

다큐멘터리 영상에서의 그래픽 자료 검출 방법

이재호^o 허일 낭종호

서강대학교 컴퓨터 공학과

leejh01623@naver.com, heol@sogang.ac.kr, jhnang@sogang.ac.kr

Graphic Shot Detection Method in Documentary Video

Jae-ho Lee^o Il Heo Jong-ho Nang

Department of Computer Science and Engineering, Sogang University

요 약

TV 프로그램 중 건강 다큐멘터리인 <생로병사의 비밀>은 전문적인 지식과 다양한 사례를 보여줌으로써 시청자들에게 많은 도움을 주는 프로그램이다. 해당 영상의 Shot들의 대해 주석을 입력할 때 기존 방법은 시간이 오래 걸리고, 주관적이라는 단점을 가지고 있었다.

본 논문에서는 <생로병사의 비밀>의 Shot Type을 일반화 시키고, 그 중 비중이 높은 그래픽 장면을 자동으로 검출하는 시스템을 제안하고, 칼라히스토그램(Color Histogram), 정렬(Sorting), 가우시안 분포(Gaussian Distribution)을 통하여 실제 장면과 그래픽 장면을 구별해 내는 시스템을 설계 및 구현 하였다. 본 논문에서 구현한 시스템을 기반으로 주석을 입력할 때 시간을 단축시키고, 객관적인 내용을 입력하여 효과적으로 주석을 할 수 있을 것으로 기대한다.

1. 서 론

TV 프로그램에서 방영하고 있는 방송들에는 다양한 장르가 존재한다. 그 중 하나인 다큐멘터리 장르도 시사, 휴먼, 자연 등으로 더 세분화 할 수 있다. 이러한 다큐멘터리는 여러 샷 유형으로 구성된다. 이러한 샷 유형을 일반화시키기 위하여 국내 외 50회 이상 방영한 시리즈물을 중점적으로 분석하였다. 그 중 가장 일반화시키기 용이하고, 명확한 KBS 건강 다큐멘터리 <생로병사의 비밀>[1]을 선택하여 실험 및 분석을 진행하였다. KBS 1-TV에서 제작한 <생로병사의 비밀>은 2002년부터 꾸준히 500회 이상이 방영된 건강 다큐멘터리이다. 여러 질병과 예방법 등에 대한 사례를 소개하고, 전문가의 인터뷰를 통해 전문적인 지식을 전달해주기 때문에 시청자들에게는 유익한 방송 프로그램 중의 하나로 꼽힌다. 이러한 <생로병사의 비밀>의 샷 유형을 일반화시키기 위하여 분석한 결과 [표 1-1]과 같이 크게는 인터뷰 장면과 자료화면, 기타(오프닝, 클로징, MC 장면 등)로 나눈다. 인터뷰 장면은 전문가 인터뷰(의사), 관계자 인터뷰(헬스 트레이너), 일반인 인터뷰(환자)로 세분화한다. 자료화면은 그래픽 장면, 인용 장면으로 나누고, 그래픽 장면은 표와 차트를 보여주는 장면, 신체의 장기나 인체의 형상, 세포들의 움직임을 보여주는 애니메이션 장면으로 나눈다. 인용은 문서(논문)를 인용한 장면, 사진을 인용한 장면으로 나눈다. 이러한 장면들 중 Annotator가 영상에 대해 주석을 할 때 도움을 줄 수 있는 그래픽 장면을 검출하기 위하여 다음과 같은 목표를 세워 진행하였다.

- 일반 장면과 그래픽 장면의 차이점 도출

- 장면에 대한 히스토그램과 정렬 시스템 구현
- 가우시안 분포를 통해 일반 장면과 그래픽 장면을 구분
본 논문에서는 <생로병사의 비밀>이 가진 방대한 건강 정보에 대해 Annotator가 효과적으로 주석을 하기 위한 노력을 기술하고, 샷 유형 중 많은 비중을 차지하는 그래픽 자료를 자동으로 검출하기 위한 방법과 구현 시스템을 소개한다. 이러한 검출 방법과 시스템을 통하여 <생로병사의 비밀>이 가진 정보들에 대해 주석할 때 시간을 절약하고, 더 나아가 원하는 정보만을 요약할 수 있는 시스템 등으로 활용할 수 있을 것이라 기대한다.

2. 배 경

2.1 기존 그래픽 샷 검출 방법

최근 논문에서는 실사 이미지와 유사한 컴퓨터 그래픽 이미지를 구별해 내기 위하여 기존의 방법들을 조합하여 새로운 프레임워크를 개발하는 것을 목표로 두고 있다. 먼저 기존에 컴퓨터 그래픽과 실제 이미지를 구별하기 위해 사용되었던 방법들을 나열하고, 각각의 특징들을 하나로 종합하여 정확도를 높이고자 하였다. (1) 실제 이미지에서 만화를 구별하기 위한 방법 (2) HSV Color Space를 이용하여 구별하는 방법 등이 있다. (1)의 경우, 실제 이미지에 삽입되어 있는 만화를 구별해 내기 위한 방법으로 배경색과 확연히 구분되는 Color와 Edge 히스토그램, 평균 채도 등을 활용한다. (2)의 경우, 색상(H), 채도(S), 명도(V) 3가지의 성분 중 H와 S를 이용하여 히스토그램 값을 출력, 비교하여 이미지를 구별한다[4].

2.2 샷 유형 자동 검출 시스템을 통한 방법

다큐멘터리에서 일반적으로 [표 1-1]과 같이 샷 유형을 구분 지을 수 있다. 각각의 샷 유형의 종류에 따라 영상에서 시간대비 차지하는 비중 또한 다르다. <생로병사의 비밀>을 분석해본 결과 인터뷰 장면, 그래픽 장면, 인용 장

□ 본 연구는 미래창조과학부 및 정보통신기술연구진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행하였음.
[R0101-15-0097,클라우드 기반 개방형 소셜 방송미디어 콘텐츠 융합 생성, 편집 및 재생을 위한 미디어 제작 및 전송 시스템 개발]

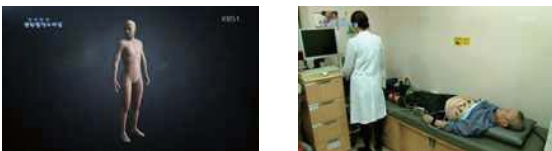
[표 1-1] 다큐멘터리 샷 유형 일반적 분류

번호	분류	상세분류	예시
1	인터뷰	전문가 인터뷰	의사(전문의)
2		관계자 인터뷰	헬스 트레이너
3		일반인 인터뷰	환자
4	자료화면	그래픽 차트	질병차트 (막대 그래프, 타원 그래프)
5		그래픽 표	질병표
6		그래픽 애니메이션	사람 (장기, 전신, 세포 애니메이션)
7	인용	인용 사진	
8		인용 문서	학술지, 관련 논문
9	기타		오프닝, 클로징, MC 장면 등

면이 전체의 약 60%이상으로 가장 큰 비중을 차지한다. 그만큼 Annotator가 주석을 입력할 때 위의 3가지 유형을 자동으로 검출해 줄 수 있다면 주석 비용을 줄일 수 있고, 추후에 재사용 면에서도 많은 도움을 줄 수 있을 것이다. 본 논문에서는 3가지의 유형 중 칼라 히스토그램과 정렬, 가우시안 분포를 이용하여 그래픽 장면을 검출하는 시스템을 제안한다.

3. Graphic 샷 자동 검출 시스템

<생로병사의 비밀>을 구조적으로 샷 유형을 나눌 때 그래픽 장면이 차지하는 비중이 높다. 일반적으로 <그림 3-1>의 (a)와 같은 그래픽 샷은 <그림 3-1>의 (b)와 같은 실제 샷과 다르게 사람에 의해 만들어진 인공적인 장면이다. 따라서 실제 장면보다 색의 수가 적을 것이다. 이 점에 착안하여 본 장에서는 칼라 히스토그램을 이용하여 그래픽 장면을 검출하는 시스템을 고안한다.



(a) 그래픽 샷 (b) 실제 샷
<그림 3-1> 그래픽 샷 및 실제 샷

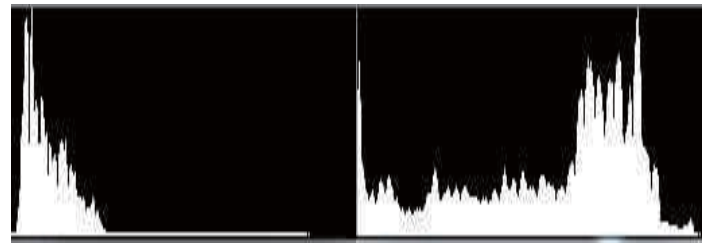
3.1 칼라 히스토그램

그래픽 장면 같은 경우 현실 세계의 모습들이 아닌 제작자가 만든 인위적인 장면이다. 이러한 장면의 경우 시청자들이 전문적인 내용이나 사례에 대해 좀 더 쉽게 접근하고, 이해할 수 있도록 도와주는 목적을 지니기 때문에, 만들어진 그림 또한 단순한 것이 일반적이다. 따라서 색의

수가 현실 세계를 나타내는 것보다 적다. 이러한 점을 이용하여 그래픽 장면을 구별한다.

칼라 히스토그램은 해당 이미지의 R, G, B값에 대해 픽셀 대비 빈도수를 나타낸다. 본 논문에서는 트루 칼라를 이용하여 히스토그램 값을 구하여 그래픽 장면을 검출하고자 한다.

<그림 3-2>는 <그림 3-1> 해당 샷에 대한 칼라 히스토그램을 출력해 본 결과이다. X축은 픽셀 값이고, Y축은 해당 픽셀 값에 대한 빈도수를 나타낸다. <그림 3-1>의 그래픽 샷에 대한 히스토그램은 낮은 픽셀 값에 치우쳐

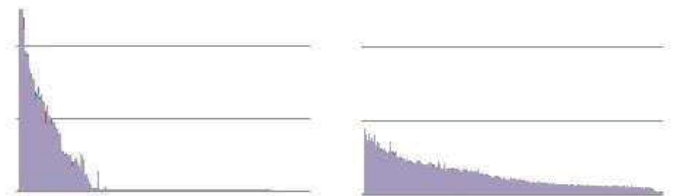


<그림 3-2> <그림 3-1>에 대한 히스토그램

있는 반면 실제 샷에 대한 히스토그램은 픽셀 전체에 대해 고르게 분포 한다.

3.2 정렬

각각의 픽셀에 대해 높은 값부터 낮은 값까지 내림차순으로 정렬을 진행한다. <그림 3-3>은 <그림 3-1>의 그래픽 샷과 실제 샷에 대해 정렬한 결과를 보여준다.



<그림 3-3> <그림 3-2>에 대한 정렬한 히스토그램

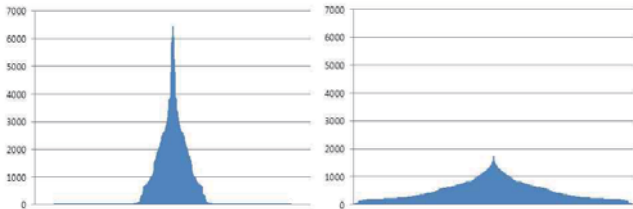
정렬한 결과를 살펴보면 그래픽 샷의 경우 일정 픽셀 값까지 기울기가 가파른 특징을 나타내고, 실제 샷의 경우 전체 픽셀 값에 대해 완만한 기울기를 가지는 특징을 나타낸다. 좀 더 명확하게 값을 비교하기 위하여 <그림 3-4>와 같이 단위 픽셀 값에 따른 평균 변화율을 구한다.



<그림 3-4> <그림 3-3>에 대한 평균 변화율

3.3 가우시안 분포

<그림 3-1>에 대한 정렬한 히스토그램 값을 역으로 한번 더 계산하여 가우시안 분포로 표현하면 <그림 3-5>와 같은 결과물을 얻을 수 있다. 그래픽 샷에 대한 분포는 가운데가 볼록 솟은 모양을 보이고, 실제 샷에 대한 분포는 좌우로 퍼지는 모양을 보인다.



<그림 3-5> <그림 3-3>에 대한 가우시안 분포

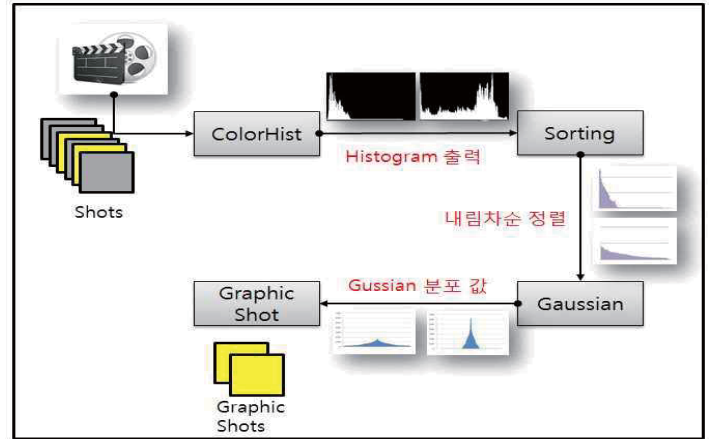
$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (3.1)$$

위 식 (3.1)에서 μ 는 평균(mean)을 의미하고, σ 는 표준편차(standard deviation)을 나타낸다. 표준 편차의 값이 클수록 좌우로 펼쳐진 정도가 크게 나타난다. 이 점을 활용하여 임계값(Threshold)을 정의한다. 각각 그래픽 샷과 실제 샷에 대한 표준편차를 구하게 되면 차이가 두드러지는 특징을 볼 수 있다. 여러 그래픽 샷과 실제 샷에 대한 표준편차의 평균을 내어 그 값을 임계값으로 주어 그래픽 샷과 실제 샷을 구별해 낸다.

4. 시스템 구현 및 분석

본 논문에서는 3장에서 기술한 시스템의 여러 이슈 및 특성들을 고려하여 <그림 4-1>과 같이 시스템을 구현하였다. 해당 영상에 대한 샷들에 대해 칼라 히스토그램을 출력한다. 출력한 히스토그램의 빈도수 값을 가장 높은 값부터 내림차순으로 정렬한다. 정렬한 값들을 역으로 한번 더 계산하여 가우시안 분포로 표현한다. 이 때 그래픽 샷과 실제 샷을 구분 지을 수 있는 임계값을 구하기 위하여 표준편차를 구하게 되고, 여러 샷들의 표준 편차의 평균을 임계값으로 하여 그래픽 샷과 실제 샷을 구별해 낸다.

본 논문에서 제안한 시스템을 바탕으로 <생로병사의 비밀>의 여러 에피소드들에 대해 테스트를 해본 결과 그래픽 샷으로 검출한 정답이 전체 정답에서 약 80% 이상을 보이는 것을 확인 할 수 있었다. 하지만 전체 샷들 중 잘못 검출된 부분이 조금 있어 Precision 값이 낮게 출력되었다. 예를 들어, 배경이 어두운 장면이거나, MC가 진행하는 샷에서 배경이 3D 그래픽 일 경우, 실제 X-Ray나 CT장면이 나오는 경우 등은 그래픽 장면과 같이 위로 솟은 가우시안 분포를 나타내게 되고, 그래픽 장면이라고 잘못 검출 된다.



<그림 4-1> Graphic 자료 자동 검출 시스템 흐름도

5. 결 론

본 논문에서는 <생로병사의 비밀>을 구성하고 있는 샷 유형 중 그래픽 장면을 자동으로 검출하기 위한 시스템을 제안 하였다. 기존에 주석을 입력 할 때 Annotator가 샷들을 보고 무분별하게 입력하여, 주석 비용이 높고, 객관적이지 못한 문제점이 발생하였다. 이러한 문제점을 보완하고자 특정 콘텐츠에 대해 샷 유형을 정의하고, 빈도수가 높은 샷 중 그래픽 샷 자동 검출을 위하여 칼라 히스토그램과 정렬, 가우시안 분포를 활용하였다. 이러한 시스템을 바탕으로 Annotator가 주석을 입력할 때 시간 비용을 줄이고, 주석 내용에 대한 객관성을 높일 수 있었다. 이를 통해 추후 콘텐츠 특화 검색 시스템에서도 유용하게 활용할 수 있을 것이라 기대한다.

차후 그래픽 장면에 대해 Precision과 Recall값을 올리기 위한 연구를 진행하고, 그래픽 장면과 더불어 인터뷰 장면, 인용 장면에 대해서도 자동으로 검출할 수 있는 방법을 설계 및 구현할 계획이다. 그리고 이러한 샷 유형 자동 검출을 통하여 사용자들이 원하는 샷 혹은 장면들만을 묶어서 볼 수 있는 비디오 요약에 대한 방법 또한 연구할 계획이다.

6. 참고 문헌

[1] KBS, *생로병사의 비밀*, Retrieved from <http://www.kbs.co.kr/1tv/sisa/health/>, 2014.
 [2] Bertini, M., Del Bimbo, A., Pala, P., "Indexing for Reuse of Tv News Shots", *Pattern Recognit.*, vol.35, no.3, pp.581-591, 2002.
 [3] B. S. Manjunath, Philippe Salembier, "Introduction to MPEG 7(Multimedia Content Description Language)", WILEY, pp.194-197.
 [4] G. Sankar, V. Zhao and Y. H. Yang., "Feature Based Classification of Computer Graphics and Real Images.", *Proceeding of IEEE ICASSP*, pp. 1513-1516, 2009.