

## 광범위한 지역 감시시스템에서의 물체기반 비디오 요약

권혜영<sup>o</sup> 이경미  
 덕성여자대학교 지능형멀티미디어랩  
 {khy8355<sup>o</sup>, kmlee}@duksung.ac.kr

### Object-based video summarization in a wide-area surveillance system

HyeYoung Kwon<sup>o</sup> and Kyoung-Mi Lee  
 Intelligent Multimedia Lab., Duksung Women's University

본 논문에서는 광범위한 지역을 감시하기 위해 설치된 여러 대의 카메라로부터 획득된 비디오에 대해 물체를 기반으로 한 비디오 요약 시스템을 제안한다. 제안된 시스템은 시야가 겹쳐지지 않은 다수의 CCTV 카메라를 통해서 촬영한 비디오들을 30분 단위로 나누어 비디오 데이터베이스를 구축하고 시간별, 카메라별 비디오 검색이 가능하다. 비디오에서 물체기반 키프레임을 추출하여 카메라별, 사람별로 비디오를 요약할 수 있도록 하였다. 또한 임계치에 따라 키프레임 검색정도를 조절함으로써 비디오 요약정도를 조절할 수 있다. 이렇게 검색된 키프레임에 대한 카메라별, 시간별 통계를 통해서 감시지역의 물체기반 이벤트를 간단히 확인해 볼 수 있다.

#### 1. 서 론

현대사회에서는 국가적, 사회적 안전뿐만 아니라, 개인과 재산에 대한 안전과 보호의 관심이 높아짐에 따라, 주요 시설물 및 관공서, 학교, 기업, 가정에 이르기까지 보안의 중요성과 적용범위가 넓어져 가고 있다. 특히, CCTV나 PC카메라 등을 이용한 영상 감시시스템은 영상을 통해 능동적으로 감시 대상을 구분하고 추적할 수 있는 효과적인 감시수단으로 이용되고 있다. 그러나 지금까지 개발된 대부분의 영상감시시스템은 감시요원이 카메라로 들어오는 영상을 육안으로 직접 지켜보아야 하기 때문에 비용과 인력의 소모가 크므로, 특정 감시요원 없이 자동적으로 감시 시간동안의 상황을 요약하고 검색할 수 있는 시스템이 필요하다.

비디오에서의 검색 및 요약 시스템은 이미 활발히 연구되고 있는데, 최근 들어 감시시스템에서의 비디오 요약 및 검색 시스템에 대한 연구가 관심을 받고 있다. Silva *et. al* 은 유비쿼터스 환경에 설치된 다수의 카메라로부터 획득된 비디오를 요약하는 시스템을 제안하였다 [1]. 이 시스템은 바닥에 감지센서를 장착하여 사람의 움직임을 추적하였고, 키프레임을 추출하여 시간별, 위치별로 비디오를 검색 요약하였다. 논문 [2]는 어두운 지하철 환경에서 사람추적과 군중모니터링을 통해 위험한 일들에 대해서 미리 카메라를 통한 4시간 재생비디오, 3개의 미리 행동이 녹음된 비디오 채널, 1개의 실시간 비디오 채널을 통해 행동을 분석하여 위험한 상황을 대처할 수 있도록 만든 실시간 시스템이다. 논문 [4]에서 제안한 비디오 감시시스템은 실생활 속에서 긴 시간동안 감시한 내용을 바탕으로 미리 정의된 위기 상황 스냅샷으로 키프레임을 자동적으로 뽑아 주석 처리하여 볼 수 있도록 만들었다. 위기 상황에 대한 비디오 요약 내용은 서버를 통해 인터넷 상에서도 볼 수 있도록 하였다.

본 논문에서는 광범위한 지역에 설치된 다수의 CCTV

카메라를 이용하여 시간 독립적 배경 제거와 움직임 검출을 통해 물체기반 키프레임을 추출하고 비디오를 요약하는 시스템을 제안한다. 제안하는 비디오 요약 시스템은 감시지역에서 촬영한 비디오들에 대해서 요약정도를 임계치에 따라 대략적으로 또는 세밀하게 검색할 수 있도록 하여 방대한 양의 감시지역 비디오를 보다 효율적으로 요약 검색할 수 있도록 하였다.

#### 2. 광범위한 지역 감시시스템

본 논문에서 광범위한 지역을 감시하기 위해 비검침 시야를 가진 다수의 카메라가 설치된 환경에서 다수의 사람 추적이 가능한 감시시스템을 대상으로 한다. 다수의 비검침 카메라는 서버와의 연결을 통하여, 넓은 지역에 걸쳐 설치된 다른 카메라와 추적된 사람에 대한 정보를 주고받을 수 있다. 사람은 사람의 위치, 색상, 움직임 방향, 중심점, 인체를 구성하는 경계영역 등의 정보를 포함한 사람모형을 이용하여 추적 대상의 정보를 전달할 수 있다. 생성된 사람모형의 정보는 각 카메라들과 서버와의 정보전달을 위하여 서버의 사람리스트에 등록, 저장

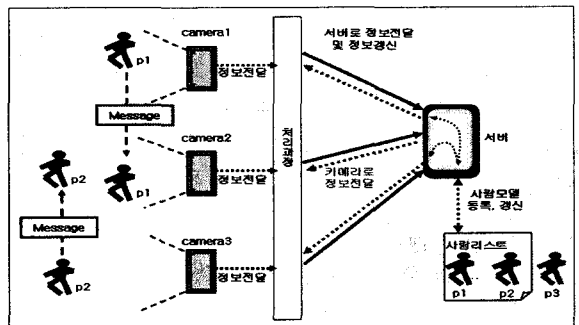


그림 1 서버와 각 카메라간의 정보전달 과정

장된다. 사람리스트에 저장된 사람모델의 정보는 매 프레임마다 사람의 움직임에 따라 사람모델을 갱신시켜 추적대상의 정보변화의 손실을 최소화시킴으로써 정확한 추적이 가능토록 한다.

그림 1은 서버와 각 카메라 사이의 정보전달 과정을 보여주고 있다. camera1은 지나가는 사람을 추출하여 사람모델로 모델링 된 P1로 서버의 사람리스트에 저장되고, 그 후 P1이 camera1에서 나가 camera2에 다시 들어오면 서버는 인식된 사람의 정보가 사람리스트에 등록이 되어 있는가를 살펴본다. 만약 P1이 사람리스트에 등록되어 있지 않았다면, P1을 새로이 등록한다. P1이 사람리스트에 이미 등록되어 있다면, 동일인으로 판단하여 계속 추적한다. 따라서 camera2에 들어온 P1은 camera1에 이어서 P1이라는 모델로 계속 추적되어질 수 있는 것이다. 이와 같은 방법으로 각 카메라에서는 카메라에 들어오는 사람이 기존에 등록되어 있는 사람인가, 새로운 사람이 비디오 프레임 안으로 들어오는가, 또 이전에 비디오에서 나갔던 사람이 다시 들어오는가를 계속해서 확인하여 추적할 수 있다[3].

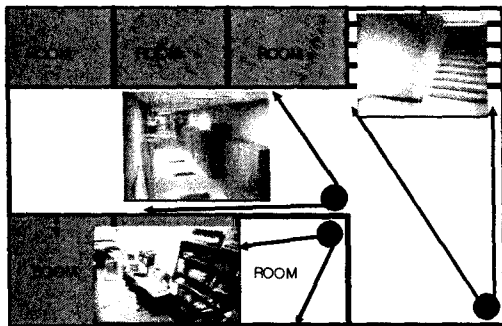


그림 2 CCTV 카메라 설치 영역

본 논문에서는 그림 2에서 보는 바와 같이 3대의 비검침 CCTV 카메라를 연결한 감시시스템에서 얻은 비디오에서 물체기반 키프레임을 검출하고 요약하여 검색할 수 있는 시스템을 제안하였다. 제안하는 시스템은 사람별 요약물체를 통해 여러 대의 카메라를 걸쳐 지나가는 사람들에 대한 검색 및 요약이 가능하다.

### 3. 키프레임 검출

키프레임은 비디오에서 중요한 이벤트나 내용을 담고 있는 프레임으로써, 좋은 비디오 요약 시스템을 구축하기 위해선 좋은 키프레임을 검출하는 것이 매우 중요하다. 본 논문에서는 물체의 움직임이 큰 프레임을 키프레임으로 정의하였고, 올바른 물체 검출을 위해 시간 독립적 배경 제거를 통한 물체를 검출하였다.

다수의 비디오카메라에서 획득된 비디오는 시간의 경과에 따라 빛과 조명에 의해 끊임없이 변화된다. 이러한 입력영상의 조명 잡음은 물체의 정보를 왜곡, 분실시키기 쉬우므로 올바르게 물체를 검출하기 매우 어렵게 된다. 본 논문에서는 시간에 따라 적응적 조명을 모델링하여 다음 프레임에서 나타날 조명과 음영을 미리 예측하

고 제거하는 시간 독립적 배경 제거법을 이용한다 [3]. 고정된 카메라에서 획득된 영상은 배경에 대한 기본프레임으로 적용할 수 있다는 점을 가정하여, 빛의 영향을 가장 받지 않은 영상을 본래영상으로 두었다. 입력된 비디오의 매 프레임마다 이전 프레임에서 계산된 조명의 값으로 현재영상을 보정하게 된다. 첫 번째 프레임을 제외한 두 번째 프레임부터는 아래의 식에 의해 보정된다.

조명 보정된 영상은 본래영상을 이용하여 배경을 분리한다. 배경영상을 분리하기 위해서 평균 배경영상을 계산한다. 새로운 프레임이 들어올 때마다, 평균 배경영상과 비교하여 차이가 보다 적은 픽셀을 배경으로 제거하고 평균 배경영상을 갱신시킨다. 그림 3(a)은 낮 시간의 강한 햇빛으로 조명의 영향을 많이 받은 영상을, 그림 3(b)은 조명 보정 후 배경 제거된 영상을 보여주고 있다.

배경이 제거된 영상에 대해, 물체를 검출하기 위해 움직임이 있는 픽셀들을 연결 요소로 그룹 화시켰다. 그룹화를 마치면, 주어진 영상에 움직임이 있는 연결 영역들의 리스트가 만들어진다. 이 연결 영역들 중에서 크기가 너무 작거나, 길거나, 이질적으로 비조밀한 불필요한 영역을 제거하고 나면 그림 3(c)에서 보이는 것처럼 물체를 검출할 수 있다.

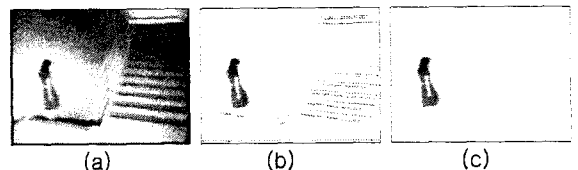


그림 3 (a) 원영상, (b) 배경 제거, (c) 움직임 영상

키프레임은 배경이 제거된 전경 영상에서의 크기가 임계치(th1)보다 큰 프레임으로 간단하게 정의될 수 있다. 그러나 이런 경우 움직임 없이 한 자리에 오래 서 있는 물체를 포함한 모든 프레임들이 키프레임으로 검출되어 키프레임의 개수가 너무 크며 중복되는 프레임을 포함하고 있어 좋은 키프레임들이 검출되었다고 보기 어렵다. 따라서 배경과의 차이뿐만 아니라, 물체의 움직임도 함께 고려한다면 움직임이 적은 중복 프레임을 제거함으로써 보다 적은 수의 키프레임을 검출할 수 있게 된다. 현재 프레임은 이전 프레임과의 차이를 통해 움직임을 계산하고, 계산된 움직임 값이 임계치(th2)보다 큰 경우만 키프레임으로 검출된다.

### 4. 키프레임 기반 비디오 요약시스템

본 논문에서 제안된 물체기반 비디오 요약 시스템은 pentium(R) 4 CPU 3.0GHz와 960MB RAM 메모리 사양의 windows XP상에서 UNIMO CCN-541 보안 카메라 세 대에서 획득한 영상을 처리하기 위하여 JAVA(JMF)를 이용하였다.

#### 4.1 데이터베이스

현재 본 시스템에서는 3대의 카메라에서 5시간동안

(아침 10시부터 오후 3시까지) 촬영한 비디오를 30분 단위로 나누어 비디오 데이터베이스에 날짜시간별, 카메라별로 저장하였다. cameraDB는 사용되고 있는 3대의 카메라에 대한 정보를 저장하고 있다. keyframeDB는 비디오에서 추출된 키프레임들의 전체적인 정보를 저장해 놓은 데이터베이스로 카메라 번호, 연월일, 시분초, 키프레임 번호, bgcount, prevcount, 검출된 물체 수로 구성되어 있다. bgcount는 배경 영상이 제거된 전경 영상의 크기로, 3.1절의 전경 영상에 사용된 임계치(th1)보다 커야만 한다. prevcount는 이전 영상과 차이로 구해진 움직임의 크기로, 3.1절의 움직임에 사용된 임계치(th2)보다 커야만 한다.

#### 4.2 시스템 소개

그림 4는 본 논문에서 제안하는 비디오 검색 및 요약 시스템의 전체적인 GUI를 보여주고 있다. 비디오 검색은 연월일과 시분을 입력하고 카메라 번호를 하나 또는 다수로 지정하여 검색한다. 비디오는 30분 단위로 나누어 검색되어진다. 또한 카메라가 설치된 감시 영역에 대한 정보도 확인할 수 있다. 그림 2와 같이 카메라들의 설치 위치와 시야를 볼 수 있다.

또한 제안하는 시스템은 검색된 감시 비디오에 대한 키프레임을 시간 순으로 나열함으로써 그림 5와 같은 비디오 요약을 제공하고 있다. 제안하는 시스템에서는 비디오 요약은 카메라별 요약과 더불어 사람별로 요약할

수 있다. 다수의 카메라가 선택된 경우, 카메라별 요약은 각각의 카메라에서 촬영한 비디오의 키프레임 요약 결과를 차례대로 보여준다. 예를 들어, 카메라 1,2,3 모두 선택된 경우 카메라 1의 비디오 요약, 카메라 2의 비디오 요약, 카메라 3의 비디오 요약 순으로 결과가 보인다. 이와 달리, 사람별 요약은 카메라와 관계없이 시간 순으로 나열하여 결과가 보인다. 예를 들어, 그림 5에서 보이는 것처럼 카메라 1과 카메라 2의 결과가 혼합되어 시간 순으로 나열된 것이다. 특히 그림 5의 사각형이 둘러져 있는 부분은 카메라 2에서 나온 사람이 카메라 1로 이동하는 모습이 잘 연결되어 보여주고 있다.

다수의 카메라가 연결된 감시시스템에서 촬영된 비디오에 대한 요약은 여전히 많은 분량을 차지하고 있다. 따라서 본 시스템에서는 사용자가 시스템에서 제공해 주고 있는 요약내용을 자율적으로 조절할 수 있도록 하였다. 그림 4 하단의 스크롤바를 조정함으로써 bgcount에 대한 임계치(th1)와 prevcount에 대한 임계치(th2)를 정하게 된다.

또한 30분 비디오의 내용을 한 눈에 보기 위하여 본 논문에서는 카메라에 나타난 물체들에 대한 통계 정보를 제공하고 있다. 그림 6은 그림 5의 요약 결과를 포함한 통계 정보로, 3대의 카메라에 대해 30분 동안 물체가 나타난 프레임들을 보여주고 있다. 그림 6에 그려진 화살표는 그림 5의 사각형 부분으로 다수의 카메라에 걸쳐 지나가는 사람에 대해 올바르게 추적하고 있음을 보여주고 있다.

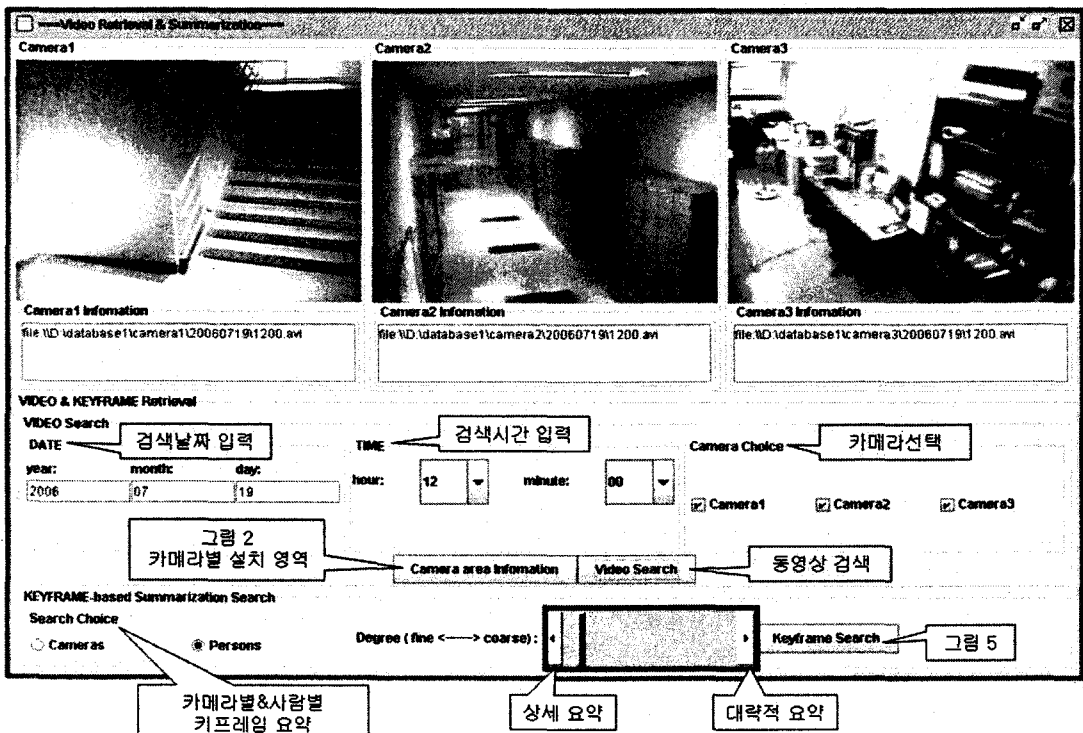


그림 4 제안하는 비디오 검색 및 요약 시스템 GUI

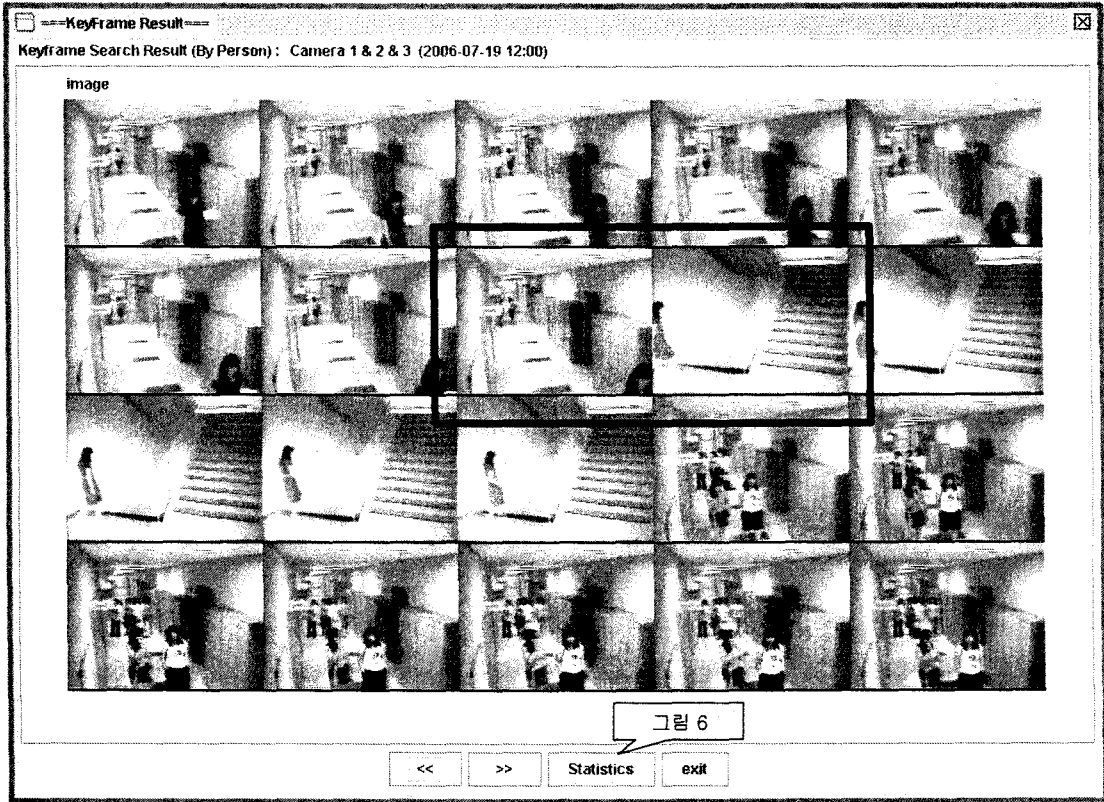


그림 5 임계치에 따른 키프레임 검색

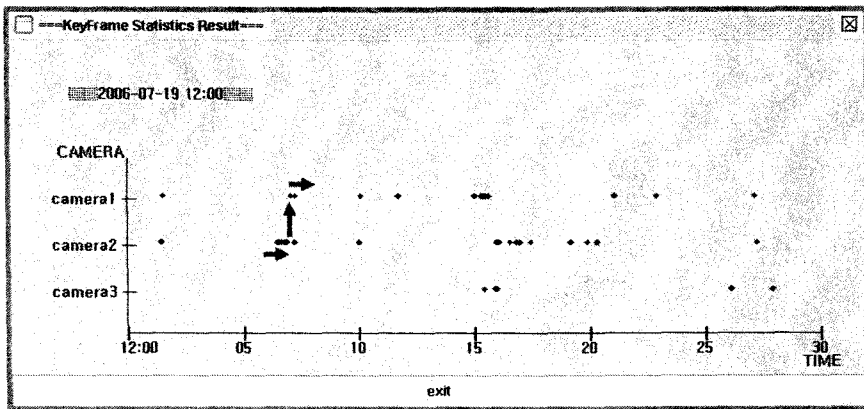


그림 6 임계치에 따른 비디오 요약 통계

5. 결론

본 논문에서 제안한 광범위한 지역 감시시스템에서의 물체기반 비디오 요약은 여러 대의 카메라를 동시에 감시해야 하는 어려움을 해결하기 위해 키프레임 요약 및 사람별 키프레임 검색을 통해 효과적으로 비디오를 요약할 수 있

었다. 또한 임계치를 이용한 키프레임 검색을 통해 시스템이 제공해 주고 있는 요약 내용을 자율적으로 요약할 수 있도록 해서 좀 더 검색의 효율을 높였다. 본 연구에 대한 향후 연구 과제는 현재의 물체 기반 키프레임 검색을 좀 더 세밀한 사람기반 키프레임 검색으로 발전시켜서 다수의 사람을 개별적으로 추적할 수 있도록 하는 것이다.

참고 논문

- [1] G.C. de Silva, T. Yamasaki and K. Aizawa, "Evaluation of video summarization for a large number of camera in ubiquitous home", In proc. ACM Multimedia, 820-828, 2005.
- [2] N.T. Siebel and S.J. Maybank, "The ADVISOR visual surveillance system" In proc. ICCV, 2004.
- [3] 이윤미, 이경미, "모델기반 다중 사람추적과 다수의 비검침 카메라를 결합한 감시시스템", 한국정보과학회논문지: 컴퓨팅 실제, 2006년 8월.
- [4] D.O. Gorodnichy, "ACE Surveillance: the next generation surveillance for long-term monitoring and activity summarization", In proc. international workshop on video processing for security, 2006.