

스포츠 중계 방송의 구조적 특성을 이용한 축구동영상 하이라이트 생성 알고리즘 (Soccer Video Highlight Building Algorithm using Structural Characteristics of Broadcasted Sports Video)

김재홍^{*} 남종호^{**} 하명환^{***} 정병희^{***} 김경수^{***}
(Jae-Hong Kim) (Jong-Ho Nang) (Myung-Hwan Ha) (Byung-Hee Jeong) (Kyung-Soo Kim)

요약 본 논문에서는 축구 동영상에서 스포츠 중계방송의 구조적 특성을 이용하여 자동적으로 하이라이트(Highlight)를 추출하는 새로운 방법을 제안하고 실험을 통하여 그 유용성을 증명한다. 일반적으로 스포츠 중계 방송에서는 중요한 이벤트(골, 반칙)가 발생하면, 그 장면을 다시 느린속도의 리플레이(Replay) 화면으로 보여주고, 리플레이가 시작되고 끝날 때 Wipe와 같은 점진적인 화면 전환 기법을 사용하는 구조적 특성을 가지고 있다. 본 논문에서는 이러한 스포츠 중계방송의 구조적 특성을 이용하여 축구 동영상의 중요한 이벤트만을 추출하여 하이라이트를 생성하는 방법을 제안한다. 이를 위해 본 논문에서는 비디오의 슬라이스 이미지(Slice Image)[8]를 이용하여 Wipe효과 장면과 줌인/아웃 장면을 검출하고, 검출된 Wipe 효과 지점을 기준으로 하여 리플레이 장면을 검출한다. 이때 이런 기본 알고리즘에 새로이 Sub-Sampling방법과 Two-Pass 방법을 더하여 보다 빠르고 정확한 Wipe검출을 하였고, 비디오의 모션 정보 및 Wipe지점의 시간간격을 이용하여 보다 정확한 리플레이 장면을 검출하였으며, 샷의 녹색영역 비율을 측정하여 보다 정확한 줌인/아웃 장면을 검출하였다. 이렇게 최종적으로 검출된 리플레이 장면과 줌인/아웃 장면을 토대로 이벤트 샷(Shot)과 선수 샷으로 구성된 하이라이트 장면을 추출하게 된다. 이러한 방법은 자동적으로 축구 중계방송을 요약해줌으로써 축구와 관련된 웹서비스나 방송용 하이라이트 제작에 유용하게 사용될 수 있을 것이다.

키워드 : 디지털 비디오 라이브러리, 비디오 인덱싱, 비디오 요약

Abstract This paper proposes an automatic highlight building algorithm for soccer video by using the structural characteristics of broadcasted sports video that an interesting (or important) event (such as goal or foul) in sports video has a continuous replay shot surrounded by gradual shot change effect like wipe. This shot editing rule is used in this paper to analyze the structure of broadcasted soccer video and extracts shot involving the important events to build a highlight. It first uses the spatial-temporal image of video[8] to detect wipe transition effects and zoom out/in shot changes. They are used to detect the replay shot. However, using spatial-temporal image alone to detect the wipe transition effect requires too much computational resources and need to change algorithm if the wipe pattern is changed. For solving these problems, a two-pass detection algorithm and a pixel sub-sampling technique are proposed in this paper. Furthermore, to detect the zoom out/in shot change and replay shots more precisely, the green-area-ratio and the motion energy are also computed in the proposed scheme. Finally, highlight shots composed of event and player shot are extracted by using these pre-detected replay shot and zoom out/in shot change point. Proposed algorithm will be useful for web services or broadcasting services requiring abstracted soccer video.

Key words : Digital Video Library, Video indexing, Video abstracting

· 본 연구는 서강대학교 산업기술연구소 지원으로 이루어졌음

* 비 회 원 : 서강대학교 컴퓨터학과
euroup@chollian.net

** 중신회원 : 서강대학교 컴퓨터학과 교수
jhnang@ccs.soqang.ac.kr

*** 비 회 원 : KBS 기술연구소 연구원

mha@kbs.co.kr

bhjung@kbs.co.kr

odyssey@kbs.co.kr

논문접수 : 2002년 10월 24일

심사완료 : 2003년 4월 4일

1. 서론

컴퓨터와 대용량 저장매체의 발달함으로써 여러 종류의 비디오 데이터들을 사용하는 많은 응용분야가 생겨났으며, 특히 초고속 인터넷의 발달로 인하여 비디오 데이터에 대한 많은 수요가 생겨나게 되었다. 이러한 비디오 데이터들은 기존의 텍스트와 이미지 위주의 정보와는 달리 비디오와 오디오 정보가 복합적으로 이용되어 단조로운 정보전달 패턴보다 진일보하여 사용자들에게 보다 쉽게 전달될 수 있는 장점을 가지고 있다. 하지만 비디오 데이터는 그 상영시간이 다른 데이터들에 비해 길고 또한 용량이 크기 때문에 이를 중요한 장면만 모아서 보다 효과적으로 사용자들에게 제공되는 방법이 필요하게 되었다. 이러한 필요성 때문에 많은 비디오 요약 방법에 대한 연구가 있어왔다. 그러나, 이러한 연구의 대부분이 그 대상을 비교적 이벤트가 적거나 없는 고정된 구조를 가지는 영화 혹은 뉴스에 집중하고 있어서 스포츠비디오 요약에 관한 연구는 이들에 비해 상대적으로 적었으며, 최근에 와서야 많은 연구가 진행되기 시작하였다.

그 동안 연구되어온 스포츠비디오 요약에 관한 연구들을 보면 거의 대부분이 단순히 샷을 나누는 것을 넘어서 샷들이 가지는 하위수준의 정보들을 이용하여 해당 샷에 대한 의미를 부여하는 단계까지 와 있음을 알 수 있다. 실제로 [1,3]의 연구에서는 기본적으로 스포츠 경기장면을 선수들의 활동이 들어가 있는 선수샷과 경기와는 구별이 되는 관중들의 샷으로 나누고 있다. D. Zhong[2]의 연구에서는 차이는 있지만 기본적으로 테니스 경기장면을 대상으로 하여 테니스 경기 자체에서 의미를 가지는 단위(서브/스트로크/발리/스매쉬 등)로 전체 비디오를 나누고 있다. H. Pan[4]에서는 모든 스포츠비디오에서 중요한 장면에서는 모두 슬로우모션의 재생(Replay)이 있다는 것에 착안하여 전체 비디오를 슬로우모션샷과 그렇지 않은 샷으로 구분하는 알고리즘을 제시하고 있다. 이들 연구를 보면 일반적인 영화나 뉴스 비디오의 단순한 샷 나누기와는 달리 샷들을 각각의 운동경기의 특성에 맞는 의미있는 단위로 나누는 것을 알 수가 있다. 본 논문에서 대상으로 하고 있는 축구의 요약에 관련된 연구들[5,7] 또한 축구 비디오에서는 경기장면과 경기장면이 아닌곳에서 녹색영역의 분포가 큰 차이를 보내는 특징을 이용하여 샷을 나누고 있다. 그리고 하이라이트 추출에 관련된 연구[4]에서도 스포츠 비디오들이 중요한 이벤트가 발생한 경우 재생(Replay)이 포함되는 슬로우모션을 가지는 특징을 이용하여 슬로우모션을 가지는 샷(SLO-MO)과 그렇지 않은 샷으로 구

분을 하고 있다. 하지만 [5,7]의 연구에서는 축구 비디오의 특징이 될 수 있는 녹색영역비율을 이용하고 있지만 단순히 전체 비디오를 샷단위로 나누는 데에 그치고 있으며, [4]의 연구에서는 중요한 이벤트장면에 대한 재생(Replay) 장면이 삽입되는 스포츠 비디오의 특성을 반영하고 있지만 반드시 슬로우 모션을 가지는 재생(Replay) 장면을 찾기 때문에 슬로우 모션을 가지지 않는 재생(Replay) 장면은 검출이 되지 않고, 재생(Replay) 장면의 대상이 되는 슬로우 모션을 가지는 샷을 찾는 과정에서 발생할 수 있는 오류에 대한 정정과정 없이 검출오류가 발생한 경우 재생(Replay)장면을 정확히 찾을 수 없고, 요약의 대상을 슬로우모션이 들어간 재생(Replay) 장면으로 하고 있기 때문에 실제 방송에서 사용되는 하이라이트 요약과 많은 차이를 가지고 있다.

본 논문에서는 실제 축구중계 비디오가 가지는 특징인 하이라이트 대상의 이벤트가 발생할 경우 영상의 진행이 줌아웃샷후 다수의 줌인샷 그리고 와이프(Wipe)효과후 재생(Replay)까지 연결된다는 점을 이용하여 실제 스포츠 뉴스 등에서 이용되는 전체 경기진행에 대한 하이라이트 영상을 생성하기 위해 개별적인 하이라이트샷 검출 알고리즘을 제안한다. 전체 시스템은 아래의 그림 1의 순서로 진행이 된다.

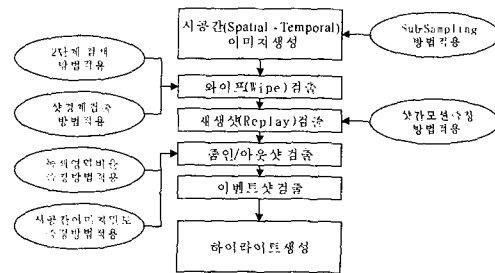


그림 1 축구비디오 하이라이트 생성 시스템

다음장에서는 기존의 스포츠비디오요약과 축구에 특화된 제반연구들에 대하여 살펴보고, 3장에서는 실제 축구비디오의 특징을 살펴본 후 그것에 맞는 하이라이트 샷 검출 알고리즘을 제안하며, 4장에서는 3장에서 제시한 알고리즘들에 대한 실험결과와 분석을 알아보면, 끝으로 5장에서 결론을 맺도록 한다.

2. 스포츠 비디오 분석에 관련된 기존의 연구

스포츠 비디오는 여타 다른 장르들 (뉴스, 영화)와는 달리 어느 특정공간에서 벌어지는 장소적 특성을 가지

며 어떠한 이벤트 단위로 인덱싱을 해야하는 특성들을 가진다. 다른 관련 연구들에서도 이러한 이벤트 단위의 인덱싱을 위해서 다수의 알고리즘을 제안한바 있다. 본 장에서는 이들 연구에서의 사용된 요약 단위, 그 요약 작업을 하기 위해 필요한 하위수준 정보들과 최종적으로 그 하위수준 정보들을 이용하여 요약을 구현한 알고리즘에 대하여 알아보도록 한다. 현재 기존의 연구들이 대상으로 하고 있는 스포츠는 관심도와 상업도가 높고 비교적 많은 샘플영상이 있고, 또한 그 영상 구조가의 패턴화가 용이한 축구와 테니스가 주종을 이루고 있다. 본 논문에서는 축구 이외의 일반적인 스포츠를 대상으로 하고 있는 연구와 본 논문이 대상으로 하고 있는 축구에 관련된 연구로 나누어 설명하도록 하겠다.

2.1 일반적인 스포츠비디오 분석에 관한 연구들

일반적으로 스포츠비디오는 영화나 뉴스등과는 달리 경기 종목마다 장소와 진행되는 패턴이 다르기 때문에 일반적인 인덱싱 방법과는 달리 각 경기의 특성을 반영한 장면의 정의가 필요하고 그것에 맞는 인덱싱 방법이 필요하게 된다. 아래의 표 1에 축구를 제외한 다른 스포츠영상에 관련된 연구들을 설명하였다.

현재 이들 연구는 전체 비디오를 각 경기의 특징에 반영하여 작은 단위로 비디오를 분할하는데 초점이 맞춰져 있지만, 실제 방송에서 필요로 하는 전체 요약생성 알고리즘 등은 제시하지 않고 있다.

2.2 축구비디오 분석에 관한 연구들

축구비디오를 대상으로 한 연구들은 기본적으로 축구가 가지고 있는 특성에 맞추어 분석을 하고 있다. 아래의 표 2에 축구비디오 분석에 관련된 연구들을 정리하였다.

위에서 살펴보았듯이 일반적인 스포츠비디오에 관련된 연구는 일부를 제외하고는 단순히 의미있는 샷단위의 분할에 그 초점이 맞춰져 있다. [4]의 연구에서는 이보다 한발 더 나가서 전체 비디오의 요약인 하이라이트를 생성하려는 연구가 있지만 실제 방송에서 만들어진 하이라이트와는 다른 문제점을 가지고 있다. 축구 비디오 분석에 관련된 연구에서도 역시 샷단위의 분할에 초점이 맞춰져 있으며 다만 [5]에서 이벤트를 찾고자 하는 시도가 있었지만 공의 빠른 움직임이 있어야만 이벤트로써 인식을 하기 때문에 축구 전체에서 일어날 수 있는 다양한 이벤트 전체에 대한 지원이 부족하다. 다음장에서는 축구비디오를 대상으로 실제적으로 축구비디오가 가지는 특징과 실제 방송에서 사용되는 하이라이트 자체에 대한 특징을 분석하여 그 특징에 맞게 실제 방송에서 사용되는 하이라이트 수준의 요약을 만들어 내는 알고리즘을 제안하도록 하겠다.

3. 축구중계비디오 하이라이트 추출 알고리즘

본 장에서는 축구중계 비디오에서 자동적으로 하이라이트를 추출하기 위하여, 방송되는 축구 비디오가 가지는 특성을 조사하고 그 특성을 이용하여 실제적으로 하

표 1 일반적인 스포츠비디오 분석에 관한 연구(축구 제외)

| 연구 | 대상영상 | 추출영상단위 | 해당영상의 특징 |
|---------------|-------|--|---|
| J. Assfalg[1] | 올림픽경기 | 일반영상 - 앵커/인터뷰샷, 스포츠샷 스포츠샷 - 경기샷, 비경기샷 | 일반적으로 경기장면샷은 샷의 길이가 길다 경기장면내에서 문자정보가 나오는 샷은 비경기샷으로 간주할 수 있다. |
| D. Zhong[2] | 테니스 | 서브샷간의 영상 | 테니스는 서브를 기본단위로 비디오가 진행되는 특징을 가진다. |
| B. Li[3] | 미식축구 | 경기샷, 비경기샷 | 경기진행중에는 전체영상이 필드가 대부분이므로 녹색이 강하게 나타나는 특징을 가진다. |
| H.Pan[4] | 일반스포츠 | 재생(Replay)샷 SLO MO샷 | 스포츠영상에서는 중요한부분이 슬로우 모션으로 재생(Replay)되는 특징을 가진다. |

표 2 축구비디오 분석에 관한 연구

| 연구 | 추출영상단위 | 해당영상의 특징 |
|-----------|------------------------------|--|
| M. Lou[5] | 원경샷, 근경샷 이벤트샷(근경샷에서 추출) | 축구경기는 녹색필드에서 진행된다. 이벤트샷의 경우 공의 속도가 빨라지게 되면서 화면상에 블러링(Blurring)이 나타나는 특징이 있다. 근경샷에서는 문자캡션이 일어나는 샷에 보다 높은 관심을 보인다. |
| P.Xu[7] | 전체화면샷 확대샷(원경/근경샷으로 다시 구분) | 축구경기는 녹색필드에서 진행된다. |

이лай트를 생성하는 알고리즘을 제안한다. 이러한 알고리즘을 통한 실험결과를 다음의 4장에서 기술하겠다.

3.1 축구중계비디오의 특성

일반적으로 축구 중계 방송에서는 경기장에 다수의 카메라를 설치한 후 동일한 경기를 여러 각도에서 촬영을 하면서, 시청자의 흥미를 유발하기 위해 방송편집자가 실제 촬영되고 있는 다수의 영상에서 가장 흥미로운 각도를 비추고 있는 소수의 카메라를 선택하고 이를 편집하여 방영하는 특징을 가지고 있다. 따라서 중계중에 특정 이벤트(슛, 골, 반칙 등)가 발생하는 경우 편집자는 이러한 이벤트를 가장 효과적으로 보여주기 위해서 일반적으로 그림 1의 (a)와 같이 이벤트 직후 줌인(Zoom-In) 화면을 통해 이벤트에 관련된 선수를 비추며, 와이프(Wipe)같은 화면전환 방법이 나온 후 이어 해당 이벤트 장면을 다시 보여주는 진행순서로 영상을 송출하는 방법을 사용하고 있다. 하지만 실제 스포츠 뉴스등에서 사용하게 되는 경기의 하이라이트는 이벤트 당시의 라이브영상과 해당 이벤트에 관련된 선수를 비추주는 샷을 묶어서 방영하는 특징을 가지고 있다. 이러한 중계방송과 하이라이트가 가지는 특징을 반영하여 중계방송 영상을 기초로 자동적으로 하이라이트를 만들기 위해서는 중계방송의 영상에서 이벤트의 라이브 영상과 관련선수샷을 정확하게 추출할 수 있는 알고리즘이 필요하게 된다. 그림 2의 (b)에서는 이벤트 발생시의 영상 진행과정과 영상전환 특징들을 보여주고 있는데, 실제 요약대상인 이벤트의 라이브샷과 관련선수를 포함하는 샷이 재생(Replay)가 발생하는 곳의 앞쪽에서 줌아웃에서 줌인으로 넘어가는 과정에 위치하고 있음을

알 수 있다. 이러한 특징을 바탕으로 본 논문에서는 이벤트샷과 관련선수샷의 정확한 위치를 찾고 이를 통해 하이라이트를 생성하는 알고리즘을 제안한다.

3.2 하이라이트 추출 알고리즘

전체적인 하이라이트를 만들기 위해서는 앞에서 언급한것처럼 이벤트샷과 관련 선수샷을 묶어주어야만 한다. 하지만 이러한 이벤트샷과 관련 선수샷 자체를 자동적으로 찾기는 그리 쉬운일은 아니다. 이러한 이벤트샷을 자동적으로 찾기 위한 하나의 방법으로써 H.Pan의 연구[4]에서는 스포츠 비디오에서 사람들이 흥미를 가질수 있는 중요한 이벤트가 발생한 경우 연속적으로 재생(Replay) 장면이 이어서 나온다는 특징을 언급하였으며, 본 논문에서도 이 특징을 이용하여 재생(Replay) 장면의 앞쪽에 있을 이벤트샷과 관련선수샷을 검출하기 위한 방법을 제시하도록 하겠다.

3.2.1 하이라이트 추출을 위한 전체 알고리즘 기본구조
본 논문에서 제안하는 알고리즘은 H.Pan의 연구[4]에서 설정했던 기본 가정과 유사한 형태인 그림 2의 (b)를 기본 설정으로 하고 있다. 이 기본설정은 모든 이벤트샷과 관련선수샷이 재생(Replay) 장면의 앞에 위치하고 있다는 것과, 재생(Replay) 장면은 2개의 와이프(Wipe) 효과 사이에 위치하고 있다는 것, 그리고 실제 이벤트샷과 관련선수샷은 재생(Replay) 장면 앞에서 줌아웃에서 줌인으로 전환되는 지점에 위치하고 있다는 것이다. 이 세가지의 기본설정에 따라 이벤트샷과 관련 선수샷의 정확한 검출을 위해서는 정확한 재생(Replay) 장면 검출과 재생(Replay)장면 검출을 위한 와이프(Wipe)검출, 그리고 줌아웃/인의 전환지점 검출방법을

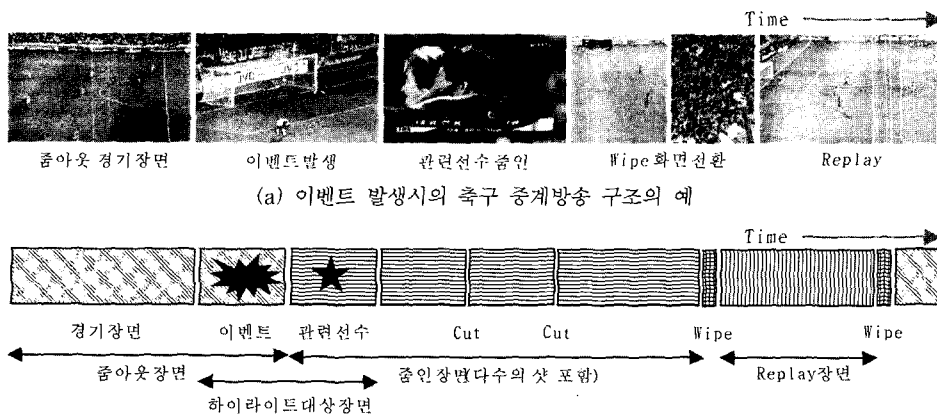


그림 2 이벤트 발생시의 축구 중계방송 비디오의 특성

필요로 하게 된다. 이러한 방법을 묶어서 전체 알고리즘의 기본구조는 아래의 단계처럼 와이프(Wipe) 검출, 재생(Replay)장면검출 후에 이벤트샷과 관련선수샷을 찾기 위한 줌인/아웃경계검출순으로 진행된다.

Step 1. 전체 동영상을 검색하면서 와이프(Wipe)지점을 모두 검출 (와이프(Wipe) 검출)

Step 2. 검출된 전체 와이프(Wipe)지점을 대상으로 두 개의 와이프(Wipe) 지점을 묶어 주면서 재생(Replay) 장면을 검출한다. 이때 전체 재생(Replay)의 길이를 저장한다.(재생(Replay)장면 검출)

Step 3. 검출된 재생(Replay) 장면의 시작지점에서 역방향으로 동영상을 검색하면서 줌아웃/인 전환지점을 검출(줌인/아웃샷 경계검출)

Step 4. 검출된 줌아웃/인 지점에서 앞쪽으로 검색하면서 첫 번째 Cut 지점 혹은 임계값지점까지 그리고 뒤쪽으로 검색하면서 첫 번째 Cut 혹은 임계값지점까지 동영상을 추출한 후 하이라이트의 대상으로 저장(하이라이트 대상검출)

Step 5. 추출되어진 하이라이트들을 재생(Replay)길이에 따라 중요도를 부여하고 사용자의 요구에 따라 중요도에 따른 동영상 재배치를 수행한 후 전체 하이라이트를 생성한다.(하이라이트 생성)

다음 절에서는 이러한 각각의 단계별 검출방법들에 대해서 보다 자세하게 설명하도록 한다.

3.2.2 와이프(Wipe) 검출

재생(Replay)영역을 찾기 위해서는 필수적으로 와이프(Wipe)경계를 찾아야만 한다. 따라서 이 와이프(Wipe)경계를 찾기 위해서 본 논문에서는 M. S. Drew의 연구[9]에서 제시한 시공간(Spatial-temporal) 영상을 이용하였다. 시공간(Spatial-temporal) 영상은 다음의 그림 3에서 설명하고 있듯이 현재의 프레임과 이전 프레임 하나에 대해 각각 일정영역의 횡축 혹은 종축의 적색과 녹색의 2차원 색상 히스토그램을 추출하여 이것들을 히스토그램 인터섹션(Histogram Intersection) 방법으로 히스토그램의 차이를 구하여 이를 하나의 선으로 표시한 것을 의미한다. 이후 연속된 프레임에 대하여 동일한 방법을 사용하여 연속적인 선들의 집합을 생성한후 이 생성된 전체 집합에 형성된 패턴형태에 따라서 Cut/와이프(Wipe)를 검출하게 된다. 본 논문에서는 횡방향의 와이프(Wipe)패턴을 검출하기 위해서 종축의 색상 히스토그램을 사용하였다. 이 시공간(Spatial-temporal) 영상을 생성하는 과정을 다음에 설명하도록 하겠다.

먼저 전체 슬라이스 이미지를 생성하는 과정은 다음과 같다. 먼저 아래의 식 (1)과 같은 방법으로 적색과 녹색을 각각 16단계씩으로 양자화한 전체 256개의 색상 bin(N)을 설정한 후 현재의 프레임에 있는 하나의 종축에 대하여 종축별로 이들 색상에 대한 히스토그램(H)

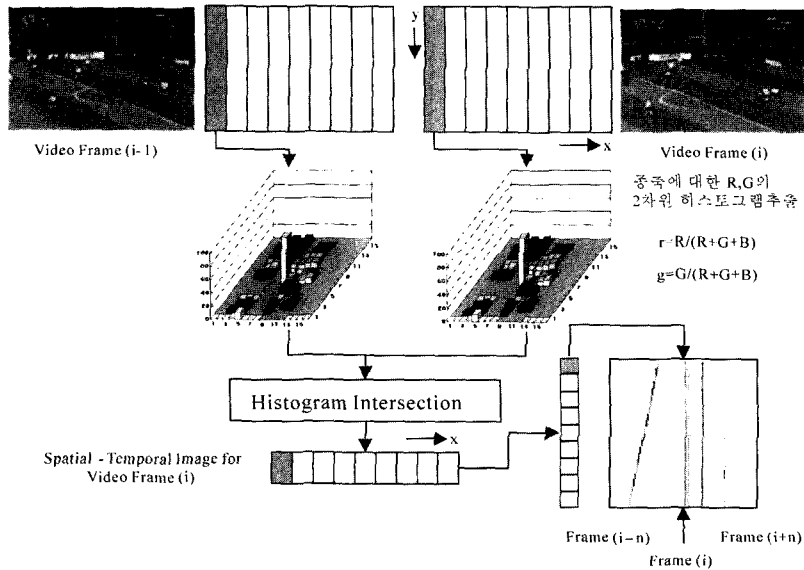


그림 3 시공간이미지(Spatial Temporal Image)의 생성과정

을 생성한다. 식 (2)에서는 같은 방법으로 이전 프레임의 동일한 종축에 대하여 히스토그램(I)를 생성한다.

$$H = \left\{ (h_1, h_2, \dots, h_n) \mid h_i \geq 0 (1 \leq i \leq n), \sum_{i=1}^n h_i = N \right\}, N = 256 \quad (1)$$

$$I = \left\{ (i_1, i_2, \dots, i_n) \mid i_i \geq 0 (1 \leq i \leq n), \sum_{i=1}^n i_i = N \right\}, N = 256 \quad (2)$$

이렇게 생성된 현재의 프레임과 이전프레임의 동일한 종축에 대한 히스토그램 H,I에 대하여 식 (3)의 방법인 히스토그램 인터섹션(Histogram Intersection) 방법을 통하여 히스토그램 차이($D_H(I, H)$)를 구한다.

$$D_H(I, H) = \frac{\sum_{i=1}^n \min(i_i, h_i)}{\min(\sum_{i=1}^n i_i, \sum_{i=1}^n h_i)} \quad (3)$$

이렇게 구해진 히스토그램 차이는 식 (4)에서 나타난 것 처럼 프레임 i에 대한 슬라이스의 종축 x를 투사하는 하나의 점이 된다. 이것을 반복하여 전체 종축에 대한 모든 히스토그램 차이를 구하면 식 (5)가 의미하게 되는 하나의 프레임 전체에 대응하는 슬라이스 하나 ($S(i)$)를 구할 수 있게 된다.

$$HI[v_{i-1}(x), v_i(x)] = D_H(I, H) \quad (4)$$

$v_i(x)$: i번째 프레임의 히스토그램

$$S(i) = HI[v_{i-1}(x), v_i(x)] \text{ for all } x \quad (5)$$

이것을 다시 동영상의 모든 프레임에 적용을 하고 생성된 이들 슬라이스를 모두 연결하게 되면 하나의 이미지를 형성하게 된다. 아래의 그림 4는 본 논문에서 검출 대상으로 하고 있는 와이프(Wipe) 효과와 와이프(Wipe)와 유사한 다른 점진편집효과들에 대한 시공간(Spatial-temporal) 영상 형태를 보여주고 있다.

위의 그림 4에서도 잘 나타나듯이 시공간(Spatial-temporal) 영상은 영상속의 특수효과가 일어나는 진행 방향에 맞추어 생성이 됨을 알수 있다. 아래의 그림에 나타나있는 시공간(Spatial-temporal) 영상은 세로 방향이 시간축이며 가로 방향이 프레임 하나에 대한 슬라이스가 된다. 본 논문에서 대상으로 하고 있는 와이프(Wipe)의 경우 거의 대부분이 왼쪽에서 오른쪽으로 진행되는 와이프(Wipe)이기 때문에 (a)의 예처럼 시공간(Spatial-temporal) 영상이 왼쪽 위에서 오른쪽 아랫방향으로 경사면이 생기고 있으며 (b)의 Motion 와이프(Wipe)의 경우 왼쪽에서 오른쪽으로 하나의 영상이 날아오는 형태이므로 역시 (a)와 비슷한 경사면을 가지고 있고, (d) Flip Over의 경우는 영상이 회전을 하기 때문에 영상이 회전을 의미하는 X자 모양으로 생성이 되며, (e) Check 와이프(Wipe)의 경우 화면안에 다수의 소규모 와이프(Wipe)가 생기는 형태이기 때문에 이를 반영하여 와이프(Wipe)와 비슷한 경사면이 다수 생기고, (f) Barn Door 와이프(Wipe)의 경우 양방향으로 영상이 펼쳐지는 형태이므로 두 개의 경사면을 가지게 된다. 그리고 (c) Cut의 경우는 이동이 없이 즉시 영상이 변하기 때문에 직선의 형태로써 나타나게 된다.

본 논문에서는 왼쪽에서 오른쪽으로 진행되는 와이프(Wipe)를 찾기 위하여 다음의 단계를 통해 와이프(Wipe)를 검출한다. 그리고 역방향의 와이프(Wipe)인 경우에는 반대의 경우를 적용한다.

Step 1. 이미지가 들어있는 윈도우내부에서 최상단 왼쪽 끝부터 그림 5에서 제시된 패턴 (a), (b), (c), (d)의 여부를 확인한다. (P_b, P_c, P_f - 상단, 가운데, 하단화소)

Step 2. 만일 패턴 (a), (b), (c), (d)중 하나가 일치하는 경우에는 바로 Step 3으로 진행을 하며, 그렇지 않고

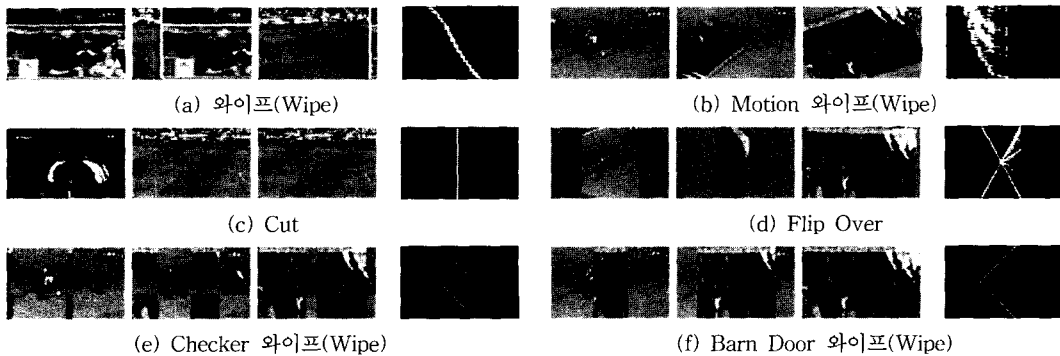


그림 4 점진편집효과 및 Cut 지점의 시공간(Spatial-temporal) 영상의 형태

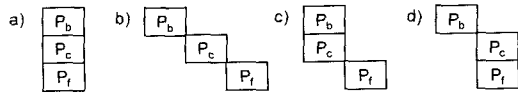
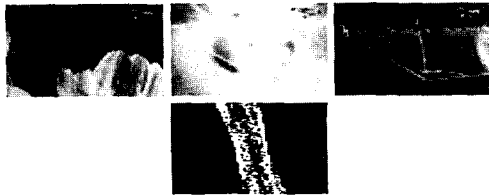
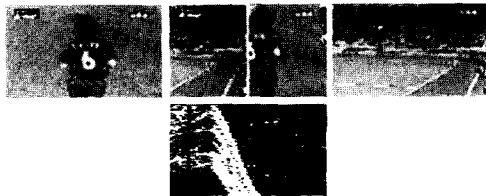


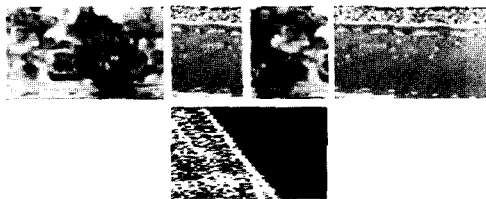
그림 5 검색대상의 축소형태



(a) 2002 피파 월드컵



(b) 컨퍼레이션컵 한국vs멕시코



(c) 컨퍼레이션컵 한국vs프랑스

그림 6 서로 다른형태의 와이프(Wipe) 패턴

패턴이 나타나지 않아서 연결이 끊어지는 경우에는 에러 확률(E_r)을 증가시킨후 Step 3으로 진행을 한다.

Step 3. 이미지의 왼쪽 끝부분으로의 도달여부를 확인한 후 왼쪽 끝부분일 경우에는 Step 4로 진행, 그렇지 않은 경우에는 해당 패턴의 좌측화소부분에서 다시 패턴여부를 조사하고 Step 2로 되돌아간다.

Step 4. 이미지의 왼쪽 끝부분에 도달한 경우 그동안 누적되었던 에러확률(E_r)과 전체적인 세로방향으로의

진행범위(WR)를 조사하여 에러확률이 에러임계값(Th_{Error})보다 작은값을 가지면서 전체의 세로 진행방향 범위가 최소 및 최대 범위 임계값(Th_{ml}, Th_{mh}) 사이에 위치할 때 이를 와이프(Wipe)로 간주한다.

하지만 위에서 제시한 단계를 거칠 경우 이미지의 종축 혹은 횡축방향으로 모든 점을 검사를 하게 되어 많은 시간이 소요되며 또한 아래의 그림 6처럼 와이프(Wipe)의 형태가 조금만 달라져도 해당 와이프(Wipe)에 대응하는 알고리즘을 다시 작성하거나 임계값을 조정해야만 하는 단점을 지니고 있다.

따라서 본 논문에서는 속도 향상을 위해서 서브샘플링(sub-sampling) 검사방법을, 또한 매년 경기마다 다른 와이프(Wipe)패턴에 대응하기 위해서 2단계(two-pass) 와이프(Wipe) 탐색방법을 사용하였다.

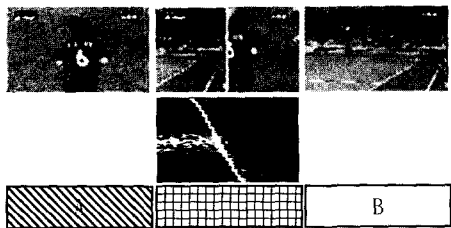
서브샘플링(sub sampling) 방법은 가로방향으로 4화소단위, 세로방향으로 8화소단위로 화소를 추출하는 방법으로 실제로 1/32 크기의 이미지를 검사하게 되는 것과 동일한 효과를 가지게 되어 많은 시간을 단축할 수가 있게 되었다. 아래의 표 3을 보게 되면 실제 검사시간은 약 3.5배정도 빠르면서도 정확도는 거의 차이가 없음을 알 수 있다. 이는 와이프(Wipe)의 시공간(Spatial-temporal) 이미지의 형태가 기울어있는 직선형태이기 때문에 서브샘플링(sub-sampling)을 하게 되어도 추출된 이미지가 직선의 형태를 그대로 유지할 수 있기 때문이다.

그리고, 2단계(two-pass) 와이프(Wipe) 탐색방법은 먼저 전체 축구동영상의 슬라이스(Slice) 영상에서 경사를 이루는 패턴은 와이프(Wipe) 혹은 대형 객체의 이동시 발생하게 되는데 이중 와이프(Wipe)가 가장 많은 빈도를 가지는 데에 기초하고 있다. 이 방법은 전체 동영상의 슬라이스(Slice) 영상을 모두 탐색하여 와이프(Wipe)에 대응될수 있는 모든 경사패턴을 검색한 후 이중 가장 빈도가 높은 경사패턴을 와이프(Wipe)로써 선택하는 방법이다. 이렇게 첫 번째 탐색을 통해 선택되어진 와이프(Wipe)패턴은 재차 두 번째 검색에서 확인절차를 거치게 되는 방법이다. 이렇게 검색을 하게 되면 유사 와이프(Wipe) 패턴을 가지는 대형 객체의 이동에

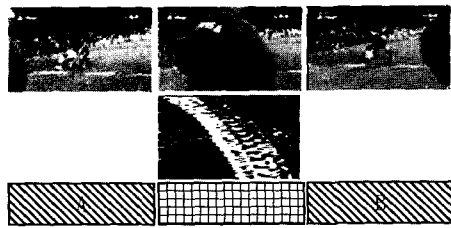
표 3 서브샘플링(sub-sampling) 검사방법의 적용결과

| 한국:멕시코(전반) (45분) | Detection Time | 와이프(Wipe) 검출 정확도 | |
|-------------------------|----------------|------------------|---------------|
| | | Recall | Precision |
| 서브샘플링(sub sampling) 사용전 | 35분 | 38/38 (100%) | 38/67 (56.7%) |
| 서브샘플링(sub sampling) 사용후 | 10분 | 37/38 (97.4%) | 37/72 (51.4%) |

대해서는 실제적으로 경사패턴이 와이프(Wipe)와는 약간씩 다른 패턴을 보이기 때문에 대부분이 2단계(two-pass) 와이프(Wipe) 탐색에서 검출이 되어 제거된다. 하지만 2단계 탐색을 통해서도 유사점이 높아 제거되지 않는 일부 패턴들에 대해서는 아래의 그림 7에서 보여 주듯이 해당 패턴의 전후에서 실제 와이프(Wipe) 패턴들은 장면과 장면을 넘기는 역할을 하기 때문에 와이프(Wipe)가 일어난 후에 장면의 변화가 감지되지만 대형 객체의 이동에서는 전후에서 장면의 변화가 감지 되지 않기 때문에 이러한 장면변화를 검출함



A의 히스토그램 ≠ B의 히스토그램
(a) 와이프(Wipe)의 경우



A의 히스토그램 = B의 히스토그램
(b) 와이프(Wipe)가 아닌 경우

그림 7 유사 와이프(Wipe) 패턴의 구분

으로써 오류의 확률을 최소한으로 줄였다. 그 결과 아래의 표 4에서 알 수 있듯이 2단계(two-pass) 방법과 샷 경계검출방법의 조합을 적용하기 전보다 Recall값이 약간 감소했지만 이전보다 높은 Precision값을 나타낼 수 있다. 이는 미리 유사 와이프(Wipe)패턴을 가지는 장면을 제거함으로써 와이프(Wipe)검출시 발생할 수 있는 잘못검출되는 경우 자체를 줄임으로써 Precision이 높아진 것이며, 또한 유사 와이프(Wipe) 패턴들을 제거할 때 실제 와이프(Wipe)패턴이지만 와이프(Wipe)가 아닌 것으로 판정되어 미리 제거됨으로써 검출이 되지 않는 경우가 발생함으로써 다소 Recall값이 떨어지게 되었다. 이렇게 서브샘플링(sub-sampling) 방법 및 2단계(two-pass) 와이프(Wipe)탐색 방법과 샷 경계검출방법의 조합을 사용하여 보다 빠르고 높은 정확도를 가지고 와이프(Wipe) 검색을 수행할 수 있게 되었다.

3.2.3 재생(Replay)장면 검출

재생(Replay)영역은 추출된 하이라이트 부분을 찾기 위한 기준점이 되면서 동시에 재생(Replay)영역의 길이에 따라서 하이라이트 부분의 중요도를 결정하는 중요한 요소이다. 재생(Replay)영역을 찾는 알고리즘은 그림 8에 나타난 것처럼 [4]에서 제시한 편집효과사이에 둘러싸인 영역이 재생(Replay)가 된다는 스포츠 중계 동영상의 구조적 특성과 한 경기안에서 일어나는 모든 편집효과는 동일하다는 점을 기본으로 하고 있다. 여기서는 그 편집효과를 와이프(Wipe)로 설정하였다.

하지만 아래의 그림 9처럼 와이프(Wipe) 자체를 잘못 검출하는 경우 재생(Replay) 영역을 설정하기 힘든 경우가 발생할 수 있다. 그림 9(a)에서 나타난 것처럼 정상적인 와이프(Wipe)의 가운데에 잘못검출된 와이프(Wipe)가 발생하는 경우 혹은 정상적인 와이프(Wipe)

표 4 2단계(two-pass) 와이프(Wipe)탐색과 샷경계검출 방법 조합의 적용결과

| 한국:멕시코(전반) (45분) | 와이프(Wipe) 검출 정확도 | |
|---------------------------------|------------------|---------------|
| | Recall | Precision |
| 2단계(two-pass)탐색과 샷경계검출방법 조합 사용전 | 38/38 (100%) | 38/67 (56.7%) |
| 2단계(two-pass)탐색과 샷경계검출방법 조합 사용후 | 37/38 (97.4%) | 37/49 (75.5%) |



그림 8 재생(Replay) 영역 설정

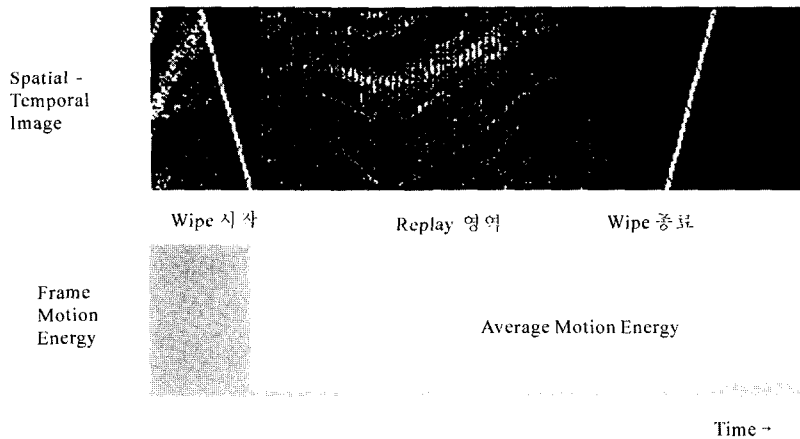
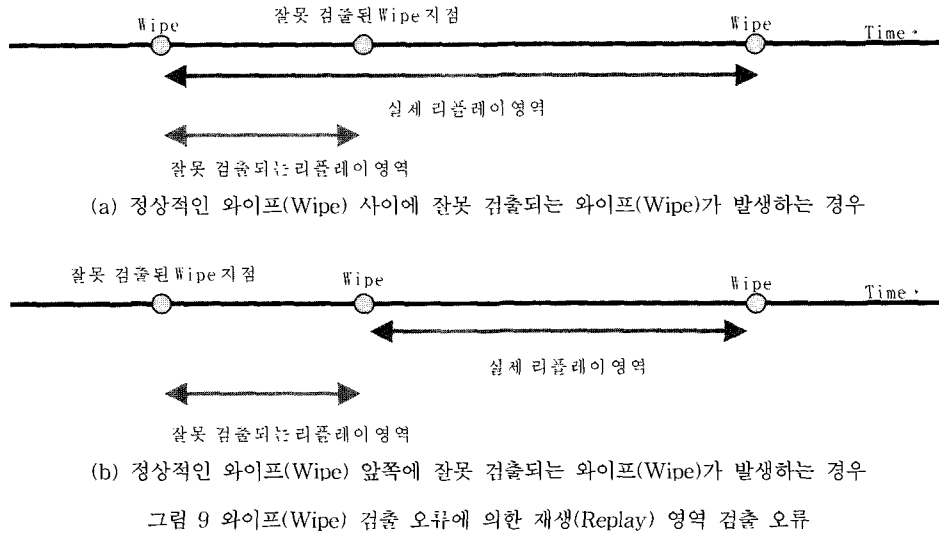


그림 10 재생(Replay) 영역의 모션변화

앞쪽에 잘못검출된 와이프(Wipe)가 검출되는 경우, 이 오류의 수정이 없다면 실제 검출되는 재생(Replay) 영역은 앞쪽의 짧은 구간이 검출되는 결과를 가져오게 된다. 만일 이렇게 재생(Replay) 영역이 잘못 검출되면 뒤에 있을 실제 하이라이트 영역 위치 검출에서도 계속해서 오류가 발생되기 때문에 반드시 이러한 오류는 반드시 최대한 수정이 되어야만 한다. 따라서 이러한 오류를 수정하기 위해서 본 논문에서는 두가지 방법을 이용하였다. 첫 번째 방법은 모션정보를 이용하는 것이고, 두 번째 방법은 와이프(Wipe)지점의 시간간격을 이용하는 방법이다.

첫 번째 모션 정보를 이용하는 방법은 실제 재생(Replay)영역은 느린 영상으로 재생되는 특징을 가지고

있기 때문에 상대적으로 주위보다 낮은 모션값을 가지게 되는 특징을 이용하고 있다. 즉, 그림 10에서 나타난 것과 같이 재생(Replay)가 시작하기 전에 높은 모션값을 가진 후 재생(Replay)가 진행되는 동안에 재생(Replay) 진행 이전보다 상대적으로 낮은 모션값을 가지게 되며 재생(Replay) 진행이 종료한 후 다시 재생(Replay) 진행 때 보다 상대적으로 높은 모션값을 가지게 된다. 아래의 그림 10에서도 재생(Replay) 영역에서 실제로 상대적으로 낮은 모션값이 나타나고 있음을 확인할 수 있으며 특히 재생(Replay) 검출을 위해 설정된 윈도우 내부의 평균 모션값보다도 낮은 값을 보이고 있음을 알 수 있다.

따라서 이러한 특징을 이용하여 와이프(Wipe)가 발생

한 지점 양쪽의 모션값(M_{bwin}, M_{awin})을 측정하여 이들의 비율이 와이프인(Wipe-In)이 될 수 있는 임계치(Th_{wipein})를 넘고, 와이프(Wipe)이후의 모션값(M_{awin})이 평균 모션값(M_{avr})이하가 되는 아래의 식 (6)의 3가지의 경우를 모두 만족하는 경우 와이프인(Wipe-In)이 된다.

$$M_{bwin} > M_{awin}, \frac{M_{bwin}}{M_{awin}} > Th_{wipein}, M_{awin} < M_{avr} \quad (6)$$

그리고 와이프아웃(Wipe-Out) 지점은 와이프인(Wipe-In)의 조건과 반대가 되는 3가지의 조건을 만족하면 된다. 즉 와이프(Wipe)가 발생한 지점 양쪽의 모션값(M_{bwout}, M_{awout})을 측정하여 이들의 비율이 와이프아웃(Wipe-Out)이 될 수 있는 임계치($Th_{wipeout}$)보다 작고, 와이프(Wipe)이후의 모션값(M_{awout})이 평균 모션값(M_{avr}) 이상이 되는 아래의 식 (7)의 조건을 모두 만족하는 시키는 경우이다.

$$M_{bwout} < M_{awout}, \frac{M_{bwout}}{M_{awout}} < Th_{wipeout}, M_{awout} > M_{avr} \quad (7)$$

이러한 조건으로 필터링을 할 경우 화면상에 매우 큰 물체가 지나가는 경우에 와이프(Wipe)와 유사한 형태의 시공간(Spatial-temporal) 영상이 생성됨으로써 발생할 수 있는 검출 오류를 방지할 수 있게 된다.

두 번째 방법은, 첫 번째 방법을 통해서도 제외되지 않는 오류부분을 수정하는 방법으로써 와이프(Wipe)가 일어나는 간격이 설정된 임계치 보다 짧은 경우에 해당 와이프(Wipe)지점을 제외하는 방법이다. 아래의 식처럼 두곳의 와이프(Wipe)지점의 시간차이($D(W_n, W_{n+1})$)를 계산하여 시간 임계치(Th_{min}, Th_{max})와의 비교를 통해 차이가 두 임계치 사이값을 가지게 되면 그대로 첫 번째와 두 번째 와이프(Wipe)발생지점 사이를 영역으로 지정하게 되며, 차이가 최대 임계치(Th_{max})를 넘는 경우는, 첫 번째 와이프(Wipe) 발생지점(W_n)과 첫 번째 발

생지점으로부터 최대 임계치를 더한 부분($W_n + Th_{max}$)까지의 영역을 선택하게 되며, 최소 임계치보다 작을 경우에는, 두 번째 와이프(Wipe)지점을 버리고 첫 번째와 세 번째 와이프(Wipe) 발생지점(W_{n+2})을 가지고 다시 비교를 하게 된다. 이 방법을 통해 첫 번째 방법으로 제외되지 않는 오류의 많은 부분을 제외 할 수가 있다. 아래의 그림 11은 위에서 언급한 두 방법을 통해 제외되는 와이프(Wipe)의 유형을 보여주고 있다.

그림 11에서 (1)-(4)의 지점은 모두 시공간(Spatial-temporal) 영상에서 와이프(Wipe)로 검출된 지역이다. 하지만 (1)의 경우는 첫 번째 방법에서 와이프(Wipe)-In의 조건중 하나인 와이프(Wipe) 발생후의 모션값이 평균 모션값보다 작아야 하는 조건($M_{awin} < M_{avr}$)을 만족하지 못하였기 때문에 제외가 되었으며, (3)의 지점은 정확히 검출된 (2)의 와이프(Wipe)지점에서 최소시간 임계치 보다 가까운 지점에 위치하였기 때문에 두 번째 방법의 3번째 조건($D(W_n, W_{n+1}) < Th_{min}$)에 의하여 제외되었다.

3.2.4 줌인/아웃샷 경계검출

줌인/아웃샷 검출은 해당 샷의 경계지점이 실제 하이라이트가 되는 지점의 기준점이 되는 부분이므로 이 역시 매우 중요하다. 본 논문에서는 와이프(Wipe) 검출에 사용되었던 [9]의 방법을 변형하여 줌인 혹은 줌아웃샷을 찾는다. [9]에서 사용된 시공간(Spatial-temporal) 영상은 프레임간의 차이를 나타내는 수치이기 때문에 이 자체가 운동성을 나타내는 척도가 된다. 아래의 그림 12를 보면 Cut 지점을 경계로 하여 왼쪽이 줌아웃샷, 오른쪽이 줌인샷이 된다.

줌아웃샷에서는 비교적 작은 물체들이 움직이기 때문에 전체적으로 운동성이 낮게 나타나서 시공간(Spatial-temporal) 영상 자체가 매우 저밀도의 특징을 가지며 전체의 모션값자체가 매우 낮게 나타난다. 그 반

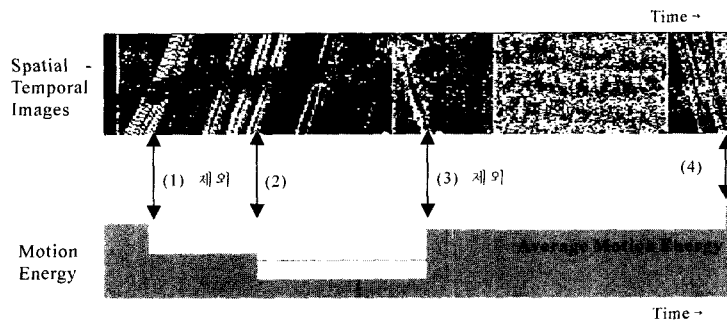


그림 11 와이프(Wipe)제외 유형

대로 줌인장면은 매우 큰 물체들이 움직이기 때문에 운동성이 높아 슬라이스 이미지가 고밀도의 특징을 가지며 모션값이 높게 나타난다. 이러한 특징을 이용하여 이러한 운동성을 구하기 위해서 본 논문에서는 시공간(Spatial temporal) 영상에 대해 다수의 윈도우를 설정하여 아래의 식(8)에서 서술한 각각의 윈도우 내부의 픽셀 밀도($D(i)$)와 평균모션값($M_{avr}(i)$)을 구한 후 인접 윈도우간의 비율을 측정한다.

$$D(i) = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} P(x, y), \quad M_{avr}(i) = \frac{\sum_{x=0}^{X-1} M(x)}{X} \quad (8)$$

(X : 윈도우의 수평길이, Y : 윈도우의 수직길이, i : 윈도우 번호) 이렇게 구해진 각각의 비율을 개별적인 최대, 최소 임계값($(Th_{dmin}, Th_{dmax}), (Th_{mmin}, Th_{mmax})$)과 비교하여 식(9)와 (10)에서 설정한 양쪽 모두의 조건을 만족하는 경우를 각각 줌아웃/인 혹은 줌인/아웃의 경계로 설정하게 된다.

$$\frac{D(i)}{D(i+1)} < Th_{dmin} \text{ and } \frac{M_{avr}(i)}{M_{avr}(i+1)} < Th_{mmin} \quad (9)$$

: 줌아웃/인 경계

$$\frac{D(i)}{D(i+1)} > Th_{dmax} \text{ and } \frac{M_{avr}(i)}{M_{avr}(i+1)} > Th_{mmax} \quad (10)$$

: 줌인/아웃 경계

하지만, 이러한 운동량만으로는 아래의 그림 13에서 보여주고 있는 실제 줌인장면이지만 아무런 움직임이 없는 샷이 발생할 경우에는 전혀 경계 검출이 이루어지지 않을 수 있기 때문에 이러한 오류를 수정하기 위해서 줌아웃 샷의 경우에는 대부분이 녹색영역이 대부분을 차지하고 있는 경기장면을 보여주고, 줌인샷에서는 대부분 사람을 강조하여 보여주는 특성을 이용하여 [5]에서 이용한 녹색영역비율 측정방법을 이용하여 이를 보완하였다.

이 방법은 실제 비디오 영상에서 각 프레임마다 전체 프레임의 가운데 영역에서 녹색으로 간주될 수 있는 화소를 추출하여 이것을 화소를 추출했던 가운데 영역에서 차지하는 비율을 구하는 방법이다. 아래의 식 (11)에서 $U_{zoom}(SF_i)$ 은 i 번째 프레임에서 녹색을 의미하는 Hue의 값이 충분히 녹색영역으로 간주될 수 있는 화소

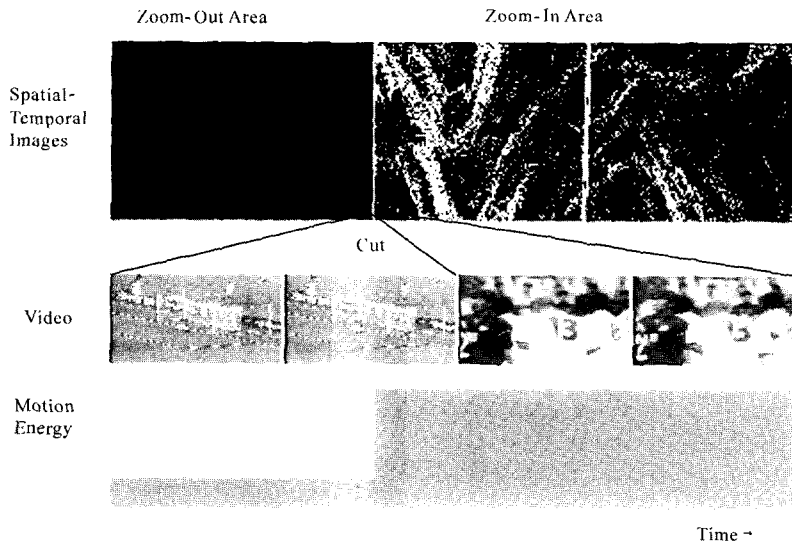


그림 12 줌인/아웃 경계 지점에서의 시공간(Spatial-temporal) 영상 및 모션의 변화



그림 13 모션이 없는 줌인샷

의 집합체를 의미하고 있다.

$$U_{ZOOM}(SF_i) = \{p_{xy}; TH_{lx} < x < TH_{hx}, TH_{ly} < y < TH_{hy}, TH_{lgreen} < Hue(P_{xy})\} \quad (11)$$

그리고 아래의 식 (12)에서 $E_{ZOOM}(SF_i)$ 는 화소를 추출한 영역의 총 화소수와 그 영역에서의 녹색화소수의 비율을 의미하게 되는데, 식 (13)이 의미하는 것처럼 이 값이 충분히 커지게 되면 녹색의 비율이 매우 높다는 것을 의미하기 때문에 해당 프레임을 줌아웃샷으로 구분 가능해진다.

$$E_{ZOOM}(SF_i) = \frac{\|U_{ZOOM}(SF_i)\|}{\| \{p_{xy}; TH_{lx} < x < TH_{hx}, TH_{ly} < y < TH_{hy}\} \|} \quad (12)$$

$$E_{ZOOM}(SF_i) > Th_{green} : \text{줌아웃 샷} \quad (13)$$

3.2.5 하이라이트 생성 알고리즘

앞에서 살펴보았던 와이프(Wipe)검출, 재생(Replay) 검출, 줌인/아웃 혹은 Out/In 장면전환 검출이 정확히 수행되었다면 실제로 하이라이트 장면들을 추출할 수 있게 된다. 앞서 3.1 절에서도 언급되었듯이 실제 하이라이트의 대상이 되는 부분은 재생(Replay) 장면 앞에서 일어나는 줌아웃/인으로의 전환지점을 중심으로 하여 그 지점의 앞쪽 일부와 뒤쪽 일부가 되기 때문에 아래의 그림 14에서 나타난 것처럼 우선적으로 재생(Replay) 시작 지점($Time_{Replay(i)}$)으로 먼저 이동을 하여야 한다. 그리고 해당 재생(Replay)장면의 길이($L(Replay(i))$)를 나중에 있을 하이라이트의 중요도로서 사용하기 위해 저장을 한다. 그 다음 재생(Replay)장면 앞쪽을 검색하면서 하이라이트의 기준점이 될 가장 가까운 줌아웃/인 전환지점($Time_{ZoomOut(k)}$)을 찾는다. 전환지점까지 이동하였다면 전환지점의 앞쪽과 뒤쪽으로 Cut이 일어난 지점을 찾는다. 이때 Cut 지점은 와이프(Wipe)지점을 찾을 때 경사면이 수직인 지점을 Cut 지점으로 검출한다. 하지만 이 Cut 지점과 줌아웃/인 전환지점간의 거리가 너무 큰 경우에는 시간임계치(TH_{Live}, TH_{People})를 설정하여 불필요하게 큰 영역을 추

출하지 않도록 한다. 따라서 Cut 지점과 줌아웃/인 지점간의 거리가 임계치보다 작을 경우에는 해당 Cut 지점이 하이라이트 추출 시작점 혹은 종료점이 되지만 그렇지 않은 경우에는 설정 임계치까지만 하이라이트를 추출하게 된다. 이렇게 해서 시작지점과 종료지점이 설정이 되면 최종적으로 시작지점과 종료지점 사이의 영상을 하이라이트로서 추출하게 된다. 추출된 개별적 하이라이트 장면들은 전체 하이라이트 요약물 생성할 때 모두 포함이 되는 것은 아니다. 추출된 하이라이트 장면들이 많은 경우 해당 하이라이트 장면과 관련된 재생(Replay) 장면의 진행시간이 긴 것부터 전체 요약본에 포함이 된다. 이는 전체 경기에서 해당 장면이 가지는 중요도는 방송 제작자가 결정을 하게 되는데 통상 중요한 장면일수록 긴 재생(Replay) 시간을 부여하기 때문에 본 논문에서도 이러한 특징을 반영하여 긴 재생(Replay) 시간을 가지는 하이라이트 장면들을 우선적으로 포함하였다.

다음장에서는 하이라이트를 찾기 위해 사용되어진 각 경계검출 알고리즘의 실험결과와 실제 하이라이트 지점 검출 결과를 보이고 그 결과에 대한 분석을 하도록 한다.

4. 실험 및 분석

본 논문에서는 두 가지 측면에서 실험을 하였다. 첫 번째는 정확한 하이라이트 지점을 찾고 또한 해당 하이라이트의 중요도를 검출하기 위해 필요한 재생(Replay) 검출과 이의 바탕이 되는 와이프(Wipe)검출, 재생(Replay)검출후 정확한 하이라이트 지점을 찾기 위한 줌인/아웃 경계검출, 그리고 최종적으로 필요한 하이라이트의 검출이다. 또한 실험을 위하여 실제 알고리즘을 구현한 하나의 시스템을 구현하였다. 본 장에서는 이 시스템의 전체 구성과 이 시스템을 통한 실험결과를 보이고 이 결과에 대한 분석을 하도록 하겠다.

3장에서 제안한 알고리즘을 구현하기 위한 시스템은

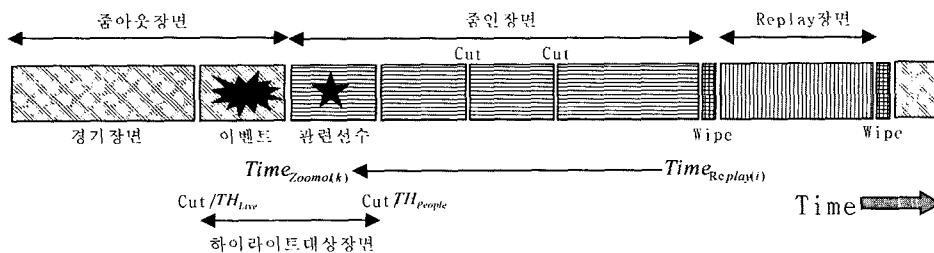


그림 14 하이라이트의 검출

윈도우즈 2000 PC환경에서 구현되었으며 비디오는 윈도우즈의 비디오파일 형식인 AVI형태의 비디오를 사용하였으며, 비디오의 내용은 2001년에 있었던 컨페더레이션컵 2경기과 이 경기에서 사용된 와이프(Wipe)효과와는 다른 형태를 가지는 2002년 피파월드컵 1경기를 대상으로 하였다. 이들 동영상들의 사양은 아래의 표 5에 정리하였다.

4.1 와이프(Wipe), 재생(Replay)샷 및 줌인/아웃샷 검출알고리즘 실험결과 및 분석

4.1.1 와이프(Wipe) 검출

본 논문에서는 와이프(Wipe)검출을 위해 기본적으로 시공간(Spatial-temporal) 영상을 이용하였다. 하지만 이 방법만을 이용할 경우에는 와이프(Wipe) 패턴의 변화 혹은 유사 와이프(Wipe) 패턴에 대해서 정확한 검출이 이루어 질 수 없는 단점이 있다. 따라서 본 논문에서는 2단계(two-pass) Search 방법과 와이프(Wipe) 전후에서의 샷 경계 검출방법을 더하여 보다 정확한 와이프(Wipe) 검출을 행하였다. 아래의 표 6은 단순히 시공간(Spatial-temporal) 영상만을 이용하여 와이프(Wipe)를 검출한 경우이고 표 7은 본 논문에서 제안한 개선된 와이프(Wipe) 검출 알고리즘을 이용한 결과이다.

위의 결과를 보면 전체적으로 높은 Precision을 보임으로써 와이프(Wipe) 전후의 샷 경계 검출방법을 통해 유사 와이프(Wipe) 패턴을 제거하는 방법이 효과가 있음을 알 수 있다. 그리고 각 경기별로 와이프(Wipe) 패턴이 조금씩 다름에도 불구하고 80% 이상의 Recall 수치를 보임으로써 2단계(two-pass) 검색방법 또한 효과

가 있음을 알 수 있다. 하지만 일부 샘플의 경우 Recall 수치가 떨어지는 결과를 보이는 경우(멕시코전 전반전, 프랑스전 전반전)가 발생을 하는데 이 경우는 유사 와이프(Wipe)패턴의 제거에서 정확한 와이프(Wipe) 지점이 유사 와이프(Wipe)패턴으로 검출되어 제거되었기 때문에 발생한 경우이다.

4.1.2 재생(Replay)샷 검출

재생(Replay)샷은 [4]의 연구에서도 사용했듯이 2개의 장면전환 사이에 위치한 장면이 그 대상이 된다. 하지만 [4]의 연구에서는 장면전환 검출에서 발생할 수 있는 오류를 전혀 고려하지 않았기 때문에 실제적인 재생(Replay)샷의 검출에서 많은 오류를 범할 수 있다. 본 논문에서는 이러한 장면전환 검출의 오류를 고려하여 재생(Replay)샷 검출의 오류를 최대한 줄이기 위해서 식 (6)과 (7)에서 처럼 검출된 와이프(Wipe)간의 간격을 측정하여 와이프(Wipe)간격이 너무 짧거나 긴 경우를 제외하는 방법과 와이프(Wipe) 지점 전후의 모션값을 측정하는 방법을 제시하였다. 아래의 표 8은 단순히 와이프(Wipe)가 일어난 지점 사이에 위치한 샷을 그대로 재생(Replay)샷으로 검출한 결과이며, 표 9는 본 논문에서 제안한 수정된 재생(Replay)샷 검출 알고리즘을 이용한 결과이다.

위의 결과를 보면 와이프(Wipe) 검출 오류에 대한 정정 및 시간간격에 대한 고려가 없다면 정확도에서 많은 차이를 보임을 알 수 있다. 특히 한국:프랑스전의 결과를 보면 개선알고리즘을 적용한 경우와 비교했을 때 와이프(Wipe) 검출 자체에서는 정확도면에서 큰 차이를

표 5 동영상 사양

| | | | | | |
|--------|-----------|-----------|-----------|-----------|-----------|
| 동영상사양 | 한국:멕시코(전) | 한국:멕시코(후) | 한국:프랑스(전) | 한국:프랑스(후) | 한국:폴란드(전) |
| 진행시간 | 48분 18초 | 48분 07초 | 46분 13초 | 47분 41초 | 47분 15초 |
| 총 프레임수 | 86877 | 86547 | 83103 | 85755 | 84957 |

표 6 시공간(Spatial-temporal) 영상만을 이용한 와이프(Wipe) 검출결과

| | | | | | | |
|-----------|--------------|--------------|--------------|--------------|---------------|----------------|
| | 한국:멕시코(전반전) | 한국:멕시코(후반전) | 한국:프랑스(전반전) | 한국:프랑스(후반전) | 한국:폴란드(전반전) | 총계 |
| Recall | 30/32(93.8%) | 34/38(89.5%) | 40/42(95.2%) | 22/32(68.8%) | 55/63(87.3%) | 181/207(87.4%) |
| Precision | 30/70(42.9%) | 34/95(35.8%) | 40/45(88.9%) | 22/23(95.6%) | 55/113(48.7%) | 181/346(52.3%) |

표 7 샷 경계 검출방법과 2단계(two-pass) Search 방법을 더한 와이프(Wipe) 검출결과

| | | | | | | |
|-----------|--------------|--------------|--------------|--------------|---------------|----------------|
| | 한국:멕시코(전반전) | 한국:멕시코(후반전) | 한국:프랑스(전반전) | 한국:프랑스(후반전) | 한국:폴란드(전반전) | 총계 |
| Recall | 29/32(90.6%) | 37/38(97.4%) | 36/42(85.7%) | 26/32(81.3%) | 63/63(100%) | 191/207(92.3%) |
| Precision | 29/38(76.3%) | 37/49(75.5%) | 36/36(100%) | 26/26(100%) | 63/108(58.3%) | 191/257(74.3%) |

보이지 않지만 재생(Replay)샷의 검출에서 검출율이 현저히 떨어짐을 알 수 있는데, 이는 그림 8에 나타난 것처럼, 단 1-2개의 와이프(Wipe) 검출 오류에 의해 선택되는 2개의 와이프(Wipe) 지점이 한 칸 앞으로 당겨지거나 뒤로 한 칸 밀려나감으로써 전체적인 재생(Replay)샷 검출 오류로 이어지는 경우이다. 그리고 표 7에서 한국:프랑스전에 비하여 멕시코전과 폴란드전에서 Precision이 떨어지는 결과가 나왔는데, 이는 와이프(Wipe) 검출 부분에서도 프랑스전에 비해 낮은 Precision을 보임으로써 이에 따른 오류의 발생확률이 높기 때문에 이의 영향을 받고 있는 것이다. 그리고 전체적으로 Recall 수치가 와이프(Wipe) 검출시보다 낮게 나타나고 있는데, 이는 와이프(Wipe)와는 별개로 실제 재생(Replay)샷의 길이가 설정된 시간 임계치보다 길게 나타나서 제외되었기 때문에 나타나는 결과이다.

4.1.3 줌 인/아웃 경계 검출

실제 하이라이트 대상의 기준점이 되는 줌아웃/인 전환지점을 찾기 위해서는 줌인 샷과 줌아웃 샷을 정확히 검출해야한다. 연구 [5,7]에서는 녹색영역검출을 통해서 줌인 샷과 줌아웃 샷을 구별하였고, 본 논문에서는 와이프(Wipe)검출에 이용했던 시공간(Spatial-temporal) 영

상을 이용하여 샷의 운동량을 측정하는 방법을 이용하였고, 거기에 부가적으로 녹색영역검출 방법을 더하여 줌인 샷과 줌아웃 샷을 검출하였다. 아래의 표 10,11은 연구[5,6]처럼 단순히 녹색영역검출 방법만을 사용하여 줌인 및 줌아웃 샷을 검출하여 줌인/아웃 경계와 줌아웃/인 경계를 찾은 것이고, 표 12,13은 본 논문에서 제시한 알고리즘을 이용하여 줌인/아웃 경계와 줌아웃/인 경계를 찾은 결과이다.

위의 결과를 볼 때 단순히 녹색영역검출 방법만을 이용한 결과보다 좋은 결과를 보이고 있는데, 이는 실제 축구 동영상에서는 줌인 샷임에도 불구하고 높은 녹색영역을 포함하고 있는 경우가 많으며, 또한 관중석을 비추는 경우에는 분명히 줌아웃 샷임에도 불구하고 녹색영역을 거의 포함하지 않는 경우도 많기 때문이다. 이렇기 때문에 본 논문에서는 실제 줌아웃 샷에서는 전체적인 모션이 작게 나타나는 특징과 [9]의 연구에서는 간과하였던 시공간(Spatial-temporal) 영상이 영상의 모션을 반영하는 특징을 이용하여 줌인 샷과 줌아웃 샷을 구별함으로써 녹색영역에 상관없이 줌인 샷과 줌아웃샷을 구별할 수 있게 되었고, 여기서 발생할 수 있는 오류에 대해 녹색영역검출방법으로 보다 정확성을 높였다.

표 8 단순 재생(Replay)샷 검출 결과

| | 한국:멕시코 (전반전) | 한국:멕시코 (후반전) | 한국:프랑스 (전반전) | 한국:프랑스 (후반전) | 한국:폴란드 (전반전) | 총계 |
|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|---------------|
| Recall | 10/16(62.5%) | 12/19(63.2%) | 3/20(15%) | 5/16(31.3%) | 11/27(40.7%) | 41/98(41.8%) |
| Precision | 10/35(28.6%) | 12/47(25.5%) | 3/22(13.6%) | 5/12(41.7%) | 11/56(19.6%) | 41/172(23.8%) |

표 9 개선된 재생(Replay)샷 검출 알고리즘을 이용한 결과

| | 한국:멕시코 (전반전) | 한국:멕시코 (후반전) | 한국:프랑스 (전반전) | 한국:프랑스 (후반전) | 한국:폴란드 (전반전) | 총계 |
|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|---------------|
| Recall | 15/16(93.8%) | 19/19(100%) | 17/20(85.0%) | 12/16(75%) | 23/27(85.2%) | 86/98(87.8%) |
| Precision | 15/24(62.5%) | 19/27(70.4%) | 17/17(100%) | 12/12(100%) | 23/33(69.7%) | 86/113(76.1%) |

표 10 녹색영역검출만을 통한 줌인 및 Out샷 검출 알고리즘을 이용한 줌인/아웃 검출결과

| | 한국:멕시코 (전반전) | 한국:멕시코 (후반전) | 한국:프랑스 (전반전) | 한국:프랑스 (후반전) | 한국:폴란드 (전반전) | 총계 |
|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|----------------|
| Recall | 89/125(71.2%) | 104/141(73.7%) | 86/123(70.0%) | 88/124(71.0%) | 110/174(63.2%) | 477/687(69.4%) |
| Precision | 89/122(73.0%) | 104/162(64.2%) | 86/112(76.8%) | 88/122(72.1%) | 110/181(60.8%) | 477/699(68.2%) |

표 11 녹색영역검출만을 통한 줌인 및 Out샷 검출 알고리즘을 줌아웃/인 검출결과

| | 한국:멕시코 (전반전) | 한국:멕시코 (후반전) | 한국:프랑스 (전반전) | 한국:프랑스 (후반전) | 한국:폴란드 (전반전) | 총계 |
|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|----------------|
| Recall | 76/91(83.5%) | 81/100(81.0%) | 71/93(76.3%) | 89/106(84.0%) | 140/168(83.4%) | 457/558(81.9%) |
| Precision | 76/104(73.1%) | 81/96(84.4%) | 71/86(82.5%) | 89/105(84.8%) | 140/266(52.6%) | 457/657(69.6%) |

4.2 하이라이트 장면검출결과

하이라이트 검출의 정확도는 하이라이트 추출의 기준점인 재생(Replay)장면앞의 줌아웃/줌인 전환점을 찾는 정확도를 기준으로 하였다. 그 검출결과는 아래의 표 14에 정리하였다.

위의 결과와 표 7,9,12,13의 결과를 놓고 보면 결국 전체적인 정확도는 줌인/아웃 혹은 줌아웃/인 검출보다는 재생(Replay)검출과 와이프(Wipe)검출의 정확도에 영향을 받고 있음을 알수 있다. [4]의 연구에서도 밝혔듯이 항상 중요한 장면인 이벤트샷이 나오는 경우 항상 재생(Replay)가 있기 때문에 하이라이트의 개수는 재생(Replay)의 개수와 동일하게 된다. 하지만 줌인/줌아웃 샷의 정확도에 따라서 그 결과가 약간 차이가 있게 된

다. 4.1절에서도 밝혔듯이 줌인샷내부에 녹색영역이 강하게 나타나는 샷이 있을 경우 줌아웃샷으로 인식하게 되어 이벤트샷이 있는 줌아웃샷까지 탐색을 하지 않고 그 지점에서 줌아웃/인 경계로 설정해버려 하이라이트 장면이 잘못검출되는 상황이 생긴다. 또한 전체 비디오에서 아래의 그림 15에서 나타나듯이 3.1절에서 언급한 축구동영상 규칙이 적용되지 않아서 찾지 못하는 부분도 있고 또한 이벤트샷지점과 재생(Replay)가 일어나는 장면사이의 시간차이가 크게 벌어지고 그 사이에 줌인 샷이 아닌 실제경기가 들어가는 줌아웃의 샷이 들어가는 경우도 발생한다. 이러한 상황에 대해서는 잘못검출된 것으로 간주하여 전체적인 정확도는 재생(Replay)검출정확도 보다 다소 떨어지는 결과가 나오게 되었다. 이

표 12 개선된 줌인 및 Out 샷 검출 알고리즘에 의한 줌인/아웃 경계 검출결과

| | 한국:멕시코 (전반전) | 한국:멕시코 (후반전) | 한국:프랑스 (전반전) | 한국:프랑스 (후반전) | 한국:폴란드 (전반전) | 총계 |
|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|----------------|
| Recall | 99/125(79.2%) | 124/141(87.9%) | 98/123(76.7%) | 100/124(80.6%) | 132/174(75.8%) | 553/687(80.4%) |
| Precision | 99/102(97.1%) | 124/132(93.9%) | 98/102(96.1%) | 100/105(95.2%) | 132/150(88.0%) | 553/591(93.6%) |

표 13 개선된 줌인 및 Out 샷 검출 알고리즘에 의한 줌아웃/인 경계 검출결과

| | 한국:멕시코 (전반전) | 한국:멕시코 (후반전) | 한국:프랑스 (전반전) | 한국:프랑스 (후반전) | 한국:폴란드 (전반전) | 총계 |
|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|----------------|
| Recall | 84/91(92.3%) | 90/100(90.0%) | 78/93(83.9%) | 99/106(93.4%) | 162/168(96.4%) | 513/558(91.9%) |
| Precision | 84/84(100%) | 90/90(100%) | 78/78(100%) | 99/99(100%) | 162/246(65.8%) | 513/597(85.9%) |

표 14 하이라이트 검출결과

| Measure | 한국:멕시코 (전반) | | 한국:멕시코 (후반) | | 한국:프랑스 (전반) | | 한국:프랑스 (후반) | | 한국:폴란드 (후반) | |
|---------|------------------|------------------|------------------|------------------|-----------------|-----------------|------------------|------------------|------------------|------------------|
| | Recall | Precision | Recall | Precision | Recall | Precision | Recall | Precision | Recall | Precision |
| 하이라이트 | 13/16 (81.2%) | 13/22 (59.1%) | 18/19 (94.7%) | 18/24 (75.0%) | 9/12 (75.0%) | 9/10 (90.0%) | 11/16 (68.7%) | 11/12 (91.6%) | 19/21 (90.5%) | 19/29 (65.5%) |

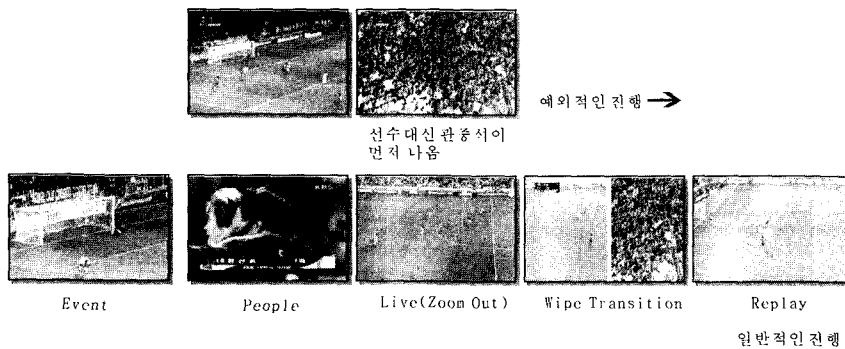


그림 15 하이라이트를 잘못 찾는 경우



그림 16 하이라이트를 찾지 못하는 경우

러한 부분에 대해서는 추후에 재생(Replay)검출 및 줌인/줌아웃샷의 검출을 보다 정확하게 하는 방법을 모색하여 전체적인 정확도를 올릴 수 있게 할 것이다. 또한 이러한 경우외에도 하이라이트를 찾지 못하는 경우가 발생할 하게 되는데 이러한 경우는 그림 16에 나타나는 것처럼 실제 이벤트에 해당되는 부분이 간혹 줌인의 상황(코너킥, 골지점에서 매우 가까운 프리킥같은 세트플레이)에서 나타나는 경우이다. 이러한 경우는 현재 알고리즘으로는 찾을 수 없는 것이 때문에 정확도를 높이기 위해선 이에 대한 별도의 방법이 필요하다.

5. 결론

축구 동영상의 요약은 현재 많은 곳에서 필요로 하고 있지만 실제 기존의 연구들에서는 단순한 샷단위의 분할에 의한 요약을 위한 연구만이 있었고 스포츠 비디오 요약에서 필요한 실제 이벤트 단위로의 요약은 아직 부족한점이 많았다. 본 논문에서는 처음부터 단순한 분할보다 의미를 내포하고 있는 요약단위를 제시하고 그 단위로의 분할과 추출 그리고 요약방법을 구현하였다.

본 논문에서 제안한 분할 및 요약단위는 실제 방송상에서 보여주고 있는 하이라이트의 단위와 동일하며 그 단위는 이벤트샷과 그 이벤트샷의 대상인 선수샷의 조합이 된다. 이 요약단위로의 검출을 위하여 색상 히스토그램 정보를 이용하는 히스토그램 인터섹션(Histogram Intersection) 방법을 이용한 슬라이스 이미지 생성방법을 통해 와이프(Wipe)를 검출하고 또한 재생(Replay) 장면의 특성과 검출된 와이프(Wipe)장면들을 이용하여 재생(Replay)샷을 추출하고, 이 재생(Replay)샷의 위치와 다시 히스토그램 인터섹션(Histogram Intersection) 방법을 다른 방법으로 응용하여 줌인/줌아웃샷을 검출한 후 이들 두 위치의 조합을 통한 정확한 하이라이트기준 위치를 검출하여 하이라이트장면을 검출하는 알고리즘을 제안하고 그 시스템을 구축하였다. 하지만 이러한 과정에서 기본 알고리즘은 몇가지 약점에 의해 긴 수행시간 및 정확도가 떨어지는 단점이 있는데, 이러한 부분에 대해서 각각 기울기 검출방법, 그리고 모션정보 이용과 녹색영역비율 방법을 추가하여 이러한 문제점을 개선하

였다.

앞으로의 연구에서는 축구만이 아닌 다른 일반적인 운동경기들에 대해서 하이라이트를 자동적으로 생성할 수 있는 알고리즘을 제안하게 될 것이다. 이를 위해서 다른 운동경기들에 대한 특성파악을 통해 해당 운동경기에서 의미를 가질 수 있는 최소한의 단위를 설정하는 연구가 계속될 것이다. 또한 본 논문에서 다루지 않았던 다른 편집효과들에 대한 알고리즘 모듈을 추가하여 특정 편집효과에만 적용이 되는 알고리즘이 아닌 일반적인 축구비디오에 모두 적용이 가능한 알고리즘이 될 수 있도록 할 것이다.

참고 문헌

- [1] J. Assfalg, M. Bertini, C. Colombo, and A. D. Bimbo, "Extracting semantic information from news and sports video," *Proceedings of 2nd International Symposium on Image and Signal Processing and Analysis ISPA 2001*, Pula, Croatia, 2001.
- [2] D. Zhong, and S. F. Chang, "Structure Analysis of Sports Video Using Domain Models," *Proceedings of IEEE Conference on Multimedia and Exhibition*, Tokyo, Japan, 2001.
- [3] B. Li, and I. Sezan, "Event Detection and Summarization in American Football Broadcast Video," *Proceedings of SPIE Conference on Storage and Retrieval for Media Databases*, San Jose, USA, 2002.
- [4] H. Pan, P. van Beek, and M. I. Sezan, "Detection of Slow-Motion Replay Segments in Sports Video for Highlight Generation," *Proceedings of ICASSP Conference*, Salt lake City, USA, 2001.
- [5] M. Lou, X. Bai, and G. Xu, "Content-based analysis and indexing of sports video," *Proceedings of SPIE Conference on Storage and Retrieval for Media Databases*, San Jose, USA, 2002.
- [6] G. Sudhir, J. C. M. Lee, and A. K. Jain, "Automatic Classification of Tennis Video for High-level Content-based Retrieval," *Proceedings of ACM Conference on Content-Based Access of*

Image and Video Databases, 1998.

- [7] P. Xu, L. Xie, and S. F. Chang, "Algorithm and System for Segmentation and Structure Analysis in Soccer Video," *Proceedings of IEEE Conference on Multimedia and Exhibition*, Tokyo, Japan, 2001.
- [8] C. W. Ngo, T.C.Pong & R.T.Chin, "Detection of Gradual Transitions through Temporal Slice Analysis," *Proceedings of Computer Vision and Pattern Recognition*, 1999.
- [9] M. S. Drew, L. Ze-Nian, and Z. Xiang, Video Dissolve and Wipe Detection via Spatio-Temporal Images of Chromatic Histogram Differences, *Proceedings of IEEE International Conference on Image Processing*, 2000.



김 재 홍

1997년 9월 서강대학교 컴퓨터학과 학사
1999년 9월 서강대학교 컴퓨터학과 석사
현재 서강대학교 컴퓨터학과 박사과정

남 중 호

정보과학회논문지 : 소프트웨어 및 응용
제 30 권 제 7 호 참조

하 명 환

정보과학회논문지 : 소프트웨어 및 응용
제 30 권 제 7 호 참조

정 병 희

정보과학회논문지 : 소프트웨어 및 응용
제 30 권 제 7 호 참조

김 경 수

정보과학회논문지 : 소프트웨어 및 응용
제 30 권 제 7 호 참조