

가사의 감정 분석을 이용한 GAN 기반 댄스 공연 배경 생성 방법

윤혜원, 곽정훈, 성연식*
동국대학교 멀티미디어공학과
hyewon@dongguk.edu, jeonghoon@dongguk.edu, sung@dongguk.edu

GAN-based Dance Performance Visual Background Generation Method using Emotion Analysis on Lyrics

Hyewon Yoon, Jeonghoon Kwak, Yunsick Sung*
Dept. of Multimedia Engineering, Dongguk University-Seoul

요약

최근 인공지능을 활용하여 예술 작품에 몰입할 수 있도록 무대 효과를 디자인하는 연구가 진행되고 있다. 무대 효과 중에서 무대 배경은 공연의 분위기를 형성한다. 춤의 장르별로 무대 배경에 사용되는 이미지를 생성하기 위해 소셜 미디어 기반 무대 배경 생성 시스템이 있다. 하지만 같은 장르 춤은 동일한 무대 배경 이미지가 제공되는 문제가 있다. 같은 장르의 춤이지만 노래의 분위기를 반영하여 차별된 무대 배경 이미지를 제공하는 것이 필요하다. 본 논문은 노래 가사의 감정을 활용하여 Generative Adversarial Network(GAN)을 통해 각 노래의 분위기를 고려한 무대 배경 이미지를 생성하는 방법을 제안한다. GAN은 노래에 포함된 단락별 감정 단어를 추출하여 스타일을 생성하도록 학습된다. 학습된 GAN은 노래 가사에 포함된 감정 단어를 활용하여 곡의 분위기를 반영한 무대 배경 이미지를 생성한다. 노래 가사를 고려하여 무대 배경 이미지를 생성함으로써 곡의 분위기가 고려된 무대 배경 이미지 생성이 가능하다.

1. 서론

춤은 감정을 표현하기 위해 신체 언어를 사용하는 예술로 춤 공연은 보통 춤 예술과 음악, 조명, 연기 및 무대 배경과 같은 다양한 종류의 무대 효과가 결합된다[1]. 무대 효과 중에서 무대 배경은 분위기를 제공하므로 가장 중요한 기능 중 하나이다. 무대 배경을 제작하기 위하여 곡의 내용, 곡의 분위기 등과 같이 복합적인 요소 고려가 필요하다.

소셜 미디어를 기반으로 춤의 장르(발레, 벨리 댄스, 스트리트 댄스, 모던 댄스, 탱고, 활츠)에 따라 무대 배경 이미지를 하는 추천하는 시스템이 있다 [1]. 무대 배경 추천 시스템은 춤의 장르별 댄서의 계시물 공유나 클릭 액션 등과 같은 장르별 댄서의 관심사를 활용하여 무대 배경 이미지가 추천된다.

* 교신저자: 성연식 (sung@dongguk.edu)

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 글로벌핵심인재양성지원사업의 연구결과로 수행되었음"(2019-0-01585)

추천된 무대 배경 이미지는 공연에서 사용되는 무대 배경 이미지로 사용 가능하다. 하지만 춤의 장르에 따라 무대 배경을 만드는 것은 각 노래 고유의 분위기를 반영하지 않아 같은 장르의 춤은 동일한 무대 배경 이미지가 제공된다. 동일한 춤의 장르일지라도 노래 고유의 분위기를 반영하여 차별화된 무대 배경 이미지를 생성하는 것이 필요하다.

노래의 고유 분위기를 파악하기 위하여 노래 가사에 포함된 감정 단어를 사용하여 각 노래의 분위기를 파악한다[2]. 노래의 가사에 포함된 감정단어로 노래의 분위기를 분석 가능하다. 하지만 노래 가사에 포함된 감정단어를 활용하여 무대 배경 이미지를 생성하는 방법이 요구된다.

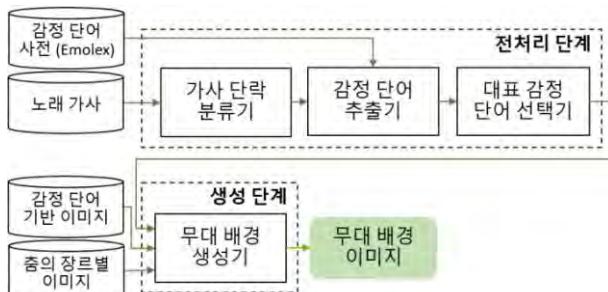
문장을 기반으로 이미지를 생성하기 위하여 Generative Adversarial Network(GAN)을 활용한다 [3]. GAN을 통하여 문장에 맞는 이미지를 생성한다. 하지만 문장에 포함된 분위기를 반영한 이미지를 제공하지 못하는 단점이 있다. GAN을 활용하여 노래 가사의 분위기를 전체 이미지에 적용하는 방법이 필요하다.

요하다.

본 논문은 GAN을 기반으로 노래 가사에 포함된 감정 단어를 이용하여 무대 배경 이미지를 생성하는 방법을 제안한다. 노래 가사에 포함되는 감정 단어 (Emotion Word)를 추출한다. GAN은 추출된 감정 단어가 나타내는 색상 특징 및 Stroke 특징을 반영한 무대 배경 이미지가 생성된다. 곡의 분위기에 맞는 새로운 무대 배경 이미지를 생성한다.

2. GAN을 활용한 노래 가사의 감정 단어 기반 무대 배경 생성 방법

노래 가사에 포함된 감정 단어를 활용하여 그림 1과 같이 무대 배경 이미지를 생성하는 방법을 제안한다. 노래 가사를 가사 단락 분류기를 통해 단락별로 구분한다. 일반적인 노래의 가사는 Verse, Chorus, Bridge로 구성된다[4]. Verse, Chorus, Bridge를 나눈 단위를 단락이라고 정의한다. Bridge는 단락이 너무 짧아 의미를 담기 어렵기 때문에 본 논문에서는 가사 단락 분류기를 통해 추출된 Chorus와 Verse에서 추출된 감정 단어를 활용한다.



감정 단어 추출기는 단락에 포함된 다수의 감정 단어를 추출한다. 하나의 단락에는 하나 이상의 감정 단어가 포함된다. 단락에 포함된 감정 단어들을 추출하기 위해 감정 단어 사전(Emolex)를 이용한다[2]. Emolex는 심리학자 Plutchik이 제시한 인간의 여덟 가지 기본 감정(anger, anticipation, disgust, fear, joy, sadness, surprise, trust)에 긍정과 부정을 더하여 10차원 수준의 감정 단어 사전을 제공한다. Emolex는 14,177개의 단어가 들어있다. 감정 단어 추출기를 통해 Emolex와 노래 가사에 포함된 단락별 감정 단어가 일치하는 감정 단어를 추출한다.

대표 감정 단어 선택기는 단락별로 대표가 되는 감정 단어를 선택한다. 단락별 추출된 다수의 감정 단어를 기반으로 TF-IDF[5]를 통해 단락별 대표하는 감정 단어를 선택한다.

무대 배경 생성기는 GAN을 활용하여 감정 단어를 기반으로 무대 배경 이미지를 생성한다. 춤의 장르별 이미지는 춤의 장르를 표현하는 이미지이다. GAN의 생성기에서는 춤의 장르별 이미지에 감정 단어 기반 이미지 스타일을 합성하여 무대 배경 이미지를 생성한다. 감정 단어 기반 이미지는 추출된 단락별 대표 감정 단어가 그림 제목 또는 설명에 포함되어 있는 이미지이다. GAN의 판별기에서는 GAN의 생성기에서 무대 배경 이미지와 감정 단어 기반 이미지가 판별 가능한지 확인한다. GAN의 생성기를 통하여 춤의 장르별 이미지에 감정 단어 기반 이미지의 색상 특징 및 Stroke 특징이 반영되어 무대 배경 이미지가 생성된다.

본 논문에서는 노래 가사의 단락별 대표 감정 단어로부터 학습된 스타일을 춤의 장르별 이미지에 적용하여 단락별 무대 배경 이미지 생성이 가능하다.

사사표기

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 글로벌핵심인재양성지원사업의 연구결과로 수행되었음”(2019-0-01585)

참고문헌

- Wen, J., She, J., Li, X., Mao, H., “Visual Background Recommendation for Dance Performances Using Deep Matrix Factorization,” ACM Transactions on Multimedia Computing Communications and Applications, Vol. 14, No. 1, pp. 1-19, 2018.
- 이재환, 임혜원, 김형주, “가사의 감정 분석과 구조 분석을 이용한 노래간 유사도 측정”, 정보과학회 컴퓨팅의 실제 논문지, Vol. 22, No. 10, pp. 479-487, 2016.
- Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., Metaxas, D.N., “Stackgan++: Realistic Image Synthesis with Stacked Generative Adversarial Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 41, No.8, pp. 1947-1962, 2019.
- “Ten Minute Master No 18: Song Structure,” MUSIC TECH magazine, pp. 62-63, 2003.
- Chowdhury, G.G., “Introduction to Modern Information Retrieval,” Facet publishing, 2010.