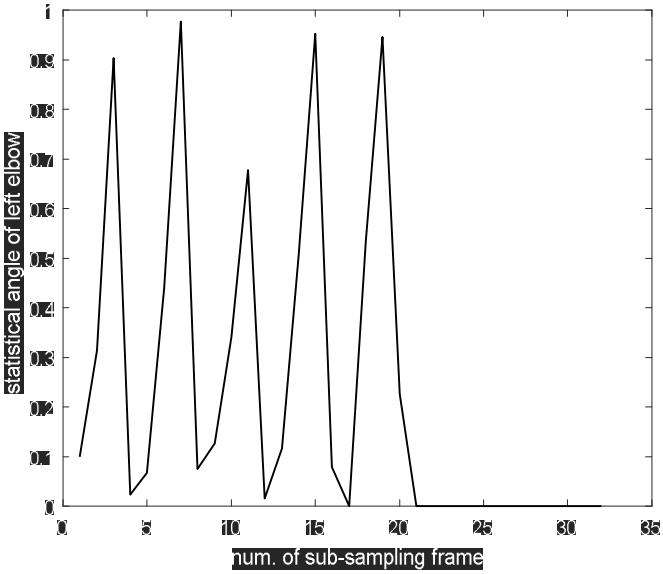


그림 3.6 좌 팔꿈치 통계적 각도

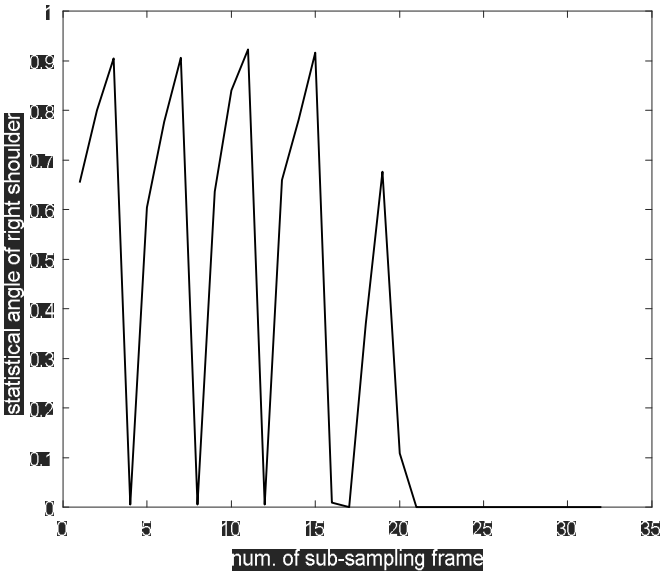


그림 3.7 우 어깨 통계적 각도

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

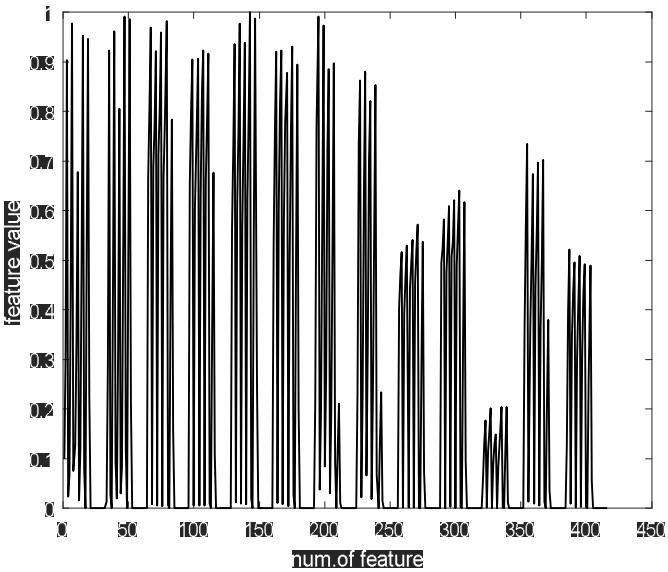


그림 3.8 K-pop 포인트 댄스 특징 데이터

제2절 PCA + LDA 차원 축소 알고리즘

K-pop 포인트 댄스는 동일한 댄스에 대해 같은 패턴의 양상을 보일 것이고 서로 다른 댄스에 대해서는 패턴이 다른 양상을 보일 것이다. 이와 같은 특성을 이용해 댄스 모션을 분류한다. 본 논문에서는 주성분 분석법(PCA)과 선형 판별 분석(LDA) 차원 축소 기법을 통해 포인트 댄스를 가장 잘 나타내는 특징 값을 추출한다.

가. 주성분 분석(PCA : Principal Component Analysis)

PCA란 고차원의 데이터를 저차원으로 환원시키는 방법이다. 데이터 각각에 대한 특성을 분석하는 것이 아니며 데이터가 모여 하나의 분포를 이룰 때 이 분포의 주성분을 분석하는 방법이다. 주성분이란 데이터들의 분산이 가장 큰 방향 벡터를 뜻한다. 그림 3.9는 임의의 3차원 데이터에서 3개의 주성분 벡터를 나타낸다. 주성분 벡터를 통해 데이터 분포의 형태를 알 수 있는데 빨간색 벡터의 방향에 따라 데이터의 분산이 가장 크고 빨간색 벡터에 수직이면서 그 다음으로 분산이 큰 방

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

향은 초록색 벡터로 알 수 있다. 그 다음의 분산이 큰 방향은 파란색 벡터이다.

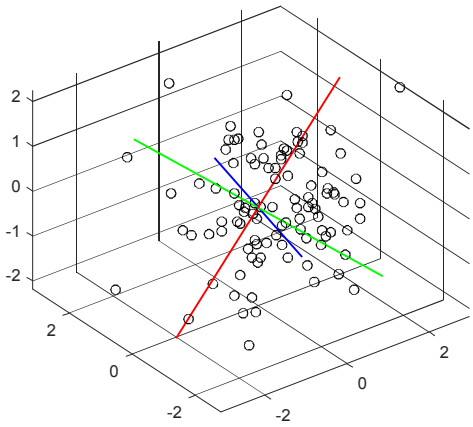


그림 3.9 임의의 3차원 데이터의 주성분 벡터

나. 선형 판별 분석(LDA : Linear Discriminant Analysis)

LDA는 각각의 클래스의 데이터가 가지는 분산과 전체 데이터가 가지는 분산을 고려하여 클래스를 나누고 데이터가 가지는 특성을 유지하여 저차원으로 축소하는 방법이다.

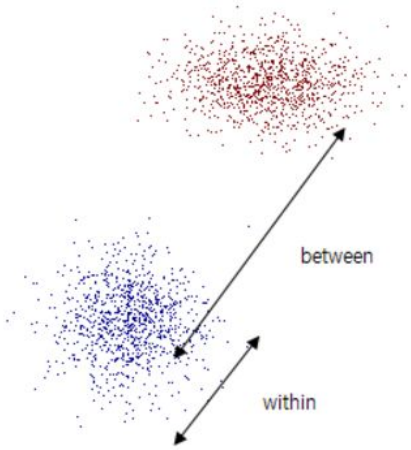


그림 3.10 클래스 내 분산과 클래스 간 분산

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

그림 3.10은 클래스 내 분산(within)과 클래스 간 분산(between)을 나타낸다. 그림에서 알 수 있듯이 클래스 내의 분산은 작을수록 클래스 간 분산은 클수록 판별이 용이해 진다. LDA는 클래스 내 분산을 나타내는 행렬(Within-Scatter Matrix)과 클래스 간 분산을 나타내는 행렬(Between-Scatter Matrix)의 비율이 최대가 되는 선형 변환법이다.

다. PCA + LDA

PCA를 사용하여 저차원으로 축소된 특징 벡터는 데이터의 특성을 잘 축약해서 표현하는데 유용하지만 클래스 간의 특징을 나타내는데 취약하다. 위의 특성을 이용해 PCA와 LDA와의 융합은 데이터의 특성을 잘 나타내면서 클래스 별 분리를 최대화 할 수 있는 장점이 있다. PCA와 LDA의 차원축소 과정은 다음과 같이 설명된다. 첫 번째 단계로 N개의 K-pop 포인트 댄스 학습 특징벡터 $z_N=(z_1,z_2,...,z_N)$와 S개의 검증 특징벡터 $y_s=(y_1,y_2,...,y_s)$가 주어질 때 학습 특징 벡터의 공분산 행렬은 식 (3.14 ~ 3.15)와 같이 정의된다.

$$\overline{z}=\frac{1}{N}\sum_{i=1}^{N}z_i \tag{3.14}$$

$$R=\frac{1}{N}\sum_{i=1}^{N}(z_i-\overline{z})(z_i-\overline{z})^T=\Phi\Phi^T \tag{3.15}$$

$\overline{z}$는 학습 특징벡터의 평균값을 나타내며 R은 공분산 행렬을 나타낸다. 식 (3.16)에서 공분산 행렬 R을 사용하여 고유 값(eigenvalue) λ_i, 고유 벡터(eigen vector)를 E_i 구한다. 이때 계산된 고유벡터는 K-pop 포인트 댄스의 고유 댄스(eigen dance)라고 할 수 있다. 고유 값의 크기가 큰 순으로 p개의 고유 벡터를 사용하여 식 (3.17)의 과정을 거치면 PCA를 사용해 p차원으로 축소된 학습 특징벡터 $X=(x_1,x_2,...,x_p)$가 계산된다.

$$R\ E_i=\lambda_i E_i(1\le i\le N) \tag{3.16}$$

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

$$X = E_i(z_i - \overline{z}) \ (1 \le i \le p) \tag{3.17}$$

p차원으로 투영된 포인트 댄스 학습 특징벡터 X를 LDA를 사용해 축소된 특징 벡터를 구한다. i번째 클래스의 학습 특징 벡터의 평균, i번째 클래스 내의 분산을 나타내는 행렬, 전체 클래스 내의 분산을 나타내는 행렬은 식 (3.18 ~ 3.20)과 같다.

$$m = \frac{1}{N} \sum_{i=1}^{n_i} X_i \tag{3.18}$$

$$S_i = \sum_{x \in X_i} (x - m_i)(x - m_i)^T \tag{3.19}$$

$$S_W = \sum_{i=1}^{c} S_i \tag{3.20}$$

n_i는 각 클래스의 학습 특징 벡터의 수를 나타낼 때 전체 특징 벡터 N개의 평균은 식 (3.21)과 같이 계산되고 식 (3.22)는 클래스 간 분산을 나타내는 행렬을 나타낸다.

$$\mu_j = \frac{1}{N} \sum_{i=1}^{N_i} C_i \tag{3.21}$$

$$S_B = \sum_{j=1}^{c} (\mu_j - \overline{\mu})(\mu_i - \mu)^T \tag{3.22}$$

식 (3.23 ~ 3.24)은 클래스 내의 분산을 나타내는 행렬과 클래스 간 분산을 나타내는 행렬의 비율이 최대가 되는 행렬을 찾아 최적화된 판별 벡터(discriminant vector)를 W를 구한다.

$$W = \arg\max \frac{|W^T S_B W|}{|W^T S_W W|} = (w_1, w_2, \ldots, w_m) \tag{3.23}$$

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

$$S_B w_i = \lambda_i S_W w_i \ , \ \ i = 1, 2, \ldots m \tag{3.24}$$

m(클래수 수-1)개의 0이 아닌 판별 값과 판별 벡터가 존재한다. 판별 벡터 W와 고유벡터 E를 사용해 식 (3.25)를 계산하면 K-pop 포인트 댄스의 피셔 댄스(fisher dance) F가 된다.

$$F_m = W_i E \ (1 \le i \le m) \tag{3.25}$$

마지막으로 식 (3.26 ~ 3.27)은 피셔댄스를 사용해 m개의 차원으로 축소된 K-pop 포인트 댄스의 학습 및 검증 최종 특징 벡터 F_{train}과 F_{test} 가 생성된다.

$$F_{train} = zF \tag{3.26}$$

$$F_{test} = yF \tag{3.27}$$

제3절 K-pop 포인트 댄스 분류

K-pop 포인트 댄스 분류를 위해 차원 축소 과정을 통해 생성된 특징 벡터는 R-ELM 분류기를 사용하여 분류된다. 전문 댄서의 3차원 관절 취득된 데이터로부터 생성된 $F_{train} = [F_{train_1} F_{train_2}, \ldots, F_{train_i}]$는 R-ELM 분류기의 학습 데이터가 되고 아마추어 댄서의 데이터로부터 생성된 $F_{test} = [F_{test_1} F_{test_2}, \ldots, F_{test_i}]$는 R-ELM 분류기의 검증 데이터로 사용된다. 입력 가중치 w, 입력 바이어스(Bias) b는 임의의 값이 할당되고 학습 데이터 출력 값 $t_{train_i} = [t_{train_1}, t_{train_2}, \ldots, t_{train_i}]$를 가지는 학습데이터 F_{train}이 입력 될 때 R-ELM 분류기의 학습 과정은 식 (3.28 ~ 3.31)을 통해 출력 가중치 β를 계산한다.

$$t_{train} = \sum_{j=1}^{n} \beta_j \cdot G_j(F_{train_i}) = \sum_{j=1}^{n} \beta_j \cdot G_j(w_{ij} \cdot F_{train_i} + b_j) \tag{3.28}$$

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

$$\mathrm{H}=\begin{bmatrix} G(w_{11}\cdot F_{train_1}+b_{11}) & \cdots & G(w_{ij}\cdot F_{train_1}+b_{ij}) \\ \vdots & \cdots & \vdots \\ G(w_{1j}\cdot F_{train_1}+b_{1j}) & \cdots & G(w_{ij}\cdot F_{train_i}+b_{ij}) \end{bmatrix} \tag{3.29}$$

$$\mathrm{H}\beta=t \tag{3.30}$$

$$\beta=(HH^t+\lambda I)^{-1}H^t t \tag{3.31}$$

계산된 출력 가중치와 학습 과정에서 사용되었던 입력 가중치, 바이어스를 사용하여 식 (3.32 ~ 3.33)의 과정을 통해 검증 데이터 F_{test}의 출력을 계산한다.

$$t_{test}=\sum_{j=1}^{n}\beta_j\cdot G_j(F_{test_i})=\sum_{j=1}^{n}\beta_j\cdot G_j(w_{ij}\cdot F_{test_i}+b_j) \tag{3.32}$$

$$D=\begin{bmatrix} G(w_{11}\cdot F_{test_i}+b_{11}) & \cdots & G(w_{ij}\cdot F_{test_i}+b_{ij}) \\ \vdots & \cdots & \vdots \\ G(w_{1j}\cdot F_{test_i}+b_{1j}) & \cdots & G(w_{ij}\cdot F_{test_i}+b_{ij}) \end{bmatrix} \tag{3.33}$$

R-ELM 분류기의 출력은 확률 값으로 계산되어지기 때문에 식 (3.34)은 가장 높은 값을 가지는 클래스를 검증 데이터의 클래스를 나타내는 $t_{test_i}=[t_{test_1},t_{test_2},\ldots,t_{test_i}]$와 비교하여 같지 않을 경우 식 (3.35)과 같이 MCR은 1씩 증가하게 된다. 이 과정을 반복하여 식 (3.36)의 분류 정확도를 계산한다.

$$z_k=\max(s_k) \tag{3.34}$$

$$MCR=MCR+1 \tag{3.35}$$

$$Classification\ \ accuracy=1-\frac{MCR}{k} \tag{3.36}$$

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

R-ELM 분류기의 성능은 일반적인 ELM과 같이 사용자에 의해 결정되는 은닉 층의 노드 개수, 활성 함수(Active function)에 따라 달라진다. 은닉 노드의 수는 입력된 학습 데이터의 수를 기준으로 설정 하고, Sigmoid, Sin, Hard-limit, Tribas, Radbas, ReLU 등 다양한 활성함수를 사용한다. 그림 3.11은 K-pop 포인트 댄스의 특징 추출, 차원 축소, R-ELM 분류기를 사용한 전체 시스템을 나타낸다.

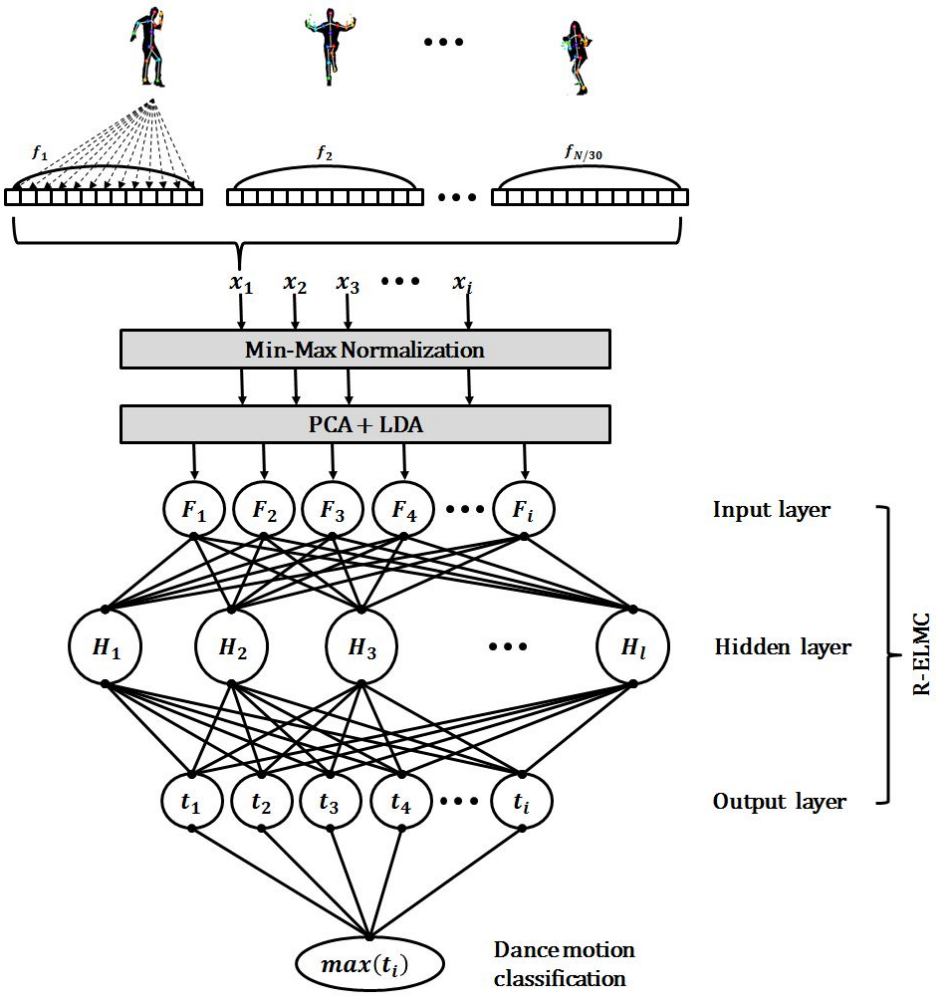


그림 3.11 PCA+LDA 과 R-ELM 분류기를 이용한 K-pop 포인트 댄스 분류 시스템

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

제4장 실험 및 결과

제1절 K-POP 포인트 댄스 DB 구축

1. Microsoft Kinect

Microsoft사에서 개발된 Kinect는 카메라를 기반으로 하는 가장 대표적인 모션 인식 하드웨어이다. Kinect는 적외선 카메라 센서를 통해 사용자의 모션을 스캔하고 거리를 계산하여 실시간으로 사용자의 모션을 인식한다. 2012년 출시된 키넥트 V1은 60일 만에 800만대가 팔리며 가장 빠르게 확산된 가전제품으로 평가받고 있으며 세계의 주목을 받았다. 2014년 Kinect V1의 구조광 구조를 TOF(Time of Flight) 방식으로 변환하여 Kinect V2를 출시했다[53]. 표 4-1은 Kinect V2는 Kinect V1과 성능 비교를 나타낸다.

표 4-1 Kinect 성능 비교

	Kinect V1	Kinect V2
Range of Depth	0.8 ~ 4.0m	0.5 ~ 8.0m
Color Resolution	640 * 480	1920 * 1080
Depth Resolution	320 * 240	512 * 424
Color fps	30 fps	30 fps
Color fps	30 fps	30 fps
Angle Depth	43 ~ 57	60 ~ 70
Joint	20	25

Kinect V2는 그림 4.1과 같이 사용자의 골격의 위치를 나타내는 25개의 3차원 관절(Joint) 좌표 데이터를 제공한다. 25개의 관절 정보 중 신체의 중심인 골반 척추를 먼저 검출한 후 상체는 중간 척추, 어깨 척추 순으로 관절 좌표를 검출하며 하체는 좌/우 엉덩이, 좌/우 무릎 순으로 관절 좌표를 검출한다. 골반 척추를 기준으로 멀어지는 관절 좌표 정보의 경우 정확도가 감소하며 지면과 가까운 좌/우 발의 관절 좌표 정보의 경우 지면과 구분이 어려워 정확도가 매우 감소한다

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

[54]. 표 4-2는 25개의 관절 번호와 정보를 나타낸다.

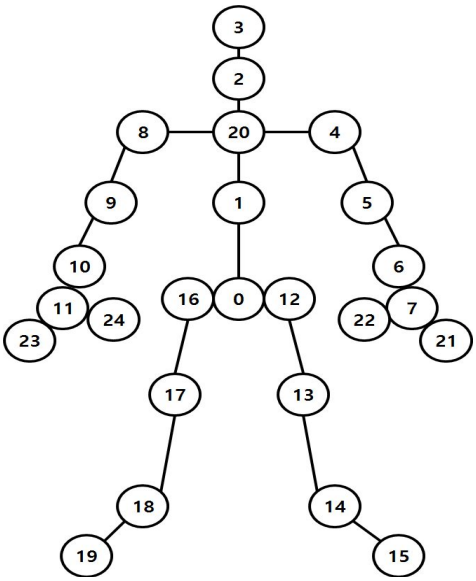


그림 4.1 키넥트 V2 Joint

표 4-2 키넥트 V2 Joint 정보

Value	Description	Value	Description
14 / 18	좌 /우 발목	13 / 17	좌 / 우 무릎
5 / 9	좌 / 우 팔꿈치	4 / 8	좌 / 우 어깨
15 / 19	좌 / 우 발	0	골반 척추
7 / 11	좌 / 우 손	1	중간 척추
6 / 10	좌 / 우 손목	20	어깨 척추
3 / 2	머리 / 목	22 / 24	좌 / 우 엄지
12 / 16	좌 / 우 엉덩이	21 / 23	좌 / 우 손끝

2. K-POP 포인트 안무 데이터 구축 환경

K-POP 포인트 안무 데이터는 4명의 전문 댄서, 40명의 아마추어 댄서들에게 취득한다. 안무 연습실에서 촬영 하며 Kinect V2와 댄서와의 거리는 약 2.0 ~ 3.8m 의 거리를 두고 취득했고 댄서의 활동 반경은 Kinect V2의 화각을 고려해 그

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

림 4.2와 같이 가로 약 2.7m 세로 약 1.8m의 공간에서 취득한다.

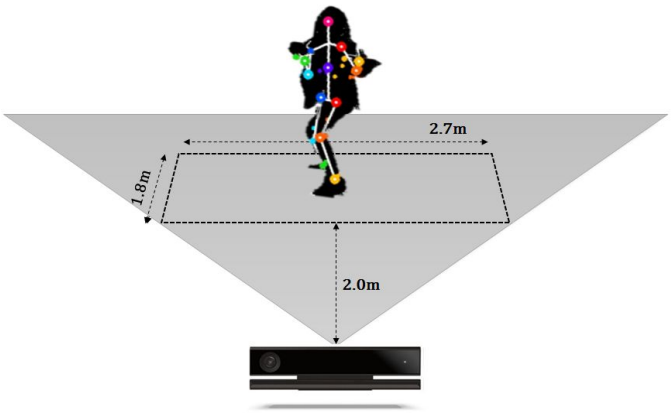


그림 4.2 데이터 구축 환경

전문 댄서는 경력 2년 이상의 전문가로 27세의 여성 전문 댄서 2명과 29세, 35세의 남성 전문 댄서 2명으로 구성되었으며 아마추어 댄서는 여자 중학생 6명, 여자 고등학생 19명, 여자 대학생 4명, 남자 고등학생 7명, 남자 대학생 4명으로 여성 아마추어 댄서 29명, 남성 아마추어 댄서 11명으로 구성되었다. 아마추어 댄서는 1년에서 2년 사이의 경력을 갖고 있다.

3. K-pop 포인트 댄스 정보

2002년에서 2015년 사이에 발매된 K-pop 중 100곡을 선정하여 곡마다 2개의 포인트 안무, 총 200개의 포인트 안무를 추출하여 데이터를 구축 한다. 표 4-3은 선정된 100곡의 정보, 그림 4.3 구축된 포인트 댄스 모션 데이터가 가장 잘 표현된 실루엣을 나타내며 그림 4.4는 하나의 포인트 댄스 실루엣의 프레임 별 변화도를 나타낸다. 전문 댄서는 개인별 200곡의 포인트 댄스를 촬영하여 총 800개의 데이터를 구축했다. 아마추어 댄서는 하나의 포인트 안무에서 2개의 데이터를 구축하기 위해 개인별 10개의 임의의 포인트 댄스를 촬영하여 총 400개의 데이터를 구축한다. 선정된 포인트 댄스는 4~9초의 시간을 가지며 구축된 데이터 중 가장 짧은 프레임 크기를 가지는 데이터는 147프레임, 가장 긴 프레임 크기를 가지는 데이터는 276 프레임의 크기를 가진다.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 4-3 K-POP 곡 정보

No	곡명	가수	발매 일자
1	까탈레나	오렌지카라멜	2014.03.12
2	상하이로맨스	오렌지카라멜	2011.10.13
3	밀크쉐이크	오렌지카라멜	2010.11.18
4	아잉	오렌지카라멜	2010.11.18
5	립스틱	오렌지카라멜	2012.09.12
6	마법소녀	오렌지카라멜	2010.06.17
7	No.1	보아	2002.04.12
8	Only one	보아	2012.07.22
9	Shadow	보아	2012.07.22
10	왜	동방신기	2011.01.03
11	주문	동방신기	2008.09.25
12	My name	보아	2004.06.11
13	문을여시오	임창정	2013.09.24
14	헐	블락비	2014.07.24
15	중독	EXO	2014.05.07
16	Darling	걸스데이	2014.07.14
17	마리오네트	스텔라	2014.02.12
18	단발머리	AOA	2014.06.19
19	Lucifer	샤이니	2010.07.19
20	누난너무예뻐	샤이니	2008.05.22
21	링딩동	샤이니	2009.10.14
22	Delicious	투하트	2014.03.10
23	산소같은너	샤이니	2008.08.28
24	Tell me why	투하트	2014.03.10
25	나혼자	씨스타	2012.04.12

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

No	곡명	가수	발매 일자
26	Give it to me	씨스타	2013.06.11
27	I swear	씨스타	2014.08.26
28	Push Push	씨스타	2010.06.03
29	Spark	보아	2004.06.11
30	Amigo	샤이니	2008.10.29
31	가식걸	씨스타	2011.08.25
32	So cool	씨스타	2011.08.09
33	Loving you	씨스타	2012.06.28
34	니가필요해	베스티	2014.08.29
35	토요일밤에	손담비	2009.03.24
36	퀸	손담비	2010.07.08
37	눈물이주르륵	손담비	2012.11.12
38	DB Rider	손담비	2010.09.16
39	니까짓게	씨스타	2010.12.03
40	Touch my body	씨스타	2014.07.21
41	Hello	샤이니	2012.02.01
42	Diamond	다나	2008.05.19
43	Shadow	비스트	2013.07.19
44	바람의유령	제국의아이들	2013.08.09
45	헤어지던날	제국의아이들	2013.03.25
46	숨	비스트	2010.09.28
47	Fiction	비스트	2011.05.17
48	Beatiful	비스트	2010.11.09
49	카페인	양요섭	2012.11.26
50	Flower	용준형	2013.12.13

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

No	곡명	가수	발매 일자
51	So good	박재범	2014.03.18
52	Like this	원더걸스	2012.06.03
53	강남스타일	싸이	2012.07.15
54	좋아	박재범	2013.04.10
55	별	박재범	2011.02.07
56	Abandoned	박재범	2011.04.27
57	아름다운밤이야	비스트	2012.07.22
58	Know your name	박재범	2012.02.07
59	Super magic	슈프림팀	2009.07.14
60	Step up	슈프림팀	2010.03.18
61	Gee	소녀시대	2009.01.05
62	소원을말해봐	소녀시대	2009.06.25
63	Shock	비스트	2010.03.01
64	빠빠빠	크래용팝	2013.06.20
65	Mr. Mr.	소녀시대	2014.02.24
66	으르렁	EXO	2013.08.05
67	Kissing you	소녀시대	2007.11.01
68	Oh!	소녀시대	2010.01.28
69	Good luck	비스트	2014.06.16
70	Goodboy	GD&태양	2014.11.21
71	No.9	티아라	2013.12.14
72	Nobody	원더걸스	2008.09.22
73	땡땡땡	슈프림팀	2010.06.03
74	괴도	태민	2014.08.18
75	LUV	에이핑크	2014.11.24

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

No	곡명	가수	발매 일자
76	맘마이아	카라	2014.08.18
77	위아래	EXID	2014.08.27
78	Sexy love	티아라	2012.09.04
79	Tell me	원더걸스	2008.06.03
80	Bad girls	이효리	2013.05.21
81	Baaam	다이나믹듀오	2013.07.01
82	Fantastic baby	빅뱅	2012.02.29
83	크레용	GD	2012.09.15
84	Hip song	비	2010.04.07
85	LA song	비	2014.01.02
86	꺼져줄게잘살아	지나	2011.01.08
87	Where u at	태양	2015.03.30
88	링가링가	태양	2013.11.08
89	웨딩드레스	태양	2009.11.13
90	내꺼하자	인피니트	2011.07.21
91	이름이뭐에요?	포미닛	2013.04.26
92	여자대통령	걸스데이	2013.06.24
93	사뿐사뿐	AOA	2014.11.11
94	Something	걸스데이	2014.01.03
95	fire	2ne1	2009.06.09
96	So hot	원더걸스	2008.06.03
97	기대해	걸스데이	2013.03.14
98	Twinkle	태티서	2012.04.30
99	전원일기	티아라	2013.04.29
100	첫사랑니	F(x)	2013.07.24

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

그림 4.3 포인트 댄스 촬영 영상

그림 4.4 포인트 댄스 실루엣

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

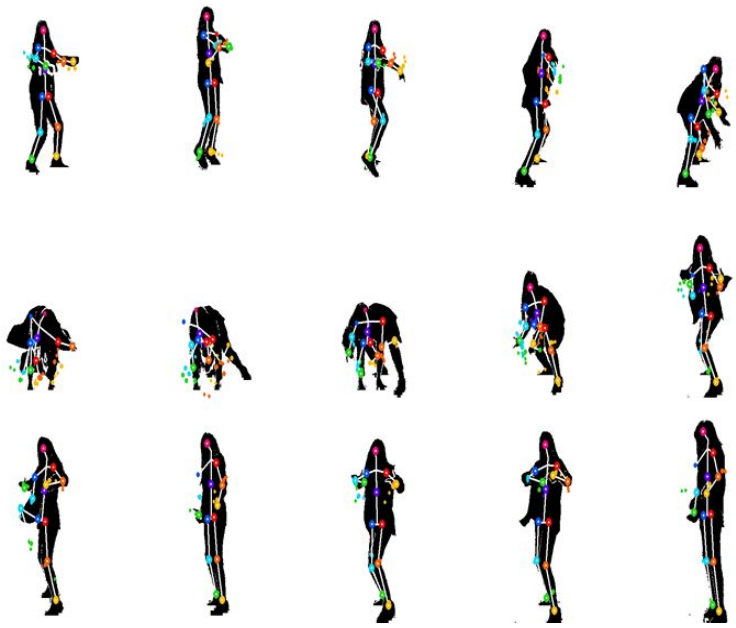


그림 4.5 포인트 댄스 프레임 별 변화도

제2절 실험 결과

4장 1절에서 설명된 K-pop 포인트 댄스 모션 데이터의 3차원 관절 좌표를 사용하여 3장의 과정을 거쳐 특징 추출, 차원 축소 과정을 거쳐 R-ELM 분류기를 사용해 포인트 댄스 모션을 분류한다. 표 4-4는 실험에 사용된 컴퓨터 사양을 나타낸다.

표 4-4 컴퓨터 사양

HW	CPU	Intel(R) Core(TM) i7-4790 CPU
	GPU	NVDIA GeForce GTX 1050
	RAM	8GB
SW	OS	Windows 10
	Program Language	Matlab R2016A

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

실험에서 사용된 K-pop 포인트 댄스 모션 3차원 관절 데이터를 사용하여 각 프레임마다 계산된 13개의 각도 데이터는 min-max 정규화를 사용해 계산된다. 포인트 댄스 데이터는 곡에 따라 크기가 다르다. 가장 긴 프레임의 크기는 276프레임이며 계산을 위해 280프레임을 기준으로 제로 패딩(zero padding) 기법을 이용해 데이터 크기를 정규화 한다. 하나의 각도의 크기는 1×280 이며 첫 프레임부터 10번째까지의 프레임을 제외하고 이후의 프레임에서 30프레임마다 계산된 통계적 수치를 계산하면 1×32의 크기를 가진다. 계산된 13개의 통계적 각도를 하나의 벡터로 이어 붙여 K-pop 포인트 댄스의 특징 벡터로 사용되며 벡터의 크기는 1×416 이다. 계산된 특징 벡터는 PCA와 LDA의 차원 축소 과정을 거쳐 생성된 피셔 댄스가 ELM 분류기의 입력 데이터로 사용되는데 이때 PCA의 고유벡터 수는 계산된 고유 값의 전체 합의 95%가 되는 지점에 대응되는 고유벡터 100개를 사용한다. 그림 4.5는 K-pop 포인트 댄스 모션 데이터에 PCA를 수행했을 때 계산된 고유 값과 고유 댄스를 나타낸다.

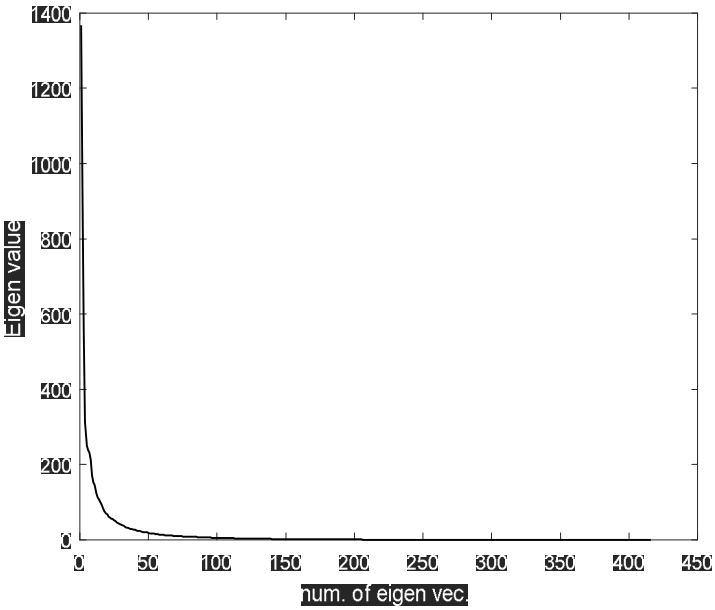


그림 4.6 포인트 댄스 데이터의 PCA 고유 값

LDA의 판별 벡터 수는 유클리디안 거리(euclidean distance)를 사용한 분류 정확도 따라 결정한다. 그림 4.6은 고유 벡터 수에 따라 축소된 데이터의 크기를 기준으로 100개의 90%인 90개까지 판별 벡터 수를 10개부터 10개씩 증가시켜 계산

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

된 분류 정확도이며 판별 벡터 수가 50, 60개 일 때 가장 높은 분류 정확도를 보이며 계산 속도를 고려하여 본 논문에서는 50개의 판별 벡터를 사용 했다. 그림 4.7은 K-pop 포인트 댄스의 피셔 댄스를 나타낸다.

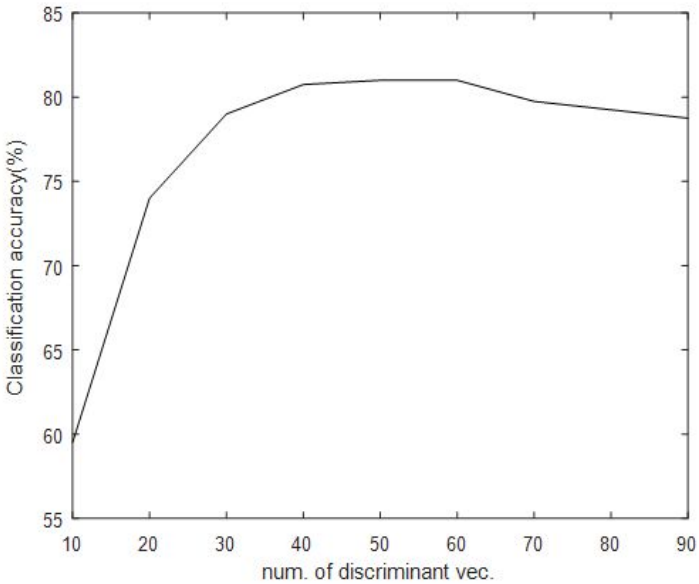


그림 4.7 판별 벡터 수에 따른 유클리디안 거리 분류 정확도

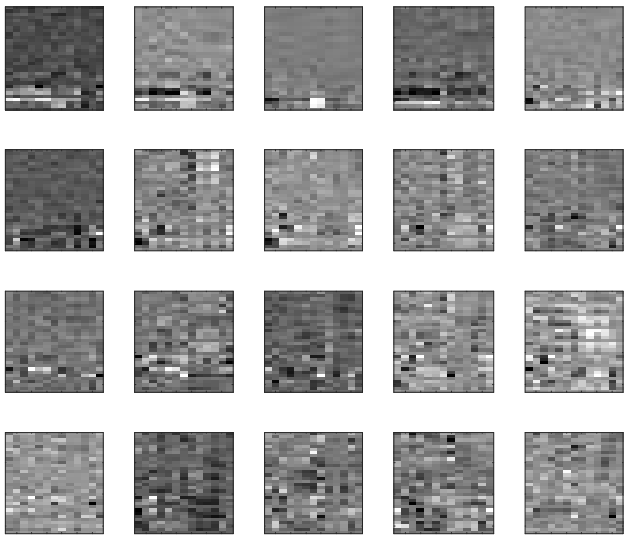


그림 4.8 K-pop 포인트 댄스의 피셔 댄스

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

최종적으로 PCA + LDA 차원 축소 과정을 수행하면 하나의 K-pop 포인트 댄스의 특징 벡터는 1×50의 크기를 갖는다. 4명의 전문 댄서 개인별 200개의 포인트 댄스를 추어 획득된 학습 데이터의 크기는 800×50의 크기를 가지며 40명의 아마추어 댄스 개인별 임의로 선정된 10개의 포인트 댄스를 추어 획득된 검증 데이터의 크기는 400×50의 크기를 가진다. 사용자에 의해 선택되는 R-ELM 분류기의 파라미터는 은닉 노드의 수와 활성 함수의 종류가 있다. 본 논문에서는 은닉 노드의 수를 10개부터 300개 까지 10개씩 증가 시키며 실험했다. R-ELM의 입력 가중치와 입력 바이어스 항은 임의의 값이 할당되어 매 계산과정에서 같은 파라미터를 사용해도 결과가 달라지기 때문에 100번의 계산과정을 같은 은닉 노드수와 활성 함수 종류를 사용 했을 때 가장 좋은 분류 성능을 표시했다. 실험에서 사용된 ELM과 R-ELM의 활성함수는 Sigmoid, Sine, Hard-limit, Tribas, Radbas 활성함수 총 5가지를 사용해 실험했다. R-ELM에서 가중치의 패널티를 결정하는 λ는 0.01의 값을 입력하여 사용했다. 제안된 방법의 우수함을 증명하기 위해 기존에 널리 알려진 기계학습 알고리즘인 k-NN(k-nearest neighbors), SVM(Support Vector Machine)과 일반적인 ELM 분류기를 사용하고 차원 축소 과정을 제외한 방법, PCA만을 사용한 방법을 PCA + LDA 방법과 비교했다.

표 4-5는 차원 축소를 하지 않고 K-pop 포인트 댄스의 통계적 각도 특징만을 사용했을 때의 분류 결과, 표 4-6은 통계적 각도 특징을 PCA를 사용해 차원 축소된 특징을 사용했을 때의 분류 결과이며, 표 4-7은 PCA+LDA를 사용해 차원 축소된 특징을 사용했을 때의 분류 결과이다.

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 4-5 통계적 각도 특징을 사용한 분류 성능

분류기	분류 성능(%)
k-NN	78.75%
SVM	82.75%
ELM_sig	76.00%
ELM_sin	79.00%
ELM_hardlim	51.25%
ELM_tribas	43.25%
ELM_radbas	58.00%
R-ELM_sig	78.75%
R-ELM_sin	79.25%
R-ELM_hardlim	52.25%
R-ELM_tribas	46.00%
R-ELM_radbas	61.50%

표 4-6 PCA를 사용해 차원 축소된 특징을 사용한 분류 성능

분류기	분류 성능(%)
k-NN	76.75%
SVM	82.75%
ELM_sig	83.00%
ELM_sin	79.75%
ELM_hardlim	73.75%
ELM_tribas	65.50%
ELM_radbas	76.00%
R-ELM_sig	83.75%
R-ELM_sin	80.75%
R-ELM_hardlim	74.00%
R-ELM_tribas	66.25%
R-ELM_radbas	77.25%

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

표 4-7 PCA+LDA를 통해 차원 축소된 특징을 사용한 분류 성능

분류기	분류 성능(%)
k-NN	81.00%
SVM	82.00%
ELM_sig	84.50%
ELM_sin	83.50%
ELM_hardlim	77.75%
ELM_tribas	82.00%
ELM_radbas	84.00%
R-ELM_sig	84.50%
R-ELM_sin	85.00%
R-ELM_hardlim	78.75%
R-ELM_tribas	82.00%
R-ELM_radbas	84.25%

그림 4.8은 PCA+LDA를 통해 차원 축소된 특징 벡터를 사용해 각각의 분류기에 대한 계산 속도를 나타낸다.

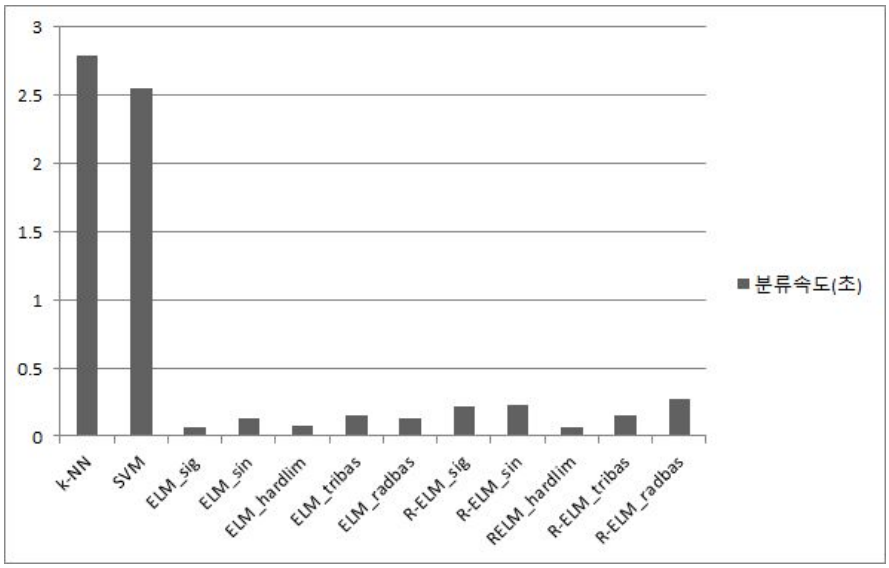


그림 4.9 PCA+LDA 특징 벡터 사용 분류기 분류 속도

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

분류 결과 및 분류 속도 결과를 종합해보면 차원축소 알고리즘을 사용하기 전에는 SVM이 더 좋은 성능을 보이지만 PCA와 PCA+LDA를 사용하여 차원 축소 과정을 거치면 R-ELM 분류기를 사용했을 때 다른 분류기의 성능 보다 더 좋은 분류 성능을 보임을 알 수 있다. 그림 4.9 ~ 4.11 은 PCA+LDA 차원축소 과정을 거친 특징벡터를 사용했을 때 높은 분류 성능을 보인 sigmoid, sin, radbas 활성 함수 기반의 R-ELM의 은닉 노드 수 변화에 따른 분류 성능 결과를 나타낸다.

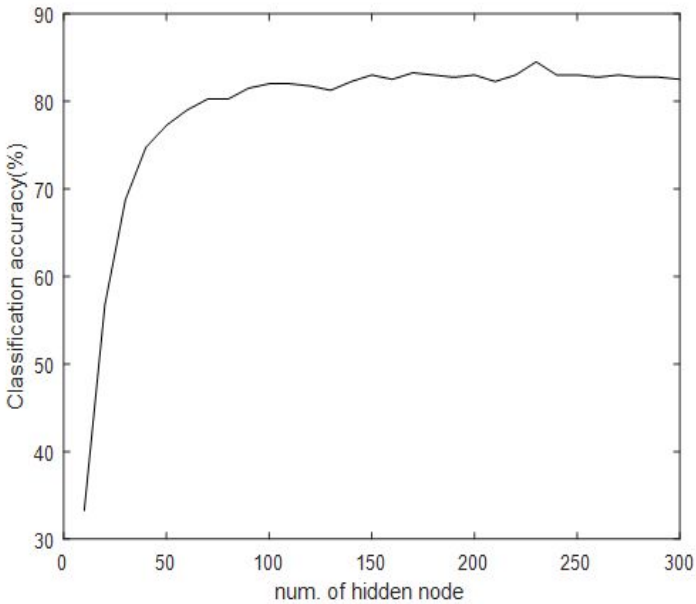


그림 4.10 은닉 노드 수 변화에 따른 sigmoid 함수 기반의 R-ELM 분류 성능

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

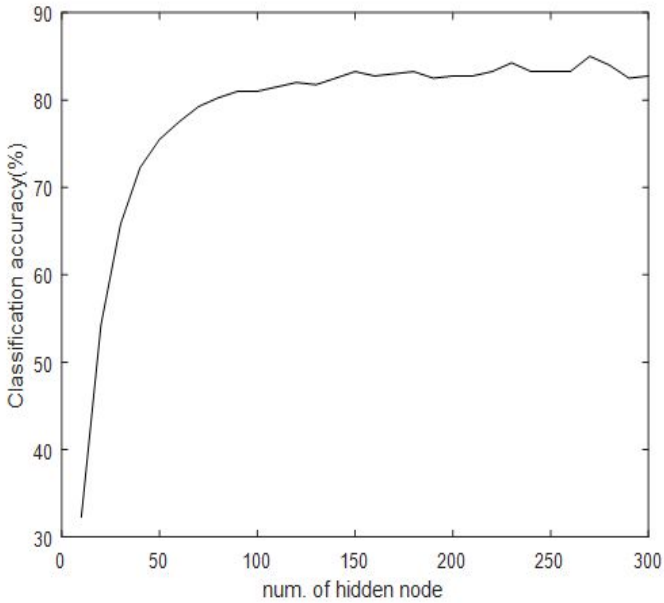


그림 4.11 은닉 노드 수 변화에 따른 sin 함수 기반의 R-ELM 분류 성능

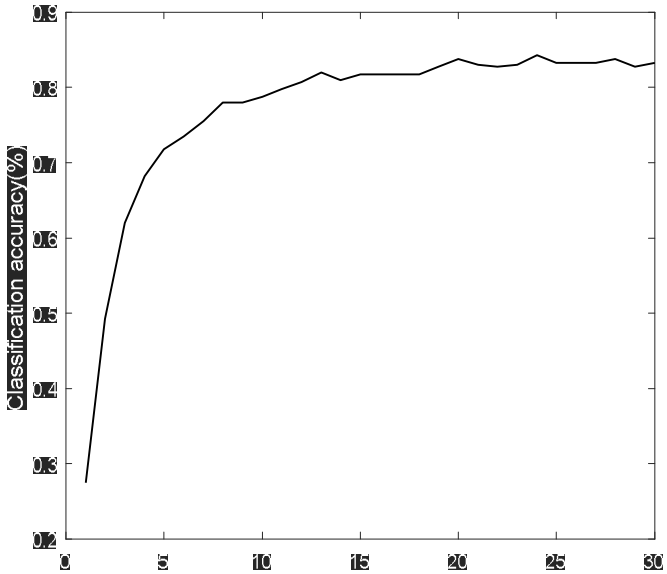


그림 4.12 은닉 노드 수 변화에 따른 radbas 함수 기반의 R-ELM 분류 성능

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

제5장 결론

본 논문에서는 안무 저작권 보호를 위한 연구로 K-pop 포인트 댄스 모션 분류 방법을 제안 했다. Kinect V2를 통해 구축된 포인트 댄스의 특징을 잘 나타내는 13개의 각도를 계산하여 PCA와 LDA 차원 축소 알고리즘을 사용해 포인트 댄스의 특성을 잘 나타내는 특징 벡터를 생성하고 다양한 활성함수 기반의 R-ELM을 사용해 분류하였다.

K-pop 포인트 댄스 모션의 3차원 관절 좌표를 사용해 매 프레임마다 좌/우 팔꿈치 각도, 좌/우 어깨각도, 좌/우 무릎 각도, 좌/우 손 벌어짐, 좌/우 발 벌어짐, 양발 각도, 양손 각도, 몸의 기울기 총 13가지의 각도를 계산했다. 포인트 댄스의 길이는 4~9초로 모두 다르기 때문에 제로 패딩 기법을 사용해 모든 곡의 프레임을 같게 하였다. 각도 값을 min-max 정규화 방법을 통해 0과 1 사이 값으로 정규화 하고 최솟값, 평균, 최댓값, 분산 총 4가지 통계적 수치를 30프레임마다 계산 후 13개의 통계적 각도를 하나로 이어 붙여 특징 벡터로 사용했다. 생성된 특징 벡터는 PCA와 LDA를 사용하여 저차원으로 축소시켰다. PCA는 데이터의 특성을 잘 축약해서 표현하는데 유용하며 클래스 간의 특징을 나타내는데 취약하며 LDA는 각각의 클래스의 데이터가 가지는 분산과 전체 데이터가 가지는 분산을 고려하여 클래스를 나누고 데이터가 가지는 특성을 유지하여 클래스 간 특징을 잘 나타내는 특성을 이용해 PCA와 LDA를 사용하여 피셔댄스를 생성해 포인트 댄스의 특성을 잘 나타내면서 댄스 별 분리를 최대화 할 수 있는 특징 벡터를 생성 했다. 생성된 K-pop 포인트 댄스 특징 벡터는 다양한 활성 함수 기반의 R-ELM 분류기를 사용해 분류했다. R-ELM 분류기는 출력 가중치의 자유도를 제약하는 규제화 방법을 사용하여 분류 성능을 높이기 위해 능형 회귀 분석법을 사용하여 출력가중치가 계산되며 R-ELM 분류기의 성능이 뛰어남을 증명하기 위해 기존에 널리 사용되는 기계학습 알고리즘인 k-NN, SVM과 최소자승법과 모어 페로스의 일반화된 역행렬로 출력가중치를 계산하는 일반적인 ELM을 사용하여 비교 분석 했다. R-ELM 분류기의 학습은 지도 학습 방법과 전 방향 학습 방법으로 다른 기계학습 방법에 비해 빠른 계산 속도를 가지며 더 높은 성능을 보였고 일반적인 ELM과 비교했을 때 계산 속도는 비슷하며 높은 성능을 보였다. 차원 축소 알고리즘을 사용하지 않고 통계적 각도만을 사용했을 경우 SVM이 82.75%의 가장 좋은 성능을 보였으나 PCA를 사용

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

하여 차원 축소된 특징 벡터를 사용 했을 때 sigmoid 함수 기반의 R-ELM이 83.75%로 가장 높은 성능을 보였으며 PCA+LDA를 사용 했을 때 sin 함수 기반 R-ELM이 85.00%의 가장 높은 성능을 보였다.

본 논문의 연구는 3차원 관절 좌표 값을 사용해 K-pop 포인트 댄스 모션을 가장 잘 나타낼 수 있는 특징을 찾는 것과 그 특징을 가장 잘 분류 할 수 있는 분류기를 사용하는 것이 주 목적이다. 피셔 댄스와 R-ELM을 사용해 높은 성능을 보이지는 못했지만 댄스는 사용자의 신체에 따라 각도 값이 크게 변할 수 있기 때문에, 다양한 신체 조건의 전문 댄서들의 데이터를 학습 한다면 좀 더 높은 분류 성능을 얻을 수 있을 것으로 예상된다.

dCollection @ chosun

조선대학교
CHOSUN UNIVERSITY

참고문헌

[1] M. Balazia, K. N. Plataniotis, “Human gait recognition from motion capture data in signature poses,” IET Biometrics, vol. 6, no. 2, pp.129-137, 2017.

[2] D. P. Barrett, J. M. Siskind, "Action Recognition by Time Series of Retinotopic Appearance and Motion Features," IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 12, pp.2250-2263, 2016.

[3] H. Eum, C. Yoon, H. Lee, M. Park, "Continuous Human Action Recognition Using Depth-MHI-HOG and a Spotter Model," sensors, vol. 15, no. 3, pp.5197-5227, 2015.

[4] O. D. Lara, M. A. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1192-1209, 2013.

[5] C. Zhu, W. Sheng, "Realtime Recognition of Complex Human Daily Activities Using Human Motion and Location Data," IEEE Transactions on Biomedical Engineering, vol. 59, no. 9, pp. 2422-2430, 2012.

[6] 김호원, 김도형, 이승욱, 최진성, 구본기, "실사 인체 아바타 생성 및 활용 동향," 전자통신동향분석, vol. 27, no. 3, pp. 1-11, 2012.

[7] 최민석 , 장백철, “전자 기기 조종을 위한 인간 동작 인식 기술 분석,” 한국컴퓨터정보학회논문지, vol. 19, no. 12, pp. 91-100, 2014.

[8] 홍석주, 이칠우, “비전 기반 제스처 인식 기술 연구동향,” 로봇과 인간, vol. 3, no. 3, pp. 15-22, 2006.

[9] J. Wang, Z. Liu, and Y. Wu, "Learning actionlet ensemble for 3D human action recognition," in Human Action Recognition with Depth Cameras, ed: Springer, pp.11-40, 2014.

[10] H. Jungong, L. Shao, D. Xu, and J. Shotton, “Enhanced Computer

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

Vision with Microsoft Kinect Sensor: A Review," IEEE Transactions on Cybernetics, vol. 43,no. 5, pp 1318-1334, 2013.

[11] http://www.intel.com/

[12] Leap Motion, "What is the Leap Motion Controller?," 2013.

[13] Vicon, Vicon System, http://www.vicon.com/

[14] 이재동, 이용환, "기계식 모션 캡처 시스템 및 응용," 한국정보기술학회논문지, vol. 7, no. 1, pp.31-38, 2009.

[15] Analogues, Gypsy, http://metamotion.com/gypsy/

[16] A. Menache, Understandin Motion Capture for Computer Animation and Video Games, Morgan Kafmann Publishers Inc, San Francisco, CA, 1999.

[17] Ascension, Motion Sar Wireless, http://www.ascension-tech.com/

[18] S. Gupta, D. Morris, S. N. Patel, D. Tan, "SoundWave: Using the Doppler Effect to Sense Gestures," CHI' 12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1911-1914, 2012.

[19] Q. Pu, S. Gupta, S. Gollakota, S. Patel, "Whole-Home Gesture Recognition Using Wireless Signals," MobiCom' 13 Proceedings of the 19th annual international conference on Mobile computing&networking, pp. 27-38, 2013.

[20] B. Yang, H. Dong, A. E. Saddik, "Development of a Self-Calibrated Motion Capture System by Nonlinear Trilateration of Multiple Kinects v2," IEEE Sensors Journal, vol. 17, no. 8, pp. 2481-2491, 2017.

[21] L. Shuai, C. Li, X. Guo, B. Prabhakaran, J. Chai, "Motion Capture With Ellipsoidal Skeleton Using Multiple Depth Cameras,"IEEE Transactions on Visualization and Computer Graphics, vol. 23, no. 2, pp. 1085-1098, 2017.

[22] R. Alazrai, M. Momani, M. I. Daoud, "Fall Detection for Elderly from Partially Observed Depth-Map Video Sequences Based on View-Invariant Human Activity Representation,"Appl. Sci, vol. 7, no. 4, 2017.

[23] Z. Liu, L. Zhou, H. Leung, H. P. H. Shum, "Kinect Posture

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

Reconstruction Based on a Local Mixture of Gaussian Process Models,"IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 11, pp. 2437-2450, 2016.

[24] Y. Du, Y. Fu, L. Wang,"Representation Learning of Temporal Dynamics for Skeleton-Based Action Recognition,"IEEE Transactions on Image Processing, vol. 25, no. 7, pp. 3010-3022, 2016.

[25] G. Zhu, L. Zhang, P. Shen, J. Song, "An Online Continuous Human Action Recognition Algorithm Based on the Kinect Sensor,"Sensors, vol. 16, no. 2, 2016.

[26] V. Bonnet, G. Venture, "Fast Determination of the Planar Body Segment Inertial Parameters Using Affordable Sensors,"IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 23, no. 4, pp. 628-635, 2015.

[27] M.C. Hu, C. W. Chen, W. H. Cheng, C. H. Chang, J. H. Lai, J. L. Wu,"Real-Time Human Movement Retrieval and Assessment With Kinect Sensor,"IEEE Transactions on Cybernetics, vol. 45, no. 4, pp. 742-753, 2015.

[28] Z. Gao, Y. Yu, Y. Zhou, S. Du, “Leveraging Two Kinect Sensors for Accurate Full-Body Motion Capture,” Sensors, vol. 15, no. 9, pp. 24297-24317, 2015.

[29] Y. Yao, Y. Fu, “Contour Model-Based Hand-Gesture Recognition Using the Kinect Sensor,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 11, pp. 1935-1944, 2014.

[30] 조성준, 강석호, "머신러닝(인공지능)의 산업 응용," ie 매거진, vol. 23, no. 2, pp. 34-38, 2016.

[31] G. B. Hwang, Q. U. Zhu, C. K. Siew, “Extreme learning machine: Theory and applications,” , Neurocomputing, Vol. 70, pp. 489-501, 2006.

[32] L. An B. Bhanu, “Image super-resolution by extreme learning machine,” 2012 19th IEEE International Conference on Image Processing, pp. 2209-2212, 2012.

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

[33] B. Saavedra-Morenoa, S. Salcedo-Sanza, L. Carro-Calvoa, J. Gascón-Morenoa, S. Jiménez-Fernándeza, L. Prietob, "Very fast training neural-computation techniques for real measure-correlate-predict wind operations in wind farms,"Journal of Wind Engineering and Industrial Aerodynamics, vol. 116, pp. 49-60, 2013.

[34] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, H. W. Ngan, "Electricity Price Forecasting With Extreme Learning Machine and Bootstrapping,"IEEE Transactions on Power Systems, vol. 27, no. 4, pp.2055-2062, 2012.

[35] H. Lee, S. J. Kim, K. Kim, M. Park, S. K. Kim, J. H. Park, S. R. Oh,"Online remote control of a robotic hand configurations using sEMG signals on a forearm,"2011 IEEE International Conference on Robotics and Biomimetics, pp. 2243-2244, 2011.

[36] R.Minhas, A. A. Mohammed, Q. M. J. Wu, "Incremental Learning in Human Action Recognition Based on Snippets,"IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 11, pp. 1529-1541, 2012.

[37] Z. Xie, K. Xu, L. Liu, Y. Xiong, "3D Shape Segmentation and Labeling via Extreme Learning Machine,"Computer graphics forum, pp. 85-95, 2014.

[38] Y. Xu, Q. Wang, Z. Wei, S. Ma, “Traffic sign recognition based on weighted ELM and AdaBoost,” Electronics Letters, vol. 52, no. 24, pp. 1988-1990, 2016.

[39] L. Oneto, F. Bisio, E. Cambria, D. Anguita, "Statistical Learning Theory and ELM for Big Social Data Analysis,"IEEE Computational Intelligence Magazine, vol. 11, no. 3, pp. 45-55, 2016.

[40] Y. Yang, Q. M. J. Wu, "Extreme Learning Machine With Subnetwork Hidden Nodes for Regression and Classification,"IEEE Transactions on Cybernetics, vol. 46, no. 12, pp. 2885-5898, 2016.

[41] X. Liu, R. Li, C. Zhao, P. Wang,"Robust signal recognition algorithm based on machine learning in heterogeneous

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

networks,"Journal of Systems Engineering and Electronics, vol. 27, no. 2, pp. 333-342, 2016.

[42] L. F.S. Cambuim, R. M. Macieira, F. M.P. Neto, E. Barros, T. B. Ludermir, C. Zanchettin, "An efficient static gesture recognizer embedded system based on ELM pattern recognition algorithm,"Journal of Systems Architecture, vol. 68, pp. 1-16, 2016.

[43] A. Iosifidis, A. Tefas, I. Pitas, "Minimum Class Variance Extreme Learning Machine for Human Action Recognition,"IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 11, pp. 1968-1979, 2013.

[44]Q. Yua, Y. Michea, E. Eirolaa, M. V. Heeswijka, E. Séverinb, A. Lendassea, "Regularized extreme learning machine for regression with missing data," Neurocomputing, vol. 102, pp. 45-51, 2013.

[45]G. B. Huang, H. Zhou, X. Ding, R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 42, no. 2, 2012.

[46] S. Saha, A. Konar, "Topomorphological approach to automatic posture recognition in ballet dance," IET Image Processing, vol. 9, no. 11, pp. 1002-1011, 2015.

[47] P. Muneesawang, N. M. Khan, M. Kyan, R. B. Elder, N. Dong, G. Sun, H. Li, L. Zhong, L. Guan,"A Machine Intelligence Approach to Virtual Ballet Training,"IEEE MultiMedia, vol. 22, no. 4, pp. 80-92, 2015.

[48] T. Han, H. Yao, C. Xu, X. Sun, Y. Zhang, J. J. Corso, "Dancelets Mining for video recommendation based on dance styles," IEEE Transactions on Multimedia, vol. 19, no. 4, pp. 712-724, 2017.

[49] S. Karavarsamis, D. Ververidis, G. Chantas, S. Nikolopoulos, Y. Kompatsiaris, "Classifying Salsa dance steps from skeletal poses," 2016 14th International Workshop on Content-Based Multimedia Indexing (CBMI), pp. 1-6, 2016.

[50] E. Hegarini, Dharmayanti, A. Syakur, "Indonesian traditional dance

Collection @ chosun

조선대학교
CHOSUN UNIVERSITY

motion capture documentation," 2016 2nd International Conference on Science and Technology-Computer (ICST), pp. 108-111, 2016.

[51] M. Turk, A. Pentland,"Eigenfaces for Recognition",Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[52] P. N. Belhumeur, J. P. Hespanha,D. J. Kriegman, “Eigenfaces vs. Fisherfaces: recognition using class specific linear projection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, 1997.

[53] https://developer.microsoft.com/enus/windows/kinect

[54] E. Auvinet, F. Multon, C. E. Aubin, J. Meunier, and M. Raison, “Detection of gait cycles in treadmill walking using a Kinect,” Gait & posture, vol. 41, No. 2, pp.722-725, 2015.

dCollection @ chosun