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

공 학 석 사 학 위 논 문

실시간 장면 전환 검출을 위한 적응적 임계값과 가중치 설정에 관한 연구

2009년 2월

부 경 대 학 교 대 학 원

컴 퓨 터 공 학 과

김 원 희

공 학 석 사 학 위 논 문

실시간 장면 전환 검출을 위한 적응적 임계값과 가중치 설정에 관한 연구

지도교수 김 종 남

이 논문을 공학석사 학위논문으로 제출함

2009년 2월

부 경 대 학 교 대 학 원

컴 퓨 터 공 학 과

김 원 희

김원희의 공학석사 학위논문을 인준함

2009년 2월 25일

주　　심　공학박사　송 하 주　인

위　　원　공학박사　신 봉 기　인

위　　원　공학박사　김 종 남　인

목 차

표목차 ····· iv

그림목차 ····· v

Abstract ····· vi

Ⅰ. 서론 ····· 1

Ⅱ. 관련연구 ····· 3

1. 장면 전환 검출 기술 분야의 기본 정의 ····· 4

가. 장면 전환 검출 기술의 정의 ····· 4

나. 장면의 정의 ····· 5

다. 장면 전환의 종류 ····· 6

라. 장면 전환 검출 기술의 사용분야 ····· 11

2. 장면 전환 검출을 위한 프레임 특징 측정 방법 ····· 12

가. 화소 기반의 방법 ····· 12

나. 히스토그램 기반의 방법 ····· 16

다. 블록 기반의 방법 ····· 20

라. 특징 기반의 방법 ····· 23

마. 움직임 기반의 방법 ····· 26

바. 다중 척도 기반의 방법 ····· 28

사. DC 계수 기반의 방법 ····· 29

아. 비트열 기반의 방법 ····· 30

자. 픽처 기반의 방법 ····· 31

3. 장면 전환 검출을 위한 임계값 설정 방법 ····· 32
가. 실험적 방법 ····· 32
나. 분석적 방법 ····· 34
다. 적응적 방법 ····· 35
4. 장면 전환 검출에서의 고속 계산을 위한 방법 ····· 37
가. 공간적 관점의 계산량 감소 방법 ····· 37
나. 시간적 관점의 계산량 감소 방법 ····· 39
5. 장면 전환 검출 기술의 성능 평가 ····· 43
가. 시간 평가 ····· 43
나. 정확도 평가 척도 ····· 44
다. 성능 평가 척도를 이용한 알고리즘의 평가 ····· 46
6. 장면 전환 검출 기술의 문제점 분석 ····· 48
가. 기존 연구들의 문제점 ····· 48
나. 본 연구의 목적 ····· 51

Ⅲ. 제안사항 방법 ····· 52
1. 프레임 평균값을 이용한 장면 전환 검출 방법 ····· 53
가. 차분 프레임 평균값 ····· 53
나. 적응적 임계값 설정 방법 ····· 54
2. 가중치 분산을 이용한 장면 전환 검출 방법 ····· 56
가. 가중치 분산 ····· 56
나. 적응적 임계값 설정 방법 ····· 60
다. 임계값 가중치 결정 방법 ····· 61
3. 히스토그램을 이용한 완전 자동적 장면 전환 검출 ····· 63
가. 차분 히스토그램 평균값 ····· 63
나. 적응적 임계값 설정 방법 ····· 64
다. 임계값 가중치의 자동적 결정 방법 ····· 65
4. 실시간 장면 전환 검출 모델 ····· 66
가. 특징값 ····· 67

나. 임계값 ····· 68
다. 가중치 ····· 70
5. 예제를 통한 제안하는 모델의 설명 ····· 72
가. 참조구간 ····· 72
나. 가중치 ····· 75
다. 전체 흐름도 ····· 78
6. 개선된 장면 전환 검출 방법 ····· 80
가. 빈 데이터 영역 처리 ····· 80
나. 페이드 아웃 효과에 의한 장면 변화 처리 ····· 82

Ⅳ. 실험 및 결과 분석 ····· 83
1. 실험 환경 ····· 84
가. 사전 정의되는 사항 ····· 84
나. 실험 환경 ····· 87
다. 실험 내용 ····· 89
2. 실험 결과 및 분석 ····· 90
가. 차분 프레임의 평균을 이용한 방법의 성능 평가 ····· 90
나. 가중치 분산을 이용한 방법의 성능 평가 ····· 92
다. 차분 히스토그램의 평균을 이용한 방법의 성능 평가 ····· 95
라. 개선된 방법의 성능 평가 ····· 98
3. 제안한 방법의 PMP에서의 구현 ····· 101
가. PMP 구현의 특징 ····· 101
나. PMP의 구성과 동작 ····· 103
다. PMP에서의 장면 전환 검출 구현 ····· 105
라. PMP에서의 장면 전환 검출 결과 ····· 107

Ⅴ. 결론 ····· 108

참고문헌 ····· 109

표 차 례

표 1. 화소 기반 방법들의 계산복잡도와 계산시간 비교 ····· 15
표 2. 히스토그램 기반 방법들의 계산복잡도와 계산시간 비교 ····· 19
표 3. 블록 기반 방법들의 계산복잡도와 계산시간 비교 ····· 22
표 4. 특징 기반 방법들의 계산복잡도와 계산시간 비교 ····· 25
표 5. 움직임 기반 방법들의 계산복잡도와 계산시간 비교 ····· 27
표 6. 다중 척도 기반 방법들의 계산복잡도와 계산시간 비교 ····· 28
표 7. 장면 전환 검출 결과의 경우의 수 ····· 46
표 8. 임계값 가중치의 필요성 ····· 77
표 9. 제안하는 방법들의 내용 ····· 84
표 10. 비교하는 방법들의 내용 ····· 86
표 11. 실험에 사용된 영상 종류와 장면 전환 개수 ····· 87
표 12. 장면 전환 프레임의 종류와 분포 ····· 88
표 13. 실험 내용 ····· 89
표 14. 실험 결과 : E1-1 ····· 91
표 15. 실험 결과 : E2-1 ····· 93
표 16. 실험 결과 : E2-2 ····· 93
표 17. 실험 결과 : E2-3 ····· 94
표 18. 실험 결과 : E3-1 ····· 95
표 19. 실험 결과 : E3-2 ····· 96
표 20. 실험 결과 : E3-3 ····· 97
표 21. 실험 결과 : E4-1 ····· 98
표 22. 실험 결과 : E4-2 ····· 99

그 림 차 례

그림 1. 장면 전환 검출 기술의 정의 ····· 4
그림 2. 동영상의 계층적 구조 ····· 5
그림 3. 샷과 신의 구성 ····· 6
그림 4. 일반적인 급진적인 장면 전환 ····· 7
그림 5. 점진적 특성을 가지는 급진적인 장면 전환 ····· 8
그림 6. 페이드아웃 효과 ····· 9
그림 7. 페이드인 효과 ····· 9
그림 8. 디졸브 효과 ····· 10
그림 9. 와이프 효과 ····· 10
그림 10. 장면 전환 검출 기술의 사용분야 ····· 11
그림 11. 공간적 계산량 감소 방법 ····· 38
그림 12. 선형 검색 방법 ····· 40
그림 13. 비선형 검색 방법 ····· 41
그림 14. 이진 검색 방법 ····· 41
그림 15. 가중치 분산과 일반적 분산의 분포 비교 ····· 58
그림 16. 실시간 장면 전환 검출 모델의 구성 ····· 66
그림 17. 프레임 이동에 따른 가변 참조 구간의 변경 ····· 73
그림 18. 참조 구간에 따른 장면 전환 검출 비교 ····· 74
그림 19. 가중치에 따른 장면 전환 검출 과정 ····· 75
그림 20. 제안한 모델의 흐름도 ····· 78
그림 21. 비디오 프레임의 시각적 구성 ····· 80
그림 22. 가변 가중치 설정 코드 ····· 82
그림 23. 구현에 사용된 PMP ····· 102
그림 24. PMP의 내부 스레드 ····· 103
그림 25. PMP에서의 두 가지 장면 전환 검출 모드 ····· 105
그림 26. PMP에서 장면 전환 검출 결과 ····· 107

A Study on Setting Adaptive Thresholds and Weighting Factors
for Real-time Shot Change Detection

Won Hee Kim

Department of Computer Engineering, The Graduate School,
Pukyong National University

Abstract

Shot change detection is a main technique for automatic temporal segmentation of video. In this paper, we survey conventional researches which include various measurement of feature extraction, threshold setting schemes, techniques of computation reduction, evaluation measurements of algorithms, and various problems. Conventional methods require still solutions of the following problems: exact frame feature extractor, adaptive threshold setting scheme, fast and real-time processing, and implementation on hardware device. To solve above problems, we propose three algorithms to detect shot change. Particularly, we propose the weighting variance to extract frames feature, adaptive threshold setting scheme using average of feature value on variable sliding window, and automatic decision technique of threshold weighting factor. Lastly, we present the real-time shot change detection model based on proposed algorithms. Our experiments obtained the detection rate of about 0.9783 in overall accuracy, demonstrating a higher detection rate and a computational amount less than the conventional algorithms. We verified the real-time operation of shot change detection by implementing our algorithm on the PMP of Homecast. The proposed algorithm will be helpful in searching video data on portable media players such as PMPs and cellular phones.

Ⅰ. 서 론

장면 전환 검출 기술은 비디오 데이터의 장면 변화가 발생하는 경계를 검출하는 기술이다. 장면은 물리적 또는 의미적 단위의 프레임들의 집합으로 정의된다. 장면의 변화는 연속되는 장면이 급격하게 변화하는 급진적 변화와 서서히 변화하는 점진적 변화로 구분된다. 장면 전환 검출 기술은 효과적인 비디오 데이터 관리를 위한 기반 기술로 사용되며, 비디오 데이터의 분할, 색인, 검색, 요약, 브라우징과 같은 작업을 수행하기 위해서 필수적으로 사용되는 기술이다.

장면 전환 검출을 위해서는 프레임의 특징을 추출하는 추출자와 추출된 특성에 따라 변화 여부를 판단하는 판단자가 필요하다. 프레임 특징을 측정하는 방법에는 공간 영역의 화소 기반, 히스토그램 기반, 블록 기반, 특징 기반, 움직임 기반, 다중 척도 기반 등의 6가지로 방법으로 분류할 수 있으며, 변환 영역의 DC 계수 기반, 비트열 기반, 픽처 기반의 3가지 방법으로 구분할 수 있다. 이런 추출된 특징을 바탕으로 변화 여부를 판단하는 역할을 하는 것이 임계값이다. 사람이 동영상 데이터를 시청하면서 장면이 변화된 것을 인지하는 것과 같이 판단의 기능을 하는 요소로써 장면 전환 검출 알고리즘의 정확도를 결정하게 된다. 임계값은 생성 원리에 따라서 실험적, 분석적, 적응적 방법의 세 가지로 분류할 수 있다. 고속의 장면 전환 검출을 위해서 공간적 관점과 시간적 관점의 계산량 감소 기법을 적용할 수 있다. 이렇게 만들어진 알고리즘은 정확도와 시간의 측면에서 여러 가지 척도를 통해서 평가된다.

1990년대 이후로 많은 연구가 이루어지면서 여러 가지 문제점들이 제시되고 해결되어왔다. 본 논문에서는 장면 변화가 발생하는 프레임의 특

징을 정확하게 추출해 내는 특징 추출 척도와 다양한 영상에 대해서 적응적으로 사용될 수 있는 임계값 설정 방법, 그리고 고속 실시간 처리를 위한 알고리즘 개선과 실제 하드웨어 장치에서의 구현 등 3가지 요소에 대한 문제 해결을 목표로 하여 연구를 진행한다. 최종적으로는 다양한 영상에서 정확한 장면 전환 검출이 가능하며 휴대용 장치에서도 적용할 수 있는 실시간처리가 가능한 적응적인 장면 전환 검출 모델을 제안하고자 한다.

본 논문에서는 크게 3가지의 장면 전환 검출 방법을 제안하였다. 정확한 프레임 특징 추출을 위한 가중치 분산을 이용하는 방법, 가변 참조 구간을 이용한 적응적인 임계값 설정 방법, 자동적 임계값 가중치 결정을 통한 완전 자동적 장면 전환 검출 방법 등을 제안하였고, 이를 바탕으로 한 실시간 장면 전환 검출 모델을 제안하였다. 9가지의 실험을 통해서 제안한 방법이 기존 방법들보다 정확하면서도 고속으로 처리되는 알고리즘임을 확인하였다. 그리고 상용 휴대용 미디어 재생 장치에 제안한 방법을 적용하여 실시간 장면 전환 검출이 가능함을 검증하였다.

본 논문의 구성은 다음과 같다. Ⅱ장에서는 장면 전환 검출과 관련된 배경 지식과 기존의 다양한 관련 연구 및 기존 연구들의 문제점과 본 논문에서 해결하고자 하는 문제점들을 제시하였다. Ⅲ장에서는 제안하는 장면 전환 검출 알고리즘과 이를 토대로 설계된 실시간 장면 전환 검출 모델, 그리고 개선된 알고리즘에 대해서 기술하고 있다. Ⅳ장에서는 제안한 방법들의 성능 평가를 위한 실험 환경과 실험 내용 및 실험 결과들을 정리하였고, 휴대용 미디어 재생 장치에 구현한 방법과 결과를 서술하였다. 마지막으로 Ⅴ장에서는 결론을 맺으면서 논문을 마무리한다.

Ⅱ. 관련연구

본 장에서는 장면 전환 검출과 관련된 기존의 개념과 다양한 연구들에 대해서 서술하였다. 장면 전환 검출 기술의 연구를 위해서 사전에 알고 있어야 하는 제반사항과 용어의 정의를 설명하고 기술의 활용분야를 기술하였다. 또 장면 전환 검출을 위한 특징값 측정 방법, 임계값 설정 방법에 대해서 서술하고, 고속의 장면 전환 검출을 위한 방법들을 설명하였다. 그리고 장면 전환 검출 기술의 성능 평가를 위한 척도들을 알아보고, 기존의 연구들의 분석과 해결하지 못한 문제점 등을 서술하고 제안하는 연구의 목적을 서술하였다.

본 장의 구성은 다음과 같다. 1절에서는 장면 전환 검출 기술과 관련된 기본 정의와 개념들에 대해서 서술하였고, 2절에서는 장면 전환 검출을 위한 프레임 특징값을 측정하는 기존의 다양한 연구들에 대해서 정리하였다. 3절에서는 장면 전환 검출을 위한 임계값 설정 방법에 대해서 기술하였고, 4절에서는 고속 장면 전환 검출을 위한 계산량 감소와 관련된 기존 연구들을 서술하였다. 5절에서는 장면 전환 검출 기술을 성능 평가 척도들을 알아보고, 마지막으로 6절에서는 기존의 연구들을 토대로 분석된 아직까지 해결되지 못한 문제점들을 제시하고, 문제점과 관련된 제안하는 연구의 목적을 정리하여 나타내었다.

1. 장면 전환 검출 기술 분야의 기본 정의

가. 장면 전환 검출 기술의 정의

장면 전환 검출 기술은 비디오 데이터의 장면 변화가 발생하는 경계를 검출하는 기술이다[1]. 장면은 프레임들의 집합으로 구성되어 있으므로 장면 전환 검출 기술은 이전 장면과 다음 장면이 바뀌는 지점의 프레임을 검출해내는 기술을 의미한다. 일반적으로 장면 전환 발생 프레임은 다음 장면이 시작되는 첫 번째 프레임으로 정의된다. 장면 전환 검출 기술의 정의를 도식화하여 그림 1에 나타내었다. 그림 1에서와 같이 장면 전환 검출은 동영상 데이터를 장면 전환 검출 알고리즘을 통해서 자동적으로 장면이 변화하는 프레임을 검출하는 기술이다.

그림 1. 장면 전환 검출 기술의 정의

Figure 1. The definition of shot change detection technique

나. 장면의 정의

‘가’항에서 장면 전환 검출 기술의 정의를 설명하면서 ‘장면’이라는 용어가 언급되었다. 장면의 정의를 알아보기에 앞서 동영상의 계층적 구조를 살펴보면 다음과 같다[2].

하나의 비디오(video) 데이터는 시퀀스(sequence)들의 집합으로 이루어진다. 시퀀스는 동영상 구조의 최상위 계층으로 신(scene)들의 집합으로 구성된다. 하나의 신은 다수의 샷(shot)으로 이루어지며, 샷은 동영상 구조의 최소 단위인 프레임(frame)의 집합으로 구성된다. 동영상의 계층 구조를 그림 2에서 도식화하여 나타내었다. 그림 2에서와 같은 동영상 데이터는 시퀀스, 신, 샷, 프레임의 4단계의 계층적 구조를 가지는 것을 알 수 있다.

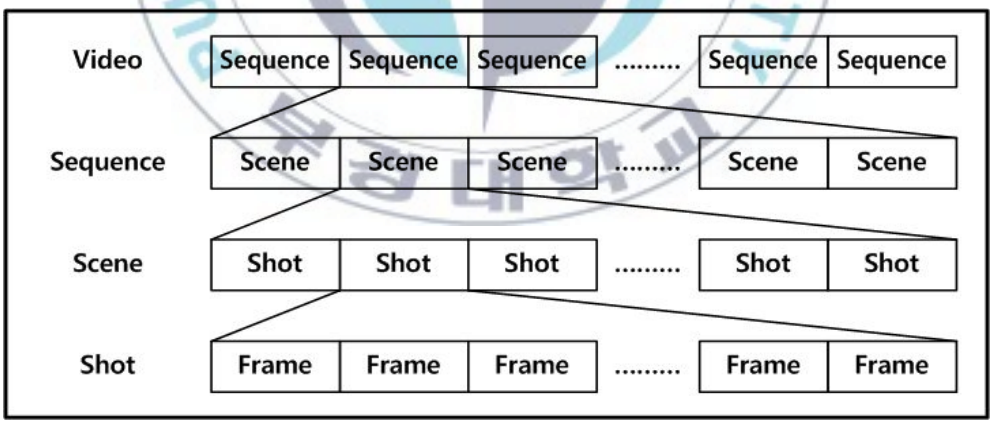


그림 2. 동영상의 계층적 구조

Figure 2. Hierarchical structure of moving picture

장면은 동일한 특성을 지니는 프레임들의 집합으로 정의할 수 있으며 두 가지로 분류된다[3]. 첫 번째로 장면은 비디오 영상을 획득하는 장치로부터 녹화되는 시점에서 한 번의 물리적인 녹화의 시작과 정지로 획득된 연속되는 프레임들의 집합으로 정의된다. 이것은 샷(shot)이라는 용어가 지니는 의미와 동일하다. 두 번째로 장면은 의미적으로 동일한 특성을 가지는 연속되는 프레임들의 집합으로 정의된다. 이것은 신(scene)이라는 용어의 의미와 동일하며 하나의 신은 하나 이상의 샷으로 구성된다.

샷과 신의 차이는 다음의 예를 통해 쉽게 알 수 있다. 먼저 동일한 공간에서의 두 화자의 대화하는 장면을 가정하면, 신은 두 화자의 대화가 시작해서 끝나는 프레임들의 전체 집합을 의미하고, 샷은 화자의 얼굴이 바뀌거나 두 명의 화자를 동시에 화면에 담을 때와 같이 카메라의 조작이 발생하는 구간내의 프레임들의 집합을 의미하게 된다. 샷과 신의 차이를 도식화하여 그림 3에서 나타내었다.

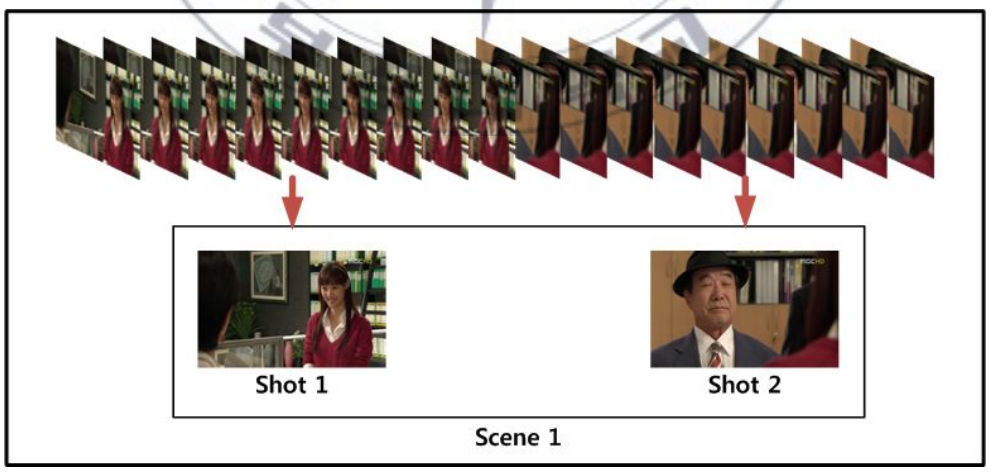


그림 3. 샷과 신의 구성
Figure 3. Structure of shot and scene

다. 장면 전환의 종류

장면이 전환되는 부분의 유형은 일반적으로 두 가지로 분류된다[4]. 첫 번째는 연속되는 장면이 급격하게 변화하는 급진적인 장면 전환이다. 급진적인 변화는 두 장면 사이의 단순한 시간적 연결로 구성된다. 일반적으로 변화하는 정도가 크기 때문에 비교적 장면 전환 검출이 쉽다. 두 번째는 연속되는 장면이 서서히 변화하는 점진적인 장면 전환이다. 점진적인 변화는 두 장면 사이의 일정한 프레임들이 영상처리 기술에 의해서 중복되어 있다. 일반적으로 변화하는 정도가 작고 여러 프레임들에 걸쳐서 나타나기 때문에 급진적인 변화보다 검출이 어렵고, 검출하는 방법도 복잡하다. 동영상의 구성은 대부분 급진적인 장면 전환으로 이루어지지만 장르의 특성이나 편집자의 의도에 따라서 점진적인 장면 전환의 비중이 높아지기도 한다.

(1) 일반적인 급진적인 장면 전환

가장 일반 적인 장면 전환의 유형으로 장면이 바뀌는 부분의 변화가 명확하다. 장면의 변화가 1프레임 만에 발생하는 것이 특징이며, 두 장면의 중복되는 부분이 존재하지 않는다.

그림 4. 일반적인 급진적인 장면 전환
Figure 4. General abrupt transition

(2) 점진적 특성을 가지는 급진적인 장면 전환

기존의 장면의 유형은 급진적인 전환과 점진적인 전환으로 분류되지만 두 가지 장면 전환의 특성을 동시에 가지는 장면 전환 유형도 존재한다. 이와 같은 장면의 전환은 본 논문에서 최초로 제시하는 것으로 점진적 특성을 가지는 급진적인 장면 전환으로 정의할 수 있다. 이런 장면 전환은 점진적 전환과 같이 영상처리에 의해 중복된 부분을 포함하고 있기 때문 급진적 장면 전환으로 정의할 수 없다. 반면 점진적 변화로 정의하기에는 그 변화의 정도가 매우 크기 때문에 서서히 점차적으로 변화하는 점진적인 전환의 특성에도 일치하지 않는다. 따라서 이런 장면 전환은 점진적 특성을 가지는 급진적 장면 전환으로 정의할 필요가 있다. 이렇게 따로 분류하는 것은 이런 장면 전환 유형이 동영상 편집자의 경향에 따라서 전체 동영상에서 높은 비중을 차지하는 경우가 있기 때문이다. 이런 장면 전환 유형은 급진적인 장면과 마찬가지로 비교적 쉽게 검출이 가능하지만 중복되는 프레임의 수와 변화 정도에 따라서 미검출이 발생할 확률이 있다. 일반적으로 중복되는 프레임은 1~3프레임 정도로 분포하기 때문에 장면 전환 검출 알고리즘의 특성에 따라서 장면 전환 프레임의 선택이 달라질 수 있다.

그림 5. 점진적 특성을 가지는 급진적인 장면 전환
Figure 5. Semi-abrupt transition

(3) 점진적인 장면 전환 - 페이드인/아웃(fade-in/out)

영상의 밝기가 점점 낮아져 화면이 어두워졌다가 다음 장면의 밝기가 증가하여 장면이 전환되는 효과를 의미한다. 일반적으로 페이드인과 페이드아웃은 연속적으로 발생되며 점진적 장면 전환 중에서는 비교적 검출이 용이하다.

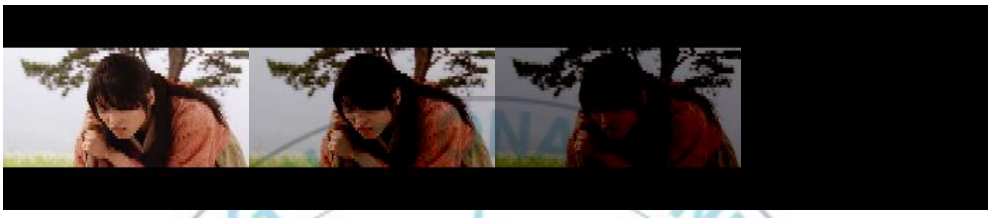


그림 6. 페이드아웃 효과
Figure 6. Fade-out effect

그림 7. 페이드인 효과
Figure 7. Fade-in effect

(4) 점진적인 장면 전환 - 디졸브(dissolve)

디졸브는 서로 다른 두 영상이 겹쳐 이전 영상의 밝기는 점점 감소하는 동시에 다음 영상의 밝기가 서서히 증가하면서 장면이 전환되는 효과를 의미한다. 2개의 화면을 합성하는 겹치기(overlapping)와 혼동될 수 있는데, 겹치기는 화면이 합성된 그대로 고정되는 반면 디졸브는 최종적으로 하나의 화면으로 종료된다. 디졸브 효과의 경우 변화 정도가 작고

중첩되는 프레임의 수가 늘어날수록 검출이 어려워지는 특성이 있다. 아래의 그림 8에서 디졸브 효과를 나타내었는데, 그림에서는 중첩되는 프레임이 2프레임에 불과하지만 실제로는 이보다 많은 프레임의 중첩으로 구성되어 있다.

그림 8. 디졸브 효과
Figure 8. Dissolve effect

(5) 점진적인 장면 전환 - 와이프(wipe)

상하, 좌우, 원형, 마름모형, 타원형 등의 다양한 형상으로 이전 영상을 밀어내면서 다음 영상이 등장하거나 커튼을 여는 것과 같이 화면을 전환하는 효과이다. 편집 효과에 따라서 다양하게 연출될 수 있으며 점진적 전환 중에서도 비교적 검출이 어렵다. 일반적으로 디졸브보다는 짧은 프레임 중첩으로 구성된다.

그림 9. 와이프 효과
Figure 9. Wipe effect

라. 장면 전환 검출 기술의 사용분야

장면 전환 검출 기술은 효과적인 비디오 데이터 관리를 위한 기반 기술로 사용된다. 비디오 데이터의 분할, 색인, 검색, 요약, 브라우징과 같은 작업을 수행하기 위해서 필수적으로 사용된다[5].

장면 전환 검출이 실제 응용환경에서 사용될 수 있는 경우는 다음과 같다. 분할 기능에서는 워터마킹이나 영상압축 기술에 사용되어 장면 단위로 워터마킹을 삽입하거나 영상을 압축하는데 장면 전환 검출이 이용된다. 색인 기능에서는 장면 전환 프레임을 기준으로 해서 미디어 재생 소프트웨어의 빨리 감기나 되감기 기능의 구현에 장면 전환 검출이 사용된다. 검색 기능에서는 감시 시스템에서 요구하는 장면을 저장된 장면 전환 프레임과의 비교를 통해서 고속으로 찾을 수 있는 기능에 주요 기술로 사용된다. 요약 기능에서는 장면 전환 프레임들만을 추출하여 스토리 보드를 구성하여 전체 비디오를 파악하기 쉽도록 하는데 이용할 수 있다. 브라우징 기능은 휴대용 미디어 재생 장치나 셋탑박스 등에서 장면이 변화된 지점들만 출력하는 기능을 위해서 장면 전환 검출 기술이 사용된다.

그림 10. 장면 전환 검출 기술의 사용분야

Figure 10. Application of shot change detection technique

2. 장면 전환 검출을 위한 프레임 특징 측정 방법

장면 전환 검출을 위해서는 프레임의 특징을 정량화할 필요가 있다. 사람의 인지 능력과는 다르게 장면을 판단하기 위해서는 영상의 특징을 수치화하는 과정이 필요하다. 1990년대 이후로 프레임의 특징을 측정하는 다양한 방법들이 연구되어 왔다[6]. 이런 방법들은 공간 영역과 변환 영역에서의 방법들로 구분할 수 있다. 공간 영역에서는 화소 기반, 히스토그램 기반, 블록 기반, 특징 기반, 움직임 기반, 다중 척도 기반 등의 6가지로 방법으로 분류할 수 있고, 변환 영역에서는 DC 계수 기반, 비트열 기반, 픽처 기반의 3가지 방법으로 구분할 수 있다. 본 절에서는 기존의 프레임 특징 측정 방법을 9가지로 구분하여 정리하였다.

가. 화소 기반의 방법

화소 기반의 방법은 연속되는 두 프레임의 상관관계를 나타내는 가장 간단한 방법이다. 화소 기반의 방법도 다양하게 연구되었는데, 1991년 Nagasaka와 Tanaka는 연속되는 프레임에서 각 프레임의 전체 화소값의 합의 절대값 차이(simple interframe difference)를 프레임의 특징으로 측정하였다[7]. 이렇게 추출된 특징값과 임의의 임계값과의 비교를 통해서 장면 전환을 판단한다. 이 방법은 식 (1)로 정의된다.

$$\left(\left| \sum_{i=1}^{X} \sum_{j=1}^{Y} P(I_t, i, j) - \sum_{i=1}^{X} \sum_{j=1}^{Y} P(I_{t-1}, i, j) \right| \right) > T \qquad (1)$$

여기에서, X와 Y는 프레임의 가로와 세로의 크기, $P(I_t,i,j)$는 I_t번째 프레임에서의 좌표 (i,j)의 밝기값을 나타내며, 이때 각각 $1 \le i \le X$, $1 \le j \le Y$ 이다.

또, Nagasaka와 Tanaka는 화소 기반의 다른 방법을 제안하였다[7]. 연속되는 두 프레임의 대응되는 화소의 절대값 차이의 합(template matching)을 프레임의 특징값으로 정의하고 임계값과의 비교를 통해서 장면 전환을 검출하였다. 이 방법은 아래의 식 (2)에서 정의하였다.

$$\left(\sum_{i=1}^{X} \sum_{j=1}^{Y} |P(I_t,i,j) - P(I_{t-1},i,j)| \right) > T \tag{2}$$

컬러 영상에 대한 화소 기반의 방법(color template matching)도 적용할 수 있는데 식 (3)에서 정의하고 있다[7].

$$\left(\sum_{i=1}^{X} \sum_{j=1}^{Y} \sum_{k=1}^{3} |P(I_t,C_k,i,j) - P(I_{t-1},C_k,i,j)| \right) > T \tag{3}$$

여기에서 $P(I_t,C_k,i,j)$는 컬러 영상에서의 I_t번째 프레임에서의 좌표 (i,j)의 k번째 컬러 공간의 밝기값을 나타내며, 이 때 $k=1,2,3$ 이다. 각 컬러 요소는 RGB 컬러 공간을 의미한다.

Zhang 등도 1993년 이진대수를 사용하는 화소 기반의 방법(boolean difference)을 제안하였다[8]. 이 방법은 연속되는 두 프레임에서 대응되는 화소의 값이 동일한 경우에는 0, 다른 경우 1을 값으로 가지는 이진 차프레임을 계산하여 이진 차프레임의 합을 프레임의 특징값으로 사용한다. 이 방법은 식 (4)로 나타낼 수 있다.

$$D(I_t, I_{t-1}, i, j) = \begin{cases} 1 & if\ |P(I_t, i, j) \neq P(I_{t-1}, i, j)| \\ 0 & otherwise \end{cases} \quad (4)$$

이 방법은 화소 밝기값의 일정량의 변화를 고려해서 더 좋은 측정 방법(improved boolean difference)으로 개선할 수 있으며 식 (5)에서 개선된 방법을 나타내었다[8].

$$D(I_t, I_{t-1}, i, j) = \begin{cases} 1 & if\ |P(I_t, i, j) - P(I_{t-1}, i, j)| > T_D \\ 0 & otherwise \end{cases} \quad (5)$$

여기에서 T_D는 허용 한계 값을 의미한다. 개선된 방법은 T_D보다 낮은 값의 차이는 동일한 값으로 간주하여 이진 차프레임을 생성한다. 이렇게 생성된 이진 차프레임의 합을 최종적으로 프레임의 특징값으로 정의하고 임계값 T와 비교하여 장면 전환을 판단한다. 이 과정은 식 (6)을 통해서 나타내었다.

$$\left(\sum_{i=1}^{X} \sum_{j=1}^{Y} D(I_t, I_{t-1}, i, j) \right) > T \quad (6)$$

Zhang 등은 이 방법에서 움직임으로부터 발생하는 오검출을 줄이기 위해 D값을 계산하기 전에 3×3 필터를 통해서 이진 차프레임을 평탄하게 만드는 개선된 방법을 제안하였다.

2001년 Ren 등은 장면 전환 검출을 위해서 연속되는 두 프레임의 정규화된 차분 에너지(normalized difference energy)와 절대값 차이의 정규화합(normalized sum of a absolute differences)을 이용하였다[9]. 이 방법을 식 (7)과 식 (8)에서 정의하고 있다.

$$\left(\frac{\sum_{i=1}^{X}\sum_{j=1}^{Y}(P(I_t,i,j)-P(I_{t-1},i,j))^2}{\left(\sum_{i=1}^{X}\sum_{j=1}^{Y}P(I_t,i,j)^2\right)\left(\sum_{i=1}^{X}\sum_{j=1}^{Y}P(I_{t-1},i,j)^2\right)} \right) > T \quad (7)$$

$$\left(\frac{\sum_{i=1}^{X}\sum_{j=1}^{Y}|P(I_t,i,j)-P(I_{t-1},i,j)|}{\sum_{i=1}^{X}\sum_{j=1}^{Y}P(I_t,i,j)+\sum_{i=1}^{X}\sum_{j=1}^{Y}P(I_{t-1},i,j)} \right) > T \quad (8)$$

이상 언급한 화소 기반의 방법들은 계산량이 적고 구현이 간단하지만 작은 영역에서의 큰 변화나 큰 영역에서의 작은 변화를 구별할 수 없는 단점이 있다[10]. 카메라와 객체의 동작, 명암 변화에 민감하게 반응하여 오검출이 발생할 수 있다[11]. 아래의 표 1에서는 화소 기반의 방법들의 계산복잡도와 계산시간을 비교하여 나타내었다.

표 1. 화소 기반 방법들의 계산복잡도와 계산시간 비교

Table 1. Comparison of computation complexity and running-time of methods based on pixel

Methods and reference	Complexity	Time
Simple interframe difference[7]	$O(p)$	1.00
Template matching[7]	$O(3p)$	2.11
Color template matching[7]	$O(9p)$	5.29
Boolean difference[8]	$O(2p)$	1.80
Improved Boolean difference[8]	$O(4p)$	2.60
Normalized difference energy[9]	$O(5p)$	3.00
Normalized sum of a absolute differences[9]	$O(5p)$	2.58

나. 히스토그램 기반의 방법

히스토그램은 장면 전환 검출의 대표적인 방법으로 영상의 밝기나 컬러 정보의 확률 분포도인 히스토그램을 구성한 후 비교를 통해서 장면 전환 여부를 결정하는 방법이다. 히스토그램 기반의 방법은 화소 기반의 방법에서 문제가 되었던 움직임에 민감하다는 단점을 어느 정도 극복한다. 이것은 영상의 전체적인 분위기를 파악하는 것으로 영상을 이해하는데 있어서 영상 내에 물체나 카메라의 움직임이 존재하는 경우라도 밝기 분포나 색상 분포는 거의 일정하기 때문이다.

Tonomura와 Abe는 화소 밝기값 히스토그램에 기반한 방법(histogram difference)을 제안하였다[12]. 영상은 히스토그램 사이의 거리 계산에 의해서 비교된다. 식 (9)에서 정의를 나타내고 있다.

$$\left(\sum_{v=0}^{V} |H(I_t, v) - H(I_{t-1}, v)| \right) > T \tag{9}$$

Nagasaka와 Tanaka는 위 방법과 유사하지만 RGB 공간의 컬러 히스토그램을 위한 64개의 빈을 사용하는 방법을 제안하였다[7]. 이 방법은 식 (10)에서 정의하고 있다.

$$\left(\sum_{v=0}^{63} |H_{64}(I_t, v) - H_{64}(I_{t-1}, v)| \right) > T \tag{10}$$

Gargi와 Kasturi는 HSV나 YIQ 등의 다른 컬러 공간에 히스토그램을 적용한 방법(chrominance histogram difference)을 제안하였다[13]. 이 방

법에서는 두 가지의 컬러 콤포넌트를 사용하며 장면 전환은 식 (11)과 같이 정의된다.

$$\left(\sum_{k=1}^{2}\sum_{v=0}^{V}|H(I_t,C_k,v)-H(I_{t-1},C_k,v)|\right)>T \tag{11}$$

Pye 등은 RGB 공간의 3가지 컬러 콤포넌트의 개별적인 차이를 계산하기 위한 방법(highest histogram difference)을 제안한다[14]. 장면 전환 검출을 위해서 3가지 값의 최대값과 임계값을 비교하는데 식 (12)에서 이 방법을 정의하였다.

$$\left(\max_{k\in\{R,G,B\}}\sum_{v=0}^{V}|H(I_t,C_k,v)-H(I_{t-1},C_k,v)|\right)>T \tag{12}$$

Gargi와 Kasturi는 가중치가 적용된 히스토그램의 평균 컬러의 차이에 기반한 방법을 제안하였다[13]. 아래의 식 (13)과 같이 정의된다.

$$\left(\sum_{k=1}^{3}\left(\sum_{v=0}^{V}H(I_t,C_k,v)v-\sum_{v=0}^{V}H(I_{t-1},C_k,v)v\right)^2\right)>T \tag{13}$$

히스토그램의 특성을 이용하는 다른 방법으로 히스토그램 교차법이 있다. 히스토그램 교차법(histogram intersection)은 두 히스토그램의 임의의 연산 함수를 이용해서 연속되는 프레임의 유사도를 계산한다[15]. 최소값 함수를 이용한 정의를 식 (14)에 나타내었다.

$$\left(1-\frac{1}{XY}\sum_{v=0}^{V} min(H(I_t,v),H(I_{t-1},v))\right)> T \tag{14}$$

식 (15)에서는 최소값 함수와 최대값 함수를 이용한 다른 방법의 히스토그램 교차법(normalized histogram intersection)을 정의하고 있다[16].

$$\left(1-\frac{1}{XY}\sum_{v=0}^{V} \frac{min(H(I_t,v),H(I_{t-1},v))}{max(H(I_t,v),H(I_{t-1},v))}\right)> T \tag{15}$$

히스토그램을 기반으로 하는 다른 방법으로 Nagasaka와 Tanaka가 제안한 Chi-square 검정(original chi-square test)이 있다[7]. 이 방법은 두 프레임 사이의 차이 값을 강조할 뿐만 아니라 카메라나 객체의 움직임을 강조할 수 있다. 식 (16)에서 이 방법을 정의하고 있다.

$$\left(\sum_{v=0}^{63} \frac{(H_{64}(I_t,v)-H_{64}(I_{t-1},v))^2}{H_{64}(I_t,v)}\right)> T \tag{16}$$

여기에서 만약 $H_{64}(I_t,v)=0$인 경우에는 $H_{64}(I_{t-1},v)=0$로 대체하여 사용한다. Dailianas 등은 개선된 chi-square 검정(chi-square test normalized by max)을 제안하였는데 식 (17)과 같이 나타낼 수 있다[17].

$$\left(\sum_{v=0}^{V} \frac{(H(I_t,v)-H(I_{t-1},v))^2}{max(H(I_t,v),H(I_{t-1},v))}\right)> T \tag{17}$$

하지만 이런 chi-square 검정은 기존의 히스토그램 방법에 비하여 정확성이 낮고 계산량이 증가하는 단점이 있다.

이상 언급한 히스토그램 기반의 방법 외에도 블록 단위의 지역적 히스토그램을 이용하는 방법, 누적 히스토그램을 이용하는 방법 등의 다양한 방법들이 존재한다[18]. 히스토그램의 특징을 이용하는 방법들은 일반적으로 계산량이 작으면서도 프레임의 특성을 잘 표현해서 비교적 정확한 장면 전환 검출을 수행할 수 있다[19]. 아래의 표 2에서는 히스토그램 기반의 방법들의 계산복잡도를 비교하여 나타내었다.

비교적 응용 환경에서도 많이 사용되고 있지만 히스토그램 기반의 방법에도 몇 가지 문제점이 존재한다. 우선 영상에서 차지하는 물체의 비율이 낮고 장면 간에 배경의 변화가 없는 경우 또는 서로 다른 장면에 존재하는 프레임이더라도 비슷한 밝기 분포를 가지는 경우에는 히스토그램으로 구분하기 모호하다. 따라서 이런 부분에서 미검출이 발생하게 된다. 이 외에도 큰 물체가 갑자기 등장하는 장면이나 밝기가 급격하게 변화하는 부분에서 유사한 장면에 대한 오검출이 발생된다.

표 2. 히스토그램 기반 방법들의 계산복잡도와 계산시간 비교
Table 2. Comparison of computation complexity and running-time of methods based on histogram

Methods and reference	Complexity	Time
Histogram difference[12]	$O(3n)$	4.29
Chrominance histogram difference[13]	$O(6n)$	55.62
Highest histogram difference[14]	$O(9n)$	16.47
Histogram intersection[15]	$O(2n)$	4.29
Normalized histogram intersection[16]	$O(4n)$	12.18
Original chi-square test[7]	$O(5n)$	12.29
Chi-square test Normalized by max[17]	$O(5n)$	12.27

다. 블록 기반의 방법

비디오 프레임의 블록 단위 샘플링은 장면 전환의 정확도를 향상시키지만 계산량을 증가시키는 특성을 가진다. 블록 단위의 비교에서는 분할된 블록에서 화소값이나 히스토그램을 이용한 알고리즘을 적용할 수 있다. 블록 기반의 방법의 장점은 잡음과 카메라 또는 객체 움직임에 대한 둔감함이다[20].

Kasturi와 Jain은 블록화된 영상들의 유사도 검정(similarity ratio)을 제안하였는데, 이 방법은 연속되는 프레임의 대응되는 블록 쌍들을 비교하였다[21]. 유사도는 식 (18)과 같이 블록에서의 평균과 특징을 이용해서 계산된다.

$$L(I_t, I_{t-1}, b) = \frac{((\sigma_{t,b}^2 + \sigma_{t-1,b}^2)/2 + ((\mu_{t,b} - \mu_{t-1,b})/2)^2)^2}{\sigma_{t,b}^2 \sigma_{t-1,b}^2} \tag{18}$$

여기에서 $\mu_{t,b}$와 $\sigma_{t,b}^2$는 영상 I_t의 b블록의 평균과 분산을 각각 의미한다. 위 방법으로 구해진 블록의 유사도는 임의의 임계값을 이용해서 이진 프레임을 생성한다. 이 과정을 식 (19)에서 나타내었다.

$$L_d(I_t, I_{t-1}, b) = \begin{cases} 1 & if\ \ L(I_t, I_{t-1}, b) > T_D \\ 0 & in\ other\ cases \end{cases} \tag{19}$$

여기에서 T_D는 허용 한계값을 의미하며 최종적인 장면 전환 판단은 식 (20)에 의해서 이루어진다.

$$\sum_{b=1}^{B} c_b L_D(I_t, I_{t-1}, b) > T \tag{20}$$

여기에서 c_b는 블록의 중요도를 나타내는 가중치로 사용되며 일반적으로 모든 블록에서 1로 적용되어 사용되어진다.

다른 잘 알려진 방법으로는 Yakimovsky가 제안한 유사도 측정 방법(Yakimovsky likelihood ratio)이 있다[9]. 각 블록의 유사도는 식 (21)과 같이 계산된다.

$$L'(I_t, I_{t-1}, b) = \frac{(\sigma^2_{\{t,t-1\},b})^2}{\sigma^2_{t-1,b}\sigma^2_{t,b}} \tag{21}$$

여기에서 $\sigma^2_{\{t,t-1\},b}$는 연속되는 두 프레임의 블록으로부터 함께 계산된 분산을 나타낸다.

Swanberg 등은 블록 기반 방법에서 히스토그램(color histogram comparison)을 사용하였다[22]. 프레임의 블록에서 계산된 컬러 히스토그램을 이용해서 장면 전환을 검출하였다. 식 (22)에서 정의를 나타내었다.

$$\left(\sum_{k=1}^{3} \sum_{b=1}^{B} \sum_{v=0}^{V} \frac{(H(I_t, b, C_k, v) - H(I_{t-1}, b, C_k, v))^2}{H(I_t, b, C_k, v) + H(I_{t-1}, b, C_k, v)} \right) > T \tag{22}$$

이 외에도 4×4 블록에서 chi-square 검정을 이용하는 방법(selective color histogram comparison), 히스토그램의 차이와 유사율을 동시에 사용하는 방법(combination of histogram difference and likelihood rate), 이웃 컬러 비율을 이용하는 방법(neighborhood color ratio), 블록의 비유

사도를 개선시키는 방법(evolution of block dissimilarity) 등 다양한 방법들이 존재한다[23-26]. 아래의 표 3에서는 블록 기반의 방법들의 계산 복잡도를 비교하여 나타내었다.

블록 기반의 방법들은 객체의 움직임이나 잡음에 대해서 비교적 강인한 검출을 하지만 상대적으로 많은 연산시간이 소요된다. 또한 블록간 비교에 사용되는 임계값과 프레임간 비교에 사용되는 임계값 등 이중 처리가 많기 때문에 적절한 임계값 설정이 어렵다는 단점이 있다.

표 3. 블록 기반 방법들의 계산복잡도와 계산시간 비교

Table 3. Comparison of computation complexity and running-time of methods based on block

Methods and reference	Complexity	Time
Similarity ratio[21]	$O(3p+15b)$	85.29
Yakimovsky likelihood ratio[9]	$O(3p+13b)$	88.67
Color histogram comparison[22]	$O(15nb+3b)$	206.42
Selective RGB histogram comparison[7]	$O(2nb+0.5b)$	208.42
Selective HSV histogram comparison[27]	$O(6nb+2b+4p)$	300.00
Neighborhood color ratio[25]	$O(5p+8b)$	45.93
HSV block dissimilarity[26]	$O(2p+8b)$	114.25

라. 특징 기반의 방법

이전까지의 방법들 역시 프레임의 특징을 추출해서 장면 전환 검출을 수행하지만 대부분의 방법들이 프레임의 평범한 특징만을 측정할 수 있었다. 반면 특징 기반의 방법들은 보다 정교하게 프레임의 특징을 추출할 수 있다. 특징 기반의 방법들은 영상의 적률(moment), 윤곽선(contour lines), 특징점(feature points), 평면좌표(planar points)나 베이지안(Bayesian) 방법, 통계적 접근, Hidden Markov 모델 등을 고려하여 프레임의 특징을 계산한다[28].

특징 기반의 대표적인 방법으로 Zabih 등이 제안한 에지 변화 비교법이 있다[29]. 전역 움직임 보상은 연속된 두 프레임에서 수행되며 에지의 추출은 Canny 알고리즘에 의해서 이루어진다. 삽입된 에지와 존재하는 에지에 대한 정규화는 아래의 식 (23)과 (24)와 같이 계산된다.

$$P(C_{out}, I_t) = 1 - \frac{\sum_{i=1}^{X}\sum_{j=1}^{Y} E(I_{t-1}, i+\alpha_{t-1,t}, j+\beta_{t-1,t}) E_d(I_t, i, j)}{\sum_{i=1}^{X}\sum_{j=1}^{Y} E(I_{t-1}, i, j)} \tag{23}$$

$$P(C_{in}, I_t) = 1 - \frac{\sum_{i=1}^{X}\sum_{j=1}^{Y} E_d(I_{t-1}, i+\alpha_{t-1,t}, j+\beta_{t-1,t}) E(I_t, i, j)}{\sum_{i=1}^{X}\sum_{j=1}^{Y} E(I_{t-1}, i+\alpha_{t-1,t}, j+\beta_{t-1,t})} \tag{24}$$

여기에서 E와 E_d는 각각 윤곽선 영상과 그것의 팽창된 영상을 나타낸다. 그리고 $(\alpha_{t-1,t}, \beta_{t-1,t})$는 연속되는 두 프레임의 전역 움직임 벡터를

나타낸다. 이때 에지 변화 파편으로 정의되는 비유사도 척도 $ECF(I_t)$는 아래의 식 (25)에 의해서 계산된다.

$$ECF(I_t) = max((C_{out}, I_t), P(C_{in}, I_t)) \tag{25}$$

그리고 식 (26)과 같이 비유사도 척도와 임계값과의 비교를 통해서 장면 전환을 판단한다.

$$ECF(I_t) > T \tag{26}$$

Eickeler와 Müller는 장면 전환 검출을 위해서 Hidden Markov Model을 사용하였다[30]. 이 방법에서는 식 (27), (28), (29)의 3가지 척도를 이용해서 프레임의 특징을 계산한다.

$$d_1(I_t) = min\left\{ \frac{\sum_{i=1}^{X}\sum_{j=1}^{Y}|P(I_t,i,j) - P(I_{t-1},i,j)|}{XY}, \frac{\sum_{i=1}^{X}\sum_{j=1}^{Y}|P(I_{t+1},i,j) - P(I_{t-2},i,j)|}{XY} \right\} \tag{27}$$

$$d_2(I_t) = \sum_{v=0}^{V}\left(|H(I_t,v) - H(I_{t-1},v)| - median\begin{pmatrix} |H(I_{t-1},v) - H(I_{t-2},v)| \\ |H(I_t,v) - H(I_{t-1},v)| \\ |H(I_{t+1},v) - H(I_t,v)| \end{pmatrix} \right) \tag{28}$$

$$d_3(I_t) = \sum_{i=1}^{X}\sum_{j=1}^{Y} \frac{P(I_t,i,j) - P(I_{t+1},i,j)}{P(I_t,i,j) - 0.5P(I_{t-1},i,j) + P(I_{t+1},i,j)} \tag{29}$$

여기에서 $d_1(I_t)$, $d_2(I_t)$, $d_3(I_t)$는 각각 움직임의 강도, 차분 히스토그램의 강도의 중간값 필터된 결과, 차분 화소와 보간된 차분 화소의 비율을 나타낸다. 이와 같이 계산된 값들은 Baum-Welch 알고리즘으로 학습된 후에 Viterbi 알고리즘에 의해서 장면 전환을 검출하게 된다.

이 외에도 특징 기반의 방법으로는 적률 불변량을 이용하는 Arman 등의 방법[31], 에지 변화량 방법을 개선한 Lienhart의 방법[32], Ardebilian 등이 제안한 특징점 방법[33], Silva 등이 제안한 평면 좌표 방법[34], Bayesian 접근법 등 다양한 방법이 연구되었다[35]. 특징 기반의 방법들은 비교적 정확하면서도 자세한 프레임 특성을 도출할 수 있지만 상대적으로 많은 계산량을 요구하는 단점이 있다[36]. 아래의 표 4에서는 특징 기반의 방법들의 계산복잡도를 비교하여 나타내었다.

표 4. 특징 기반 방법들의 계산복잡도와 계산시간 비교
Table 4. Comparison of computation complexity and running-time of methods based on feature

Methods and reference	Complexity
Moment invariants[31]	$O(23p)$
Edge change ratio[36]	$O(26p)$
Edge-based contrast[32]	$O(25p)$
Planar points[34]	$O(51p)$
Hidden Markov Model with three features[30]	$O(13p+11n)$

마. 움직임 기반의 방법

연속되는 비디오 시퀀스에서 움직임은 자연적으로 발생하게 되며, 이 움직임은 장면 전환을 판단하는 중요한 요소가 된다. 이러한 사실에 기인하여 움직임 정보를 이용한 몇 가지 장면 전환 판단 방법이 제안되었다[37]. 움직임 기반 방법들은 카메라의 움직임, 움직임 벡터, 빛의 흐름 등을 이용한다.

움직임 벡터는 연속된 두 프레임간의 각 화소의 위치 변위를 나타내는 것이다. 서로 다른 장면에 존재하는 두 프레임인 경우에는 움직임 벡터의 방향이나 크기의 변화가 크기 때문에 이런 특성이 장면 전환 검출에 사용된다. 움직임 벡터가 장면 전환 검출에 유용한 이유는 카메라 움직임을 나타내는 좌우 선회(panning), 상하 선회(tilting), 확대 축소(zooming) 등의 정보를 얻어낼 수 있기 때문이다. 움직임 벡터 필드 상에 나타나는 일정한 패턴은 카메라 움직임을 잘 묘사한다. 카메라 움직임이나 영상 내의 물체의 움직임은 장면 전환 검출 시 오검출을 발생시키는 요인이다. 따라서 먼저 상관관계(correlation)를 이용하여 각 프레임에 존재하는 카메라 움직임을 필터링하고 장면 검출을 수행한다.

Cherfaoui와 Bertin은 아핀 변환 모델을 이용하여 전역 움직임 파라미터를 추정하고 이 추정된 움직임을 이용해서 장면의 고정, 이동, 줌 등을 분류하였다[38]. 단위 시간 동안의 움직임이 일정하지 않다면 장면 전환으로 판단한다.

Bouthemy 등은 우세 다해상도 움직임 추정을 장면 전환 검출에 이용하였다[39]. 이 방법은 여섯 개의 파라미터의 조합으로 생성된 전역 2차원 매개 모델을 이용하여 움직임을 추정한다.

Akutsu 등은 움직임 몽롱화 척도를 장면 전환 검출에 이용하였고[40],

Fatemi 등은 장면 전환 검출을 위한 정보로써 빛의 흐름 정보를 사용했다[41]. Porter 등은 연속되는 두 프레임의 움직임 벡터의 상관도를 이용해서 장면 전환을 측정하였다[42]. 아래의 표 5에서는 움직임 기반의 방법들의 계산복잡도를 비교하여 나타내었다.

표 5. 움직임 기반 방법들의 계산복잡도와 계산시간 비교
Table 5. Comparison of computation complexity and running-time of methods based on motion

Methods and reference	Complexity
Global motion[38]	$O(20p)$
Motion vectors[40]	$O(75p)$
Optical flow[41]	$O(105p)$
Frequency domain correlation[42]	$O(75p)$

PUKYONG NATIONAL UNIVERSITY 부경대학교

바. 다중 척도 기반의 방법

보다 정확한 장면 전환 검출을 위해서 복수의 척도를 사용하는 방법도 연구되었다. Quénot와 Mulhem은 간소화된 윤곽선 분석 알고리즘에 따른 chi-square 검정의 컬러 히스토그램에 기반한 방법을 제안하였다[43]. Gauch 등은 평균 밝기, 연속되는 프레임의 픽셀 변화, 컬러 분포의 변화 등 3가지 척도를 이용해서 장면 전환 검출을 수행하였다[44].

Yusoff 등은 다섯 가지 알고리즘을 조합한 방법을 제안하였다. 이 방법에서는 평균 밝기, 기하학적 거리, 히스토그램 비교, 유사도 비교, 움직임 추정 비교 등의 다섯 가지 방법이 사용되었다[45]. Naphade 등은 프레임 쌍에 대해서 컬러 히스토그램과 화소 밝기 차이를 이용한 K-means 클러스터링 알고리즘으로 장면 전환을 판단하였다[46]. 또 Lee 등은 이와 유사하게 밝기 히스토그램과 화소 밝기 차이를 이용한 K-means 클러스터링 알고리즘으로 장면 전환 검출을 수행하였다[47]. 아래의 표 6에서는 다중 척도 기반의 방법들의 계산복잡도를 비교하여 나타내었다.

표 6. 다중 척도 기반 방법들의 계산복잡도와 계산시간 비교

Table 6. Comparison of computation complexity and running-time of methods based on multiple measure

Methods and reference	Complexity
Chi-square test and edge analysis[43]	$O(26p+5n)$
Average brightness, pixel value change and color distribution change[44]	$O(4p+3n)$
Combination of five algorithms[45]	$O(85p+9n+17b)$
K-means using color histogram and pixel intensity difference[46]	$O(4p+9n)$
K-means using gray-level histogram and pixel intensity differences[47]	$O(3p+3n)$

PUKYONG NATIONAL UNIVERSITY
부경대학교

사. DC 계수 기반의 방법

한 장의 영상은 일정한 크기의 정방영역으로 나뉘어져 각 영역에 대해 변환 처리를 행함으로써 영역내의 평균값으로부터 매우 정교한 최고주파수의 영상성분에 이르기까지 여러 주파수의 영상 성분으로 분할된다. 이런 영상의 직교변환은 다양한 방법이 제안되었는데 DCT(Discrete Cosine Transform)가 기존의 영상 압축 분야에서 많이 사용되었다. DCT의 장점은 변환 전에는 화면에 불규칙하게 퍼져 있던 화소값이 변환 후에는 저주파 항 쪽으로 집중되는 분포를 가진다는 것이다[48]. 따라서 고주파 항들을 버리는 조작을 통해 정보손실이 거의 없이도 정보압축을 할 수 있게 되었다. MPEG이나 Motion JPEG의 경우 8×8블록의 크기에 대해서 2차원 DCT 변환을 사용한다. 이 때 DC 계수는 영상의 평균적인 밝기의 8배에 해당하는 값을 가지게 된다. DC 계수는 화면 전체에 대한 정보를 담고 있으므로 이러한 DC 계수가 프레임의 특성을 대체할 수 있다. 따라서 DC 계수를 이용한 장면 전환 검출이 가능해진다.

Meng 등은 DC 계수와 모션 벡터의 분산을 이용해서 장면 변화를 분석하였다[49]. Zhang 등은 P픽처와 B픽처의 유효 모션 벡터의 수를 이용한 장면 전환 검출 기술을 제안하였고[8], Yeo 등은 DC 영상의 시퀀스를 이용해서 장면 전환을 분석하였다[50]. 이런 접근 방법들은 연속된 프레임들의 DC 영상의 히스토그램의 차이 또는 화소 차이에 기반한 방법들이다. 위 방법들은 실제 저전력을 요구하는 하드웨어에 적용하기에는 적합하지 않다.

Fernando 등은 B 픽처의 메크로블럭 예측의 통계적 특성을 이용해서 장면 전환을 검출하는 알고리즘을 제안하였고[51], Seong 등은 Fernando의 방법을 디지털 위성 수신기가 삽입된 하드디크스에 적용하였다[52].

이런 방법들은 I픽처와 P픽처로만 구성된 고화질 비디오 스트림에는 적용할 수 없는 단점이 있으며, 휴대용기기에서 구현하기에는 여전히 어려운 점들이 존재한다.

Sethi 등은 연속되는 I 픽처의 DC 계수의 밝기 히스토그램의 차이를 사용했다[53]. 고속의 장면 검출이 가능하지만 정확도가 저하되는 단점이 있다. Kim 등은 MPEG 압축 영역에서 인트라 픽처의 축소 영상을 이용한 알고리즘을 PVR에 적용하였다[54].

아. 비트열 기반의 방법

MPEG 비트열로 부터 DCT DC Size와 DCT DC Difference에 해당하는 값을 얻어내어 각 블록간의 밝기 차에 관한 정보를 얻을 수 있다. DC 계수는 내포된 정보가 크기 때문에 DCT DC Size와 DCT DC Difference 두 부분으로 나누어져 VLC(variable length coding) 하게 된다. DCT DC Size(1-11비트 VLC)는 DC 계수의 밝기 차이인 DCT DC Difference에 따라서 코드의 길이가 결정된다. 따라서 매크로블록 내에 존재하는 4개의 밝기 성분을 나타내는 Y 블록들 간의 차이를 차례대로 d1, d2, d3라하고 이 세 값들 사이의 관계를 조사한다. 그리고 매크로블록을 단순 영역인지 에지 등이 있는 복합 영역인지를 구분하고 전체 영상에 대하여 특징점들을 구성하여 MSE(mean square error)에 따라 장면 전환 여부를 판단한다. 하지만 이 방법은 I픽처에서 발생하는 비트열 길이 정보만을 이용하기 때문에 전체적인 검출 성능이 떨어지는 단점이 있다[55].

자. 픽처 기반의 방법

움직임을 추정하는데 사용되는 참조 픽처에 따라 앞의 프레임을 참조하는 순방향 예측(Forward Prediction)과 뒤의 프레임을 참조하는 역방향 예측 (Backward Prediction)이 있다. MPEG에서는 세 가지 픽처 형식을 지원하는데, 움직임 추정방법을 사용하지 않는 I픽처(Intra-coded Picture)와 순방향 예측만을 사용하는 P픽처(Predictive-coded Picture), 양방향 예측을 이용하는 B픽처(Bidirectionally Predictive-coded Picture)가 있다.

이런 픽처들의 특성을 이용해서 장면 전환을 검출할 수 있다[56]. 현재 GOP에 있는 I픽처와 다음 GOP에 있는 I픽처를 비교함으로써 장면 전환을 검색할 수 있다. 만약 의심되는 픽처가 발견되지 않는다면 GOP에 있는 P, B픽처를 검색하지 않고 다음 GOP를 검색한다. 만약 I픽처에서 비트율이 증가하면 현재 GOP안에 있는 P픽처를 검색하게 된다. 장면 전환으로 의심되는 픽처가 없다면 현재 GOP와 다음 GOP I픽처 사이에서 장면 전환이 발생할 것이다. 현재 GOP에 있는 P 픽처에서 비트율이 증가하면 B 픽처를 검색하게 된다. 이때 여러 가지 상황이 발생하게 되는데 우선 peak가 하나 발생하면 장면 전환으로 선언한다. 그리고 peak가 여러 개의 B픽처에서 발생하면 장면 전환으로 맨 처음의 peak를 가지는 B픽처를 선언한다. 하지만 B픽처에서 장면 전환으로 의심되는 픽처가 발견되지 않는다면 위에서 발견된 P 픽처를 장면 전환 픽처로 선언한다. 만약 P 픽처에서도 비트율이 큰 픽처가 발견되지 않는다면 다음 GOP에 있는 I픽처를 장면 전환 픽처로 정의한다.

3. 장면 전환 검출을 위한 임계값 설정 방법

2절에서는 공간 및 변환 영역에서 프레임의 특성을 측정하는 다양한 방법을 알아보았다. 이렇게 측정된 프레임의 특징값들은 대부분의 방법에서 임계값과 비교를 통해서 장면 전환 유무를 판단한다. 즉, 사람이 동영상 데이터를 시청하면서 장면이 변화된 것을 인지하는 것과 같이 판단의 기능을 하는 요소가 바로 임계값이다. 프레임의 특성을 잘 추출하더라도 판단 기준이 적절하지 못하다면 정확한 장면 전환 검출이 불가능하다. 따라서 적절한 임계값의 설정이 장면 전환 검출 알고리즘의 우수성을 결정하게 된다. 다시 말해서 임계값은 장면 전환 검출의 정확도를 결정하는 중요한 요소라 하겠다.

임계값을 생성하는 방법도 다양하게 연구되어 왔는데, 실험적 방법, 분석적 방법, 적응적 방법의 세 가지로 분류할 수 있다[1]. 본 절에서는 임계값의 설정 방법과 기존의 임계값 생성을 위한 여러 가지 연구들을 정리하였다.

가. 실험적 방법

실험적 방법은 고정된 수치를 임계값으로 사용하는 방법이다[1]. 사전 정의된 값을 임계값을 설정하고 실험을 수행해서 장면 전환 검출 결과를 구한다. 사전 조사된 실제 장면 전환 결과와 비교해서 다시 임계값을 수정한다. 이런 과정을 최적의 검출 결과를 얻을 때까지 반복한다[57]. 예를 들어, 100개의 장면 전환이 존재할 경우 실험 결과에서 90개를 찾은

경우에는 임계값을 더 낮게 수정하고 반대의 경우에는 높게 수정한다.

실험적 방법의 경우에는 임계값 설정에 필요한 계산량이 없기 때문에 별도의 계산복잡도가 추가되지 않는 장점이 있다. 하지만 위에서 언급한 바와 같이 결과에 따라서 임계값을 수정해야하는 단점이 있다. 또 영상이 달라짐에 따라서 이전에 사용했던 임계값은 대부분 그대로 적용할 수 없다. 최적의 결과를 얻기 위해서는 많은 회수의 반복을 요구하게 되므로 실험적 방법에는 많은 한계가 존재한다. 사용자의 개입이 매번 요구되므로 자동적 알고리즘 구현이 불가능하며 실제 응용환경에서의 사용도 불가능하다.

Yi와 Ling은 실험적 임계값 설정 방법을 이용해서 장면 전환 검출을 수행하였다[58]. 이 방법에서는 화소값 기반의 다중 척도를 사용하였는데, 세 프레임의 차의 절대값의 평균, 히스토그램 평활화된 두 프레임 차의 절대값 평균의 차, 히스토그램 평활화된 두 프레임 분산의 차의 절대값, 히스토그램 평활화된 프레임 차의 절대값의 평균 등 4가지를 사용하였다. 각각의 척도에 서로 다른 임계값을 적용해서 발생할 수 있는 오류를 최소화하였다.

실험적 방법으로 장면 전환 검출 성능을 테스트하는 방법은 두 가지로 분류할 수 있다. 첫 번째는 완전 고정 임계값 방법으로 하나의 비디오 시퀀스에서 얻어질 최적 임계값을 이용해서 다른 모든 영상에 대해서 실험을 수행한다. 따라서 영상에 따라서 검출결과의 정확도 변화의 폭이 크다. 두 번째 방법은 가변 고정 임계값 방법으로 모든 비디오 시퀀스에 대해서 최적의 결과를 얻을 때까지 반복 실험을 하는 것이다. 완전 고정 임계값 방법에 비해서 검출 정확도는 향상되지만 반복 실험 횟수는 상당히 증가하게 된다.

나. 분석적 방법

분석적 방법은 실험적 방법의 많은 단점을 보완할 수 있다. 이 방법은 입력 비디오 시퀀스 전체를 분석하여 그 결과로부터 임계값을 생성하는 방법이다. 입력 비디오 시퀀스의 프레임 내 정보나 프레임 간 정보를 종합하여 변화가 심한 부분에 대한 정보를 생성하고 이를 토대로 임계값을 설정한다[59].

분석적 방법은 최적 임계값을 얻기 위한 반복 실험이 필요 없으며 비디오 시퀀스를 한번 검사하는 것으로 임계값을 설정할 수 있다. 또한 이 과정에서 사용자의 개입이 없이 동작할 수 있다. 그리고 입력 비디오 시퀀스의 정보를 이용하기 때문에 비교적 신뢰도가 높은 임계값을 도출할 수 있는 장점이 있다. 하지만 이 방법은 임계값을 얻기 위한 별도의 과정과 시간을 요구하기 때문에 계산복잡도의 증가가 발생한다. 또한 분석 과정이 필요하기 때문에 실시간 처리를 요구하는 응용환경에서의 사용도 어렵다. 그리고 생성된 임계값 역시 입력 영상에 종속적이기 때문에 입력 영상이 변화할 때마다 분석이 필요하게 되는 단점이 있다[60].

고경철과 이양원은 chi-square 검정과 자동 임계치 결정 알고리즘을 이용한 장면 전환 검출을 제안하였다[61]. 변형된 chi-square 검정으로 비디오 시퀀스의 특징을 추출하고 2단계를 걸쳐서 임계값을 결정한다. 먼저 주어진 전체 차이 값들로부터 평균과 표준편차를 구하며, 이 평균값을 만족하는 차이 값들로부터 다시 평균과 표준편차를 계산하며, 이러한 연속적인 평균값 및 표준편차의 계산으로부터 표준편차가 가장 큰 시점에서의 평균값을 기준으로 임계값을 결정하는 방법이다.

분석적 방법은 다양한 영상의 통계를 이용해서 영상에 독립적으로 사용할 수 있는 통계 분석적 방법으로 개선된 연구도 진행되었다[62].

다. 적응적 방법

적응적 방법은 기존의 실험적 방법이나 분석적 방법의 단점을 개선한 발전된 방법이다. 기존의 방법들은 입력 영상에 대해서 하나의 임계값으로 장면 전환을 판단하기 때문에 영상 내의 예외적인 변화에 대해서 강인하지 못하다. 적응적 방법은 비디오 프레임의 정보에 따라서 임계값이 가변적으로 설정되어 보다 정확하게 장면 전환을 검출 할 수 있다. 이 방법은 비디오 시퀀스의 지역적 특징으로부터 임계값을 설정하되 이 임계값은 비디오 데이터의 특성에 따라 변화하여 보다 정확하게 임계값을 생성한다[63].

Yusoff 등은 사전 정의된 참조 구간에서의 프레임들의 차의 평균을 이용하는 방법을 제안하였다[64]. 이 방법에서는 사전 정의된 임계값에 프레임들의 차의 평균을 더하는 방법, 곱하는 방법, 평균을 더하고 표준편차를 곱하는 세 가지의 임계값 결정 모델을 제시하고 있다.

Zhi와 Cai는 고정 참조 구간과 비트율을 이용한 장면 전환 검출 방법을 제안하였다[65]. 이 방법에서는 연속되는 프레임의 비트율의 차이를 이용해서 특징을 추출하고 차분 히스토그램의 평균과 분산을 이용해서 적응적 임계값을 생성한다.

Miene 등은 chi-square 검정을 이용해서 적응적 임계값을 설정하였다[66]. 이 방법에서는 프레임의 밝기값의 chi-square 히스토그램으로 프레임 특징값을 정의하고, 참조 구간 내에서의 프레임들의 chi-square 히스토그램의 최대값에 가중치를 적용시킨 값을 임계값으로 생성한다.

박진형 등은 히스토그램의 평균과 표준편차를 이용한 방법을 제안하였다[67]. 이 방법은 120프레임 단위로 구분된 구간에서 밝기 히스토그램의 차이값을 구하고 차이값들의 평균과 표준편차를 이용한 신뢰구간을 벗어

나는 값을 가지는 프레임을 장면 전환 프레임으로 검출한다.

Cheng 등은 고정 참조 구간에서의 차분 히스토그램 분산과 블록 기반 히스토그램을 이용해서 장면 전환을 구분하였다[68]. 이 방법은 200프레임의 고정 참조 구간에서 차분 히스토그램의 분산과 블록 히스토그램의 평균을 계산하고 여기에 일정 상수를 곱한 결과를 임계값으로 정의한다.

적응적 임계값 방법들은 대부분 한 번의 검사를 통해서 장면 전환을 검출하기 때문에 실시간 응용 환경에서의 적용이 가능하다[69]. 또한 프레임의 변화에 따라서 임계값이 정해지기 때문에 비교적 정확한 검출이 이루어진다. 이런 방법들에서는 최적의 참조 구간을 설정하는 것이 검출 결과에 많은 영향을 주게 된다. 정확한 검출을 위해서 다중 척도를 사용하게 되면 임계값 생성을 위한 계산복잡도가 많이 증가하게 되므로 단순한 알고리즘의 개발이 요구된다. 최근에는 참조 구간의 설정을 현재 비교하는 프레임과 유사한 프레임들만 이용하도록 하는 가변 참조 구간에 관한 연구도 진행되었다[70].

PUKYONG NATIONAL UNIVERSITY 부경대학교

4. 장면 전환 검출에서의 고속 계산을 위한 방법

장면 전환 검출에서 검출의 정확도와 함께 검출 속도도 중요하게 고려되어야 한다. 오늘날에는 비디오 데이터의 생성과 획득, 저장이 손쉽게 이루어지며 따라서 대용량 비디오 데이터를 효과적으로 관리하기 위해서는 고속의 처리가 이루어져야 한다. 장면 전환 검출에서도 대용량의 비디오 데이터를 대상으로 하기 때문에 빠른 처리 속도를 요구한다.

장면 전환 검출 분야에서 계산량을 감소시킬 수 있는 방법은 크게 두 가지로 구분된다[70]. 공간적(spatial) 관점과 시간적(temporal) 관점에서의 계산량 감소이다. 여기에 한 가지 추가되는 사항이 알고리즘적 간소화이다. 본 절에서는 계산량 감소를 위한 방법들의 개념과 기존 연구들을 살펴본다.

가. 공간적 관점의 계산량 감소 방법

공간적 관점의 계산량 감소는 비디오 데이터 프레임들의 공간적 크기를 축소시키는 것을 의미한다[71]. 예를 들어, 720×480 크기의 비디오 데이터를 360×240 크기로 축소 시켜서 장면 전환 검출을 수행하는 것이다. 이와 같은 처리는 프레임의 부표본화(sub-sampling)를 통해서 이루어진다. 일반적으로 영상처리에서는 영상을 축소하는 과정에서 보간법이나 anti-aliasing filter 등을 사용하지만 장면 전환 검출에서는 시각적인 의미를 고려하지 않기 때문에 이런 후처리 과정을 필요로 하지 않는다.

부표본화를 통해서 프레임당 연산횟수를 감소시킬 수 있다. 예를 들어,

720×480의 비디오 프레임에서 장면 전환 검출을 위해서 화소당 1번씩의 연산을 수행한다고 가정하면, 360×240 크기로 부표본화 하여 축소한다면 연산 횟수를 1/4로 감소시킬 수 있다. 전자의 경우 345600번의 연산이 필요하지만 후자의 경우 86400번의 연산만 요구된다.

이런 공간적 관점의 계산량 감소는 연산횟수를 줄임으로써 시간적 절약이 가능하지만 과도한 프레임의 축소는 검출 정확도에 손실을 초래한다. 정보량의 관전에서 원본 프레임의 정보가 축소를 통해서 소실되기 때문에 그만큼의 검출률의 저하가 발생한다. 그림 11에서는 공간적 관점의 계산량 감소의 방법을 나타내고 있다.

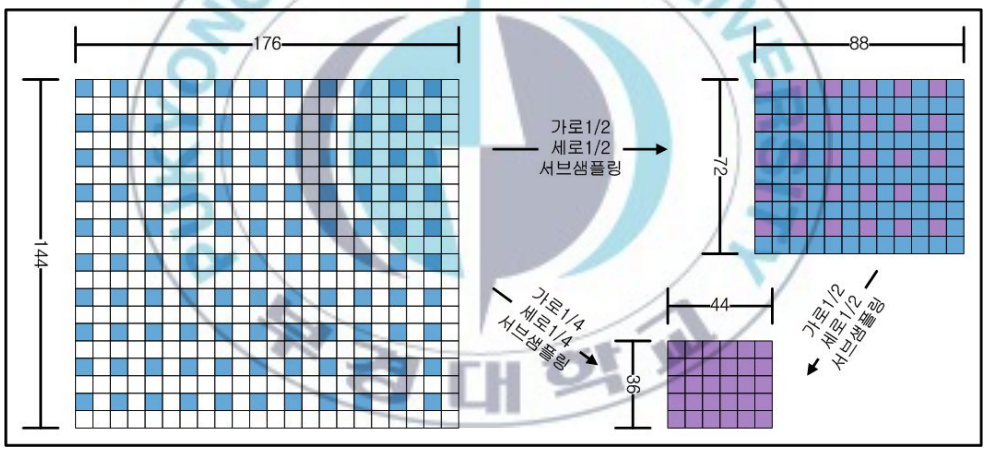


그림 11. 공간적 계산량 감소 방법

Figure 11. A reducing method of spatial computation amount

2001년 최우진 등은 대각선 방향의 화소값만을 추출하여 결과를 배열에 저장하고, 이 배열들의 MSE를 이용하여 장면 전환을 검출하는 방법을 제안하였다[72].

나. 시간적 관점의 계산량 감소 방법

시간적 관점의 계산량 감소는 비디오 데이터의 시간적 중복을 제거하여 계산량을 절감시키는 방법이다[73]. 이 방법은 연속되는 프레임은 거의 변화가 없다는 일반적인 특성을 이용한다. 즉, 연속되는 프레임을 비교하지 않고 비교 프레임들 사이의 일정한 개수의 프레임들을 생략해서 비교한다. 예를 들면 100번 프레임과 101번 프레임을 비교하는 것이 아니라 100번 프레임과 110번 프레임을 비교하는 것이다.

이와 같은 프레임 생략을 통해서 장면의 비교 횟수를 감소시킬 수 있다. 예를 들어 100개의 프레임으로 구성된 비디오를 장면 전환 검출하는데 있어서, 연속되는 프레임 비교 방법의 경우 99번의 비교가 필요하지만 프레임 생략 방법의 경우 생략 구간을 10프레임으로 하면 총 9번의 비교를 수행하게 된다. 따라서 90% 이상의 비교 횟수 감소가 발생한다.

시간적 관점의 계산량 감소를 통해서 획기적으로 비교 횟수를 줄일 수 있지만 생략 구간이 길어짐에 따라서 검출 정확도의 저하가 발생한다. 이것은 인접한 장면 간의 유사성이 구간 생략을 통해서 낮아지기 때문이다. 이와 함께 프레임 생략을 이용하게 되면 임계값의 설정도 힘들어진다. 비교 장면 간의 변화가 커지면서 장면 전환 프레임에서의 변화량의 폭이 커지기 때문이다. 또 프레임 생략의 경우에 장면 전환으로 판단할 경우 생략된 프레임들에 대한 재검색을 통해서 정확한 장면 전환 프레임을 찾아야하기 때문에 부가적인 처리가 요구된다. 아래의 그림들을 통해서 프레임 생략과 관련된 비디오 시퀀스의 검색 방법을 정리하였다.

(1) 선형 검색(linear search)

선형 검색은 비교 프레임의 생략 없이 순차적으로 비디오 시퀀스를 검

색하는 방법이다. 선형 검색에는 순차검색과 역차검색의 두 가지 방법이 있다. 순차(forward) 검색은 시작 프레임부터 검색하는 방법이며 역차(backward) 검색은 마지막 프레임부터 역으로 검색하는 방법이다. 일반적인 장면 전환 검출은 순차적 검색을 수행하며 구간 내 재검색 등 특별한 경우에 역차 검색이 사용되기도 한다. 그림 12에서 선형 검색 방법을 나타내었다.

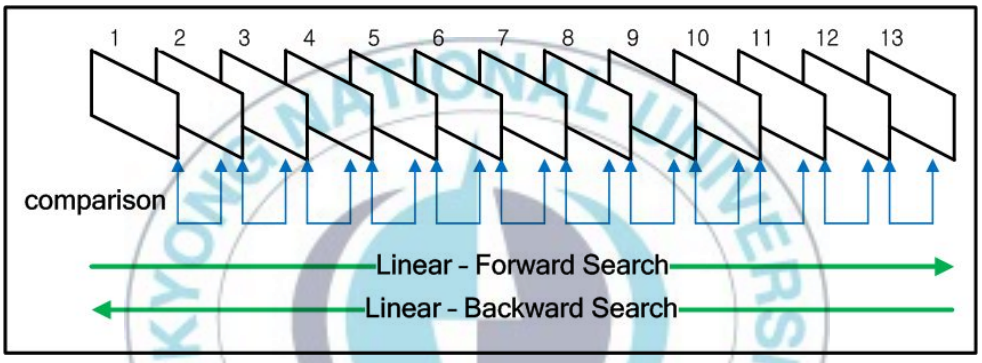


그림 12. 선형 검색 방법

Figure 12. Linear searching methods

(2) 비선형 검색(non-linear search)

비선형 검색은 비교 프레임 사이의 프레임 생략을 통해서 검색하는 방법이다. 비선형 검색에는 프레임 생략 구간의 길이가 일정한 고정 생략 검색과 프레임 생략 구간의 길이가 변화하는 가변 생략 검색 방법의 두 가지가 있다. 고정 생략 검색은 항상 동일한 구간의 프레임 생략을 가지게 되므로 프레임 생략 구간 설정에 대한 부가적인 처리가 필요 없다. 반면 가변 생략 검색은 현재 비교 결과가 다음의 생략 구간 결정에 영향을 주기 때문에 생략 구간을 계산하기 위한 처리가 필요하다. 아래의 그림 13에서 비선형 검색 방법을 도식화하여 설명하였다.

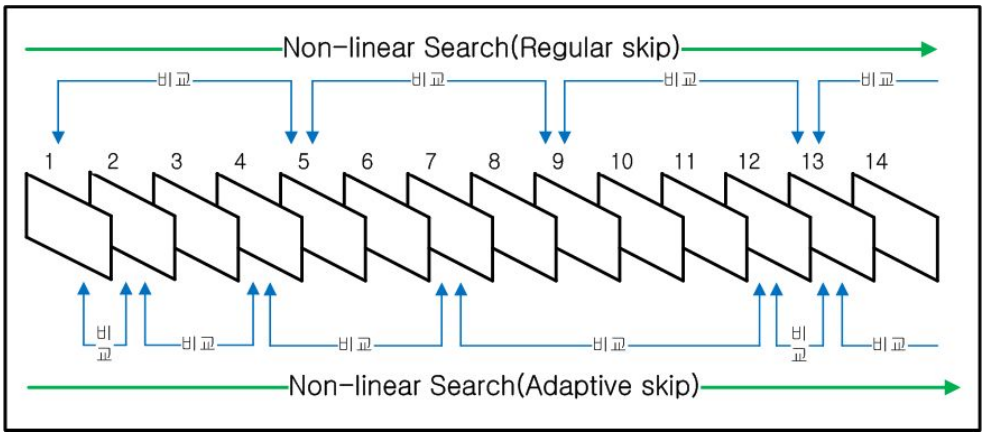


그림 13. 비선형 검색 방법

Figure 13. Non-linear searching methods

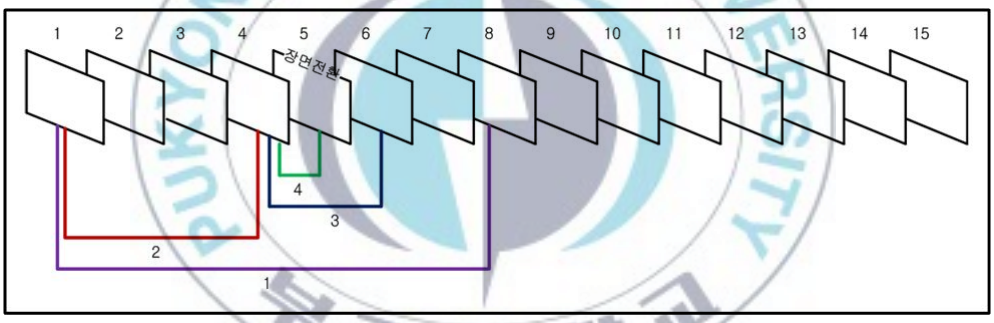


그림 14. 이진 검색 방법

Figure 14. Binary searching methods

(3) 이진 검색(binary search)

이진 검색은 기존의 순차적인 비교 방법을 탈피한 방법으로 프레임 비교 횟수를 상당히 감소시킬 수 있는 방법이다. 이 방법은 임의 개수의 프레임을 하나의 구간으로 설정하고 그 구간을 반으로 나누어서 구간의 처음과 마지막 프레임을 비교한다. 장면 전환이 있다고 판단되면 동일한 과정을 반복하게 되고, 장면 전환이 없다고 판단되면 나머지 절반의 구

간에서 동일한 과정을 반복한다. 위의 그림 14에서 이진 검색 방법을 도식화하여 나타내었다. 이 방법의 경우 불필요한 비교를 최소화할 수 있지만 이진 검색의 각 구간마다 적절한 판단 기준, 즉 임계값을 설정하는 것이 상당히 난해하다. 구간에 따라 다른 임계값이 필요하기 때문에 오히려 이로 인한 계산량의 오버헤드가 발생될 수 있다.

이와 같은 시간적 관점의 계산량 감소를 이용한 다양한 연구가 있었다. 2000년 Hua와 Oh는 가변 구간 생략 방법을 이용해서 일반적인 순차 검색보다 16배 빠른 방법을 제안하였다[74]. 이 방법에서는 프레임 비교 결과에서 유사도가 크면 다음 생략 구간을 증가시키고, 유사도가 작으면 다음 생략 구간을 감소시킨다. 그리고 장면 전환으로 판단할 경우에는 생략 구간 내에서는 고정 생략 구간 방법을 이용해서 정확한 장면 전환 프레임을 검출해낸다.

Baskar 등은 Hua의 방법을 개선시킨 프레임 생략 기반의 장면 전환 검출 방법을 제안하였다[75]. 이 방법에서는 프레임의 생략 길이는 2배 정수배 크기로 제한되며, 구간의 이전 프레임과 중앙 프레임을 먼저 비교하여 장면 전환 발생 시 구간 내에서 고정 생략 구간 재검색을 수행하고 장면 전환이 없는 경우에는 중앙 프레임과 현재 프레임을 비교한다.

권오덕 등은 가중치를 이용한 가변 길이 프레임 생략 기법을 제안하였다[76]. 이 방법에서는 프레임 생략 길이 결정에 가중치 요소를 이용해서 조절하는데 장면 전환 발생 시에는 가중치를 1/2로, 비전환시에는 가중치에 +1을 하여 프레임 생략 길이를 계산한다.

Lee 등은 적응적 생략 길이 결정 방법과 카메라 움직임을 고려한 장면 전환 검출 방법을 제안하였고[77], 나윤정 등은 비디오 시퀀스의 I 프레임만의 비교를 통한 장면 전환 검출 방법을 소개하였다[78].

5. 장면 전환 검출 기술의 성능 평가

장면 전환 검출 알고리즘의 성능 평가는 시간적 관점과 정확도 관점에서 이루어질 수 있다. 즉, 정확하게 장면 전환을 검출하는 것과 빠르게 장면 전환을 검출하는 두 가지 측면에서 알고리즘이 평가된다. 다양한 성능 평가 척도가 존재하는데 본 절에서는 이런 평가 척도들을 소개하고 평가 척도들의 의미, 그리고 이 결과를 토대로 알고리즘을 평가하는 방법 등을 정리하였다.

가. 시간 평가 척도

시간 평가를 위한 가장 직관적인 척도는 실행 시간을 꼽을 수 있다. 여기에서 실행 시간은 전처리 및 후처리 과정을 제외한 장면 전환 알고리즘이 시작되고 검출 결과가 나올 때까지의 시간을 측정한다. 여기에서의 실행 시간은 사용자의 체감 시간이 되고 하드웨어적 성능에 따라서 달라진다. 따라서 실행 시간은 하드웨어 종속적이므로 다른 하드웨어 성능을 가지는 장치에서의 알고리즘 비교는 객관적이지 못하다[79].

이와 대비되는 개념의 척도로 계산복잡도가 있다. 계산복잡도는 알고리즘의 분석을 통해서 알고리즘에서 사용되는 공간적, 시간적 소비를 이론적으로 계산하는 것이다. 계산복잡도는 하드웨어에 독립적으로 사용될 수 있는 장점이 있지만 분기문이 많이 발생하는 경우에는 신뢰도가 낮아진다. 이것은 이론적 계산복잡도와 실제 실행 시간이 반드시 비례관계에 있는 것은 아님을 의미한다[6].

나. 정확도 평가 척도

현재까지는 대부분의 장면 전환 검출 연구에서는 정확도가 우선적으로 평가된다. 일반적으로 정확도를 평가하는 척도로는 정밀도(precision)와 회수도(recall)가 사용된다[80]. 정확도를 측정하기 위해서는 실험 영상에 대한 사전 조사가 요구된다. 즉 사람의 판단을 이용한 실제 장면 전환 프레임에 대한 정보를 확보하는 것이 선행되어야 한다. 정확도의 측정은 실제 장면 전환 프레임을 알고 있다는 가정에서 수행된다. 우선 precision과 recall의 수식적 정의를 식 (30)과 (31)에서 나타내고 있다.

$$precision = \frac{N_c}{N_c + N_f} \tag{30}$$

$$recall = \frac{N_c}{N_c + N_m} \tag{31}$$

여기에서 N_c는 정확하게 검출한 장면 전환 개수, N_f는 잘못 검출한 장면 전환 개수, N_m은 검출하지 못한 장면 전환 개수를 각각 의미한다. 따라서 precision은 임의의 알고리즘을 통해서 장면 전환 검출을 수행한 결과 중에서 정확한 검출한 비율을 의미하며, recall은 실험 영상의 실제 장면 전환 결과 중에서 해당 알고리즘이 정확하게 검출한 비율을 의미한다. 역으로 설명하면 precision을 통해서는 오검출(false detection)의 정도를 알 수 있으며, recall을 통해서는 미검출(missing detection)의 정도를 알 수 있다. 여기에서는 최대값을 1, 최소값을 0으로 가지며 표현에 있어서 이해를 돕기 위해서 100을 곱한 값을 사용하기도 한다.

최근에는 정확도를 판단하는 종합적인 척도로써 F1을 사용한다[81]. F1은 precision과 recall을 종합하여 판단할 수 있는데, 식 (32)과 같이 precision과 recall의 조화 평균으로 정의된다. F1을 이용하면 precision이나 recall 중에서 한 척도만 높고 다른 척도는 낮은 경우의 알고리즘들을 객관적으로 비교할 수 있다.

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{32}$$

연구자의 특성에 따라서 위와 같은 비율로 측정하지 않고 프레임의 개수로 성능을 표현하는 경우도 있다. 즉, 전체 프레임의 개수, 실제 장면 전환 프레임의 개수, 정확하게 찾은 장면 전환 프레임 개수, 잘못 찾은 장면 전환 프레임 개수, 찾지 못한 장면 전환 프레임 개수를 각각 나타낼 수 있다. 이것은 테스트 영상의 샘플이 아주 적은 경우 또는 아주 많은 경우에 이해를 높일 수 있는 방법이다.

이 외에도 정확도의 성능을 측정하는 방법으로는 오류 검출 개수가 있다. 여기에서 오류 검출 개수는 오검출과 미검출의 개수를 합친 결과값으로 정의된다. 이것은 전체 장면 전환 프레임 중에서 몇 개의 에러가 발생되었는지에 대한 빈도를 의미한다. 오류 검출 개수가 낮을수록 알고리즘의 정확도는 높아진다. 일반적으로 잘 사용되지 않지만 알고리즘의 개선 정도가 낮은 경우, 개선의 정도를 보다 이해하기 쉽게 한다.

앞서 언급한 바와 같이 정확도 측정에 있어서는 실제 장면 전환 프레임에 대한 정보를 알아야한다. 정확한 알고리즘의 성능 측정을 위해서는 사전 정보 수집 과정에서의 실수를 최소화 하는 것도 필수적 요건이다.

다. 성능 평가 척도를 이용한 알고리즘의 평가

‘가’항과 ‘나’항에서 언급한 척도들은 어떤 값으로 계산되는데, 본 항에서는 성능 평가 척도를 이용해서 장면 전환 검출 알고리즘의 완성도를 평가하는 방법을 소개한다.

시간적 측면에서의 알고리즘 평가는 다소 주관적인 특성을 가진다. 이것은 평가하는 사람에 따라서 빠르고 느린 정도가 다를 수 있음을 의미하는데, 실시간 응용 환경에서의 사용을 고려할 때 비디오 데이터의 재생시간보다 검출 수행 시간이 짧으면 비교적 좋은 알고리즘으로 평가할 수 있다.

정확도 측면에서의 평가는 다소 복잡한 경우의 수가 발생한다. 아래의 표 7에서는 6가지 발생할 수 있는 검출 결과의 경우에 수에 대해서 나타내고 있다.

표 7. 장면 전환 검출 결과의 경우의 수

Table 7. The number of cases of shot change detection results

경우	실제개수	검출개수	정검출	오검출	미검출	precision	recall	F1
1	100	100	100	0	0	1.0000	1.0000	1.0000
2	100	100	95	5	5	0.9500	0.9500	0.9500
3	100	110	100	10	0	0.9091	1.0000	0.9524
4	100	90	90	0	10	1.0000	0.9000	0.9474
5	100	96	93	3	7	0.9688	0.9300	0.9490
6	100	104	97	7	3	0.9327	0.9700	0.9510

1번 경우와 같이 precision과 recall이 모두 1이 나올 때가 가장 이상적인 경우이지만 실제 알고리즘을 실험하게 되면 완벽한 검출이 힘든 경우

가 많다. 이것은 알고리즘의 태생적 결함, 임계값의 부적절한 설정, 비디오의 독특한 편집 등의 이유로 발생하게 된다.

2번에서 6번까지의 경우는 모두 10개의 오류 검출 개수를 가지는 경우이다. 종합적인 척도인 F1을 가지고 평가하면 0.95에 근사하는 값을 가지므로 비슷한 성능을 가진다고 평가할 수 있다. 2번의 경우에는 오검출과 미검출이 동일하게 나타나는데 이 알고리즘은 프레임의 특징 추출 척도는 부정확하지만 임계값은 정확하게 설정된 경우다. 따라서 검출 개수는 실제 검출 개수와 동일하게 나타난다. 이런 알고리즘을 사용하게 되면 정확한 장면 검출은 어렵지만 실제 장면 개수는 알 수 있다.

3번의 경우에는 미검출 없이 오검출만 발생한 경우인데 이 알고리즘은 최소한 장면 전환이 있는 부분은 모두 검출한 것을 의미한다. 감시 카메라와 같이 장면 전환 자체가 중요한 의미를 가지는 응용 환경에서 사용될 수 있다.

4번의 경우에는 오검출 없이 미검출만 존재하는 경우로써 검출 결과가 모두 장면 전환을 의미한다. 따라서 이 알고리즘은 부분적으로 장면 전환이 빠졌지만 의미적으로는 크게 중요하지 않은 비디오의 스토리 보드나 요약 같은 작업에 사용될 수 있다.

5번과 6번의 경우에는 검출 개수도 실제와 동일하지 않으며 오검출과 미검출이 동시에 나타나는 경우인데 이런 알고리즘들은 특징값 추출 척도도 부정확하지만 임계값의 설정 방법도 잘못 된 경우이다. F1의 측면에서는 다른 경우와 유사하지만 실제 응용 환경에서 사용하기에는 다소 부적절한 방법이다.

6. 장면 전환 검출 기술의 문제점 분석

1990년대 이후로 장면 전환 검출은 다양하게 연구되어 왔다. 현재까지 많은 문제점들이 제기되었고 문제점을 해결하기 위한 여러 가지 연구들이 진행되었다. 본 절에서는 장면 전환 검출 기술의 문제점들을 제시하고 아직까지 해결되지 않은 문제점들을 알아본다. 또한 본 연구를 시작하게 동기와 연구의 목적들을 정리하였다.

가. 기존 연구들의 문제점

연구 초기의 관심은 프레임의 특성을 추출하는 척도에 관한 문제들이었다. 프레임의 특징을 잘 표현할 수 있는 특징 측정 척도들이 발전되어 왔다. 화소 기반의 문제점들을 개선하기 위해서 히스토그램이 사용되었고 블록 형태의 적용을 통해서 더욱 정확한 특징을 측정할 수 있다. 여기에 특징 기반의 방법이나 움직임 기반의 방법들로 프레임의 특징을 잘 추출할 수 있다. 또한 다중 척도를 이용해서 프레임의 특징을 측정하면서 프레임 특성 추출 척도의 문제는 거의 해결되었다.

장면 전환 검출의 연구가 이루어지면서 현재까지도 활발하게 연구되는 부분이 바로 점진적 장면 변화에 대한 정확한 검출 문제이다. 디졸브나 페이드인/아웃, 와이프를 비롯한 다양한 영상 편집기술이 발달되면서 비디오 데이터에 포함되는 점진적 장면 변화의 구간도 증가하였다. 다양한 연구가 진행되어 정확한 검출 결과를 나타내는 연구도 있지만 아직도 급진적 변화에 비해서 정확도가 낮으며 검출 시간이 길다는 단점이 있다.

점진적 변화와 함께 많이 연구되는 것이 다양한 예외적인 장면 변화에 대한 정확한 검출 문제이다. 섬광이나 플래시와 같은 밝기의 급격한 증가, 지역적인 컬러 변화, 전체적으로 어두운 부분에서의 변화, 객체와 물체의 빠른 이동, 장면 내의 장면 변화 등이 주로 연구되고 있다. 이런 연구들은 변화 부분들의 영상의 특성을 분석해서 기존의 알고리즘과 결합하는 방식을 취하고 있다. 여기에서 문제가 되는 것은 부가적 처리를 통해서 기존의 알고리즘의 검출정확도에 손실을 초래하지 않도록 해야 한다는 점이다. 아직까지는 급진적 장면 변화와 점진적 장면 변화, 그리고 위에서 언급한 예외적인 장면 변화 모두에 대해서 강인한 검출 방법의 연구는 거의 없다.

비디오 데이터의 활용 및 저장이 활발해지면서 다양한 영상에 적용할 수 있는 장면 전환 검출 알고리즘의 연구도 주목받고 있다. 이것은 곧 임계값의 적응적 설정 방법과 관련된다. 다양한 영상에 적용하기 위해서는 적응적으로 영상의 특성에 따라서 임계값을 설정할 수 있어야 하기 때문이다. 참조 구간에서의 프레임의 특성을 이용하는 다양한 적응적 임계값 설정 방법들이 제안되었는데, 이런 방법들은 참조 구간의 설정에 따라서 검출 결과의 성능이 달라진다. 고정적인 참조 구간보다는 조건에 따라서 참조 구간의 길이가 달라지도록 하는 방법이 최근에 연구되고 있다. 더욱 정확한 검출을 위해서는 비교하는 장면과 유사한 장면만 참조하도록 하는 참조 구간 설정 방법이 연구되어야 한다.

장면 전환 검출 기술뿐만 아니라 모든 분야에서 중요하게 생각되는 것이 고속 처리와 실시간 처리에 관한 문제다. 장면 전환 검출은 대용량의 비디오 데이터를 이용하기 때문에 고속 처리는 필수적으로 요구된다. 공간적 해상도의 축소와 시간적 프레임의 생략을 통한 방법은 많은 연구가 진행되었다. 더욱 고속의 처리를 위해서는 간단한 특징값 추출 척도와

임계값 설정 방법이 요구된다. 다중 척도와 복잡한 임계값 설정은 정확한 검출은 보장할 수 있지만 처리 속도가 저하된다. 이런 상충 관계를 적절히 조화하면서도 알고리즘을 고속화 할 수 있는 연구가 필요하다. 이와 함께 실시간 처리 환경에서 사용할 수 있는 연구가 병행되어야 하는데 이것은 응용 환경에서의 사용을 위해서는 필수적이다. 비디오 시퀀스 전체나 넓은 구간을 이용한 처리보다 짧은 구간이나 두 프레임의 비교를 통해서 장면 전환을 검출할 수 있는 연구가 필요하다. 따라서 사용할 수 있는 프레임의 정보가 적은 경우에도 정확한 검출을 할 수 있는 연구가 요구되고 있다.

또 한 가지 언급해야할 문제점은 장면 전환 검출의 연구 환경이다. 현재 거의 대부분의 연구는 PC나 워크스테이션의 환경에서 연구되고 실험되어진다. 현재 휴대용 미디어 재생 장치나 휴대 전화 등 비디오 데이터를 재생할 수 있는 장치들이 많이 보급되어 있다. 하지만 기존에는 이런 장치들을 위한 연구가 거의 이루어지지 않고 있는 실정이다. 휴대용 장치의 경우 하드웨어 자원도 PC보다 부족하고 성능도 떨어지며, 전원 공급도 제한적이다. 또한 사용자들은 즉각적인 장면 전환 검출의 결과를 확인하고 싶어 한다. 이런 요건들을 만족시키기 위한 기존의 연구는 거의 없었으며, 따라서 휴대용 장치에서 사용할 수 있는 장면 전환 검출 연구가 필요하다.

마지막으로 고려되어지는 사항은 실험 데이터의 표준화이다. 장면 전환 검출에 있어서는 아직도 공통적으로 사용되는 표준 동영상이 없다. 따라서 연구자에 따라서 테스트에 사용하는 비디오 데이터가 다르기 때문에 알고리즘의 직접적인 비교가 힘들다. 기존의 알고리즘들을 반드시 구현해야 비교할 수 있기 때문에 연구의 시간적 낭비가 발생한다. 이와 같은 표준 데이터의 문제는 앞으로 해결되어야 하는 사항이다.

나. 본 연구의 목적

위에서 언급한 바와 같이 장면 전환 검출 기술에 있어서 아직도 해결되지 않은 문제점들이 있다. 본 논문에서는 이와 같은 문제점을 해결하여 보다 개선된 장면 전환 검출 알고리즘의 개발을 목표로 한다.

본 논문에서 다음의 3가지 문제에 대해서 초점을 맞춰서 연구를 진행한다. 먼저 장면 전환 프레임을 정확하게 구별할 수 있는 프레임 측정 척도에 대한 연구를 수행한다. 기존의 척도들을 개선해서 장면 전환 프레임의 차이를 더욱 명확하게 할 수 있는 척도에 대한 연구를 수행한다.

다음으로는 다양한 영상에 대해서 정확한 검출이 가능하도록 적응적 임계값 설정에 대한 연구를 진행한다. 기존의 참조 구간을 이용하는 방법을 개선해서 유사한 장면만 참조할 수 있는 적응적 참조 구간 설정 방법을 연구하고 가중치 요소를 적용해서 임계값의 신뢰도를 높일 수 있는 방법의 연구를 진행한다.

마지막으로는 고속 처리를 위한 알고리즘 개선과 휴대용 미디어 재생 장치에 적용하기 위한 알고리즘을 연구한다. 위에서 언급한 개선된 프레임 특징값 측정 척도와 적응적 임계값을 이용한 장면 전환 검출 알고리즘을 시간, 공간적 중복성 제거를 통한 환경에서도 정확한 검출이 가능하도록 개선하고, 휴대용 장치에 적용하기 위한 알고리즘 경량화를 연구하고자 한다.

최종적으로 본 논문에서는 다양한 영상에서 정확한 장면 전환 검출이 가능하며 휴대용 장치에서도 적용할 수 있는 실시간처리가 가능한 적응적인 장면 전환 검출 모델을 제안하고자 한다.

Ⅲ. 제안하는 방법

본 장에서는 제안하는 장면 전환 검출 알고리즘을 상세하게 기술하고 있다. 본 장에서는 먼저 장면 전환 프레임을 명확하게 구분하기 위해서 기존의 분산에 가중치를 적용한 가중치 분산을 제안한다. 그리고 제안한 알고리즘을 다양한 영상에 적용하기 위해서 임계값을 적응적으로 생성하는 방법을 제안한다. 기존의 참조 구간을 이용하는 방법에서 현재 비교하는 프레임과 유사한 장면들만 참조 하도록 하는 가변 참조 구간 방법을 제안한다. 또 적응적 임계값의 신뢰도를 높이기 위해서 가중치 개념을 도입하고 가중치를 자동적으로 생성하는 방법을 제안한다. 최종적으로 이러한 방법들을 토대로 실시간 적응적인 장면 전환 검출 모델을 제안한다.

본 장의 구성은 다음과 같다. 먼저 1절에서는 고속 장면 전환 검출을 위한 프레임 평균값과 적응적 임계값을 이용한 장면 전환 검출 방법을 설명한다. 다음 2절에서는 가중치 분산을 이용한 프레임 특징 추출 방법과 이를 이용한 적응적 장면 전환 검출 방법에 대해서 서술하였다. 3절에서는 정확도를 높이기 위해서 히스토그램을 이용한 장면 전환 검출 방법을 제안하고, 가중치 요소도 적응적으로 생성하도록 하는 완전 자동적인 장면 전환 검출 알고리즘에 대해서 기술하였다. 4절에서는 3가지 방법을 토대로 하는 실시간 적응적인 장면 전환 검출 모델에 대해서 설명하고 있다. 5절에서는 도표와 예제를 통해서 제안한 방법의 이해를 돕고, 마지막으로 6절에서는 제안한 방법의 문제점을 개선시킨 알고리즘을 기술하였다.

1. 프레임 평균값을 이용한 장면 전환 검출 방법

본 절에서는 프레임의 평균값을 이용한 장면 전환 검출 방법을 설명한다. 프레임의 평균값은 계산량이 가장 적은 방법으로 고속처리와 실시간 처리를 위해서 유용하게 사용될 수 있다. 적응적 임계값은 가변 참조 구간 동안의 프레임 평균값을 이용해서 계산하는데 이 방법에 대해서 아래에 자세하게 서술되어 있다.

가. 차분 프레임 평균값

본 절에서 제안하는 방법은 차분 프레임의 평균을 이용해서 프레임의 특징값을 추출한다. 이 방법은 계산량이 작기 때문에 응용환경에서 활용도가 높다고 판단했기 때문이다. 상대적으로 낮은 검출 정확도는 임계값의 적응적 설정으로 향상시킬 수 있다는 가정을 바탕으로 한다.

차분 프레임의 평균은 다음의 식 (33)과 같이 계산된다. 프레임의 차이의 특성을 표현하기 위해서 절대값 차를 이용한다. 여기에서 $ADFM$은 차분 프레임의 평균, M과 N은 프레임의 크기, $F(m,n)$은 (m,n)에서의 밝기값, F_i는 i번째 프레임을 각각 나타낸다. 이렇게 구해진 $ADFM$은 임계값과 비교해서 장면 전환을 판단하는데, 일반적으로 $ADFM$이 임의의 임계값보다 큰 경우에 장면 전환이 발생했다고 가정한다.

$$ADFM_i = \frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|F_i(m,n) - F_{i-1}(m,n)|}{M \times N} \tag{33}$$

나. 적응적 임계값 설정 방법

장면 전환 검출에 있어서 임계값은 중요한 역할을 한다. 특징값을 토대로 장면 전환 발생 유무를 결정하는 것은 순전히 임계값을 통해서 이루어진다.

본 논문에서는 앞서 2장에서 언급한 바와 같이 다양한 영상에 적용가능하면서도 실시간으로 동작할 수 있는 장면 전환 검출 방법을 제안한다. 따라서 임계값의 설정에 있어서도 이런 전제에 부합하도록 하는 3가지 조건이 있다. 첫 번째는 모든 영상에 적응적으로 설정되어야 하는 것이다. 두 번째는 프레임의 재생과 동일한 순차성을 가져야 한다는 것이다. 세 번째는 지연시간이 없어야 한다는 것이다.

본 논문에서 제시하는 위의 3가지 조건을 만족하는 해결방법은 다음과 같다. 첫 번째 모든 영상에 대한 적응적 설정을 위해서는 입력 영상 자체의 정보를 이용해야 한다. 두 번째 순차성을 가지기 위해서는 현재 프레임의 이전 프레임들의 정보만 이용한다. 세 번째 지연시간을 최소화하기 위해서는 이미 계산되어진 값을 이용한다.

이런 3가지 조건의 해결방법으로부터 제안하는 임계값 설정방법은 다음의 식 (34)로 나타낼 수 있다. 여기에서 THF는 임계값, S_i는 i번째 프레임의 번호, S_p는 이전 장면 전환 프레임의 번호, WF는 임계값 가중치를 각각 나타낸다.

$$THF_i = \frac{\sum_{j=S_p}^{S_i} ADFM_j}{S_i - S_p + 1} \times WF \tag{34}$$

즉, i번째 프레임에서의 적응적 임계값 THF_i는 p번째 프레임에서 i번째 프레임까지의 $ADFM$의 평균에 임계값 가중치 WF의 곱으로 정의할 수 있다. $ADFM_i$가 THF_{i-1}보다 큰 경우에 장면 전환 발생으로 간주한다. 장면 전환이 발생하면 임계값은 0으로 초기화 된다.

제안하는 임계값 설정 방법은 일정 구간 동안의 $ADFM$의 평균을 이용한다. 여기에서 구간은 비교 프레임이 옮겨감에 따라서 다르게 설정된다. 이전 장면 전환 프레임의 다음부터 현재 프레임까지, 즉 현재 프레임과 동일한 장면의 $ADFM$만을 계산에 포함시키기 위한 방법이다. 이런 참조 구간 설정은 기존의 고정된 참조 구간을 이용하는 것 보다 참조 구간 내의 프레임 특성이 균일적이다. 따라서 임계값의 정확도가 향상된다.

또한 여기에서 임계값 가중치 WF의 역할이 중요한 것을 알 수 있다. 만약 $WF=1$인 경우, 즉 가중치를 적용하지 않는 경우에 $THF \approx ADFM$이 된다. 이것은 임계값의 생성이 가변 참조 구간 내의 $ADFM$의 평균으로 이루어지기 때문에 두 값이 유사할 확률이 높다. 따라서 장면 전환이 없는 부분에서 오검출이 발생될 수 있다. 그러므로 THF에 마진을 적용해야하는데 이 과정에서 임계값 가중치 WF가 사용되는 것이다. 하지만 WF가 너무 큰 값을 가지면 실제 장면 전환 프레임에서 미검출이 발생되며, 너무 작은 값을 가지면 장면 전환이 없는 부분에서 오검출이 발생된다. 따라서 적절한 WF의 설정도 알고리즘의 성능 결정에 중요한 역할을 한다. 제안하는 방법에서는 다양한 영상의 실험을 통해서 통계적으로 10이라는 값을 사용한다.

제안한 방법을 통해서 입력 영상에 독립적인 장면 전환 검출이 가능하다. 하지만 차분 프레임의 평균값으로는 프레임의 특성을 보다 섬세하게 구분하지 못하는 단점으로 인해서 상대적으로 낮은 검출률을 가진다. 이를 위해서 개선된 프레임 특징 추출 척도를 2절에서 제안한다.

2. 가중치 분산을 이용한 장면 전환 검출 방법

1절에서 차분 프레임의 평균값을 이용한 장면 전환 검출 방법을 알아보았다. 입력 영상에 독립적으로 동작하는 영상 적응성은 만족하였지만 검출 정확도가 상대적으로 낮은 단점이 있었다. 이런 문제점을 해결하고자 본 절에서는 개선된 프레임 특징 추출 척도를 제안한다.

가. 가중치 분산

본 절에서는 차분 프레임의 분산을 이용하여 프레임의 특징값을 추출한다. 차분 프레임의 분산은 사전 실험 결과 차분 프레임 평균보다 계산량은 증가하지만 검출 정확도는 향상되는 것을 알 수 있었다. 또한 부표본화된 비디오를 대상으로 한 실험에서도 강인한 검출률을 나타내었다. 차분 프레임의 분산은 다음의 식 (35)와 (36)에 의해서 계산된다. 여기에서 DFM은 차분 프레임의 평균, DFV는 차분 프레임의 분산, D_i는 차분 프레임을 각각 나타낸다.

$$DFM_i = \frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(F_i(m,n) - F_{i-1}(m,n))}{M \times N} \tag{35}$$

$$DFV_i = \frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(D_i(m,n) - DFM_i)^2}{M \times N} \tag{36}$$

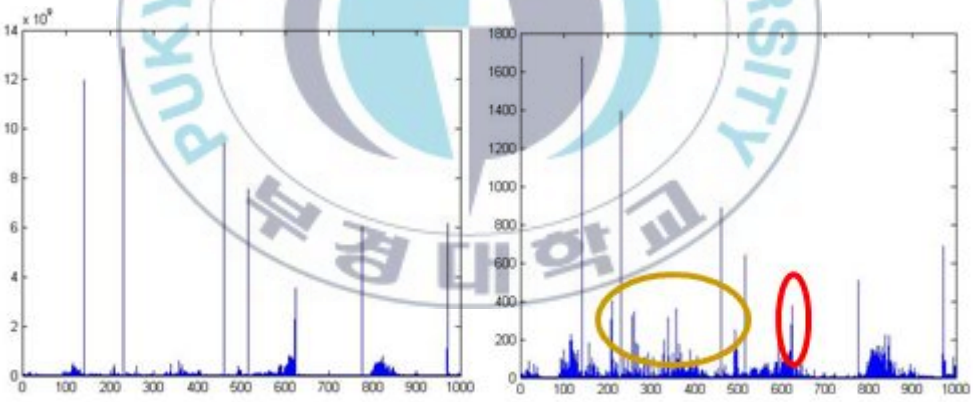
본 절에서는 식 (35)와 (36)으로부터 계산된 분산값에 가중치를 적용하는 방법을 제안한다. 가중치를 부여하는 것은 장면 전환 프레임에서의 분산값을 증폭시키기 위해서이다. 하지만 단순히 식 (36)의 분산 계산에서 임의의 상수를 곱하게 되면 모든 분산의 값이 커지게 되므로 무의미한 연산만 부가될 뿐이다. 장면 전환이 없는 부분의 낮은 분산값은 소폭의 상승만 발생하게 하고 장면 전환이 있는 부분에서의 분산값은 대폭 상승이 발생하도록 계산되어져야 한다. 이를 위해서 평균 계산식인 식 (35)에도 가중치를 적용한다. 이와 같은 가중치가 적용된 분산은 식 (37)과 (38)을 통해서 유도된다.

$$WDFM_i = \frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(F_i(m,n) - F_{i-1}(m,n))}{(M/w_v)\times(N/w_v)} \tag{37}$$

$$WDFV_i = \frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(D_i(m,n) - WDFM_i)^2}{(M/w_v)\times(N/w_v)} \tag{38}$$

여기에서 $WDFM$은 가중치가 적용된 평균, $WDFV$는 가중치가 적용된 분산을 의미한다. 분모에 가중치를 나누어줌으로써 가중치가 적용된 평균 $WDFM$은 $w_v^2 \times DFM$이 된다. 따라서 $WDFM$을 이용한 가중치 분산 $WDFV$는 일반적인 분산 DFV보다 상당히 큰 값으로 계산되어진다. 여기에서 중요한 것은 이 값을 증폭이 선형적으로 증가하는 것이 아니라 장면 전환이 없는 부분은 작게, 장면 전환이 발생하는 부분에서는 비약적으로 크게 적용된다는 점이다. 이런 분산값의 분포를 통해서 장면 전환이 발생하는 지점을 더욱 정확하게 검출할 수 있다.

그림 15에서는 제안한 가중치 분산과 일반적 분산의 분포를 비교하여 나타내었다. 동일한 프레임에 대한 비교를 통해서 프레임 특징의 양극화가 제안한 가중치 분산을 통해서 더욱 두드러짐을 확인할 수 있다. 그림 15를 분석하면 (a)와 (b) 모두 전체적으로는 장면 전환을 잘 표현하고 있는 것 같지만, (b)에서는 황색 원으로 표시된 부분과 같이 장면 전환이 없는 프레임에서 값이 높게 형성되는 것을 알 수 있다. 이런 부분들은 우측의 적색 원으로 표시된 장면 전환 프레임의 값과 비슷하게 되어 오검출이 발생될 확률이 높다. 반면 (a)에서는 오검출의 위험성이 배제되어 명확하게 장면 전환과 비전환 프레임들을 구분할 수 있다. 이와 같은 가중치 분산의 특징을 이용하면 장면 전환 검출의 정확도를 높일 수 있으며 적응적 임계값을 설정하는데도 유리하게 작용한다.



(a) (b)

그림 15. 가중치 분산과 일반적 분산의 분포 비교
((a) 가중치분산, (b) 일반적분산)

Figure 15. Distribution comparison of the weighting variance with a general variance
((a) weighting variance, (b) general variance)

지금까지 설명한 가중치 분산에서 가중치의 결정은 다음과 같은 사항을 고려해야 한다. 가중치 w_v의 변화에 따른 분산의 분포는 $w_v = 4$에서부터 달라진다. 즉, $w_v = 1$인 경우가 가중치가 적용되지 않은 분산을 의미하고, $1 < w_v < 4$인 경우에서는 분포가 조금씩 변화하지만 그림 15와 같이 명확하게 달라지지는 않는다. 하지만 $w_v > 4$ 이상이 되더라도 분산값의 변화는 발생하지만 분포는 거의 일정하게 유지된다. 이런 가중치에 따른 분산값 분포의 변화는 프레임의 크기, 즉 화소의 개수와도 관련이 있다. 프레임의 크기가 증가하게 되면 적용되는 가중치의 값도 비례하여 커져야한다.

가중치의 값이 클수록 장면 전환 검출에 유리한 요소가 될 수 있지만 때로는 장면 전환 프레임에서의 분산값이 데이터 형(type)의 한계를 넘어가는 경우가 발생하여 검출에 악영향을 주기도 한다. 따라서 프레임의 크기, 분산값의 분포, 데이터 형의 한계 등을 고려해서 가중치를 설정해야 한다. 본 논문에서는 이 같은 요소들을 고려해서 $w_v = 4$를 알고리즘에 적용한다.

PUKYONG NATIONAL UNIVERSITY 부경대학교

나. 적응적 임계값 설정 방법

본 절에서도 영상에 독립적으로 적용할 수 있도록 1절에서 제안한 적응적 임계값 설정 방법을 이용한다. 여기에서는 1절의 차분 프레임의 평균이 아닌 가중치 분산의 평균값을 이용한다. 하지만 본 절에서 사용되는 임계값 계산식은 1절의 식 (34)와는 구간의 길이 설정 방법이 약간 다르다. 1절에서 제안한 임계값 설정 방법은 이전 장면 전환 프레임과 현재 프레임을 포함하게 되는데 이것이 검출률이 낮아지는 원인으로 분석되었다. 그 이유는 Ⅱ장의 1절에서 언급한 점진적 특성을 가지는 급진적 장면 전환 때문이다. 이전 장면 전환 프레임을 임계값 계산에 포함하게 되면서 다음 장면이 아닌 이전 장면의 특징값도 반영되게 되고 따라서 임계값의 정확도가 저해된 것이다. 이런 문제를 개선하기 위해서 본 절에서는 이전 장면 전환 프레임을 제외하고 임계값을 계산한다. 식 (39)에서 가변 참조 구간에서의 가중치 평균을 이용하는 임계값 설정 방법을 정의하고 있다.

$$THV_i = \frac{\sum_{j=S_p+1}^{S_i-1} WDFV_j}{S_i - S_p + 1} \times WV_i \tag{39}$$

여기에서 THV_i는 i번째 프레임에서의 임계값, WV는 임계값 가중치를 각각 나타낸다. i번째 프레임에서의 적응적 임계값 THV_i는 이전 장면 전환의 다음 프레임부터 현재 프레임의 이전 프레임까지의 $WDFV_j$의 평균값에 임계값 가중치 WV를 곱한 값으로 정의된다. $WDFV_i$ 가 THV_{i-1} 보다 큰 경우에 장면 전환을 판단하며 그때 임계값은 초기화 된다.

다. 임계값 가중치 결정 방법

앞선 임계값 결정 수식에서도 가중치가 사용되며 그 역할은 1절에서 언급한 바와 같다. 테스트 비디오를 대상으로 하여 통계적으로 $WV=10$ 일 때 가장 좋은 검출 결과를 얻을 수 있었다.

하지만 이런 고정된 값을 가중치로 사용하는 것을 완전한 의미의 적응적 임계값이 아니다. 입력 영상이 달라짐에 따라서 10이 아닌 다른 값에서 더 좋은 검출 결과가 나올 수도 있다. 실제로 영상에 따라서 10보다 작은 값 또는 큰 값에서 검출 결과가 더 좋게 나오기도 한다. 결국 임계값 가중치 역시 입력 영상의 정보를 토대로 해서 생성되어야 한다는 결론에 이른다.

본 절에서는 임계값 가중치를 자동적으로 생성하는 방법을 제안한다. 임계값 가중치를 생성하는데 소모되는 시간을 최소화하기 위해서 이미 계산되어진 값을 이용하도록 한다. 제안하는 임계값 가중치의 결정 방법은 다음의 식 (40)과 (41)에서 정의하고 있다.

$$NZPM_i = \frac{\sum_{j=S_p+1}^{S_i-1} NZP_j}{S_i - S_p + 1} \tag{40}$$

$$WV_i = \frac{(NZPM_i \times 10) + (((M \times N) - NZPM_i) \times 10)}{M \times N} \tag{41}$$

여기에서 $NZPM$은 가변 참조 구간 동안의 NZP의 평균, ZNP는 차분 프레임의 0이 아닌 픽셀의 개수를 각각 의미한다. i번째 프레임에서의

임계값 가중치 WV_i는 가변 참조 구간 동안의 차분 프레임의 화소값의 변화가 발생한 양과 변화가 발생되지 않는 양의 비율을 이용해서 설정할 수 있다.

이런 가중치 결정 방법은 프레임의 변화 정도를 이용해서 설정하게 되므로 영상에 적응적일 수 있다. 실제로 테스트를 통해서 적응적 가중치의 값을 분석해 보면 10을 중심으로 형성되는 것을 확인할 수 있다. 프레임의 변화 정도에 따라서 가중치 값이 달라지지만 검출 결과에 영향을 줄 수 있는 수준의 변화는 발생하지 않는 것을 알 수 있었다.

본 절에서 제안하는 가중치 분산을 이용한 적응적 장면 전환 검출 방법은 1절에서 제안했던 프레임 평균을 이용한 방법보다 검출 정확도가 향상되었다. 따라서 제안한 가중치 분산이 프레임의 특성을 더욱 잘 구분하는 것을 알 수 있었다. 하지만 분산을 이용하기 때문에 그만큼의 계산량 증가가 발생한다. 또 적응적 가중치 계산을 위한 부분도 추가되었기 때문에 실제 장면 전환 검출에서의 실행 시간이 크게 증가되었다. 다음 절에서는 비교적 검출 정확도가 높으면서 계산복잡도가 낮은 히스토그램을 이용한 장면 전환 검출 방법을 제안한다. 이와 함께 적응적 가중치 설정에서도 더욱 간소화된 방법을 제시한다.

3. 히스토그램을 이용한 완전 자동적 장면 전환 검출

2절에서는 가중치 분산을 이용한 장면 전환 검출을 제안했었는데, 검출의 정확도는 향상되었지만 시간복잡도가 증가하는 단점이 있었다. 따라서 본 절에서는 비교적 정확한 장면 전환 방법으로 사용되는 히스토그램을 프레임 특징 추출을 척도로 이용한다. 히스토그램은 분산을 계산하는 과정보다 간단하기 때문에 계산량의 감소 측면에서도 효과적인 방법이라는 가정에서 시작한다.

가. 차분 히스토그램 평균값

본 절에서는 프레임의 특징값을 측정하기 위해서 차분 히스토그램의 평균값을 이용한다. 식 (42)에서는 차분 히스토그램의 평균을 정의하고 있다.

$$DHM_i = \frac{\sum_{j=0}^{255} |H_i(k) - H_{i-1}(k)|}{256} \tag{42}$$

여기에서 DHM은 차분 히스토그램의 평균, H_i는 i번째 프레임에서의 히스토그램을 각각 나타낸다. 일반적으로 DHM의 값이 임의의 임계값보다 큰 경우에 장면 전환이 일어난 것으로 판정한다.

나. 적응적 임계값 설정 방법

본 절에서는 2절에서 제안한 방법을 이용해서 적응적 임계값을 생성한다. 다만 여기에서는 가변 참조 구간 동안의 차분 히스토그램의 평균값을 계산에 이용한다. 적응적 임계값 설정 방법을 식 (43)에 정의하고 있다.

$$THH_i = \frac{\sum_{j=S_p+1}^{S_i-1} DHM_j}{S_i - S_p + 1} \times WH_i \tag{43}$$

여기에서 THH는 적응적 임계값, WH_i는 임계값 가중치를 각각 의미한다. i번째 프레임에서의 임계값 THH_i는 가변 참조 구간 동안, 즉 이전 장면 전환 프레임부터 현재 프레임의 이전 프레임까지의 DHM_j의 평균에 임계값 가중치 WH_i를 곱한 값으로 정의된다. 현재 프레임의 DHM과 임계값 THH를 비교해서 DHM이 더 큰 경우에 장면 전환 검출이 발생한 것으로 판단하다. 그리고 장면 전환이 발생된 경우에는 임계값을 0으로 초기화한다. 그리고 장면 전환의 다음 프레임에서는 장면 전환 비교를 생략하고 임계값만 생성한다. 이때의 초기 임계값은 DHM 값을 그대로 이용하게 된다.

2절에서와 같이 가중치 WH에 대한 테스트 결과 2.7에서 가장 좋은 검출 결과가 나타나는 것을 알 수 있었다. 하지만 통계적 가중치는 영상에 적응성을 가지지 못하므로 아래에 적응적 가중치 결정 방법을 제안하고 있다.

다. 임계값 가중치의 자동적 결정 방법

임계값 가중치의 역할은 앞선 2절에서 설명하였다. 2절에서 제안한 임계값 가중치 결정식은 차분 프레임 화소 개수를 이용하기 때문에 계산복잡도가 높았다. 본 절에서 제안하는 방법은 프레임의 특징 추출 척도로써 히스토그램을 이용하기 때문에 가중치 결정에서도 히스토그램의 빈의 개수를 이용한다. 식 (44)와 (45)에서 이 방법을 표현하고 있다.

$$NZBM_i = \frac{\sum_{j=S_p+1}^{S_i-1} NZB_j}{S_i - S_p + 1} \tag{44}$$

$$WH_i = \frac{512}{NZBM_i} \tag{45}$$

여기에서 $NZBM$은 NZB의 가변 참조 구간 동안의 평균, NZB는 차분 히스토그램의 0이 아닌 빈의 개수를 각각 나타낸다. 즉, 임계값 가중치 WH_i는 차분 히스토그램에서 변화가 발생한 빈들의 개수의 참조 구간 동안에서의 누적 평균을 512에 나누어준 값으로 정의된다. 여기서 512로 설정된 것은 두 개의 히스토그램이 사용되기 때문이다. 실제 가중치의 값을 분석해보면 2.5에서 3.0을 중심으로 분포되는 것을 확인할 수 있다.

이와 같은 임계값 가중치 결정은 2절에 비해서 계산량이 감소된 방법이다. 제안하는 임계값 설정 방법과 가중치 설정 방법을 통해서 영상에 독립적이며 실시간으로 처리가 가능한 완전 자동적인 장면 전환 검출 알고리즘을 구성하였다.

4. 실시간 장면 전환 검출 모델

앞선 1, 2, 3절에서는 각각 프레임의 평균, 가중치 분산, 히스토그램 등을 이용한 적응적인 장면 전환 검출 방법에 대해서 알아보았다. 제안한 장면 전환 검출 방법은 공통적인 알고리즘의 구성을 가지고 있다. 프레임 특징 추출 척도의 특성을 제외한 알고리즘의 일반적인 구성은 아래의 그림 16과 같이 3가지로 이루어져있다.

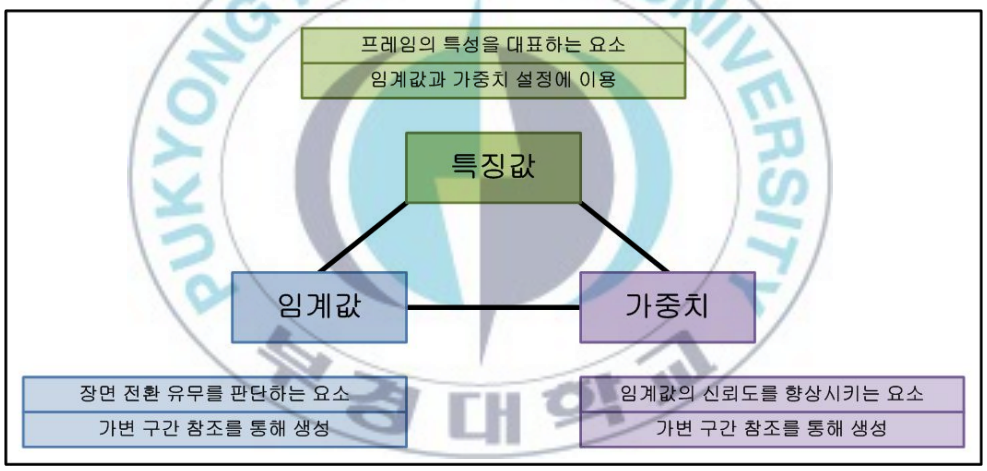


그림 16. 실시간 장면 전환 검출 모델의 구성

Figure 16. Structure of real-time shot change detection model

본 절에서는 앞서 제안한 장면 전환 검출 방법들의 공통점들을 이용하여 실시간 장면 전환 검출 모델을 제안한다. 실시간 장면 전환 검출 모델은 특징값, 임계값, 가중치의 세 가지 요소로 구성되는데 아래에서 이 세 가지 요소들에 대해서 알아본다.

가. 특징값

제안하는 실시간 장면 전환 검출 모델의 구성 요소 중에 특징값은 프레임의 특성을 나타내는 요소로 정의된다. 제안하는 모델에서는 이 특징값을 FV로 정의한다.

여기서 특징값은 프레임의 특성을 대표할 수 있는 값이면 어떠한 요소도 가능하다. 예를 들어, 앞서 제안하는 방법에서 사용했던 차분 프레임의 평균, 가중치 분산, 차분 히스토그램의 평균 등이 될 수 있다. 기존의 장면 전환 검출 알고리즘들은 이러한 특징값에 종속적으로 생성되어 왔다. 즉 특징값에 따라서 장면 전환 판단을 위한 알고리즘들이 만들어진 것이다. 이런 경우 알고리즘은 특정한 특징값에만 사용될 수 있으므로 다른 특징값을 사용하게 되면 알고리즘을 수정해야하는 단점이 있다.

반면 제안하는 모델은 위에서 언급한 특징값에 독립적으로 사용할 수 있다. 프레임의 특성이 어떤 형태로든지 스칼라 값으로만 유도된다면 제안하는 방법에 적용할 수 있다. 제안하는 방법이 특징값에 독립적인 이유는 특징값을 프레임을 대표하는 대표값으로만 사용할 뿐 특징값이 의미하는 프레임의 특성은 장면 전환 판단의 기준으로 사용하지 않기 때문이다. 예를 들어 프레임의 평균값이나 히스토그램의 평균값은 각각이 의미하는 프레임의 특성은 다르지만 제안하는 방법에서는 프레임을 대표하는 값 이상의 의미는 없다. 따라서 제안하는 방법이 특징값으로부터 독립적일 수 있다. 이런 특성의 장점은 기술의 응용환경에서 공간영역의 값이나 주파수 영역의 값에 구애받지 않고 사용할 수 있어 여러 응용분야에서 사용이 가능하다는 점이다.

나. 임계값

다음으로 고려되어야 하는 요소는 임계값이다. 프레임의 특징값으로부터 장면 전환을 판단하는 것은 순전히 임계값의 역할이다. 따라서 적합한 임계값 설정이 정확한 장면 전환 검출 알고리즘을 결정하게 된다.

가장 일반적으로 사용되는 임계값 결정 방법은 고정된 값을 사용하는 것으로, 단순히 임의의 값을 임계값으로 설정해서 장면 전환을 검출하는 것이다. 이 방법은 임계값 설정의 근거가 없으며, 실제 응용 환경에서는 사용할 수 없다. 또 영상마다 다른 임계값을 필요로 하는 단점이 있다. 다른 방법으로는 비디오 시퀀스 전체의 특징을 분석하여 그것을 토대로 임계값을 설정하는 방법이다. 이 방법은 최소한 2번 이상의 전체 시퀀스에 대한 반복이 필요하기 때문에 실시간 응용 환경에서는 사용할 수 없다. 또한 상당한 연산 시간을 필요로 하기 때문에 부적합하다.

따라서 실시간 장면 전환 검출을 위한 임계값 결정을 위해서는 다음 3가지 조건이 만족되어야 한다. 첫 번째는 모든 영상에 적응적으로 설정되어야 하는 것이다. 두 번째는 프레임의 재생과 동일한 순차성을 가져야 한다는 것이다. 세 번째는 지연시간이 없어야 한다는 것이다.

본 논문에서 제시하는 위의 3가지 조건을 만족하는 해결방법은 다음과 같다. 첫 번째 모든 영상에 대한 적응적 설정을 위해서는 입력 영상 자체의 정보를 이용해야 한다. 두 번째 순차성을 가지기 위해서는 현재 프레임의 이전 프레임들의 정보만 이용한다. 세 번째 지연시간을 최소화하기 위해서는 이미 계산되어진 값을 이용한다.

이런 3가지 조건의 해결방법으로부터 제안하는 모델의 임계값 설정방법은 다음의 식 (46)으로 나타낼 수 있다.

$$TH_i = \frac{\sum_{j=S_p+1}^{S_i-1} FV_j}{S_i - S_p + 1} \times WE_i \tag{46}$$

여기서 S_i는 i번째 프레임, S_p는 이전 장면 전환 프레임, W_i는 i번째 프레임의 임계값 가중치를 각각 나타낸다. i번째 프레임에서의 임계값 TH_i는 $p+i$번째 프레임에서 $i-1$번째 프레임까지의 FV_j의 평균에 i번째 프레임의 임계값 가중치 WE_i의 곱으로 정의할 수 있다. 즉, 제안하는 모델의 적응적 임계값은 가변 참조 구간에서의 특징값의 평균을 이용해서 계산된다. i번째 프레임에서의 특징값 FV_i가 적응적 임계값 TH_{i-1}보다 큰 경우에 장면 전환을 검출한다. 장면 전환 검출이 발생하면 TH는 0으로 초기화한다. 식 (46)을 적용하지 않는 예외적인 임계값 설정은 장면 전환이 발생한 다음 프레임과 비디오 시퀀스의 맨 처음 프레임에서만 이루어진다. 이 경우에는 장면 전환 판단은 생략되고, 초기 임계값을 생성하는 연산만 수행한다. 초기 임계값은 해당 프레임의 FV의 값을 이용한다.

제안한 임계값 설정 방법을 통해서 실시간을 만족하며 영상에 적응성을 가지는 임계값 생성이 가능해진다. 비교 프레임의 이동에 따라서 참조 구간이 변화하게 되는데, 이와 같은 가변 참조 구간을 이용하는 것은 현재 비교하는 프레임과 동일한 장면 내의 프레임들의 정보만 이용하기 위해서이다. 제안하는 가변 참조 구간 기법은 참조 프레임들의 정보가 균일성을 가지기 때문에 임계값의 정확도가 향상된다.

임계값은 가변 참조 구간 동안의 평균 특징값에 임계값 가중치 WE를 적용해서 최종적인 값을 계산한다. 임계값의 신뢰도를 높이기 위한 역할을 하는 WE에 대해서는 다음 항에서 기술하고 있다.

다. 가중치

앞서 언급한 바와 같인 제안하는 실시간 모델에서 임계값을 생성하는 과정에서 WE의 역할이 중요하다. 가변 참조 구간을 이용하는 임계값 설정 방법은 개념적으로는 성립되지만 실제 적용을 위해서는 마진을 적용해야 하는데, 이 마진이 WE가 된다.

만약 가중치를 사용하지 않는다면, 즉 $WE_i = 1$을 만족할 경우 임계값과 특징값의 관계는 $TH_i \approx FV_i$이 된다. 이것은 TH가 참조 구간에서의 프레임들의 FV의 평균으로 이루어지기 때문이다. 따라서 두 값이 비슷한 값을 가지게 되면 장면 전환이 발생하지 않는 부분에서 오검출을 일으키게 된다. 그러므로 적당히 TH를 증가시켜야 하는데 여기에서 임계값 가중치 WE를 이용한다. 임계값 가중치를 적용함으로 해서 임계값의 신뢰도를 향상시킬 수 있다.

본 논문에서는 WE를 설정하는 두 가지 방법을 제시한다. 첫 번째 방법은 통계적인 수치를 이용하는 것이다. 영상의 크기를 고정된 변수로 설정하면 서로 다른 영상에 대해서도 가중치 WE가 거의 동일한 값 또는 좁은 범위 내에서 최적의 장면 전환 검출을 수행하는 것을 실험적으로 알 수 있었다. 따라서 표본 집단의 충분한 개수의 샘플을 만족한다면 통계적인 임계값 가중치 WE를 구할 수 있다. 예를 들어 히스토그램을 특징값으로 이용했을 경우 176×144 크기의 비디오 영상에서는 $WE = 2.7$ 또는 $2.5 < WE < 3.0$에서 가장 정확한 장면 전환 검출을 수행하는 것을 알 수 있었다. 하지만 통계적 수치를 이용하는 방법은 언제든지 예외가 발생할 수 있다. 입력 영상의 정보와는 무관한 특성을 가지기 때문에 영상에 대한 적응성을 만족하지 못할 수 있다.

이런 단점을 개선하기 위해서 본 절에서는 임계값 가중치도 적응적이며 자동적인 설정이 가능하도록 하는 방법을 제안한다. 이 방법은 비교하는 두 프레임의 특징값의 변화량 중에서 변화가 생긴 부분과 생기지 않은 부분의 비율을 이용하는 것을 기초로 한다. 이 개념을 수식으로 표현하면 식 (47)과 (48)로 정의할 수 있다.

$$NZM_i = \frac{\sum_{j=S_p+1}^{S_i-1} NZ_j}{S_i - S_p + 1} \tag{47}$$

$$WE_i = WF(NZM_i) \tag{48}$$

여기에서 NZM_i은 가변 참조 구간 동안의 NZ_j의 평균, NZ_j는 프레임의 변화량 중에서 변화가 생긴 요소의 개수를 의미하며, WE_i는 NZM_i를 함수 WF에 적용시킨 가중치를 나타낸다. 즉, 제안하는 적응적 가중치 WE는 가변 참조 구간 동안의 프레임의 특징값의 변화가 발생한 양을 함수 WF에 적용시킨 결과값으로 정의한다. 여기에서 함수 WF는 프레임의 특징값에 따라서 다르게 적용될 수 있다. 식 (41)과 (45)가 함수 WF의 구체화된 형태를 의미한다.

두 번째 제안하는 가중치 설정 방법을 통해서 제안하는 실시간 장면 전환 검출 모델을 완전 자동적인 알고리즘으로 완성된다. 사용자의 개입 없이 입력 영상에 따라서 정확한 장면 전환 검출을 수행할 수 있다. 가변 참조 구간이라는 개념을 이용하기 때문에 제안하는 내용에 대한 이해가 어려울 수 있다. 다음 절에서는 그림과 예제를 통해서 제안하는 실시간 장면 전환 검출 모델을 설명하고 있다.

5. 예제를 통한 제안하는 모델의 설명

제안하는 모델은 비교 프레임이 변화함에 따라서 참조 구간이 달라지는 가변 참조 구간 방식을 이용한다. 본 절에서는 그림을 통해서 제안한 모델을 설명하고 예제를 통해서 실제 알고리즘을 수행하는 과정을 알아본다. 그리고 전체 흐름도를 통해서 제안하는 모델의 설명을 마무리한다.

가. 참조구간

제안하는 모델은 가변 참조 구간에서의 평균 특징값을 이용해서 임계값과 가중치를 설정한다. 그림 17에서는 비교 프레임의 이동과 이에 따른 참조 구간이 변화하는 과정을 도식화하여 나타내었다.

1단계에서는 두 프레임을 비교하여 프레임 특징값의 평균값 FV_1을 계산하고, 초기 임계값 $TH_1 = FV_1$을 생성한다. 이 단계에서는 장면 전환 판단 과정은 생략된다. 2단계에서는 먼저 FV_2를 계산하고, TH_1과 FV_2를 비교하여 장면 전환을 판단하게 된다. 여기서는 $FV_2 < TH_1$이므로 장면 전환이 발생하지 않으며, 따라서 $TH_2 = (FV_1 + FV_2)/2$가 된다. 3단계도 먼저 FV_3을 계산하고, TH_2와 FV_3을 비교한다. 여기서도 $FV_3 < TH_2$가 되므로 장면 전환이 발생하지 않으며 $TH_3 = (FV_1 + FV_2 + FV_3)/3$이 된다. 다음으로 4단계에서도 먼저 FV_4를 계산하고, TH_3과 FV_4를 비교하게 된다. 여기서는 장면이 바뀌는 부분으로 $FV_4 > TH_3$이 되고, 장면 전환 프레임의 번호가 저장 또는 출력된다. 이 때 임계값은 $TH_4 = 0$으로 초기화된다. 이후의 장면 전환 검출의 진행은 다시 위의 단계를 반복하게 된다.

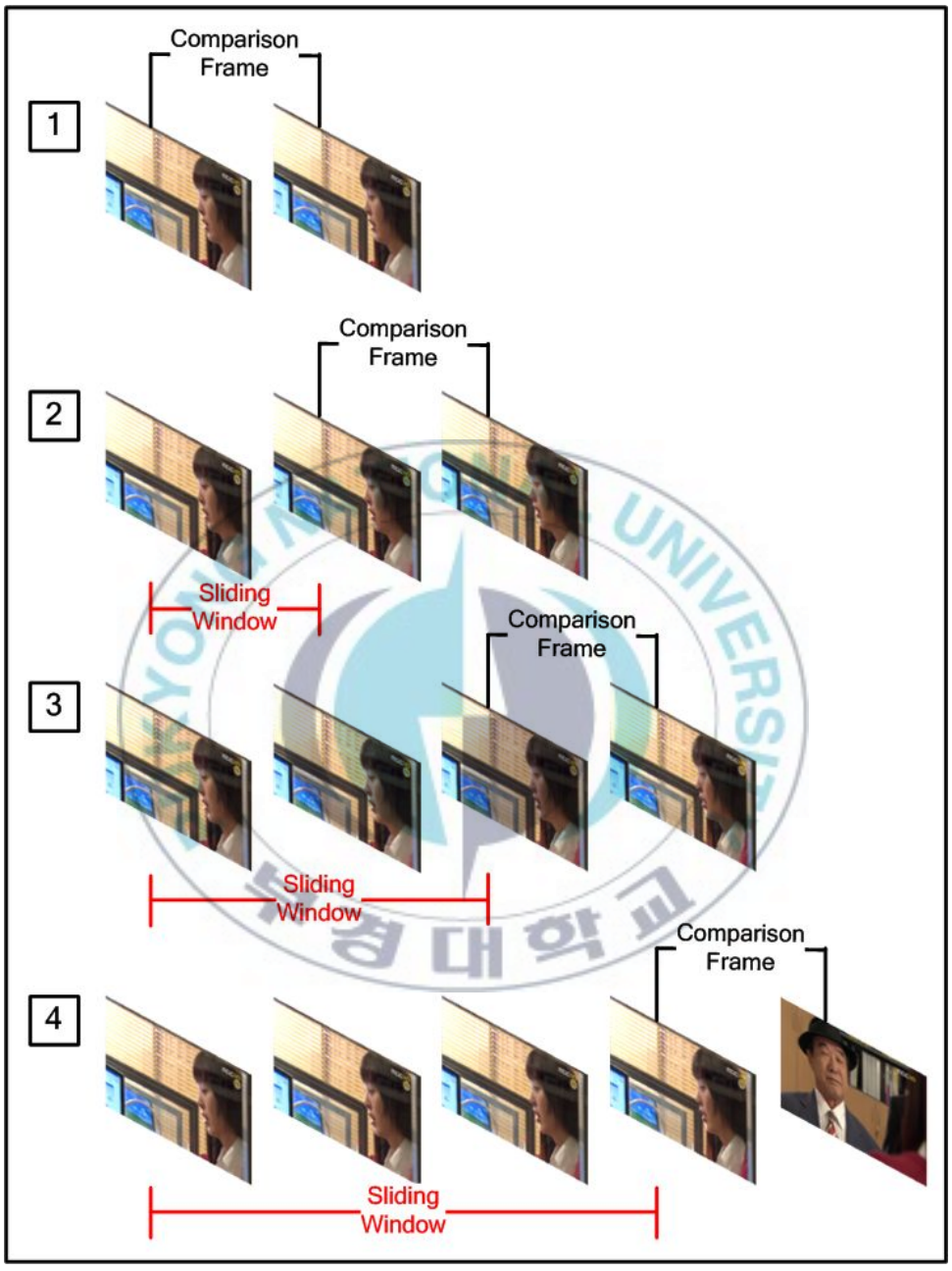
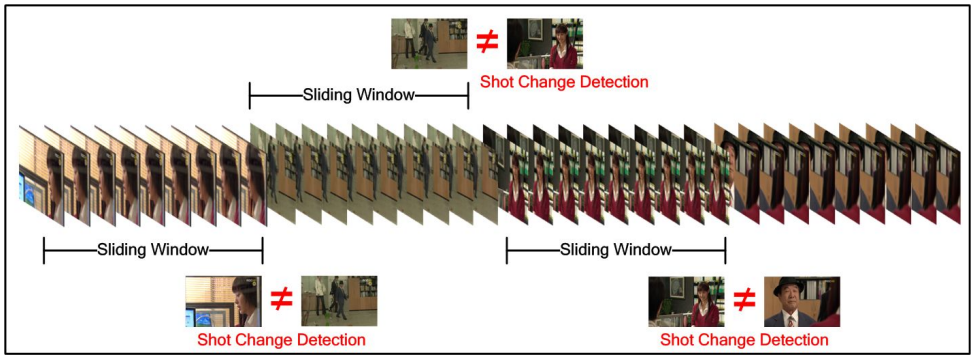
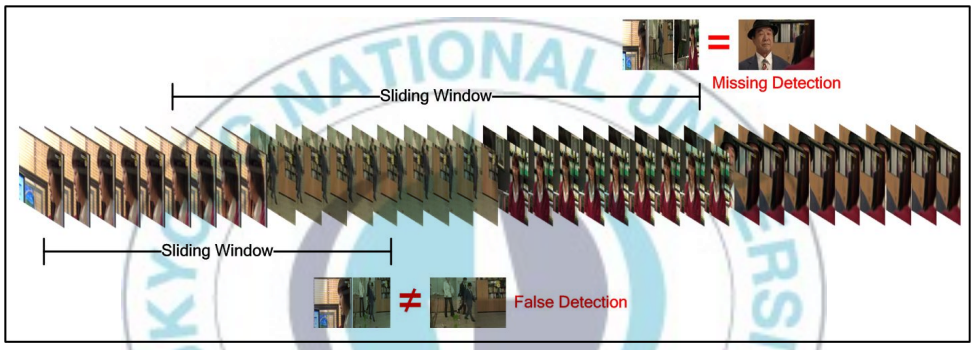


그림 17. 프레임 이동에 따른 가변 참조 구간의 변경

Figure 17. Alteration of variable sliding window according frame move



(a)



(b)

그림 18. 참조 구간에 따른 장면 전환 검출 비교((a) 가변, (b)고정)

Figure 18. Comparison of shot change detection according sliding window((a) variable, (b) fixed)

그림 18에서는 참조 구간 설정에 따른 장면 전환 검출 결과를 비교하였다. (a)에서는 제안하는 가변 참조 구간을 이용한 경우를 나타내었다. 가변 참조 구간 방법을 사용하면 동일 장면의 프레임들의 특징값만을 모아서 다음 장면과 비교할 수 있기 때문에 정확한 검출이 가능해진다. 반면 (b)와 같이 고정된 길이의 참조 구간을 사용하게 되면, 여러 장면의 특징들이 섞이게 되므로 특징값의 균일성을 저해시킨다. 결과적으로 장면 비교에서 오검출이나 미검출을 발생시킬 확률이 높아진다.

나. 가중치

앞서 언급한 바와 같이 가중치는 임계값의 신뢰도를 향상시키는 역할을 한다. 그림 19에서는 가중치를 사용한 경우와 사용하지 않은 경우를 가정하여 각각의 장면 전환 검출 과정을 나타내었다.

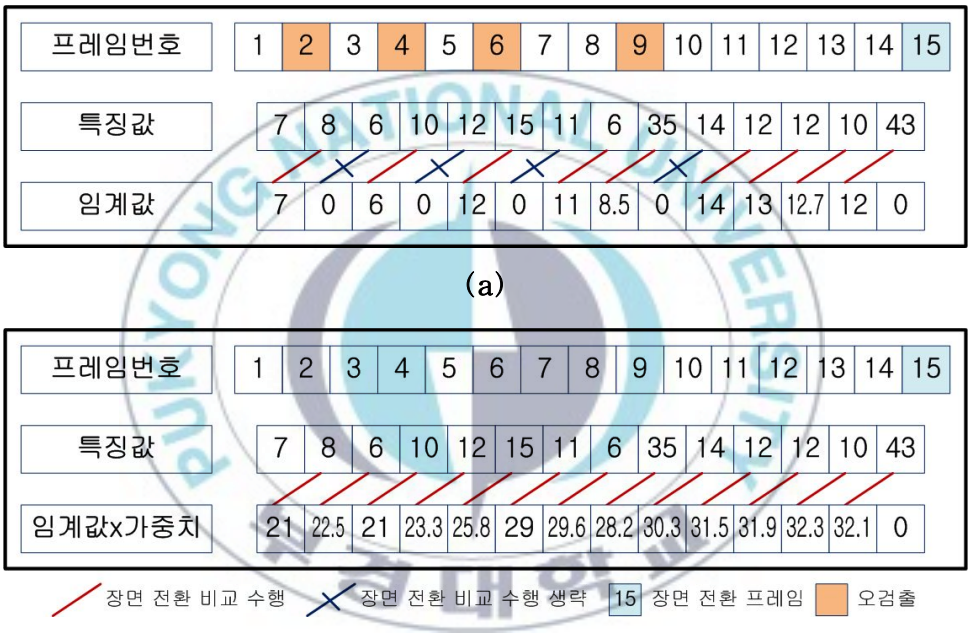


그림 19. 가중치에 따른 장면 전환 검출 과정
((a) 가중치 비적용, (b) 가중치 적용)

Figure 19. Scheme of shot change detection by weighting factor
((a) weighting not applied, (b) weighting applied)

그림 19의 (a)의 경우에 가중치를 적용하지 않고 임계값만으로 장면 전환 검출을 수행하는 절차를 나타내고 있다. 적색으로 표시된 프레임 번호, 2, 4, 6, 9번 프레임에서 오검출이 발생하는 것을 확인할 수 있다.

대부분 임계값과 특징값의 크기가 거의 유사하지만 특징값보다 임계값이 작기 때문에 오검출이 발생된다. 9번과 10번 프레임과 같이 동일한 장면 내에서도 특징값의 차이가 큰 경우가 예외적으로 발생할 수 있다. 이런 경우에도 가중치를 적용하지 않은 경우에서는 잘못된 장면 검출 결과를 나타낸다.

반면 (b)에서와 같이 가중치를 3으로 적용한 경우 오검출 없이 정확한 장면 전환 검출을 하는 것을 확인 할 수 있다. 9번과 10번 프레임에서도 특징값이 갑자기 증가하였지만 적절한 가중치의 사용으로 오검출이 발생되지 않는다. 따라서 임계값에 가중치를 적용하는 것이 바람직함을 확인할 수 있다.

아래의 표 8에서는 가중치를 사용한 경우와 사용하지 않은 경우에 대한 실제 검출 결과를 바탕으로 임계값 가중치의 필요성을 나타내고 있다. 특징값으로 차분 히스토그램의 평균값을 사용하여 동일한 비디오 데이터의 동일 구간에서의 특징값, 임계값, 검출 결과를 비교하였다.

가중치가 적용된 표의 좌측에서는 28번 프레임에서 장면 전환을 검출하였다. 이것은 실제로도 장면 전환이 발생한 프레임으로써 정확하게 장면 전환을 찾아낸 것을 확인할 수 있다. 반면 가중치를 적용하지 않은 표의 우측에서는 많은 오검출이 발생하게 되어 장면 전환 검출을 제대로 수행할 수 없음을 확인할 수 있다. 임계값이 특징값의 평균으로 계산되므로 이런 현상이 발생하게 된다. 따라서 가중치를 적용하지 않은 경우 임계값의 분포가 전체적으로 낮아져서 오검출이 발생하게 되므로 적당한 크기의 가중치가 요구된다. 만약 너무 높은 가중치가 적용된다면 오검출은 없겠지만 상대적으로 실제 장면 전환을 검출하지 못하는 미검출의 빈도가 상승될 것이다.

표 8. 임계값 가중치의 필요성

Table 8. Necessity of threshold weighting factor

번호	가중치 적용			가중치 미적용		
	DHM	*THH*× *WH*	상태	*DHM*	*THH*×1	상태
1	8.210938	0.000000	start	8.210938	0.000000	start
2	11.414063	11.581267	compare	11.414063	8.210938	false
3	8.859375	18.335766	compare	8.859375	0.000000	start
4	12.132813	19.923497	compare	12.132813	4.429688	false
5	7.921875	22.703057	compare	7.921875	0.000000	start
6	8.625000	22.592727	compare	8.625000	3.960938	false
7	8.960938	22.794393	compare	8.960938	0.000000	start
8	9.492188	23.031293	compare	9.492188	4.480469	false
9	9.164063	23.379227	compare	9.164063	0.000000	start
10	11.015625	23.527371	compare	11.015625	4.582031	false
11	10.414063	24.161576	compare	10.414063	0.000000	start
12	7.984375	24.595206	compare	7.984375	5.207031	false
13	8.351563	24.361667	compare	8.351563	0.000000	start
14	15.031250	24.253576	compare	15.031250	4.175781	false
15	11.101563	25.411255	compare	11.101563	0.000000	start
16	10.046875	25.761083	compare	10.046875	5.550781	false
17	7.593750	25.914541	compare	7.593750	0.000000	start
18	9.406250	25.557023	compare	9.406250	3.796875	false
19	11.015625	25.633048	compare	11.015625	0.000000	start
20	7.726563	25.939230	compare	7.726563	5.507813	false
21	10.218750	25.661856	compare	10.218750	0.000000	start
22	11.164063	25.724104	compare	11.164063	5.109375	false
23	12.492188	26.028262	compare	12.492188	0.000000	start
24	9.445313	26.337463	compare	9.445313	6.246094	false
25	12.679688	26.421007	compare	12.679688	0.000000	start
26	8.851563	26.716667	compare	8.851563	6.339844	false
27	9.726563	26.596955	compare	9.726563	0.000000	start
28	39.039063	26.572531	SCD	39.039063	4.863281	SCD
29	11.632813	0.000000	start	11.632813	0.000000	start

다. 전체 흐름도

제안하는 실시간 장면 전환 검출 모델은 특징값, 임계값, 가중치로 구성된다. 그림 20에서는 제안하는 모델의 장면 전환 검출 알고리즘의 전체 흐름도를 나타내고 있다.

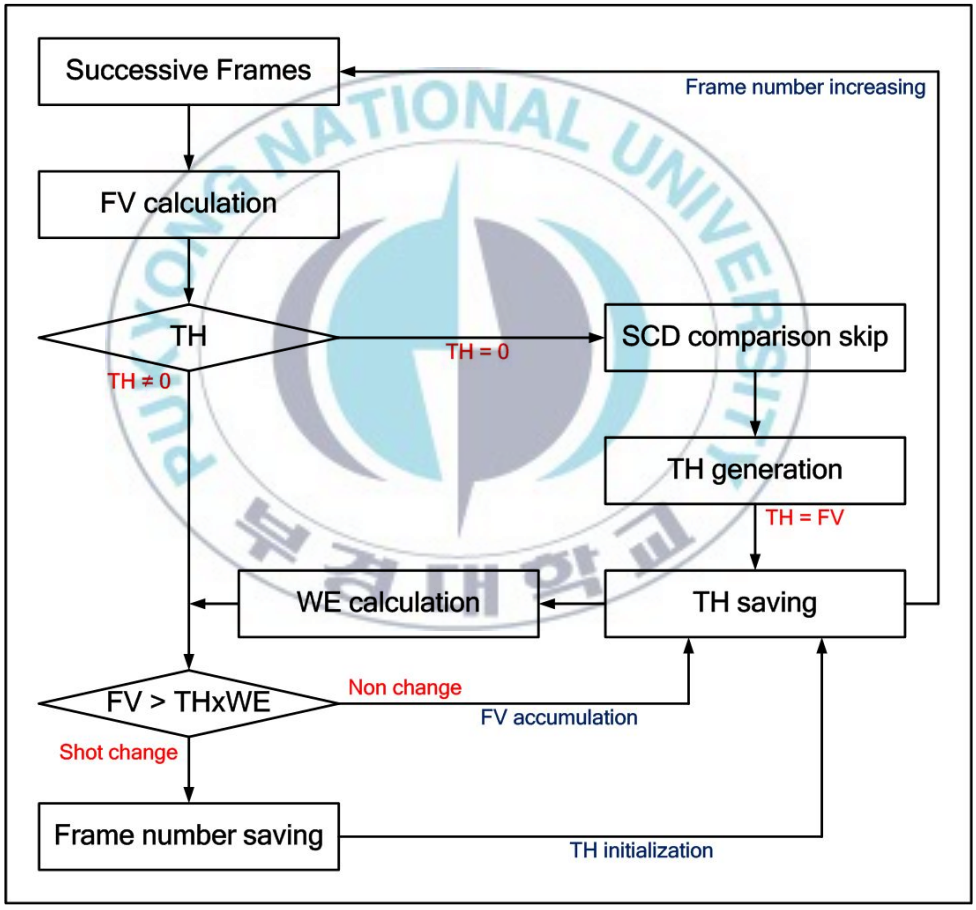


그림 20. 제안한 모델의 흐름도

Figure 20. Flow chart of proposed model

제안하는 모델은 연속된 두 프레임의 입력으로 시작된다. 먼저 두 프레임으로부터 특징값 FV를 계산한다. 다음에는 현재의 임계값 TH의 값을 검사해서 0인 경우에는 장면 전환 검출 비교를 생략하고 초기 임계값만 생성한 뒤에 다음 프레임의 입력으로 진행한다. 여기에서 초기 임계값은 앞서 계산한 특징값을 그대로 저장한다.

앞서 TH의 값이 0이 아닌 경우에는 장면 전환 검출 비교 과정으로 넘어간다. 여기에서는 특징값 FV와 임계값에 가중치를 적용한 $TH \times WE$를 비교하여 장면 전환을 판단한다. 장면 전환으로 판정된 경우에는 장면 전환 프레임의 번호를 저장하고 임계값을 0으로 초기화 시킨다. 그리고 다음 프레임의 입력으로 진행한다.

앞서 장면 전환 판단 과정에서 장면 전환이 아니라고 판정된 경우에는 현재 FV를 임계값에 누적해서 새로운 임계값을 생성하고 다음 프레임의 입력으로 진행한다.

제안하는 실시간 장면 전환 검출 모델은 위의 일련의 처리 과정을 반복하면서 장면 전환 검출을 진행한다.

PUKYONG NATIONAL UNIVERSITY 부경대학교

6. 개선된 장면 전환 검출 방법

본 장에서는 1, 2, 3절을 통해서 실시간 적응적인 장면 전환 검출 방법을 제안하였고, 4절에서 제안한 방법들을 토대로 한 실시간 장면 전환 검출 모델을 제시하였다. 본 절에서는 제안한 방법의 몇 가지 문제점을 해결하여 개선된 장면 전환 검출 알고리즘을 제안한다.

알고리즘의 개선은 3절에서 제안한 차분 히스토그램을 이용한 장면 전환 검출에서 시작한다. 본 절에서는 빈 데이터 영역의 처리, 페이드아웃 효과에 의한 장면 변화 등 3가지에 대한 처리 방법을 설명하고 있다.

가. 빈 데이터 영역 처리

비디오 데이터의 시각적인 구성을 살펴보면 영상 영역과 빈 데이터 영역으로 이루어진 것을 알 수 있다. 그림 21에서와 같이 비디오 데이터의 시각적 구성은 내용을 전달하는 영상 영역과 흑색으로 표현되는 빈 데이터 영역으로 나누어진다.

그림 21. 비디오 프레임의 시각적 구성
Figure 21. Visual constitution of video frame

본 논문에서는 영상 영역(image area)을 IA, 빈 데이터 영역(blank data area)을 BDA로 각각 정의한다. BDA는 비디오 데이터의 가로와 세로 비율에 따라서 화면에 IA로 나타낼 수 없는 부분을 대체하는 역할을 한다. 따라서 아무런 정보도 없는 빈 영역이다. BDA는 비디오 데이터 전체에 걸쳐서 모든 프레임에서 나타나지만 프레임을 대표하는 고유의 특성으로 간주하기는 힘들다. 즉, BDA의 크기만큼 영상영역이 더 있다고 간주하면 프레임의 변화량으로 가중치를 계산하는 방법에 대한 수정이 필요하게 된다. 정확하게는 3절의 식 (44)에서 NZB의 값이 새롭게 조절되어야 한다.

이를 위해서 프레임의 크기에서 IA와 BDA의 크기를 고려한 수식을 적용해서 NZB의 값을 새롭게 계산할 수 있다. 다른 방법으로는 일정한 실수의 곱을 통해서 가장 좋은 검출 결과를 나타내는 수치를 통계적으로 찾는 방법이 있다. 위의 두 가지 방법을 비교해본 결과 기존의 NZB를 $NZB \times 1.1$로 대체하는 것이 가장 효율적임을 알 수 있었다.

따라서 식 (44)에서 BDA를 고려하여 개선한 임계값 가중치 계산식을 식 (49)로 정의할 수 있다.

$$NZBM_i = \frac{1.1 \times \sum_{j=S_p+1}^{S_i-1} NZB_j}{S_i - S_p + 1} \tag{49}$$

나. 페이드아웃 효과에 의한 장면 변화 처리

다음으로 고려하는 요소는 페이드아웃(fade-out, FO) 장면 변화에 대한 검출 방법이다. Ⅱ장에서 언급한 FO 효과는 영상에서 화면이 점점 어두워져서 암흑으로 변하고 이후 새로운 장면이 나타내는 영상 편집 기술을 의미한다. 전체 비디오 프레임에 차지하는 비중은 적지만 장면 전환 검출의 정확도를 높이기 위해서 예외적인 처리를 해주어야 한다. 이를 위해서 본 절에서는 가변적 가중치 설정 방법을 이용한다.

FO 프레임들에서는 화면이 점점 암흑으로 변하기 때문에 차츰 프레임의 특징값이 낮아진다. 반면 임계값은 크게 변화하지 않기 때문에 엉뚱한 지점에서 오검출이 발생하고, 이로 인해서 실제 장면 전환 지점에서는 검출을 하지 못하게 된다. 이것을 해결하기 위해서 프레임의 특징값인 *DHM*이 1이하인 프레임의 개수에 따라서 가변적 가중치를 적용한다. 즉, 프레임의 특징값이 작기 때문에 임계값의 크기를 강제로 낮게 설정되도록 조절하는 것이다. 자세한 코드는 그림 22에서 나타내고 있다.

```
// Fade-out transition count
if (DHM[z] < 1)      w_c++;

// Variable weighting factor setting
if (w_c < 10 && w_c > 0)                    w = 5;
else if (w_c < 20 && w_c >= 10)             w = 4;
else if (w_c < 30 && w_c >= 20)             w = 3;
else if (w_c >= 30)                         w = 2;
else                                        w = 512/NZBM;
```

그림 22. 가변 가중치 설정 코드

Figure 22. Setting codes of variable weighting factor

Ⅳ. 실험 결과 및 분석

본 장에서는 제안한 방법에 대한 실험 결과와 분석을 기술하고 있다. 기존의 다양한 방법들과 비교를 통해서 제안한 방법의 우수성을 증명하고자 한다. 제안한 방법의 성능을 평가하기 위한 실험 방법과 내용, 실험의 과정, 실험의 결과와 의의 등을 기술하였다. 또한 PMP에서의 구현을 통해서 제안한 방법이 실제 하드웨어 장치에서도 실시간으로 장면 전환 검출이 가능한 것을 알아본다.

본 장의 구성은 다음과 같다. 먼저 1절에서는 실험 환경에 대해서 설명한다. 실험 전에 정의되는 여러 가지 사항들에 대해서 기술하고, 실험 환경과 실험 내용들을 서술하고 있다. 2절에서는 실험 결과와 분석을 기술하였다. 차분 프레임의 평균을 이용한 방법, 가중치 분산을 이용한 방법, 차분 히스토그램의 평균을 이용한 방법과 개선된 방법 등 총 4가지 제안한 방법을 기존의 방법들과 비교하여 나타내었다. 3절에서는 제안한 방법의 PMP에서의 구현에 대해서 설명하고 있다. 장면 전환 검출의 하드웨어 장치에서의 구현의 의의와 PMP의 구성 및 동작을 설명하고 제안한 방법의 PMP에서 구현하는 절차와 과정, 실험 결과를 분석하였다.

1. 실험 환경

본 절에서는 제안한 방법의 실험을 위한 환경에 대해서 알아본다. 사전 정의되는 사항들을 알아보고 실험을 수행한 환경과 실험 방법의 내용들을 정리하였다.

가. 사전 정의되는 사항

실험에 앞서 정의되는 몇 가지 정의되어야 할 요소들이 있다. 먼저 성능 평가를 위해서 실험된 제안한 방법은 크게 3가지로 분류되며 세부적인 개선정도에 따라서 총 10가지의 방법들이 있다. 표 9에서 이 방법들의 이름과 약어, 구체적으로 사용된 요소들을 각각 정의하고 있다.

표 9. 제안하는 방법들의 내용

Table 9. Contents of the proposed methods

방법	약 어	내 용(특징값/임계값/참조구간/가중치)
P1	FM_AT_FS_SW	프레임평균 / 적응적 / 가변 / 통계적
P2	WV_AT_FS_OW	가중치분산 / 적응적 / 가변 / 최적
P3	WV_AT_FS_SW	가중치분산 / 적응적 / 가변 / 통계적
P4	WV_AT_FS_AW	가중치분산 / 적응적 / 가변 / 적응적
P5	IH_AT_FS_SW	히스토그램 / 적응적 / 가변 / 통계적
P6	IH_AT_FS_AW	히스토그램 / 적응적 / 가변 / 적응적
P7	IH_AT_FS_HW	히스토그램 / 적응적 / 가변 / 혼합
P8	IH_AT_FS_AW+BD	히스토그램/적응적/가변/적응적 + BDA
P9	IH_AT_FS_AW+BD+FO	히스토그램/적응적/가변/적응적+BDA+FO
P10	CS_AT_FS_AW	chi-square test / 적응적 / 가변 / 적응적

표에서 제시한 제안한 방법들 중에서 P1은 Ⅲ장 1절에서 제안했던 차분 프레임의 평균을 특징값으로 사용하고, 가변 참조 구간의 평균 특징값을 이용하는 적응적 임계값을 이용한다. 또 임계값 가중치는 실험을 통해서 통계적으로 얻어진 값을 사용하는 방법이다.

P2~P4는 Ⅲ장 2절에서 제안했던 가중치 분산을 특징값으로 사용하는 방법으로 역시 가변 참조 구간을 이용하는 적응적 임계값을 이용한다. 차이점으로는 P2에서는 가중치로 최적의 결과를 보이는 값을, P3에서는 실험을 통한 통계적 가중치를, P4에서는 가변 참조 구간을 이용한 적응적 가중치를 각각 사용한다.

P5~P9는 Ⅲ장 3절에서 제안했던 차분 히스토그램을 특징값으로 사용하는 방법이다. 가변 참조 구간을 이용하는 적응적 임계값을 공통적으로 이용한다. 하지만 P5에서는 통계적 가중치를 사용하고, P6에서는 적응적 가중치를 이용한다. P7에서는 통계적 가중치와 적응적 가중치를 혼합한 방법을 쓰며, P8은 P6에서 BDA 처리를 통한 개선된 알고리즘이다. P9는 P8에서 FO 처리를 추가시킨 알고리즘이다.

마지막으로 P10은 히스토그램과 유사한 방식으로 계산되는 chi-square 검정을 이용한 방법으로 P6 방법에서 특징값만 chi-square 검정으로 바꾸어서 사용하였다.

제안하는 방법과 함께 정의되어야 하는 사항이 비교방법에 관한 것이다. 아래의 표 10에서는 실험에서 사용된 비교방법들의 약어와 내용들을 정의하고 있다.

C1은 화소값의 차이를 특징값으로 하며 최적임계값을 사용하는 방법이다. 여기에서 최적임계값이라는 것은 임계값으로 고정된 방법을 사용하여 모든 테스트 비디오가 최적의 검출 결과가 나올 때까지 반복해서 실험하여 검출 결과가 최고가 나온 경우의 임계값들의 집합이다. 따라서

앞서 Ⅱ장의 3절에서 설명한 가변 고정 임계값 방법으로 실험한 결과가 되겠다. C2~C9까지의 방법들은 표에 언급한 특징값에 최적 임계값을 사용한 방법들이 되겠다. C10~C12는 각각의 특징값에 적응적 임계값을 사용하는데, 제안하는 방법과는 다른 고정 구간 참조를 한다. 이때의 고정 참조 구간은 122프레임 간격으로 한다. 122프레임은 실험 데이터의 평균적 장면간 거리를 의미한다. C13~C15의 방법은 참조한 논문의 방법들을 구현한 것이다.

표 10. 비교하는 방법들의 내용

Table 10. Contents of comparison methods

방법	약 어	내 용
C1	PW_OT[8]	화소값 / 최적임계값
C2	MS_OT[72]	MSE / 최적임계값
C3	FV_OT[82]	분산값 / 최적임계값
C4	HV_OT[45]	히스토그램 분산값 / 최적임계값
C5	IH_OT[12]	밝기 히스토그램 / 최적임계값
C6	CH_OT[14]	컬러 히스토그램 / 최적임계값
C7	CS_OT[7]	chi-square test / 최적임계값
C8	LH_OT[18]	블록 유사도 / 최적임계값
C9	DC_OT[48]	DC 계수값 / 최적임계값
C10	FV_AT_FS[6]	분산값 / 적응적임계값 / 고정참조구간
C11	IH_AT_FS[16]	밝기 히스토그램 / 적응적임계값 / 고정참조구간
C12	CH_AT_FS[14]	컬러 히스토그램 / 적응적임계값 / 고정참조구간
C13	CH_ME[68]	Cheng의 방법
C14	KO_ME[61]	Ko의 방법
C15	KI_ME[54]	Kim의 방법

실험 결과를 효과적으로 이해하고 용어의 혼동을 막기 위해 표 9와 10에서 여러 가지 방법들을 정리하였다.

나. 실험 환경

제안한 방법들의 성능 평가를 위한 실험 환경은 다음과 같다. 사용한 PC는 Core Duo 2.4GHz의 CPU와 2G의 RAM으로 구성되었고, 장면 전환 검출 프로그램은 VC++ .net 7.1을 이용해서 구현하였다. 장면 전환 검출 결과와 육안으로 구분한 장면 전환 결과와의 비교는 MATLAB 7.3을 통해서 이루어졌다. 모든 방법들에 대한 실험에 공통적으로 사용된 비디오 데이터는 4가지 장르의 10개 시퀀스를 이용하였다. 공정한 실험을 위해서 실험 데이터는 랜덤하게 선택되었으며, 공통적으로 전체 시퀀스 중에서 시작 5분에서부터 5분 동안의 데이터를 추출하여 사용하였다. 포맷에 대한 공정성을 기하기 위해서 추출된 비디오 데이터를 YUV 4:2:0 포맷의 9000프레임의 시퀀스로 변화하여 실험하였다. 또 비디오 데이터의 크기는 원본 영상의 크기에 상관없이 176×144로 축소하여 생성하였다.

표 11. 실험에 사용된 영상 종류와 장면 전환 개수

Table 11. The number of shot change detection and video title

장 르	제 목(장면 전환 개수)
드라마	쩐의전쟁(52), 커피프린스(93), 프리즌브레이크(88)
영 화	프리덤라이터스(111), 복면달호(32), 올드미스다이어리(52)
뉴 스	KBS뉴스(69), SBS뉴스추적(63)
오 락	무한도전(98), 개그콘서트(78)

표 11에서는 실험에 사용된 영상의 장르와 제목, 각 영상별 장면 전환 개수를 나타내고 있다. 총 90000프레임의 데이터에 737개의 장면 전환 프레임이 존재한다. 표 12에서는 737개의 장면 전환 중에서 급진적 장면

전환과 점진적 장면 전환, 점진성을 가지는 급진적 장면 전환의 개수와 비율을 나타내고 있다. 점진성을 가지는 급진적 장면 전환을 일반적인 급진적 전환으로 포함하면 전체에서 98% 이상의 비중을 차지하는 것을 알 수 있다. 실제 대부분의 비디오 시퀀스를 살펴보면 점진적 장면 전환이 차지하는 비중이 아주 낮은 것을 알 수 있다.

표 12. 장면 전환 프레임의 종류와 분포

Table 12. Kinds and distribution of shot change frames

구 분	급진적변화	점진성 급진적변화	점진적변화	합 계
개 수(개)	409	315	13	737
비 율(%)	55.5	42.7	1.8	100

제안한 방법의 성능 평가 척도는 앞의 Ⅱ장 5절에서 언급하였던 precision, recall, F1, 오류개수, 실행시간을 사용한다. 실험 결과를 나타내는 표에서는 precision, recall, F1은 소수점 4자리까지 표현하였고, 실행시간은 초단위 이하 소수점 2자리까지 표시하였다. 표에 나타난 실험 결과는 10개의 테스트 비디오의 결과의 평균값이다. 또 표에서의 FS는 검출한 장면 전환 개수, CD는 정확하게 검출한 장면 전환 개수, FD는 잘못 검출한 장면 전환 개수, MD는 검출하지 못한 장면 전환 개수, P는 precision, R은 recall, FM은 오류개수, RT는 실행시간을 각각 의미하는 용어다.

다. 실험 내용

제안한 방법들의 성능을 평가하기 위한 실험은 크게 4가지로 진행되었다. 첫 번째 실험은 최적 임계값을 사용하는 9가지 기존의 방법들과 제안하는 프레임 평균을 이용한 장면 전환 검출 방법을 176×144 크기의 영상에서 비교하였다. 두 번째 실험은 최적 임계값을 사용하는 9가지 기존의 방법들과 제안하는 가중치 분산을 이용한 장면 전환 검출 방법을 176×144, 11×9 크기의 영상에서 비교하였고, 고정참조구간을 이용하는 기존의 3가지 방법과도 비교하였다. 세 번째 실험에서는 제안하는 히스토그램을 이용한 장면 전환 검출 방법과 기존의 고정임계값 기반의 9가지 방법, 고정참조구간 기반의 3가지 방법, 적응적 임계값 기반의 3가지 방법과 각각 176×144 영상에 비교하였다. 네 번째 실험에서는 개선된 히스토그램을 이용한 장면 전환 검출 방법을 기존의 히스토그램 기반의 6가지 방법과 비교하였고, 제안한 10가지 방법의 결과를 검토하였다. 표 13에서는 각 실험의 내용을 나타내고 있다.

표 13. 실험 내용

Table 13. Contents of experiments

실험	비교방법들	제안방법들	실험영상크기
E1-1	C1~C9	P1	176×144
E2-1	C1~C9	P2~P4	176×144
E2-2	C10~C12	P2~P4	176×144
E2-3	C1~C9	P4	11×9
E3-1	C1~C9	P5~P6	176×144
E3-2	C10~C12	P5~P6	176×144
E3-3	C13~C15	P5~P6	176×144
E4-1	C5~C7, C11~C12	P5~P10	176×144
E4-2	-	P1~P10	176×144

2. 실험 결과 및 분석

본 절에서는 제안한 방법과 기존의 여러 방법들을 비교한 결과를 나타내고 결과의 분석을 통해서 제안한 방법의 우수성을 확인하고자 한다. 제안하는 방법들을 4가지로 분류하여 성능을 평가하였으며, 세부적 총 9가지의 실험을 진행하였다. 표에 사용된 실험 방법 및 비교 방법, 제안 방법들은 모두 1절에서 정의한 약어로 표현하였다.

가. 차분 프레임의 평균을 이용한 방법의 성능 평가

Ⅲ장 1절에서 제안한 차분 프레임을 이용한 장면 전환 검출 방법 P1을 기존의 고정 임계값 기반의 9가지 방법과 비교하였다. 이 실험은 표 13에서 E1-1로 정의하였고, 표 14를 통해서 실험 결과를 나타내었다.

실험 E1-1의 결과 제안한 방법 P1은 precision 0.8918, recall 0.8955, F1 0.8937, FM 153개, RT 23.84초의 성능을 나타낸다. 정확도에서는 F1을 기준으로 6번째, 시간에서는 RT를 기준으로 평균치를 나타내고 있다. 즉 제안한 방법 P1은 정확도와 시간적 측면에서 C1~C9의 중간 정도 수준을 나타낸다.

P1은 화소 기반의 방법인 C1, C2와는 비슷한 수행시간을 가지지만 정확도는 유사하거나 높게 나타나며, 분산 기반의 방법인 C3, C4 보다는 정확한 것을 알 수 있다. 하지만 히스토그램 기반의 방법인 C5~C7 보다는 검출 정확도가 0.04 이상 낮으며, 블록 기반의 방법인 C8보다 역시 0.05이상의 낮은 검출률을 보인다.

결과적으로 제안한 방법 P1은 전체적인 검출 정확도에서는 기존의 방법들보다 좋지 않은 방법이다. 하지만 동일한 프레임 특징 추출 척도를 사용하는 방법들에 비해서는 향상된 검출률을 나타낸다. 또한 비교 방법들이 최적 임계값을 사용하여 실제 응용환경에서는 적용할 수 없는 단점이 있지만 제안한 방법은 적응적 임계값을 사용하기 때문에 응용 환경에서의 사용성은 높다.

제안한 방법 P1에서는 무엇보다도 낮은 검출 정확도가 우선적으로 해결되어야할 문제로 지적된다. 이와 함께 임계값 가중치 역시 적응적으로 설정하여 입력 영상에 대한 알고리즘의 독립성을 높이는 문제도 해결되어야 한다.

표 14. 실험 결과 : E1-1

Figure 14. Experimental results : E1-1

방법	FS	CD	FD	MD	P	R	F1	FM	RT
C1	736	622	114	115	0.8451	0.8440	0.8445	229	23.29
C2	738	664	74	73	0.8997	0.9009	0.9003	147	23.82
C3	733	648	85	89	0.8840	0.8792	0.8816	174	30.12
C4	736	614	122	123	0.8342	0.8331	0.8337	245	22.90
C5	734	686	48	51	0.9346	0.9308	0.9327	99	22.28
C6	737	711	26	26	0.9647	0.9647	0.9647	52	131.35
C7	737	704	33	33	0.9552	0.9552	0.9552	66	23.11
C8	735	705	30	32	0.9592	0.9566	0.9579	62	15.60
C9	735	552	183	185	0.7510	0.7490	0.7500	368	13.21
P1	**721**	**643**	**78**	**75**	**0.8918**	**0.8955**	**0.8937**	**153**	**23.84**

나. 가중치 분산을 이용한 방법의 성능 평가

제안한 방법 P2~P4는 가중치 분산을 이용한 장면 전환 검출 방법으로 Ⅲ장 2절에서 제시한 방법들이다. 첫 번째 실험 E2-1은 P2~P4를 기존의 고정 임계값 기반의 방법 C1~C9와 비교하였다. 실험 결과는 표 15에 제시하였다. 실험 E2-1의 결과 제안한 방법은 P4를 기준으로 precision 0.9444, recall 0.9457, F1 0.9451, FM 81개, RT 32.61초로 평가되었다. 이 결과값들은 정확도는 4번째, 시간은 기존 분산 기반의 방법과 유사한 수준이다. 상세한 결과를 살펴보면, P4는 화소나 분산 기반의 방법들보다는 최소 0.04 정도 높은 검출율을 보이는 반면 히스토그램 기반의 방법이나 블록 기반의 방법들 보다는 0.1 정도 낮은 검출 정확도를 나타낸다. 기존의 분산을 이용한 방법인 C3보다는 현저하게 검출률이 향상된 것을 알 수 있으며 실행시간의 증가는 미미한 수준이다. 최고 수준은 아니지만 P1보다는 검출률의 향상 정도가 0.05 정도로 매우 높다.

두 번째 실험 E2-2는 P2~P4를 고정 참조 구간을 이용하는 방법 C10~C12와 비교하였다. 표 16에 제시된 실험 결과에서 알 수 있듯이 분산값을 이용하는 C10에 비해서 0.05 이상 높은 검출 정확도를 나타낸다. 이것은 제안하는 가중치 분산이 프레임의 분포를 명확하게 양극화 시키며, 가변 참조 구간의 임계값 설정이 고정 참조 방법보다 정확함을 증명하는 것이다. 히스토그램을 기반으로 하는 C11, C12와도 유사한 결과를 보이는 것을 알 수 있다.

세 번째 실험 E2-3은 P4와 C1~C7을 비교하는데 있어서 비디오 데이터의 크기를 11×9로 부표본화한 영상을 사용하였다. P4를 기준으로 precision 0.9286, recall 0.9185, F1 0.9235, FM 112개, RT 9.05초로 성능 평가되었다.

표 15. 실험 결과 : E2-1

Figure 15. Experimental results : E2-1

방법	FS	CD	FD	MD	P	R	F1	FM	RT
C1	736	622	114	115	0.8451	0.8440	0.8445	229	23.29
C2	738	664	74	73	0.8997	0.9009	0.9003	147	23.82
C3	733	648	85	89	0.8840	0.8792	0.8816	174	30.12
C4	736	614	122	123	0.8342	0.8331	0.8337	245	22.90
C5	734	686	48	51	0.9346	0.9308	0.9327	99	22.28
C6	737	711	26	26	0.9647	0.9647	0.9647	52	131.35
C7	737	704	33	33	0.9552	0.9552	0.9552	66	23.11
C8	735	705	30	32	0.9592	0.9566	0.9579	62	15.60
C9	735	552	183	185	0.7510	0.7490	0.7500	368	13.21
P2	**736**	**703**	**33**	**34**	**0.9552**	**0.9539**	**0.9545**	**67**	**30.48**
P3	**729**	**690**	**39**	**46**	**0.9465**	**0.9375**	**0.9420**	**85**	**30.29**
P4	**738**	**697**	**41**	**40**	**0.9444**	**0.9457**	**0.9451**	**81**	**32.61**

표 16. 실험 결과 : E2-2

Figure 16. Experimental results : E2-2

방법	FS	CD	FD	MD	P	R	F1	FM	RT
C10	733	648	85	89	0.8840	0.8792	0.8816	174	32.33
C11	720	689	31	46	0.9569	0.9374	0.9471	77	21.48
C12	693	670	23	65	0.9668	0.9116	0.9384	88	144.10
P2	**736**	**703**	**33**	**34**	**0.9552**	**0.9539**	**0.9545**	**67**	**30.48**
P3	**729**	**690**	**39**	**46**	**0.9465**	**0.9375**	**0.9420**	**85**	**30.29**
P4	**738**	**697**	**41**	**40**	**0.9444**	**0.9457**	**0.9451**	**81**	**32.61**

표 17에서 제시된 것과 같이 비교 방법들은 표 15에 비해서 검출률의 손실이 아주 큰 것을 알 수 있다. 화소 기반의 방법들은 0.15 내외, 히스토그램 기반의 방법들은 0.5 내외의 검출률 손실을 나타낸다. 반면 제안한 방법 P4의 검출률 손실은 0.02로 비교 방법들에 비해서 미미한 손실

이 발생한 것을 확인할 수 있다.

결과적으로 제안한 가중치 분산을 이용한 장면 전환 검출 방법 P2~P4는 정확도 측면에서는 P1보다 크게 향상되었으며, 동일한 분산을 이용하는 방법들보다도 정확하게 장면 전환을 검출한다. 또한 축소된 프레임, 즉 부표본화된 비디오 데이터에서도 검출률의 손실이 적은 것을 확인할 수 있었다. 이런 결과를 토대로 제안한 가중치 분산이 일반적 분산에 비해서 프레임의 특징을 더욱 정확하게 추출한다는 것을 알 수 있었으며, 제안한 가변 참조 구간의 적응적 임계값이 고정 참조 구간 방법 보다 신뢰도가 높음을 검증할 수 있었다. 그리고 부표본화된 환경에서 다른 방법들과 비교해서 더욱 강력한 장면 전환 검출을 하는 것을 알 수 있다.

하지만 제안한 방법 P2~P4는 P1의 방법이나 비교 방법들 보다 비교적 계산시간이 많이 소요되는 것을 알 수 있다. 그리고 검출 결과의 정확도가 기존의 방법들을 넘어서지 못하고 있다. 따라서 계산량을 줄이면서도 더욱 정확한 장면 전환 검출이 가능한 방법이 요구된다.

표 17. 실험 결과 : E2-3

Figure 17. Experimental results : E2-3

방법	FS	CD	FD	MD	P	R	F1	FM	RT
C1	730	555	175	182	0.7603	0.7531	0.7566	357	7.81
C2	737	583	154	154	0.7910	0.7910	0.7910	308	8.38
C3	739	532	207	205	0.7199	0.7218	0.7209	412	8.19
C4	739	318	421	419	0.4303	0.4315	0.4309	840	8.06
C5	729	256	473	481	0.3512	0.3474	0.3492	954	7.94
C6	685	271	414	466	0.3956	0.3677	0.3812	880	17.51
C7	730	555	175	182	0.7603	0.7531	0.7566	357	8.11
P4	**728**	**676**	**52**	**60**	**0.9286**	**0.9185**	**0.9235**	**112**	**9.05**

다. 차분 히스토그램의 평균을 이용한 방법의 성능 평가

차분 히스토그램의 평균을 이용한 장면 전환 검출 방법은 Ⅲ장 3절에서 제안하였다. 첫 번째 실험으로는 제안한 P5~P6을 기존의 최적 임계값 방법 C1~C9와 비교하여 표 18에 나타내었다. 적응적 가중치를 사용한 P6을 기준으로 precision 0.9833, recall 0.9607, F1 0.9719, FM 41개, RT 22.83초로 평가되었다. 이결과는 비교한 모든 방법들보다 높은 검출률로써 제안한 방법이 최적 임계값을 사용한 방법보다도 더 정확한 검출이 가능한 것을 검증한 결과이다. 계산량 측면에서도 기존의 히스토그램 방법과 거의 동일한 수준으로 나타난 것을 알 수 있다. 특히 동일한 밝기 히스토그램을 이용하는 C5에 비해서 0.04의 검출 정확도가 향상된 것을 알 수 있는데, 이것은 임계값의 설정 방법이 검출 정확도에 많은 영향을 주는 것을 증명하는 것이다.

표 18. 실험 결과 : E3-1
Figure 18. Experimental results : E3-1

방법	FS	CD	FD	MD	P	R	F1	FM	RT
C1	736	622	114	115	0.8451	0.8440	0.8445	229	23.29
C2	738	664	74	73	0.8997	0.9009	0.9003	147	23.82
C3	733	648	85	89	0.8840	0.8792	0.8816	174	30.12
C4	736	614	122	123	0.8342	0.8331	0.8337	245	22.90
C5	734	686	48	51	0.9346	0.9308	0.9327	99	22.28
C6	737	711	26	26	0.9647	0.9647	0.9647	52	131.35
C7	737	704	33	33	0.9552	0.9552	0.9552	66	23.11
C8	735	705	30	32	0.9592	0.9566	0.9579	62	15.60
C9	735	552	183	185	0.7510	0.7490	0.7500	368	13.21
P5	**737**	**719**	**18**	**18**	**0.9756**	**0.9756**	**0.9756**	**36**	**22.86**
P6	**720**	**708**	**12**	**29**	**0.9833**	**0.9607**	**0.9719**	**41**	**22.83**

두 번째 실험으로는 제안한 방법 P5~P6과 고정 참조 구간을 이용하는 C10~C12를 비교하였다. 표 19에서 제시된 것과 같이 히스토그램을 이용하는 고정 참조 구간 방법 C11보다 F1을 기준으로 최소 0.3 이상 높아진 것을 알 수 있다. 따라서 제안한 방법의 가변 참조 구간 설정이 고정된 참조 구간 설정을 이용하는 것 보다 더욱 정확한 검출이 되는 것을 알 수 있었다. 그리고 RT를 비교하여 보면 고정 참조 구간 방법에 비해서 가변 참조 구간을 이용하면서 사용된 부가적 계산량이 극히 적은 것을 알 수 있다. 고정 참조 방법의 경우 참조 길이의 결정 방법, 시작 프레임부터 참조 길이를 만족할 프레임까지의 예외 처리 등의 해결되어야 할 문제점들이 있다.

표 19. 실험 결과 : E3-2

Figure 19. Experimental results : E3-2

방법	FS	CD	FD	MD	P	R	F1	FM	RT
C10	733	648	85	89	0.8840	0.8792	0.8816	174	32.33
C11	720	689	31	46	0.9569	0.9374	0.9471	77	21.48
C12	693	670	23	65	0.9668	0.9116	0.9384	88	144.10
P5	**737**	**719**	**18**	**18**	**0.9756**	**0.9756**	**0.9756**	**36**	**22.86**
P6	**720**	**708**	**12**	**29**	**0.9833**	**0.9607**	**0.9719**	**41**	**22.83**

세 번째 실험으로는 제안한 방법 P5~P6과 적응적 임계값을 사용하는 방법 C13~C15를 비교하였다. C13과 C14의 경우에는 recall에서만 좋은 결과를 나타내는데, F1을 기준으로 비교하게 되면 제안한 방법이 최소 0.06이상 개선된 결과를 보이고 있다. 또한 실행 시간에서도 기존의 방법들을 넘어서지 않는 것을 알 수 있다. 따라서 제안한 임계값 설정 방법을 통해서 계산복잡도가 더 낮으면서 기존의 방법보다 정확하게 장면 전

환 검출을 하는 것을 알 수 있다. 표 20에서는 실험 E3-3의 결과를 나타내고 있다.

결과적으로 제안한 히스토그램을 이용한 장면 전환 검출 방법 P5~P6은 모든 비교 방법들보다 정확하게 장면 전환 검출을 수행할 수 있으며 실행 속도도 기존의 방법들 수준임을 알 수 있었다. 고정 참조 구간 방법과 비교하여 가변 참조 구간을 이용하는 것이 더욱 정확한 검출이 되는 것을 확인하였고, 기존의 적응적 임계값 기반의 방법들보다 임계값 설정의 신뢰도가 더욱 향상된 것을 알 수 있었다. 이런 비교들에서도 실행시간이 기존의 방법보다 짧은 것을 확인하였다.

하지만 제안한 P5~P6에서도 개선되어야 할 사항들이 있다. 통계적 가중치보다 적응적 가중치를 사용할 경우 검출률이 약간 저하되었고, 점진적 변화에 대한 검출, 그리고 예외적인 장면 변화에 대한 오류 검출 등의 문제점들을 해결해야 한다.

표 20. 실험 결과 : E3-3

Figure 20. Experimental results : E3-3

방법	FS	CD	FD	MD	P	R	F1	FM	RT
C13	806	694	112	43	0.8610	0.9417	0.8995	155	28.83
C14	864	729	135	8	0.8438	0.9891	0.9107	143	26.81
C15	723	649	74	88	0.8976	0.8806	0.8890	162	23.34
P5	**737**	**719**	**18**	**18**	**0.9756**	**0.9756**	**0.9756**	**36**	**22.86**
P6	**720**	**708**	**12**	**29**	**0.9833**	**0.9607**	**0.9719**	**41**	**22.83**

라. 개선된 방법의 성능 평가

Ⅲ장 6절에서 제안한 개선된 히스토그램 방법인 P8과 P9의 성능평가와 함께 히스토그램 기반의 기존의 방법들과 제안한 방법들을 비교하여 표 21에 나타내었다. 가장 성능이 좋은 P9를 기준으로 precision 0.9796, recall 0.9769, F1 0.9783, FM 32개, RT 22.87로 성능 평가 되었다. 기존 방법들과 제안 방법의 성능을 비교하기 위해서 검출률의 평균을 계산하여 본 결과 F1을 기준으로 기존 방법들이 0.9477, 제안한 방법이 0.9751로 0.03 정도 향상된 것을 확인할 수 있다. 평균 오류 개수 FM도 50%이상 감소한 것을 알 수 있으며 평균 RT도 30% 수준인 것을 알 수 있다.

표 21. 실험 결과 : E4-1

Figure 21. Experimental results : E4-1

방법	FS	CD	FD	MD	P	R	F1	FM	RT
C5	734	686	48	51	0.9346	0.9308	0.9327	99	22.28
C6	737	711	26	26	0.9647	0.9647	0.9647	52	131.35
C7	737	704	33	33	0.9552	0.9552	0.9552	66	23.11
C11	720	689	31	46	0.9569	0.9374	0.9471	77	21.48
C12	693	670	23	65	0.9668	0.9116	0.9384	88	144.10
평균					**0.9555**	**0.9400**	**0.9477**	**76.4**	**68.46**
P5	737	719	18	18	0.9756	0.9756	0.9756	36	22.86
P6	720	708	12	29	0.9833	0.9607	0.9719	41	22.83
P7	742	722	20	15	0.9730	0.9796	0.9763	35	23.02
P8	733	718	15	19	0.9795	0.9742	0.9769	34	22.84
P9	735	720	15	17	0.9796	0.9769	0.9783	32	22.87
P10	743	719	24	18	0.9677	0.9756	0.9716	42	23.92
평균					**0.9764**	**0.9738**	**0.9751**	**36.6**	**23.06**

이런 결과를 토대로 제안한 히스토그램을 이용한 적응적 장면 전환 검출 방법이 임계값 설정, 가중치 설정, 참조 구간 설정 등에서 기존의 방법들보다 정확한 것을 증명할 수 있다. 적응적 가중치를 사용하는 P6에서부터 BDA처리와 FO처리를 개선한 P9까지 개선 내용에 따라서 검출률이 향상되는 것을 확인할 수 있으며 개선된 내용이 정확하게 반영되고 있음을 알 수 있다. F1에서는 향상정도를 판단하기 어렵지만, FM을 비교하여 보면 상대적 검출률 향상을 쉽게 확인할 수 있다. P10은 히스토그램과 비슷한 성향을 가지는 chi-square 검정을 이용한 방법으로 역시 좋은 검출 성능을 나타내는 것을 알 수 있다. 참고적으로 P9를 기준으로 최적화된 소스와 하드웨어 성능이 높은 컴퓨터를 기준으로 하여 실제 RT가 12초대로 측정되는 것을 알 수 있었다. 따라서 제안한 방법이 고속 및 실시간으로 동작하면서 영상에 따라서 적응적으로 장면 전환 검출을 하는 것을 알 수 있다.

표 22. 실험 결과 : E4-2

Figure 22. Experimental results : E4-2

방법	FS	CD	FD	MD	P	R	F1	FM	RT
P1	721	643	78	75	0.8918	0.8955	0.8937	153	23.84
P2	736	703	33	34	0.9552	0.9539	0.9545	67	30.48
P3	729	690	39	46	0.9465	0.9375	0.9420	85	30.29
P4	738	697	41	40	0.9444	0.9457	0.9451	81	32.61
P5	737	719	18	18	0.9756	0.9756	0.9756	36	22.86
P6	720	708	12	29	0.9833	0.9607	0.9719	41	22.83
P7	742	722	20	15	0.9730	0.9796	0.9763	35	23.02
P8	733	718	15	19	0.9795	0.9742	0.9769	34	22.84
P9	735	720	15	17	0.9796	0.9769	0.9783	32	22.87
P10	743	719	24	18	0.9677	0.9756	0.9716	42	23.92

두 번째 실험에서는 제안한 모든 방법들을 비교하여 표 22에 나타내었다. 최초 프레임 평균값을 이용한 P1에서는 낮은 정확도가 문제점으로 분석되었고, 이를 개선하기 위해서 가중치 분산을 도입한 P2~P4를 제안하였다. P1에 비해서 검출 정확도가 상승되었지만 상대적으로 실행시간이 증가한 단점이 있었다. 이런 문제점 때문에 히스토그램 기반의 P5~P6을 제안하였고 비교적 정확한 장면 전환 검출을 수행하는 것을 확인하였다. 여전히 몇 가지 남아있는 문제점들을 개선하여 P7~P9의 방법들을 제안하였고 소폭의 검출률 상승을 유도하였다. 마지막으로는 제안한 모델을 적용을 위해서 테스트 되지 않았던 특징값 chi-square 검정을 이용하여 실험하여 역시 높은 수준의 검출률을 확인할 수 있었다.

다양한 실험들을 통해서 임계값 설정 원리와 방법, 가변 참조 구간에 대한 신뢰성, 자동적 임계값 가중치 설정 원리와 방법 등 제안한 여러 가지 요소들이 모두 검출 정확도를 향상시키는데 올바르게 적용된 것임을 알 수 있다. 또한 실제 응용 환경 적용을 위한 영상에 대한 알고리즘의 독립성도 만족하는 것을 알 수 있었고, 프레임의 재생과 동일한 순차성을 가지고 고속으로 동작하기 때문에 실시간 처리도 가능한 것을 확인하였다.

다음 절에서는 PC에서의 시뮬레이션 결과를 토대로 제안한 방법을 실제 PMP에 구현하는 방법과 결과를 알아본다.

3. 제안한 방법의 PMP에서의 구현

본 절에서는 제안한 장면 전환 검출 방법을 PMP에서 구현하고 구현된 결과를 정리하였다. PMP는 portable multimedia player로써 휴대 장치에서 멀티미디어 데이터의 재생과 저장 등이 가능한 장치로 다양한 제품들이 상용화되고 그 보급률이 증가하고 있다. 본 절에서는 먼저 PMP에서의 구현의 특징들을 알아보고, PMP의 구성과 동작하는 과정을 설명한다. 그리고 PMP에서 제안한 방법을 구현하는 과정을 기술하고 검출 결과를 분석한다.

가. PMP 구현의 특징

본 논문에서 제안한 장면 전환 검출 알고리즘을 실제 사용되는 PMP에 적용하여 실시간 환경에서의 장면 전환 검출 기법 활용 가능성을 알아보았다.

PMP에서는 PC와는 달리 모든 알고리즘을 테스트하는 것은 상당히 힘든 일이다. 일단 개발이 완료된 상품에 대한 소스 코드의 수정이나 기능의 추가는 개발자가 아닌 상황에서는 상당히 어려운 일이다. 따라서 PC에서의 실험을 통해서 가장 적합한 방법을 구현하고자 했으며, 제안한 방법이 기존의 방법들과 비교하여 가장 PMP에서의 구현이 적합하다는 것이 검증되었다. 컬러히스토그램 비교법의 경우에는 연산량이 너무 많아서 실시간 처리에는 부적절하며, 고정된 임계값을 사용하는 방법은 임계값의 수정을 요하기 때문에 자동성을 요구하는 장치에서의 사용은 불

가능하다. 또한 기존의 적응적 임계값 설정 방법은 전체 비디오 시퀀스의 정보를 필요하기 때문에 실시간 처리에서는 사용할 수 없다.

PMP는 PC와는 다르게 연산 장치의 속도나 메모리의 저장 공간, 전력 소비 등의 하드웨어 성능이 낮기 때문에 이런 요소들을 고려한 알고리즘이 필요하다. 제안하는 방법은 실험을 통해서 기존에 방법들과 비교해도 계산량이 증가하지 않는 것을 알 수 있었고, 적응적 임계값을 사용하여 자동적인 임계값 설정이 가능하므로 실제 응용환경에서의 사용도 가능하다. 또한 기존의 적응적 임계값 설정 방법과는 다르게 전체 비디오 시퀀스의 정보를 필요로 하지 않으며, 이전 장면 전환 이후의 프레임 특징값 정보만 알고 있으면 장면 전환 판단이 가능하다. 따라서 제안한 방법은 실제 PMP에 적용이 가능하다고 판단된다.

그림 23에서는 구현에 사용된 PMP를 나타내고 있다. 구현에 사용된 PMP는 Homcast사의 TVUS HM 900 PLUS 모델을 사용한다.

그림 23. 구현에 사용된 PMP
Figure 23. PMP used in our implementation

나. PMP의 구성과 동작

실험에 사용된 PMP의 중앙처리장치는 Texas Instrument사의 DM320으로서 100MHz의 DSP를 내장하고 있다. 비압축된 SD급의 비디오를 화면에 나타내기 위해서는 27.5MHz의 클럭(clock)신호가 필요하고, 압축된 영상을 디코딩하기 위해서는 더 많은 클럭이 필요하기 때문에 비디오 재상중간에 다른 처리를 하기 위한 리소스는 부족하게 된다. 또한 전원 공급이 배터리에 의해서 이루어지기 때문에 처리 과정이 많을수록 많은 전력을 소비하게 된다. 그리고 연산 수행을 위한 저장 공간이 충분하지 않기 때문에 공간적 계산복잡도도 고려되어야 한다. 따라서 본 논문에서 제안한 방법과 같이 계산복잡도가 낮은 알고리즘이 사용되어야 한다.

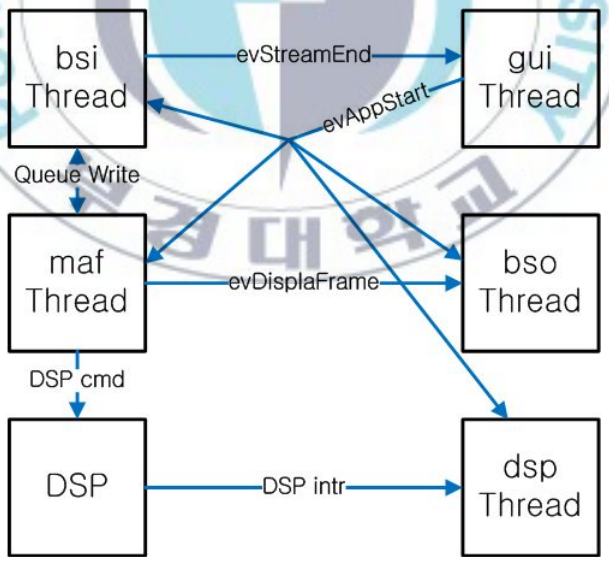


그림 24. PMP의 내부 스레드

Figure 24. Internal threads of a PMP

실험에 사용된 PMP는 그림 24와 같이 내부적으로 5개의 스레드(GUI, BSI, BSO, MAF, DSP) 간의 이벤트 처리를 통해서 비디오 데이터 처리가 이루어진다. GUI는 PMP에서 버튼 조작이 있을 경우 그 조작에 대한 핸들과 미디어 파일을 불렀을 때 시작 이벤트를 BSI, BSO, MAF에 각각 보내어 동작의 시작을 알린다. BSI는 GUI에서 시작된 이벤트가 지시하는 동작에 따라서 비디오의 프레임 정보를 MAF에 전달하고 MAF는 DSP를 이용하여 동영상을 디코딩하여 공유 버퍼에 저장한다. 이 때 저장되어 있는 비디오 프레임의 영상형태는 YCbCr(4:2:2)의 형태로 저장되어 있고, BSO는 저장된 비디오 프레임을 RGB의 형태로 변환하여 재생 가능하도록 만들어 준다. 그림 24에서는 위에서 설명한 5개의 스레드간의 이벤트 처리에 대한 흐름을 나타내었다.

저장된 비디오 데이터를 재생하는 동안에는 PMP의 왼쪽 가운데에 위치한 원형 버튼을 이용해서 제어를 한다. 따라서 제안하는 방법을 PMP의 기능으로 구현하기 위해서는 원형 버튼의 입력으로부터 동작하도록 만들어져야 한다. 다음 항에서 PMP에서의 구현 절차와 방법을 알아보도록 한다.

PUKYONG NATIONAL UNIVERSITY 부경대학교

다. PMP에서의 장면 전환 검출 구현

본 논문에서는 두 가지 모드로 PMP에서 실시간 장면 전환 검출 기능을 구현하였다. 그림 25에서는 원형 버튼의 입력으로부터 장면 전환 검출 결과를 처리하는 과정까지의 두 가지 모드에 해당하는 PMP에서 장면 전환을 검출하는 흐름도를 나타내었다.

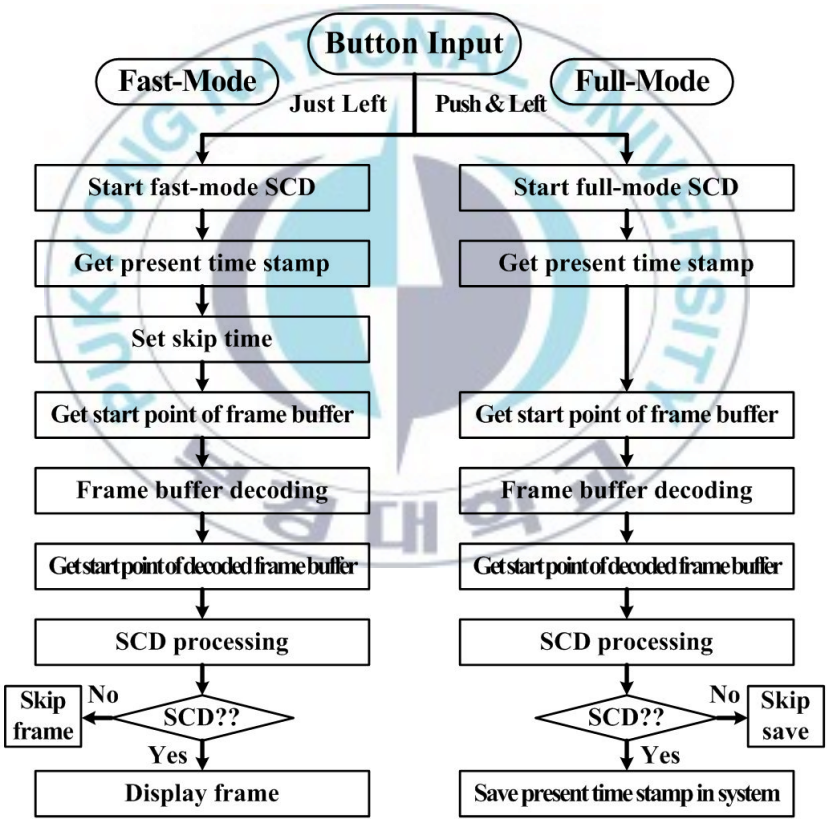


그림 25. PMP에서의 두 가지 장면 전환 검출 모드
Figure 25. Two SCD modes on a PMP

구현에 사용된 방법은 제안한 방법 중에서 가중치 분산을 이용한 적응적 장면 전환 검출 방법인 P4를 이용하였다. P4를 이용한 이유는 제안한 방법 중에서 가장 계산복잡도가 높기 때문이다. P4를 이용해서 실시간 장면 전환 검출이 된다면 다른 방법들을 적용했을 때에도 실행이 가능하다고 판단된다.

첫 번째 모드는 full-mode로써 제안한 방법 P4를 그대로 적용시킨 모드이다. 이 모드는 동영상의 재생 중에 실행되는 모드로써 동영상의 재생 중에 원형 버튼을 누르고 왼쪽으로 밀면 작동이 된다. 재생과는 별도로 백그라운드에서 장면 전환 검출을 수행하면서 장면 전환이 발생할 때마다 로그파일에 해당 시간 정보를 기록한다. 차후 동일한 영상을 재생할 경우 장면 전환 검출이 된 프레임만 재생할 수 있는 기능을 지원한다.

두 번째 모드는 fast-mode로써 제안한 방법 P4에 프레임 생략 기법을 적용시켜서 구현한 모드이다. 재생 중에 원형 버튼을 왼쪽으로 밀면 fast-mode가 실행되는데, 이 경우에는 바로 다음의 장면 전환 프레임을 표현하게 된다. 즉, 실시간으로 장면 전환 검출을 해서 화면에 표시하는 작업을 반복하는 것이다. 사용자는 장면 전환된 프레임들만 화면을 통해서 볼 수 있는 기능이다.

실험에 사용된 PMP는 WinCE 5.0 운영체제를 사용하고 있으며 Embedded Visual C++ 4.0을 이용해서 구현하였다.

라. PMP에서의 장면 전환 검출 결과

제안한 방법을 실제 PMP에 구현하여 본 결과 주관적으로 확인되는 오검출은 거의 발생되지 않았다. 장면 전환 검출과정에서 프레임 비교로 발생되는 지연도 없었으며 버퍼링을 비롯하여 재생에 지장을 주는 요소들은 없었다. 실제로 전체 비디오를 장면 검출하는 과정은 검사를 시작하는 과정에서의 1프레임 지연 발생이 전부였다. 따라서 제안한 방법이 실제 PMP에서 잘 구현되었음을 검증할 수 있었고 검출 결과도 PC에서의 실험과 거의 일치하는 것을 알 수 있었다. 그림 26에서는 실제 PMP를 이용해서 장면 전환 검출을 실시한 예를 나타내었다.

그림 26. PMP에서 장면 전환 검출 결과
Figure 26. Results of shot change detection on a PMP

Ⅴ. 결론

장면 전환 검출은 비디오 데이터의 장면 변화가 발생하는 경계를 검출하는 기술로써 효과적인 비디오 데이터 관리를 위한 기반 기술이다.

본 논문에서는 장면 전환 검출을 위한 다양한 특징 추출 방법과 임계값을 설정하는 방법, 계산량 감소를 위한 방법, 알고리즘의 성능 평가 방법과 여러 가지 문제점 등 기존의 연구들을 체계적으로 정리하였다.

본 논문에서는 기존에 해결되지 않았던 문제들 중에서 장면 변화가 발생하는 프레임의 특징을 정확하게 추출해 내는 특징 추출 척도와 다양한 영상에 대해서 적응적으로 사용될 수 있는 임계값 설정 방법, 그리고 고속 실시간 처리를 위한 알고리즘 개선과 실제 하드웨어 장치에서의 구현 등 세 가지 요소에 대한 문제 해결을 목표로 하여 연구를 진행하였다.

이를 위해서 세 가지의 장면 전환 검출 방법을 제안하였고 세부적으로는 정확한 프레임 특징 추출을 위한 척도인 가중치 분산, 가변 참조 구간의 평균 특징값을 이용한 적응적인 임계값 설정 방법, 자동적 임계값 가중치 결정을 통한 완전 자동적 장면 전환 검출 방법을 제안하였다. 최종적으로는 제안한 방법을 바탕으로 한 실시간 장면 전환 검출 모델을 제시하였다.

동일한 비디오 데이터를 이용한 9가지 실험을 통해서 제안한 방법이 기존의 방법들보다 정확하면서도 고속으로 처리되는 알고리즘임을 확인하였고, 상용 PMP에서 제안한 알고리즘의 구현 및 동작을 확인하였다.

제안한 장면 전환 검출 기술은 다양한 소프트웨어 및 하드웨어 플랫폼과 PMP나 휴대폰 같은 휴대용 장치에서의 고속 비디오 데이터 관리를 위해서 유용하게 사용될 것이다.

참 고 문 헌

[1] C. Cotsaces, N. Nikolaidis, and I. Pitas, "Video Shot Detection and Condensed Representation a Review," *IEEE Signal Processing Magazine*, Vol. 23, Issue 2, pp. 28-37, Mar. 2006.

[2] M. Cooper, T. Liu, and E. Rieffel, "Video Segmentation via Temporal Pattern Classification," *IEEE Trans. Multimedia*, Vol. 9, Issue 3, pp. 610-618, Apr. 2007.

[3] N. Goela, K. Wilson, F. Niu, A. Divakaran, and I. Otsuka, "An SVM Framework for Genre-Independent Scene Change Detection," *IEEE Int. Conf. Multimedia and Expo*, pp. 532-535, Jul. 2-5, 2007.

[4] G. Camara-Chavez, F. Precioso, M. Cord, S. Phillip-Foliguet, and A. de A. Araujo, "Shot Boundary Detection by a Hierarchical Supervised Approach," *Int. Works. Systems, Signals and Image Processing and EURASIP Conf. Speech and Image Processing, Multimedia Communications and Services*, pp. 197-200. Jun. 27-30, 2007.

[5] J. Yuan, H. Wang, L. Xiao, W. Zheng, J. Li, F. Lin, and B. Zhang, "A Formal Study of Shot Boundary Detection," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 17, Issue 2, pp. 168-186, Feb. 2007.

[6] S. Lefèvre, J. Hollera, and N. Vincent, "A Review of Real-time Segmentation of Uncompressed Video Sequences for

Content-based Search and Retrieval," *Real-Time Imaging*, Vol. 9, No. 1, pp. 73-98, 2003.

[7] A. Nagasaka and Y. Tanaka, "Automatic Video Indexing and Full-video Search for Object Appearances," *in Proc. IFIP Conf. Visual Database Systems*, Vol. A-7, pp. 113-127, Oct. 1991.

[8] H. J. Zhang, A. Kankanhalli, and S. W. Smoliar, *Automatic Partitioning of Full-Motion Video*, Multimedia Systems, New York, pp. 10-28, 1993.

[9] W. Ren, M. Sharma, and S. Singh, "Automated Video Segmentation," *Int. Conf. Information, Communication, and Signal Processing*, Oct. 2001.

[10] Y. Taniguchi, A. Akutsu, and Y. Tonomura, "Panorama Excerpts : Extracting and Packing Panoramas for Video Browsing," *in ACM Int. Conf. Multimedia*, pp. 427-436, Nov. 1997.

[11] S. Lawrence, D. Ziou, M. Auclair-fortier, and S. Wang, "Motion Insensitive Detection of Cuts and Gradual Transitions in Digital Videos," *Technical Report 266, DMI, University of Sherbrooke, Canada*, May, 2001.

[12] Y. Tonomura and S. Abe, "Content Oriented Visual Interface Using Video Icons for Visual Database Systems," *Journal of Visual Languages and Computing*, Vol. 1, No. 2, pp. 183-198, Oct. 1990.

[13] U. Gargi and R. Kasturi, "An Evaluation of Color Histogram-Based Methods in Video Indexing," *Int. Works.*

Image Databases and Multimedia Search, pp. 75-82, Aug. 1996.

[14] D. Pye, N. J. Hollinghurst, T. J. Mills, and K. R. Wood, "Audio-Visual Segmentation for Content-Based Retrieval," *Int. Conf. Spoken Language Processing*, Dec. 1998.

[15] N. Haering, N. Vitoria, R. Lobo, J. Qian, and M. Ibrahim Sezan, "A Framework for the Design of Event Detectors," *in Asian Conf. Computer Vision*, Jan. 1999.

[16] O. Javed, S. Khan, Z. Rasheed, and M. Shah, "A Framework for Segmentation of Interview Videos," *IASTED Int. Conf. Internet and Multimedia Systems and Applications*, Nov. 2000.

[17] A. Dailianas, R. B. Allen, P. England, "Comparison of Automatic Video Segmentation Algorithms," *SPIE Conf. Integration Issues in Large Commercial Media Delivery Systems*, Vol. 2615, pp. 2-16, Jan. 1996.

[18] B. Gunsel, A. M. Ferman, and A.M. Tekalp, "Temporal Video Segmentation Using Unsupervised Clustering and Semantic Object Tracking," *Journal of Electronic Imaging*, Vol. 7, No. 3, pp. 592-604, 1998.

[19] S. Kim and R. H. Park, "A Novel Approach to Scene Change Detection Using a Cross Entropy," *IEEE Int. Conf. Image Processing*, Vol. 3, pp. 937-940, Sep. 2000.

[20] M. S. Lee, Y. M. Yang, and S. W. Lee, "Automatic Video Parsing Using Shot Boundary Detection and Camera Operation Analysis," *Pattern Recognition*, Vol. 34, No. 3, pp. 711-719, 2001.

[21] R. Kasturi and R. C. Jain, *Computer Vision : Principles*, IEEE Computer Society Press, Washington DC, pp. 469-480. 1991.

[22] D. Swanberg, C. F. Shu, and R. Jain, "Knowledge Guided Parsing and Retrieval in Video Databases," *SPIE Conf. Storage and Retrieval for Image and Video Databases*, Vol. 1908, pp. 13-24, Feb. 1993.

[23] H. Ueda, T. Miyatake, and S. Yoshizawa, "An Interactive Natural-Rotion-Picture Dedicated Multimedia Authoring System," ACM Conf. Human Factors in Computing Systems, pp. 343-350, Apr. 1991.

[24] R. Dugad, K. Ratakonda, and N. Ahuja, "Robust Video Shot Change Detection," *IEEE Works. Multimedia Signal Processing*, pp. 376-381, Dec. 1998.

[25] D. A. Adjeroh, M. C. Lee, and C. U. Orji, "Techniques for Fast Partitioning of Compressed and Uncompressed Video," *Int. Journal of Multimedia Tools and Applications*, Vol. 4, No. 2, pp. 225-243, 1997.

[26] S. Lefèvre, J. Holler, and N. Vincent, "Real Time Temporal Segmentation of Compressed and Uncompressed Dynamic Colour Image Sequences," *Int. Works. Real Time Image Sequence Analysis*, pp. 56-6262, Aug. 2000.

[27] C. M. Lee and M. C. Ip, "A Robust Approach for Camera Break Detection in Color Video Sequences," *IAPR Int. Works. Machine Vision Applications*, pp. 502-505, Dec. 1994.

[28] B. T. Truong, C. Dorai, and S. Venkatesh, "New Enhancements

to Cut, Fade, and Dissolve Detection Processes in Video Segmentation," ACM Int. Conf. Multimedia, pp. 219-27, Oct. 2000.

[29] R. Zabih, J. Miller, and K. Mai, "A Feature-Based Algorithm for Detecting and Classifying Production Effects." *Multimedia Systems,* Vol. 7, Issue 2, pp. 119-128, 1999.

[30] S. Eickeler and S. Müller, "Content-Based Video Indexing of TV Broadcast News Using Hidden Markov Models," *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 2997-3000, Mar. 1999.

[31] F. Arman, R. Depommier, A. Hsu, and M. Y. Chiu, "Content-Based Browsing of Video Sequences," *ACM Int. Conf. Multimedia*, pp. 97-103, Nov. 1994.

[32] R. Lienhart, "Comparison of Automatic Shot Boundary Detection Algorithms," *SPIE Conf. Storage and Retrieval for Image and Video Databases*, Vol. 3656, pp. 290-301, Jan. 1999.

[33] M. Ardebilian, X. Tu, and L. Chen, "Robust 3D Clue-Based Video Segmentation for Video Indexing," *Journal of Visual Communication and Image Representation*, Vol. 11, No. 1, pp. 58-79, 2000.

[34] R. Silva, J. Gomes, and L. Velho, "Segmentation of Video Sequences Using Volumetric Image Processing," Eurographics Works. Multimedia, Sep. 1999.

[35] N. Vasconcelos and A. Lippman, "Statistical Models of Video Structure for Content Analysis and Characterization," *IEEE*

Trans. Image Processing, Vol. 9, No. 1, pp. 3-19, 2001.

[36] A. F. Smeaton, G. Gormley, J. Gilvarry, B. Tobin, S. Marlow, and N. Murphy, "An Evaluation of Alternative Techniques for Automatic Detection of Shot Boundaries in Digital Video," *Irish Machine Vision and Image Processing Conf.*, pp. 45-62, Sep. 1999.

[37] S. Mann and R. W. Picard, "Video Orbits of the Projective Group : A Simple Approach to Featureless Estimation of Parameters," *IEEE Trans. Image Processing,* Vol. 6, No. 9, pp. 1281-1295, 1997.

[38] M. Cherfaoui and C. Bertin, "Temporal Segmentation of Videos: A New Approach," *SPIE Conf. Digital Video Compression: Algorithms and Technologies*, Vol. 2419, pp. 38-47, Feb. 1995.

[39] P. Bouthemy, M. Gelgon, and F. Ganansia, "A Unified Approach to Shot Change Detection and Camera Motion Characterization," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 9, No. 7, pp. 1030-1044, 1999.

[40] A. Akutsu, Y. Tonomura, H. Hashimoto, and Y. Ohba, "Video Indexing Using Motion Vectors," *SPIE Conf. Visual Communications and Image Processing*, Vol. 1818, pp. 1522-1530, Nov. 1992.

[41] O. Fatemi, S. Zhang, and S. Panchanathan, "Optical Flow Based Model for Scene Cut Detection," *Canadian Conf. Electrical and Computer Engineering*, Vol. 1, pp. 470-473, May, 1996.

[42] S. V. Porter, M. Mirmehdi, and B. T. Thomas, "Video Cut

Detection Using Frequency Domain Correlation," *IAPR Int. Conf. Pattern Recognition*, Vol. 3, pp. 413-416, Sep. 2000.

[43] G. Qu!enot and P. Mulhem, "Two Systems for Temporal Video Segmentation," *European Works. Content Based Multimedia Indexing*, pp. 187-194, Oct. 1999.

[44] J. M. Gauch, S. Gauch, S. Bouix, and X. Zhu, "Real Time Video Scene Detection and Classification," *Information Processing and Management*, Vol. 35, No. 3, pp. 401-420, 1999.

[45] Y. Yusoff, J. Kittler, and W. Christmas, "Combining Multiple Experts for Classifying Shot Changes in Video Sequences," *IEEE Int. Conf. Multimedia Computing and Systems*, Vol. 2, pp. 700-704, Jun. 1999.

[46] M. R. Naphade, P. Mehrotra, A. M. Ferman, J. Warnick, T. S. Huang, and A. M. Tekalp, "A High-Performance Shot Boundary Detection Algorithm Using Multiple Cues," IEEE Int. Conf. Image Processing, Vol. 2, pp. 884-887, Oct. 1998.

[47] H. C. Lee, C. W. Lee, and S. D. Kim, "Abrupt Shot Change Detection Using an Unsupervised Clustering of Multiple Features," *Int. Conf. Acoustics, Speech, and Signal Processing*, Vol. 3, pp. 2015-2018, Jun. 2000.

[48] J. Feng, K. T. Lo, and H. Mehrpour, "Scene Change Detection Algorithm for MPEG Video Sequence," *Int. Conf. Image Processing*, Vol. 1, pp. 821-824, Sep. 1996

[49] J. Meng, Y. Juan, and S. F. Chang, "Scene Change Detection in a MPEG Compressed Video Sequence," *SPIE Conf. Digital*

Video Compression: Algorithms and Technologies, Vol. 2419, pp. 14-25, Feb. 1995.

[50] B. Yeo and B. Liu, "Rapid Scene Analysis on Compressed Video," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 5, Issue 6, pp. 533-540, Dec. 1995.

[51] W. A. C. Fernando, C. N. Canagarajah, and D. R. Bull, "Scene Change Detection Algorithms for Content-Based Video Indexing and Retrieval," *Electronics and Communication Journal*, Vol. 13, No. 3, pp. 117-126, 2001.

[52] Y. K. Seong, Y. Choi, J. Park, and T. Choi, "A Hard Disk Drive Embedded Digital Satellite Receiver with Scene Change Detector for Video Indexing," *IEEE Trans. Consumer Electronics*, Vol. 48, No. 3, pp. 776-782, 2002.

[53] I. K. Sethi and N. Patal, "A Statistical Approach to Scene Change Detection," *Storage and Retrieval for Image and Video Databases*, Vol. 2420, pp. 329-338, Feb. 5-10, 1995.

[54] J. R. Kim, S. J. Suh, and S. H. Sull, "Fast Scene Change Detection for Personal Video Recorder," *IEEE Trans. Consumer Electronics*, Vol. 49, No. pp. 683-688, Aug. 2003.

[55] N. V. Patel and I. K. Sethi, "Compressed Video Processing for Cut Detection," IEE Proc. Vision, Image and Signal Processing, Vol. 143, Issue 5, pp. 351-323, Oct. 1996.

[56] S. Antani, R. Kasturi, and R. Jain, "A Survey on the Use of Pattern Recognition Methods for Abstraction, Indexing and Retrieval of Images and Video," *Pattern Recognition*, Vol. 35,

No. 4, pp. 945-965, 2002.

[57] Z. Cernekova, C. Kotropoulos, and I. Pitas, "Video Shot Segmentation Using Singular Value Decomposition," *IEEE Int. Conf. Multimedia and Expo*, Vol. 2, pp. 301-302, Jul. 2003.

[58] X. Yi and N. Ling, "Fast pixel-Based Video Scene change Detection," *IEEE Int. Symp. Circuits and Systems*, Vol. 4, pp. 3443-3446, 2005.

[59] D. Lelescu and D. Schonfeld, "Statistical Sequential Analysis for Real-Time Video Scene Change Detection on Compressed Multimedia Bitstream," *IEEE Trans. Multimedia*, Vol. 5, No. 1, pp. 106-117, Mar. 2003.

[60] A. Hanjalic, "Shot-Boundary Detection: Unraveled and Resolved," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 12, No. 2, pp. 90-105, Feb. 2002.

[61] 고경철, 이양원, "비디오 분할을 위한 자동 임계치 결정 알고리즘," 한국컴퓨터정보학회논문지, 제10권, 제6호, pp. 65-73, 2005.

[62] S. Y. Shin, G. R. Sheng, and K. H. Park "A Scene Change Detection Scheme Using Local x^2-Test on Telematics," *Int. Conf. Hybrid Information Technology*, Vol. 1, pp. 588-592, Nov. 9-11, 2006

[63] J. Yu and M.D. Srinath, "An efficient Method for Scene Cut Detection," *Pattern Recognition*, Vol. 22, No. 13, pp. 1379-1391, Nov. 2001.

[64] Y. Yusoff, W. Christmas, and J. Kittler, "Video Shot Cut Detection Using Adaptive Thresholding," *Conf. British Machine*

Vision, Sep. 11-14, 2000.

[65] M. Zhi and A. N. Cai, "Shot Change Detection with Adaptive Thresholds," *IEEE Int. Works. VLSI Design and Video Technology*, pp. 147-149, 2005.

[66] A. Miene, A. Dammeyer, Th. Hermes, and O. Herzog, "Advanced and Adaptive Shot Boundary Detection," *Proc. ECDL WS Generalized Documents*, pp. 39-43, 2002.

[67] 박진형, 장동식, 송광섭, 유헌우, "통계적 기법을 이용한 장면 전환 검출," 한국경영과학회 학술대회논문집, 제1권, pp. 676-678, Jan. 2000.

[68] Y. Cheng, X. Yang, and D. Xu, "A Method for Shot Boundary Detection with Automatic Threshold," *Proc. IEEE Region 10 Conf. on Computers, Communications, Control and Power Engineering*, Vol. 1, pp. 582-585, Oct. 28-31, 2002.

[69] H. Li, G. Liu, Z. Zhang, and Y. Li, "Adaptive Scene-Detection Algorithm for VBR Video Stream," *IEEE Trans. Multimedia*, Vol. 6, Issue 4, pp. 624-633, Aug. 2004.

[70] 김원희, 정용재, 김종남, "PMP에서 실시간 동작을 위한 적응적인 장면 전환 검출 기법," Telecommunication Review, 제18권, 제4호, pp. 746-759, Aug. 2008.

[71] 안현수, 오일석, "크고 적응적인 생략인자에 의한 비디오 영상 장면의 고속 추출," 한국정보과학회 추계학술발표논문집, 제22권, 제2호(A), pp. 481-484, 1995.

[72] 최우진, 김단환, 정기봉, 오무송, "MSE를 이용한 동영상의 장면전환검색에 관한 연구," 한국정보과학회 춘계학술발표논문집, 제28

권, 제1호(b), pp. 171-179, 2001.

[73] 김신형, 김중헌, 전병태, 박두영, 장종환, "프레임 비교 횟수를 최소화하는 장면전환 검출 알고리즘 연구," 한국통신학회논문지, 제27권, 제5호(A), pp. 506-516, 2002.

[74] K. A. Hua and J. H. Oh, "Detecting Video Shot Boundaries up to 16 Times Faster," *Proc. of ACM Int. Conf. Multimedia*, pp. 396-387, 2000.

[75] A. Baskar, M. Prabukumar, and S. Sathishkumar, "Improved Adaptive Skip Algorithm for Video Shot Boundary Detection," *Int. Conf. Signal Processing, Communications and Networking*, pp. 492-496, Feb. 22-24, 2007.

[76] 권오덕, 최창규, 장용석, 김승호, "가변 길이 프레임 생략 방법을 이용한 장면 전환 검출," 한국정보과학회 추계학술발표논문집, 제28권, 제2호, pp. 268-270, 2001.

[77] M. S. Lee, B. W. Hwang, S. H. Sull, and S. W. Lee, "Automatic Video Parsing Using Shot Boundary Detection and Camera Operation Analysis," Int. Conf. Pattern Recognition, Vol. 2, pp. 1481-1483, Aug. 16-20, 1998.

[78] 나윤정, 하명환, 이상길, "동영상 검색을 위한 고속 장면 전환 검출," KOBA 방송기술 워크샵, Jun. 1999.

[79] X. Ling, L. Chao, L. huan, and X. Zhang, "A General Method for Shot Boundary Detection," *Int. Conf. Multimedia and Ubiquitous Engineering*, pp. 394-397, Apr. 24-26, 2008.

[80] J. Wang and W. Luo, "A Self-Adapting Dual-Threshold Method for Video Shot Transition Detection," *IEEE Int. Conf.*

Networking, Sensing and Control, pp. 704-707, Apr. 6-8, 2008.

[81] G. Boccignone, A. Chinaese, V. Moscato, and A. Picariello, "Foveated Shot Detection for Video Segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 15, Issue 3, pp. 365-377, Mar. 2005.

[82] J. Bescos, G. Cisneros, G. Martinez, J. M. Menendez, and J. Cabrera, "A Unified Model for Techniques on Video-Shot Transition Detection," *IEEE trans. Multimedia*, Vol. 7, Issue 2, pp. 293-307, Apr. 2005.

[83] H. Lu and Y. Tan, "An Effective Post-Refinement Method for Shot Boundary Detection," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 15, Issue 11, pp. 1407-1421, Nov. 2005.

[84] J. Bescos, "Real-Time Shot Change Detection Over Online MPEG-2 Video," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 14, Issue 4, pp. 475-484, Apr. 2004.

[85] 김영선, 김원희, 문광석, 김종남, "축소된 프레임에서 적응적 임계값을 이용한 장면 전환 검출," 한국멀티미디어학회 춘계학술발표대회논문집, 제10권, 제1호, pp. 21-24, May, 2007.

[86] 김원희, 정용재, 유태경, 문광석, 김종남, "가중치 분산을 이용한 적응적 장면 전환 검출," 한국멀티미디어학회 추계학술발표대회논문집, 제10권, 제2호, pp. 275-278, Nov. 2007.

[87] 정용재, 김원희, 정태일, 문광석, 김종남, "계산량 간소화 기법을 이용한 PMP에서의 장면 전환 검출 기법," 한국멀티미디어학회 추계학술발표대회논문집, 제10권, 제2호, pp. 86-89, Nov. 2007.

[88] 김원희, 정용재, 유태경, 문광석, 김종남, "가변 구간 참조를 이용

한 적응적 장면 전환 검출," 한국신호처리시스템학회 추계학술발표대회논문집, 제8권, 제2호, pp. 179-182, Dec. 2007.

[89] 김종남, 김원희, 정용재, 문광석, "평균 특징값을 이용한 적응적 장면 전환 검출 기술," 제 20회 영상처리 및 이해에 관한 워크샵, Feb. 2008.

[90] 김원희, 문광석, 김종남, "실시간 SCD 모델을 이용한 장면 전환 검출과 PMP에서의 구현," 한국멀티미디어학회 춘계학술발표대회 논문집, 제11권, 제1호, pp. 225-228, May, 2008.

[91] 김원희, 문광석, 김종남, "실시간 장면 전환 검출을 위한 자동적인 임계값 설정 방법," 한국신호처리시스템학회 추계학술발표대회논문집, 제9권, 제1호, pp. 196-199, Jun. 2008.

[92] 김원희, 김종남, "하드웨어 적용을 위한 실시간 장면 전환 검출," 제21회 신호처리합동학술대회 논문집, 제21권, 제1호, pp. 45, Sep. 2008.

[93] W. H. Kim, Y. J. Jeong, K. S. Moon, and J. N. Kim, "Adaptive Shot Change Detection Technique for Real-time Operation on PMP," *Int. Conf. Convergence and Hybrid Information Technology*, Vol. 1, pp. 295-298, Nov. 11-13, 2008.

감사의 글

2년 전 석사과정을 시작했던 날, 5년 전 연구실에 들어왔던 날, 그리고 이 논문을 마무리하게 된 지금, 결코 짧지 않았던 시간동안에 저를 보살펴주시고 도와주신 많은 분들에게 이 페이지를 빌어 한없는 감사의 마음을 보냅니다.

가장 먼저 감사의 마음을 전하고 싶은 분은 바로 지도교수님이시자 삶의 멘토이신 김종남 교수님이십니다. 부족한 제자에게 연구가 무엇인지를 가르쳐주시고, 꿈을 가르쳐주신 교수님께 깊은 감사의 마음을 전하고 싶습니다.

아울러 오랜 기간 연구실의 가장 큰 어른으로써 항상 좋은 조언을 해주신 문광석 교수님께도 감사의 마음을 전합니다. 부족한 논문의 심사를 해주시느라 수고하여주신 주심 송하주 교수님과 신봉기 교수님께도 감사의 마음을 드립니다. 아울러 대학원 재학동안에 멋진 강의와 도움을 주신 학과장 조우현 교수님 이하 모든 컴퓨터 공학과 교수님들께도 감시의 뜻을 전합니다.

연구실에도 고마움을 전하고 싶은 식구들이 많습니다. 언젠간 꼭 한번 진심어린 감사의 맘을 전하고 싶었던 경원 선배, 오랜 기간 연구실에서 함께 해왔던 태경 선배, 가장 많은 시간을 부대꼈던 용재선배 등에게 우선 감사의 글을 전합니다. 그리고 맘이 잘 통했던 -그래서 친구 같았던- 혁기선배, 먼저 고생하고 졸업했던 진호, 같은 주제로 고민했던 -괴롭히고 또 물심양면 도와주었던- 영선선배 등에게도 감사의 마음을 전합니

다. 또 앞으로 연구실에서 고생해야할 후배 성연, 영웅, 재한, 인준과 아직은 연구실 새내기 동주, 유종, 성준, 찬훈, 영탁, 학부동안 같이 공부했던 승출, 형택, 두경에게도 감사의 글을 전합니다.

그 밖에도 2년 동안 같이 고생했던 대학원 선배님과 후배들 그리고 동기들에게도 감사의 뜻을 전하고, 학부를 같이 보냈던 정민과 보곤에게도 감사의 맘을 전합니다. 또 삶의 활력이 되었던 우리 GS 패밀리 재한, 재민, 민식, 민철에게도 감사의 글을 전하고 싶고, 그리고 지금은 자주 못 보는 베프 태수, 오랜 시간 늘 친구의 자리를 지켜준 -때론 형제처럼 때론 친구처럼- 친구 우영, 친구이자 나의 스폰서였던 -그래서 늘 고마운- 친구 태현에게 특별한 감사의 마음을 전하고 싶습니다.

이 밖에 순간의 실수로 기억하지 못하여 다 언급하지 못하는 많은 분들에게도 감사의 마음을 전합니다.

마지막으로 -사실 가장 처음 떠올랐지만- 언제나 나를 지지해주는 응원군이자 버팀목, 하지만 언제나 감사함을 잊고 지내는 우리 가족, 아버지, 어머니, 형님께 가장 많이 감사의 마음을 전하고 싶습니다.

정말 마지막으로 어둠과도 같았던 시간동안 폭발하지 않고 잘 버텨온, 그리고 더 어두운 앞으로의 시간들과 싸워야 하는 나에게 격려와 칭찬과 감사를 전합니다.

2009年 2月

한참 추운 겨울 어느 날, 깊은 감사의 마음을 담아서

金 元 熙 드림