

## 카메라 흔들림을 고려한 배경 모델 생성 방법

\*이재훈, \*\*김형민, \*\*\*박종일<sup>1</sup>, \*\*\*\*김유경, \*\*\*\*김광용  
 \*한양대학교 컴퓨터공학부, \*\*한양대학교 전자컴퓨터통신공학과,  
 \*\*\*한양대학교 컴퓨터소프트웨어학과, \*\*\*\*한국전자통신연구원  
 \*ljh0128@hanyang.ac.kr, \*\*hmkim@mr.hanyang.ac.kr, \*\*\*jipark@hanyang.ac.kr  
 \*\*\*\*yk.kim@etri.re.kr, \*\*\*\*kwangyk@etri.re.kr

### The Background Modeling Method under Camera Shaking

\*Jaehoon Lee, \*\*Hyungmin Kim, \*\*\*Jong-Il Park, \*\*\*\*Yookyung Kim,  
 \*\*\*\*Kwang-Yong Kim  
 \*Dept of Computer Science Engineering, Hanyang University  
 \*\*Dept of Electronics and Computer Engineering, Hanyang University  
 \*\*\*Dept of Computer Software, Hanyang University  
 \*\*\*\*Intelligent Convergence Media Research Dept, ETRI

## 요 약

본 논문에서는 고정된 카메라 환경에서 카메라의 흔들림에 강인한 배경 영상을 생성할 수 있는 배경 모델링 방법을 제안한다. 흔들리지 않은 영상을 기준 영상으로 설정하고 기준 영상에서 해리스 코너 검출기를 이용하여 특징점들을 검출한다. 이후 입력 영상에 대해 동일한 방식으로 특징점을 추출한 뒤 템플릿 매칭과 거리 비교를 이용하여 공통적으로 나타나는 배경 영역들에 대한 특징점만을 선별한다. 기준 영상에서의 특징점과 목표 영상에서의 대응되는 특징점 쌍을 이용하여 보정을 위한 호모그래피 행렬을 계산한다. 이렇게 계산된 보정 행렬을 이용하여 흔들린 목표 영상을 보정하게 된다. 흔들린 영상들을 보정한 후 보정된 영상들로 배경 모델을 생성하게 되면 정확한 배경 모델을 생성할 수 있다.

## 1. 서론

CCTV, 보안 카메라와 같은 고정된 카메라 환경에서의 객체 추적 기법들 중에서 배경 모델링 방법은 객체를 식별할 수 있는 가장 간단하면서 중요한 기법이다. 대표적인 배경 모델링 방법은 가우시안 모델을 이용하는 방법 [1], 커널을 이용하는 방법 [2], 코드북 [3]을 이용하는 방법 등이 있다. 이러한 방법들은 일반적으로 시간의 흐름에 따른 객체의 움직임을 분석하여 움직이는 객체를 제거한 배경 모델을 생성하게 된다. 하지만 바람이나 지진 등으로 카메라가 흔들리게 될 경우 객체가 아닌 영상 자체가 움직이기 때문에 전체적인 배경 모델이 변화하여 배경 모델링에 문제가 발생한다. 배경 모델링에 문제가 발생하게 되면 움직이는 객체를 인식하는데 어려움이 생기고 이는 전반적인 시스템의 성능 하락을 야기한다. 따라서 본 논문에서는 정확한 배경 영상을 생성하여 객체 추적을 용이하게 하기 위해 특징점을 이용하여 영상을 보정함으로써 카메라의 흔들림에 강인한 배경 모델링 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2 장에서는 제안하는 흔들린

영상을 보정하는 배경 모델 생성 방법에 대해 설명하고, 3 장에서는 제안한 방법에 대해 실험을 수행한 결과를 나타내고 분석한다. 마지막으로 4 장에서는 결론을 맺는다.

## 2. 배경 모델 생성 방법

카메라가 흔들리게 될 경우 크게 두 가지 상황에서 문제점을 야기한다. 첫 번째로 배경 모델링을 하는 과정에서 흔들린 영상이 입력으로 들어올 경우 그림 1 과 같이 부정확한 배경 영상을 학습하게 되어 정확한 객체 검출을 할 수 없게 된다. 그리고 두 번째로는 배경 모델링이 잘 되었다 하더라도 객체 검출을 하는 과정에서 고정된 영상이 입력으로 들어오면 객체를 잘 추출하지만 흔들린 영상이 입력으로 들어올 경우 그림 2 와 같이 순수한 객체가 아닌 배경의 일부분들이 검출되어 성능이 떨어지는 현상이 발생한다.

따라서 이러한 문제점들을 해결하여 영상이 흔들려도 배경 모델링이 잘 이루어질 수 있도록 하고 배경 모델링이 된 후에도 객체 검출 과정에서 정확한 객체가 검출될 수 있도록 이 논

<sup>1</sup> 교신 저자

문에서는 흔들린 영상을 고려한 배경 모델링 생성 방법을 제안한다.

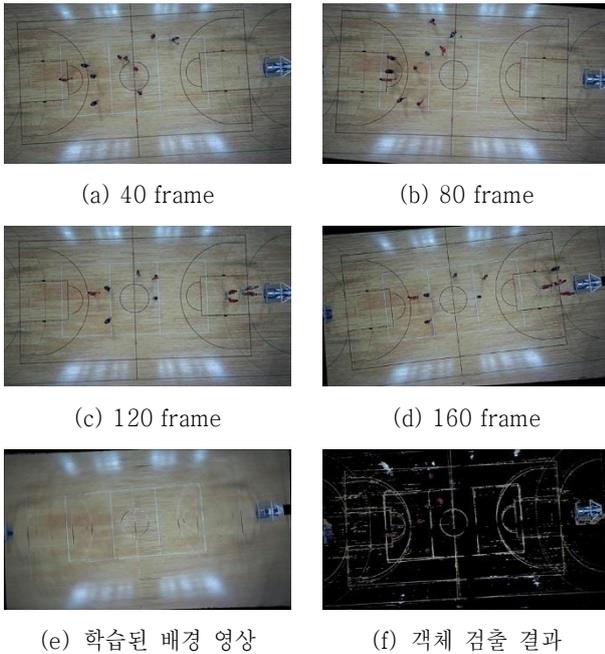


그림 1. 흔들린 영상으로 배경 모델링을 한 결과

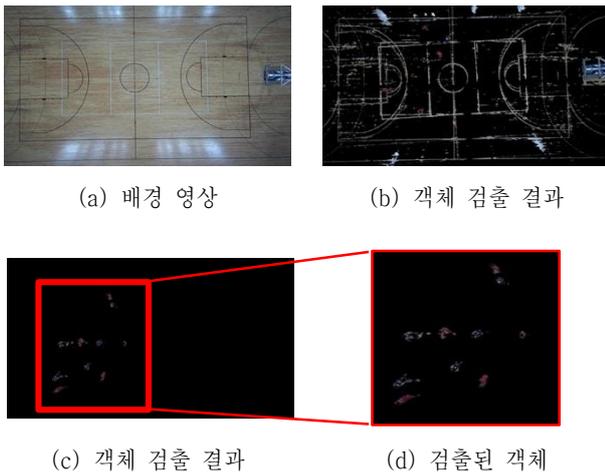


그림 2. 배경 모델링 후 객체 검출 과정에서 흔들린 영상이 입력으로 들어왔을 경우의 객체 검출 결과

그림 3 은 제안하는 방법의 전체적인 흐름도이다. 정상적인 영상과 흔들린 영상을 구분하기 위해서는 각 영상에서 공통적으로 나타나는 부분에서의 특징점을 추출하는 것이 중요하다. 우선 흔들리지 않은 영상을 기준 영상으로 설정하고 기준 영상에서 특징점들을 추출하고 목표 영상에서도 특징점을 추출한다. 기준 영상이 배경 영상이 아니기 때문에 특징점이 객체 영역에 포함될 수 있다. 이러한 문제를 해결하기 위해 템플릿 매칭 (Template matching)을 통해 공통적으로 나타나는 배경 영역의 특징점만을 찾아낸다. 이후 두 특징점 쌍을 이용하여 호모그래피 계산 (Homography)을 통해 보정 행렬을 계산한다. 보정 행렬을 이용하여 흔들린 영상을 보정한 후 배경 모델링을 이용하여 배경 모델을 학습하여 배경 모델을 생성한다.

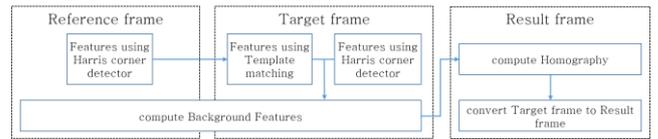


그림 3. 배경 모델 생성 방법 흐름도.

## 2.1 특징점 추출

입력 영상으로부터 헤리스 코너 검출기(Harris corner detector) [4]를 이용하여 특징점들을 추출한다. 영상 내에 윈도우를 이동시켰을 때 모든 방향으로의 영상 변화가 크면 코너로 판단하는 방식으로 검출된 코너를 특징점으로 판단할 수 있다. 그리고 나서 검출된 코너 위치의 정확도를 높이기 위해 수식 1 과 같이 부화소(Sub Pixel)단위로 탐색하여 코너 위치를 보정한다.

$$\sum_{i \in N} \Delta p_i * (q - p_i) = 0 \quad (1)$$

여기서  $q$ 는 부화소 단위로 찾은 코너점이고,  $p_i$ 는 코너점 주변에  $N$ 개의 이웃 픽셀,  $\Delta p_i$ 는  $p_i$ 에서의 영상 기울기이다.  $q$ 를 찾기 위해 식 1 을 식 2 처럼 변환한 후 계산한다.

$$q = \left( \sum_{i \in N} \Delta p_i * \Delta p_i^T \right)^{-1} \sum_{i \in N} (\Delta p_i * \Delta p_i^T * p_i) \quad (2)$$

식 2 를 이용하여 특징점을 추출하여 저장한 후 이후 입력 영상들에 대해서도 동일한 방식으로 특징점을 추출하여 특징점 매칭을 수행한다.

## 2.2 특징점 매칭

2.1 절에서 추출된 특징점은 움직이는 객체 영역들도 포함되어 있다. 특징점 매칭을 수행하기 전에 기준 영상으로부터 추출한 특징점들 중 객체 영역의 특징점들을 제외시켜야 한다. 객체는 영상 내에서 수시로 움직이고 기준 영상과 목표 영상에서 동일한 위치에 있지 않다. 그래서 객체 영역의 특징점으로 영상 보정을 수행할 경우 제대로 된 보정 영상을 얻을 수 없다. 객체 특징점을 제외시키기 위해 식 3 과 같이 기준 영상에서 추출한 특징점과 흔들린 목표 영상에서 추출한 특징점 사이의 거리 비교를 통해 객체 영역의 특징점을 제외한다.

$$\sqrt{(x - x')^2 + (y - y')^2} < \alpha \quad (3)$$

여기서  $(x, y)$ 는 기준 영상에서의 특징점 위치를 나타내고,  $(x', y')$ 는 흔들린 목표 영상에서의 특징점 위치,  $\alpha$ 는 흔들림으로 써 나타날 수 있는 배경 특징점 쌍의 위치 차이에 대한 임계치

(Threshold)값이다. 기준 영상에서의 특징점  $(x,y)$  에 대해서 목표 영상에서의 특징점  $(x',y')$  의 거리가 임계치보다 작은 점이 존재하는지 찾아낸다. 두 특징점 사이의 거리가 흔들림으로써 나타날 수 있는 오차 값보다 클 경우 관련이 없는 객체 특징점으로 판단할 수 있고 반대로 오차 값보다 작을 경우 배경 영역의 특징점 쌍으로 간주할 수 있다.

하지만 거리 비교를 통해 배경 특징점 쌍을 찾아냈지만 객체 영역의 특징점과 배경 영역의 특징점 사이의 거리가 가까워서 임계치보다 작았을 경우 거리 비교 만으로는 구분할 수 없는 경우가 있다. 이러한 문제를 해결하기 위해 SSD(Sum of Squared Difference)를 이용한 템플릿 매칭(Template matching)을 수행하여 정확한 배경 특징점 쌍을 찾는다. 기준 영상의 특징점을 중심으로 일정한 크기의 탐색 범위를 지정하고 흔들린 목표 영상에서 식 4 를 이용하여 가장 유사한 위치를 찾는다.

$$R(x,y) = \sum_{x',y'} (T(x',y') - I(x+x',y+y'))^2 \quad (4)$$

여기서  $R(x,y)$ 는 흔들린 목표 영상에서의 특징점 위치이고,  $T$ 는 기준 영상의 특징점 템플릿,  $I$ 는 흔들린 목표 영상이다.  $T$ 와 유사한 특징점을 가질 경우 해당 위치에서 위 식 4 의 결과 값이 작게 나타난다. 따라서 거리 비교를 통해 찾게 된 특징점 쌍을 대상으로 일정 영역에 대해 템플릿 매칭을 수행하게 되면 객체 영역의 특징점을 제외한 정확한 배경 영역의 특징점 쌍을 찾아낼 수 있다. 그림 5 는 식 4 를 이용하여 특징점 매칭을 수행한 결과이다. 빨간색 점은 기준 영상에서 추출한 특징점들로서 흔들리지 않았을 때 특징점의 위치를 나타낸다. 파란색 점은 흔들린 영상, 즉 보정할 목표 영상에서 추출한 특징점들이다. 거리 비교와 템플릿 매칭을 통해 영상들에서 공통적으로 나타나는 배경 부분에 대해 특징점 쌍을 추출해 낸 것을 확인할 수 있다.

### 2.3 영상 보정

2.2 절에서 추출된 특징점 쌍으로 영상 보정을 위한 보정 행렬을 계산한다. 보정 행렬은 평면 물체의 2D 영상 변환 관계를 나타내는 호모그래피(Homography)를 이용하여 계산하며 계산 방법은 식 5 와 같다.

$$w \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (5)$$

여기서  $w$  는 동차좌표(Homogeneous coordinate)의 변인자(Scale factor)이고  $(x',y')$ 는 기준 영상의 특징점,  $(x,y)$ 는 목표 영상의 특징점 쌍을 나타낸다. 호모그래피 행렬을 구하기 위해서는 최소 4 쌍 이상의 매칭 쌍이 필요하다. 따라서 4 쌍 이상의 특징점이 나올 수 있도록 특징점 추출에서 코너점의 개수를 적절하게 조절한다. 계산된 호모그래피 행렬을 이용해 흔들린 목표 영상에 대해 변환을 시켜주면 흔들리지 않은 영상과 동일한 각도로 보정이 된다. 이렇게 보정된 영상으로

배경 학습 알고리즘을 이용하면 더 정확한 배경 영상을 얻을 수 있게 된다.

### 3. 실험 결과 및 분석

실험 영상은 실제 농구 경기장 천장에 카메라를 설치하여 농구 경기를 촬영한 영상을 가지고 흔들리는 환경을 고려하기 위해 연속해서 임의의 각도로 회전을 시켜서 만든 영상을

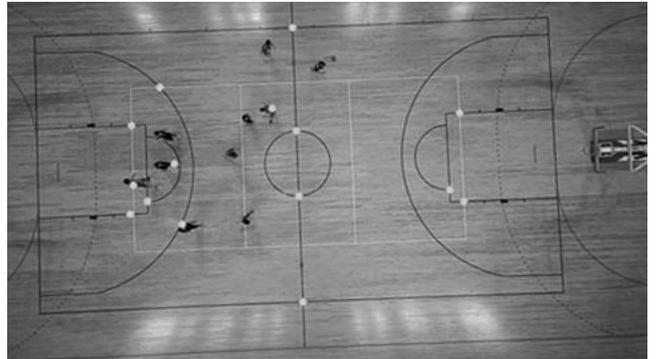


그림 4. 기준 영상에서의 특징점 추출 결과.

이용하였다. 그림 4 는 2.1 절의 방법으로 특징점을 추출한 결과 영상이다.

추출된 특징점들을 배경 영역에서만 아니라 객체 영역에서도 추출된다. 객체의 경우 영상마다 위치가 다를 수 있기 때문에 2.2 절의 방법으로 배경 영역에서만 특징점을 선정한다. 그림 5 는 최종 선택된 특징점들의 결과를 나타낸다.

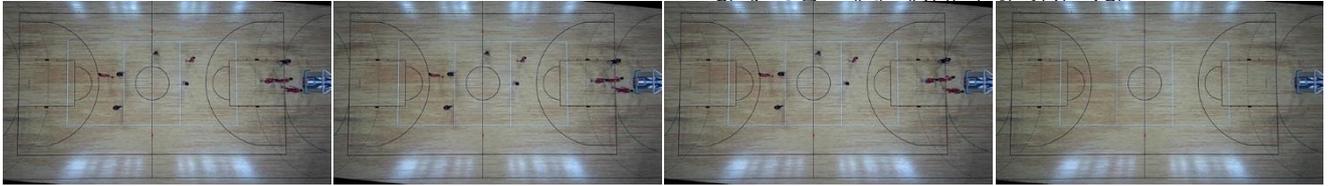


그림 5. 템플릿 매칭 통해 선정된 최종 특징점들.

최종 선택된 특징점들을 이용하여 2.3 절의 보정 행렬을 계산한 뒤 목표 영상을 보정한다. 그림 6 은 흔들린 목표 영상을 템플릿 매칭으로 찾은 배경 특징점 쌍을 이용해 호모그래피 계산을 하여 보정한 결과 영상이다. 왼쪽 영상은 흔들린 프레임 영상이고 오른쪽 영상이 보정한 결과 영상이다.

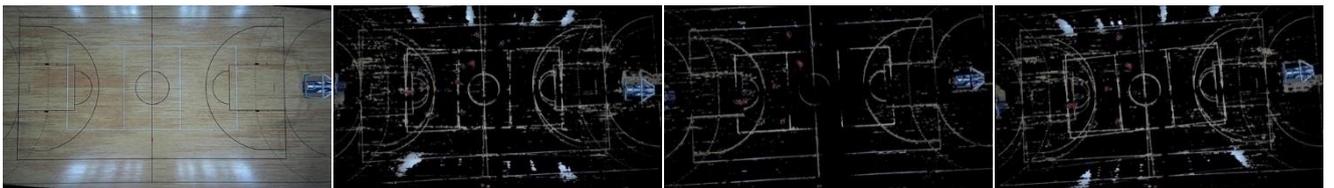


(가) 흔들린 영상들로 학습한 배경 모델 영상.

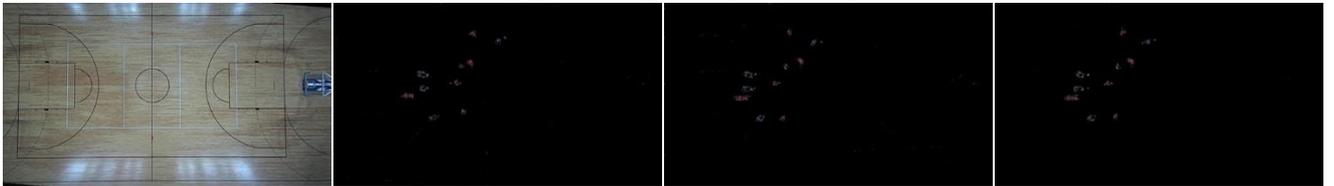


(나) 제안한 방법으로 보정한 영상들을 이용하여 학습한 배경 모델 영상.

그림 7. 제안한 방법으로 배경 모델을 생성하는 과정 (14,15,16 frame)



(가) 흔들린 영상들에 대한 객체 검출 영상.



(나) 제안한 방법으로 보정한 영상을 이용하여 객체를 검출한 영상.

그림 8. 제안한 방법으로 배경 모델링 후 객체를 검출하는 과정 (8,11,21 frame)

그림 7 의 (나)는 제안한 방법으로 생성한 배경 모델이며 배경 모델링 알고리즘[5]을 통해 학습한 배경 영상 결과이다. 흔들린 영상을 보정하지 않고 배경 모델을 생성한 결과보다 제안한 방법으로 배경 모델을 생성하였을 때 학습된 배경 영상이 더 선명하고 정확한 것을 확인할 수 있다. 그리고 그림 8 은 배경 모델을 생성하고 난 이후 객체 검출에도 이 방법을 적용시킬 수 있으며 영상이 흔들릴 경우 (가)와 같이 정확한 객체 검출이 이루어질 수 없다. 논문에서 제안한 방법을 적용시켜 객체 검출을 수행할 경우 (나)와 같이 객체만을 올바르게 검출한 것을 확인할 수 있다.

#### 4. 결론

정확한 배경 영상을 생성하기 위해서는 영상을 흔들리지 않게 고정시키는 것이 무엇보다 중요하다. 이를 위해 영상에서 공통적으로 나타나는 배경 영역의 특징점을 해리스 코너 검출기(Harris corner detector)로 검출한다. 대상 영상에서의 특징점과 템플릿 매칭(Template matching)으로 유사한 특징점 쌍들을 만들고 두 특징점들의 거리 차이를 이용하여 흔들린 정도를 계산한다. 호모그래피(Homography)로 흔들린 영상을 보정해줌으로써 고정된 영상을 생성할 수 있다. 이렇게 보정된 영상으로 배경 학습 알고리즘을 이용해 정확한 배경 영상을 만들 수 있다.

감사의 글: 본 연구는 미래창조과학부 및 정보통신기술 연구진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행하였음. [R0101-16-0293, 방송용 영상 인식 기반 객체 중심 지식 융합 미디어 서비스 플랫폼 개발]

#### 참고문헌

[1] Lee, Dar-Shyang. "Effective Gaussian mixture learning for video background subtraction." IEEE Transactions on Pattern Analysis and Machine Intelligence 27.5 (2005): 827-832.  
 [2] Elgammal, Ahmed, et al. "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance." Proceedings of the IEEE 90.7 (2002): 1151-1163.  
 [3] Kim, Kyungnam, et al. "Real-time foreground-background segmentation using codebook model." Real-time imaging 11.3 (2005): 172-185.  
 [4] Harris, Chris, and Mike Stephens. "A combined corner and edge detector." Alvey vision conference. Vol. 15. 1988.  
 [5] Yao, Jian, and Jean-Marc Odobez. "Multi-layer background subtraction based on color and texture." 2007 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2007.