

음성 정보와 DTW 알고리즘을 활용한 성인 콘텐츠 필터링

Adult Contents Filtering using Voice Information and DTW

조정익* · 이일병**
Jungik Cho and Yillbyung Lee

* 연세대학교 컴퓨터과학과

** 연세대학교 컴퓨터과학과

요 약

본 논문은 필터링 수행율을 향상시키기 위해, DTW 알고리즘을 제시한다. 여기에서 말하는 콘텐츠 필터링은 음성의 특징을 사용해서 콘텐츠를 구분하는 것을 확인하는 기술이다. 즉, 이 방법이 일반 콘텐츠와 성인 콘텐츠를 구분한다. 음성에 대한 정보를 추출하는 방법이 콘텐츠를 필터링하는데 있어서 기여를 할 수 있다. 즉, DTW 알고리즘을 사용하여 필터링 인식률을 향상하는 방법이라고 제안을 한다. 마지막으로, 본 논문에서 제안한 방법의 적용 가능성과 일반성을 평가하기 위하여 수치적인 예를 적용한다. 제안하는 성질의 정확도를 시험하기 위해서 실험을 제공하였다. 결과적으로 일반 콘텐츠와 성인 콘텐츠 특성의 차이를 알았다. 추후에 이 성질을 필터링 성능 향상에 응용할 수 있다.

키워드 : 음성 정보, 특징 추출, 성인 콘텐츠, 일반 콘텐츠, DTW 알고리즘

Abstract

This paper deals with the DTW algorithm for the filtering contents, in order to improve the filtering performance rate. Contents filtering is the technology that confirm the identification of contents by using the feature of voice. Such technique is classified into general contents and adults contents. This proposed method extracts the information of voice contribute to improvement of filtering contents. In other words, We proposed filtering identification rate can be improved by using DTW algorithm. As a result, the proposed method is utilized improvement of filtering contents. Finally, we provide contents examples to test the accuracy of the proposed feature. Consequently, We know that the difference of characteristic between general contents and adults contents. In the future, We utilize this to improve filtering performance rate.

Key Words : Voice information, Feature abstraction, Adult contents, General contents, DTW algorithm

1. 서 론

현재 전 세계는 초고속 인터넷 시대에서 Web 2.0이라는 새로운 상황을 맞이하고 있다. 이에 따라 지금까지 수동적인 입장에서 정보를 받아들이는 사용자들이 이제 능동적이고 주도적으로 Web 환경에 참여하여 능동적인 피드백을 받고 있다. 이에 편승해 사용자가 만든 멀티미디어 콘텐츠(Multimedia Contents)가 Web 환경에 쏟아지고 있다. 이런 일련의 변화는 정보의 공유라는 역할도 하고 있으나, 개인 사생활 정보의 유출, 유해한 정보의 출현 등 부작용도 만만치 않다. 그 중에서도 성인 콘텐츠의 노출은 사회적으로도 큰 문제가 되고 있다. 그래서 본 논문이 제안하는 방법은 성인 콘텐츠에 대한 필터링을 하는데 있어서 음성 정보를 활용해 보고자 한다. 먼저 성인 콘텐츠에서 나오는 음성을 패턴 정합 기법인 DTW를 이용하여, 일반 콘텐츠의 음성 정보와 비교하여 특징을 추출하는 것이다. 즉, 성인 콘텐츠에서 쓰인 후보

음성과 일반 콘텐츠에서 획득한 시험용 음성을 각각 읽어 들여 특징 추출을 행하고, 후보 음성과 시험용 음성의 특징 열을 DTW 알고리즘을 사용하여 비교하여, 거리가 최소가 되는 음성의 정보를 인식 결과로 결정한다. 실험 과정은 성인 콘텐츠에서 획득한 특정 음성을 기준으로 다른 10개의 성인 콘텐츠에서 획득한 음성과 일반 콘텐츠에서 획득한 음성에서 어떤 차이점을 보이는지를 DTW 알고리즘의 거리가 최소가 되는 음성을 결과값으로 하여 일반 콘텐츠와 성인 콘텐츠간의 특성 차이를 확인한다. 이 방법은 추후 성인 콘텐츠를 필터링을 하는데 많은 도움이 될 것으로 기대가 된다.

본 논문에서는 DTW를 이용하여 성인 콘텐츠를 필터링 하는데 있어서 성능을 향상 방안을 제시하는데 있다. DTW는 음성의 묶음 구간을 포함한 비교를 하는데 있어서 좋은 성능을 보이고 있다. 본 논문에서는 이러한 특징을 이용하여 성인 콘텐츠를 필터링하는데 있어 성능을 향상할 수 있는 방법을 제안한다. 마지막으로, 본 논문에서 제안한 DTW를 이용한 필터링 방법에 대한 효율성과 정확성을 검증하기 위해 실험을 실시한다. 추후에 DTW를 이용한 방법은 필터링 성능 향상을 위한 하나의 방법으로 진화가 될 수 있다고 판단된다.

접수일자 : 2008년 3월 1일

완료일자 : 2008년 3월 1일

2. DTW 알고리즘

DTW 알고리즘이란 무엇인가? 시간 축 상에서의 비선형 신축을 허용하는 패턴 매칭 알고리즘이라고 정의할 수 있다. 즉, 길이가 다른 두 열에서 어느 한 열을 기준으로 두 열을 비교하기 위해서는 어느 한 열이 신장(늘어남)되거나 축소되어야만 한다. 패턴열의 길이가 다를 경우에는 매핑 함수를 적용해야만 하는데 선형 매핑 함수를 적용할 경우에는 신축 과정에서는 열의 성분을 속아 내거나 신장 과정에서는 열의 성분이 보간하여 두 열의 길이를 같게 만든 후에야 비로소 두 열의 비교가 가능할 것이다. DTW 알고리즘을 이용하면 비선형 매핑 함수를 최적으로 찾아가면서 동시에 비교가 이루어진다. DTW 알고리즘은 일단 두 열의 각 성분에 대한 거리척도 값을 비용으로 설정한다. 그리고 두 열이 이루는 격자(lattice) 상에서 각 열의 시작 성분에서 시작하여 끝 성분에 이르기까지 비용 테이블에 최소 비용을 순환적으로 택하여 저장하는 점화식을 이용하는 동적 계획법으로 매핑 함수를 찾아가면서 두 열을 비교하는 알고리즘이다. 최종적으로 끝 성분에서 비용 테이블에 저장되는 비용 값이 두 열에 대한 유사도가 된다. 그래서 DTW 알고리즘은 주로 음성 인식에서 많이 이용된다. 비교 루틴이 아주 컴팩트하고 단순하여 단어 단위의 간단한 음성 인식기에 적용 가능한 알고리즘이다. DTW를 이용한 음성 인식은 PCM 디지털 데이터를 그대로 사용하는 것이 아니라, 음성이 10~20ms 지속 시간 동안은 정상적(stationary)인 구간이라고 가정하고 행하는 단구간 분석에 의하여 프레임 단위로 음성 특징벡터를 추출하는 전처리를 거친 후에 이루어진다. 그리고 인식 과정에서는 음성 인식 후보 단어 각각을 이러한 특징 추출 과정을 거쳐 기준 벡터열로 미리 준비하여 두고, 인식할 단어에 대한 특징을 추출하여 시험 벡터 열을 각 후보 단어와 DTW 알고리즘을 이용하여 비교하여 최소가 되는 후보 단어 카테고리를 인식 결과로 결정하는 비교적 간단한 음성 인식 알고리즘이다.[6]

3. 제안하는 방법

3.1 음성특징 추출

LPC 켈스트럼 계수를 Bilinear Transform을 사용하여 멜 스킵트럼으로 변환을 한다. 사용되는 Bilinear Transform은 전대역 필터를 사용하여 주파수를 변화시키는 방법이다.[5]

$$Z_{\omega}^{-1} = \frac{z^{-1}-a}{1-az^{-1}} \quad (-1 < a < 1) \quad (1)$$

$$\omega_{\omega} = \omega + 2 \tan^{-1} \left(\frac{a \sin \omega}{1 - a \cos \omega} \right)$$

여기서, a는 주파수 왜곡 파라미터를 말하며, ω_{ω} 은 ω 변화되는 주파수, ω 는 정규화된 샘플링 주파수로 정의된다.

4. 실험 및 결과 고찰

4.1 실험

본 논문에서 제안한 방법의 적용 가능성과 일반성을 평가하기 위하여 수치적인 실험을 제안한다. 즉, 제안하는 성질의 정확도를 시험하기 위해서 표본을 추출하여 거기에서 특정한 성질을 추출하는 것이다. 결과적으로 이런 실험을 일반 콘텐츠와 성인 콘텐츠의 특성의 차이를 판단할 수 있을 것이다. 추후에 이 성질을 필터링 성능 향상에 응용할 수 있다. 실험에서 동영상 정보의 오디오 영역에서 음성 인식을 위해서 사용한 것은 Matlab 프로그램이다. 제안하는 방법은 일단 음성 정보에서 콘텐츠를 가장 잘 표현할 수 있는 특징값을 추출하고, 이 추출된 음성열을 DTW에 의한 방법으로 콘텐츠를 필터링하게 된다. 음성 편집기를 이용하여 50개의 샘플링 음성 데이터를 준비를 한다. 파일은 Wav 파일로 준비를 한다. 콘텐츠의 특징을 잘 비교하기 위해서 LPC 켈스트럼 계수를 Bilinear Transform을 사용하여 멜 스킵트럼으로 변환을 한다. 그리고 여기서 무작위로 성인 콘텐츠로 판단되는 10개의 데이터와 일반 콘텐츠로 판단되는 10개의 데이터를 추출한다. 1개의 기준 이 되는 성인 콘텐츠를 정하여, 이것을 10개의 성인 콘텐츠와 일반 콘텐츠와 비교를 통하여 각각 콘텐츠의 특성을 파악한다. 20개의 데이터를 Wav 파일로 준비하여 비교 실험을 실시한다. 단계별로 구분을 해보면 1단계는 각 성인 콘텐츠와 일반 콘텐츠를 각각 읽어 들여, 특징 추출을 한다. 2단계는 추출된 20개의 콘텐츠의 특징열을 DTW 알고리즘으로 비교한다. 3단계는 거리가 최소가 되는 카테고리 후보를 각각 비교하여 어떤 콘텐츠에 속하는지 결과를 결정한다.[6]

표 1. DTW 알고리즘을 이용한 성인 및 일반 콘텐츠의 거리.

Table 1. Avg distance between adults and general contents using DTW algorithm.

성인 콘텐츠	일반 콘텐츠
2.92415e + 003	3.18856e + 003

이 표에서 보였듯이 성인 콘텐츠가 가까운 거리를 나타내고 있다. 특히, 위 표에서는 표시가 되어 있지 않지만, 여성의 음성을 많이 포함하고 있는 콘텐츠의 경우에는 거리값 차이(1.9414e + 003 ~ 1.8525e + 003)가 훨씬 줄어드는 것을 알 수 있다. 이런 성인 및 일반 콘텐츠의 차이점을 비교 분석해 보면 이런 콘텐츠들을 필터링하는데 있어서 큰 성능 향상을 이룰 수 있을 것으로 판단이 된다.

4.2 결과고찰

본 논문에서 제안한 특성 차의 적용 가능성과 일반성을 평가하기 위하여 수치적인 예를 실험에 적용을 하였다. 즉, 제안하는 성질의 정확도를 시험하기 위해서 멜 스킵트럼 및 DTW 알고리즘을 이용한 실험을 하였다. 그 결과 일반 콘텐츠와 성인 콘텐츠의 차이를 확실히 분석할 수 있었다. DTW 알고리즘을 활용한 이런 방법은 추후에 이 성질을 필터링 시스템에 적용을 하게 되면 필터링의 성능 향상에 큰 기여를 할 수 있을 것으로 판단이 된다.

참 고 문 헌

- [1] 최용선(Yong-Sun Choi), 오상훈(Sang-Hoon Oh), 이수영(Soo-Young Lee). "핵심어 검출을 위한 단일 끝점 DTW 알고리즘," 대한전자공학회, 전자공학회논문지-SP 전자공학회논문지 제41권 SP편 제3호, pp. 209-219 (11pages), 2004. 5.
- [2] 김종국(Jong-Kuk Kim), 조왕래(Wang-rae Jo), 배명진(Myung-Jin Bae). "음성의 묵음구간 검출을 통한 DTW의 성능개선에 관한 연구," 한국음성과학회, 음성과학 제10권 4호, pp. 117-124 (8pages), 2003. 12.
- [3] 우선경, 강철호. "DTW 알고리즘을 이용한 화자적응 신경망에 관한 연구," 대한전자공학회, 대한전자공학회 학술대회 논문집 (신호처리합동) 제8권 1호, pp. 1118-1122 (5pages), 1995. 1.
- [4] 정중순, 배재욱, 배명진. "윈도우 환경에서 음성을 이용한 사용자 확인에 관한 연구," 한국음향학회지, Vol.17, No.5. 1998.
- [5] 정중순, "대표 평균패턴과 가중 캡스트럼을 이용한 화자인식의 성능 향상에 관한 연구," 석사학위논문, 한국과학기술원, 1996.
- [6] 한학용, "MATLAB 실습을 통한 입체적 학습," 패턴인식 개론, 2005.
- [7] Furui S. and Sondhi. "Advance in Speech Signal Processing," Dekker.

저 자 소 개



조 정 익 (Jungik Cho)

2001년 한동대 건설도시환경공학부 졸업.
2007년~현재 연세대 컴퓨터과학과 석사과정
관심분야 : 퍼지 이론, 신경회로망, 데이터 마이닝

Phone : 02) 365 - 4598
Fax : 02) 365 - 2579
E-mail : abc97kr@yahoo.co.kr



이 일 병 (Yillbyung Lee)

1985년 Ph.D., Computer & Information Science, University of Massachusetts
1995년~현재 연세대학교 컴퓨터공학과 교수
관심분야 : 문자인식, 문서인식, 데이터 마이닝, 신경회로망, 영상처리

Phone : 02) 365 - 4598
Fax : 02) 365 - 2579
E-mail : yblee@csai.yonsei.ac.kr