

제어 가능한 카메라 환경에서 실시간 관심 보행자 검출 및 추적

이병선*, 이은주*

*한밭대학교 정보통신전문대학원

Real-Time Interested Pedestrian Detection and Tracking in Controllable Camera Environment

Byung-sun Lee*, Eun-joo Rhee*

*Graduate School of Information and Communication, HANBAT National University

E-mail: lbs@hanbat.ac.kr

요 약

본 논문에서는 실시간으로 획득된 칼라 영상에서 CMODE(Correct Multiple Object DEtection) 방법을 이용하여 움직이는 다수 물체를 검출하고, 위치 정보와 색상 정보를 이용하여 관심 보행자만을 추적하는 새로운 알고리즘을 제안한다.

다수 물체가 검출되면, 사람의 구조적 특징과 형태 정보를 이용하여 나무의 흔들림이나 차량의 움직임은 제거하고 관심 보행자만을 검출한다. 검출된 관심 보행자 추적을 위한 1차 유사성 판단은 이전 관심 보행자의 무게중심과 현재 관심 보행자의 무게중심간의 거리차를 이용한다. 1차 유사성이 판단된 영역에 대하여 k-평균 알고리즘으로 세 개의 특징점을 구하고, 각 특징점의 3×3 영역에 대한 평균 색상값으로 2차 유사성을 판단하여 추적하도록 한다. 카메라 배율은 원거리의 보행자에 대한 추적을 용이하게 하기 위해서 조정하고, 카메라 시계(FOV: Field of View)는 보행자의 위치가 화면 내의 일정 범위에 있지 않을 경우에 조정한다.

실험 결과, 제안한 CMODE 방법이 라벨링 방법보다 평균 접근 횟수가 1/4배정도 덜 접근하였으며, 평균 검출시간도 3배정도 빠르게 검출됨을 확인할 수 있었다. 나무의 흔들림으로 인한 영역이나 차량의 움직임 영역, 그림자 영역과 같이 복잡한 배경에서도 관심 보행자 검출은 평균 96.5%의 높은 검출률을 보였다. 관심 보행자 추적은 위치 정보와 색상 정보를 이용하여 평균 95%의 높은 추적률을 보였으며, 관심 보행자는 카메라 시계와 배율을 조정함으로써 연속적으로 추적할 수 있었다.

ABSTRACT

This thesis suggests a new algorithm to detects multiple moving objects using a CMODE(Correct Multiple Object DEtection) method in the color images acquired in real-time and to track the interested pedestrian using motion and hue information.

The multiple objects are detected, and then shaking trees or moving cars are removed using structural characteristics and shape information of the man, the interested pedestrian can be detected. The first similarity judgment for tracking an interested pedestrian is to use the distance between the previous interested pedestrian's centroid and the present pedestrian's centroid. For the area where the first similarity is detected, three feature points are calculated using k-mean algorithm, and the second similarity is judged and tracked using the average hue value for the 3x3 area of each feature point. The zooming of camera is adjusted to track an interested pedestrian at a long distance easily and the FOV(Field of View) of camera is adjusted in case the pedestrian is not situated in the fixed range of the screen.

As a experiment results, comparing the suggested CMODE method with the labeling method, an average approach rate is one fourth of labeling method, and an average detecting time is faster three times than labeling method. Even in a complex background, such as the areas where trees are shaking or cars are moving, or the area of shadows, interested pedestrian detection is showed a high detection rate of average 96.5%. The tracking of an interested pedestrian is showed high tracking rate of average 95% using the information of situation and hue, and interested pedestrian can be tracked successively through a camera FOV and zooming adjustment.

Keyword : CMODE방법, 관심 보행자, k-means, 다수 물체

I. 서론

최근 보안과 범죄 예방에 대한 관심이 크게 증대되면서 자동 감시 시스템이나 보안시스템에 대한 연구가 활발히 진행 중이다[1-3]. 특히 동영상에서 자동으로 움직이는 물체를 검출하고 추적하기 위한 연구가 활발히 진행되고 있다[4-7]. 움직이는 물체에는 사람, 자동차, 동물 등 다양하지만 대부분이 사람을 검출하고 추적하는 연구가 주를 이루고 있다[8-12]. 그러나 여전히 강인한 상황(예를 들어 나무들의 움직임, 짙은 그림자 등)으로 인한 분할에 대한 기술적 인 연구가 요구되고 있다.

따라서 본 논문에서는 Pan-Tilt-Zoom 카메라를 통해 획득한 영상에서 CMODE방법을 이용하여 실시간으로 움직이는 다수 물체를 검출한다. 검출된 다수 물체 영역에 대하여 사람의 구조적 특징과 형태 정보를 이용하여 관심 보행자만을 검출하고, 선택된 관심 보행자에 대해 움직임 정보와 색상 정보를 이용하여 연속적으로 추적하는 시스템을 제안하였다.

본 논문의 구성은 2장에서 관심 보행자를 검출하고, 관심 보행자를 추적하는 방법에 관하여 기술하고, 3장에서는 실험 및 고찰에 대해, 4장에서는 결론 및 향후 연구방향에 대해 기술하였다.

II. 본론

본 장에서는 CMODE 방법에 의한 다수 물체를 검출한 후 사람의 구조적 정보와 형태 정보를 이용하여 관심 보행자를 검출하고, 움직임 정보와 색상 정보를 이용하여 관심 보행자를 추적하기 위한 방법에 대해 기술한다.

그림 1은 전 과정을 나타내는 흐름도이다.

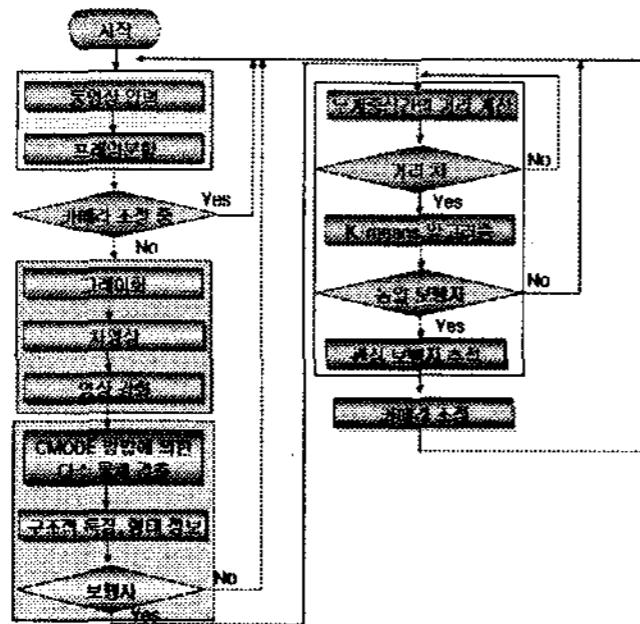


그림 1. 제안한 방법의 전체 흐름도

2.1 차영상 및 닫힘 연산

먼저 입력 영상을 프레임별로 나누고, 입력된 칼라 영상을 그레이 영상으로 변환한다. 영상내의 움직이는 물체를 검출하기 위하여 계산량이 적기 때문에 빠르고 쉽게 물체의 위치를 추출할 수 있는 차영상을 사용하였다[13]. 차영상의 잡

음을 없애고 명확한 물체의 영역만을 검출하기 위하여 실험에 의해 구해진 임계치로 이진화 과정을 수행한다. 이진화된 영상에 대하여 유사한 작은 구멍들이나 틈새를 채우게 되는 형태학적 처리의 8방향 닫힘 연산을 수행한다.

그림 2는 차영상을 구하고 실험치에 의한 임계값으로 이진화 한 후 형태학적 처리의 닫힘 연산을 수행한 결과 영상을 보인 것이다.

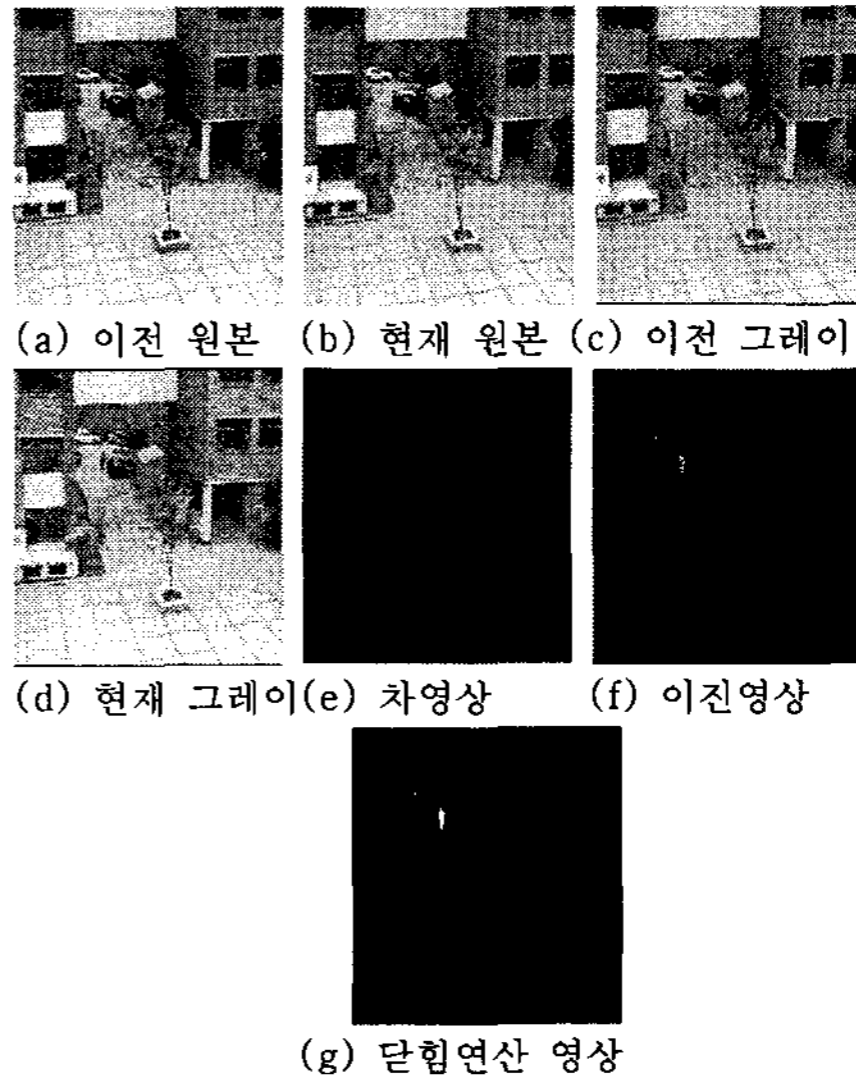


그림 2. 차영상 및 닫힘 연산 결과 영상

2.2 다수 물체 영역 검출

움직이는 물체의 영역을 검출하기 위해서 닫힘 연산된 영상의 이진변환 밀도값을 이용한다. 먼저 x축과 y축의 이진변환 밀도를 (식 1)를 이용하여 계산한 후, 연결된 구간의 이진변환 밀도의 누적값을 계산한다.

$$x\text{축: } f(x) = \sum_{i=0}^{319} i \quad x=0,1,\dots,319$$

$$y\text{축: } f(y) = \sum_{j=0}^{239} j \quad y=0,1,\dots,239 \quad (\text{식 } 1)$$

누적된 x, y축의 이진변환 밀도값을 이용하여 CMODE방법을 수행하여 다수 물체 영역을 검출한다. CMODE방법은 다음과 같이 수행된다.

- 단계 1: 파라미터 초기화
- 단계 2: Y축 수평 프로젝션 -> Y축 구간 설정
- 단계 3: Y축 구간에 대해 X축 수직 프로젝션 -> X축 구간 설정
- 단계 4: X축 구간이 1개 이상이면 X축 구간에 대해 Y축 구간 재설정
- 단계 5: 전체 영상에 대한 처리가 끝났으면 종료, 그렇지 않으면 단계3으로 간다

그림 3은 CMODE 방법을 수행하여 다수 물체 영역을 올바르게 정확하게 결정한 결과 영상으로, ①의 물체 영역 때문에 ②의 영역이 본래의

크기보다 크게 y축 영역이 설정되므로, x축 구간에 대해 y축 구간을 재설정함으로써 본래의 크기대로 영역을 검출하였다.

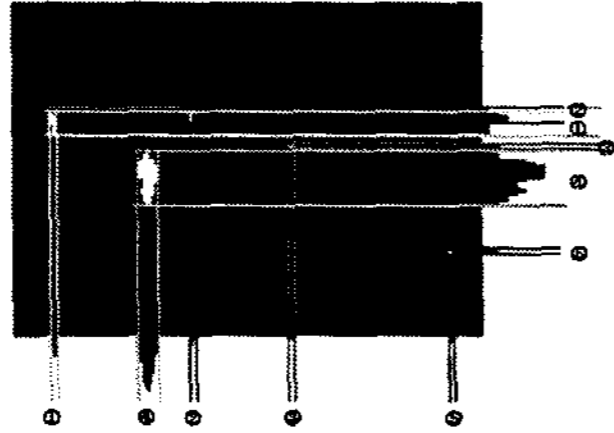
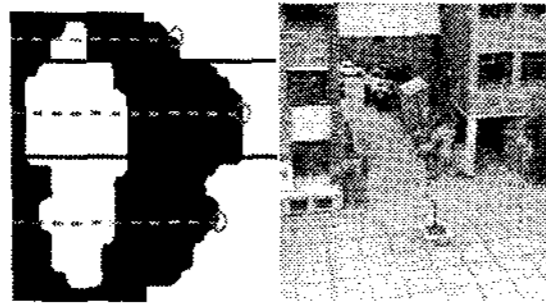


그림 3. 다수 물체 영역 검출 예

다수 물체 영역이 검출되면 각 물체 영역의 가로, 세로 비를 계산한다. 사람이 움직일 경우에는 가로의 길이보다 세로의 길이가 길다는 구조적 조건을 이용하여 가로의 길이가 세로의 길이보다 길거나 같다면 보행자 영역으로 판단하지 않고 제거한다. 남은 영역에 대해 사람의 형태 정보를 이용하여 최종 관심 보행자 영역을 검출한다.

그림 4는 (a)의 사람의 형태 정보를 이용하여 최종 관심 보행자 영역을 검출한 영상이다.



(a) 사람의 형태정보 (b) 관심보행자 검출
그림 4. 최종 관심 보행자 영역 검출

2.3 관심 보행자 추적

관심 보행자 추적을 위해서는 검출된 보행자 영역의 무게중심을 이용한다. 현재 프레임의 보행자 영역과 이전 프레임의 보행자 영역의 중심을 계산한 후, 두 중심의 거리 차를 계산한다. 그 거리가 실험에 의해 정한 임계치 이상이면 다른 보행자로 판단하여 무시하고, 거리가 임계치 이하이면 동일 보행자로 판단한다. 무게중심만으로 동일 보행자인지를 판단하면 가까운 거리의 다른 보행자도 동일 보행자로 판단하는 오류가 발생하므로 사람은 머리, 몸통, 다리 부분으로 크게 나눌 수 있으므로 k-means 알고리즘을 이용하여 세 개의 무게중심점을 구한다. 각각의 무게 중심점과 그 주변 3X3 영역의 Hue값의 평균값을 이용하여 동일 보행자인지를 판단한다. 이전 프레임의 세 점의 평균 Hue값과 현재 프레임의 세 점의 평균 Hue값이 임계치 범위에 있으면 동일 보행자로 판단하여 추적을 계속한다.

2.4 카메라 조정

동일한 관심 보행자로 판단되면, 추적을 계속하기 위하여 카메라 조정을 결정한다. 추적된 관심 보행자의 중심점이 화면의 일정 영역에 있지 않을 경우, 카메라의 방향을 조정하여 관심 보행

자의 위치를 화면의 일정 영역에 위치하게 한다. 또한 멀리 있는 보행자의 인식률을 높이기 위해 배율을 조정하여 영역을 확대한다. 단, 카메라를 조정할 경우 조정에 따른 오차를 최소화하기 위하여 물체 검출 및 추적을 잠시 중지한다.

III. 실험 및 고찰

본 논문은 실험을 위해, Pan-Tilt-Zoom과 칼라 영상을 출력하는 카메라로 320 x 240 사이즈의 동영상을 초당 10 프레임씩 캡처하였고, 알고리즘 구현은 Visual C++ 6.0을 사용하였다.

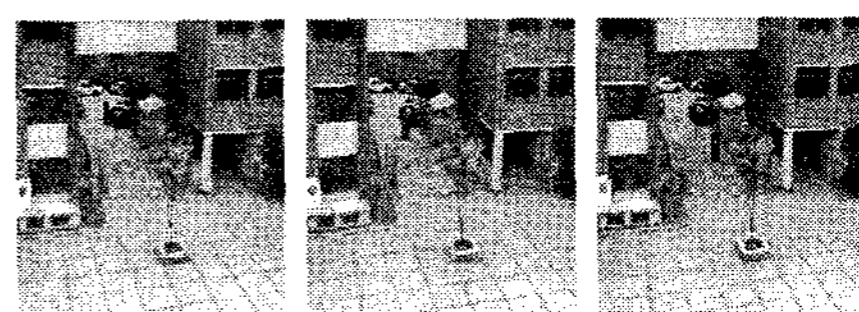
카메라와의 통신을 위하여 RS-232 Serial포트를 사용하였으며, 카메라로부터 영상을 받아들이기 위하여 영상 캡처카드를 사용하였다. 프로그램에서 동영상의 크기와 영상의 RGB값을 가진 이미지를 받아들일 수 있도록 조정하였다.

3.1. 다수 물체 검출 및 관심 보행자 추적

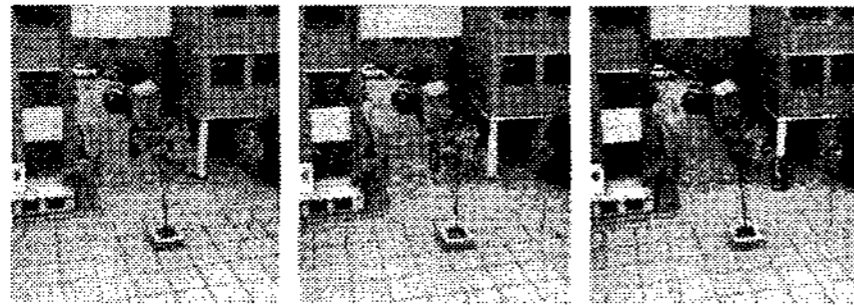
다수 물체 검출에서는 실시간 처리를 위하여 차영상과 실험치에 의해 얻어진 임계치로 이진화를 수행하였다. 이진화된 영상에서 유사한 작은 구멍들이나 틈새를 채우며 물체의 외곽선을 평활하게 하고 형상이나 크기는 유지시키는 형태학적 처리인 닫힘 연산으로 물체의 형태를 명확하게 하였다. 이진변환 밀도를 이용한 CMODE방법으로 다수 물체 영역을 결정하고, 다수 물체 영역 중에서 사람의 구조적 특징과 사람의 형태 정보를 이용하여 관심 보행자 영역을 검출하였다.

검출된 관심 보행자 영역에 대하여 이전 보행자 영역의 중심간의 거리를 이용하여 동일 보행자인지를 1차적으로 판단하고, k-평균 알고리즘으로 구한 세 개의 중심점의 3x3 영역에 대한 평균 hue값을 이용하여 2차적으로 동일 보행자인지를 판단하였다. 동일 보행자로 판단되면 추적하도록 하였다. 멀리 있는 보행자의 인식률을 높이기 위해 카메라 배율을 조정하여 일정 크기로 확대하였으며, 보행자의 중심이 화면내의 범위에 있지 않을 경우 카메라의 시계를 조정하여 연속적으로 추적할 수 있도록 하였다.

그림 5는 검출된 보행자에 대해 중심간의 거리를 이용한 움직임 정보와 k-평균 알고리즘으로 얻어진 색상 정보를 이용하여 관심 보행자에 대해 추적한 결과를 보여주고 있다. (a)~(f)는 시간 흐름 순서대로 나열한 것이다.

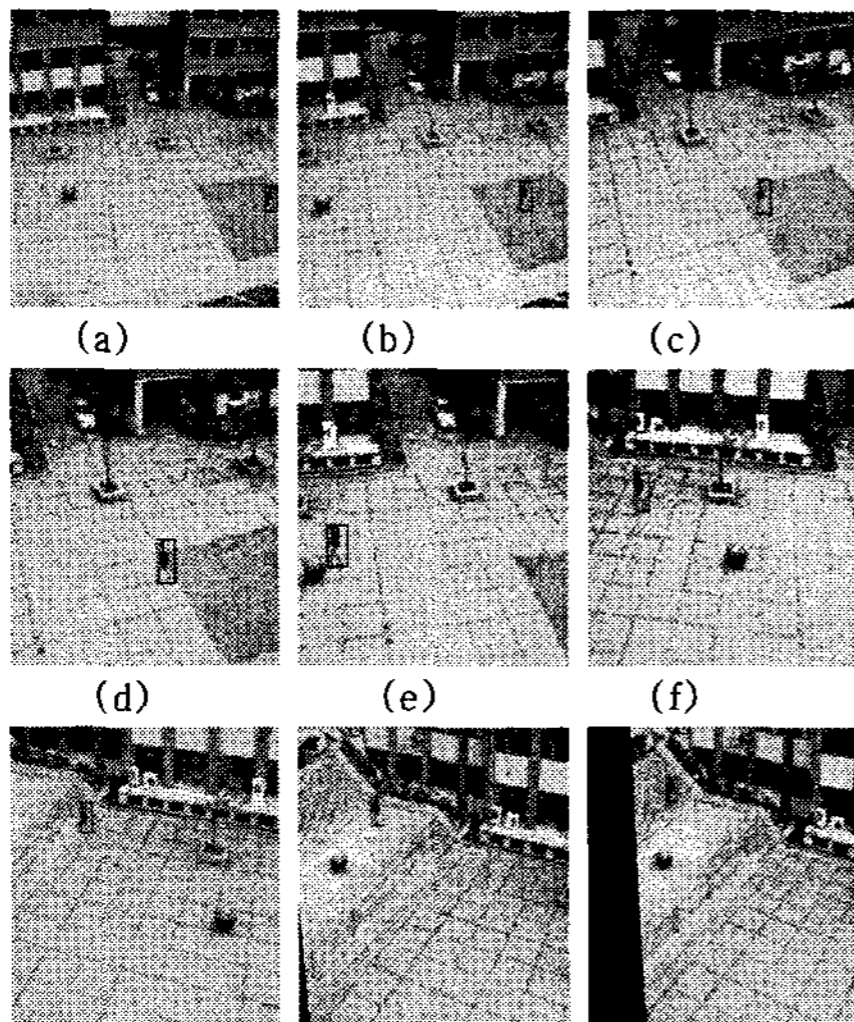


(a) (b) (c)



(d) (e) (f)
그림 5. 관심 보행자 추적 결과 영상

그림 6은 보행자의 중심이 화면내의 임의의 위치에 있지 않을 경우 카메라의 시계를 조정하고, 보행자의 인식률을 높이기 위하여 배율을 조정한 영상이다. (a)~(d)는 배율을 조정한 영상을, (e)~(i)는 보행자의 위치에 따라 카메라 시계를 조정한 것을 시간 흐름대로 표시한 것이다.



(a) (b) (c)
(d) (e) (f)
(g) (h) (i)
그림 6. 카메라 배율과 시계 조정 영상

3.2. 제안 방법의 성능 평가

관심 보행자 영역을 결정하는 부분에 대한 성능을 알아보기 위하여, 연속된 실시간 영상에서의 의도된 움직임을 보이는 다수 물체의 영역에서 관심 보행자 영역만을 검출할 수 있는지를 조사하였다. 표 1은 다수 물체 영역을 결정하는 부분에 대한 성능을 알아보기 위해 연속된 실시간 영상에서 움직이는 물체의 영역을 검출하는 시간과 접근 횟수를 조사하여 테스트 횟수에 따라 나타낸 것이며, 표 2는 연속된 영상에서 움직이는 관심 보행자의 영역을 검출한 결과에 대한 성능을 나타낸 표이다. 표 3은 검출된 관심 보행자에 대한 추적이 성공하는지에 대한 결과를 나타낸 것이다.

표 1. 다수 물체 검출 성능 비교

방법 횟수	제안한 CMODE 방법		라벨링 방법	
	시간(ms)	접근 횟수(회)	시간(ms)	접근 횟수(회)
1회	30	48,425	47	80,432
2회	16	20,044	47	82,032
3회	1	6,512	31	77,008
4회	1	5,923	31	77,264
5회	16	16,323	31	77,664
평균	13	19,445	37	78,880

표 2. 관심 보행자 영역 검출

횟수	1회	2회	3회	4회	5회
전체frame수	150	200	130	150	150
성공frame수	146	192	124	144	147
실패frame수	4	8	6	6	3
성공률[%]	97.3	96.0	95.4	96.0	98.0

표 3. 관심 보행자 추적

횟수	1회	2회	3회	4회	5회
전체frame수	146	192	124	143	147
성공frame수	141	186	115	135	139
실패frame수	5	6	9	8	8
성공률[%]	96.6	96.9	92.7	94.4	94.6

관심 보행자 영역을 검출하는 과정에서는 검출된 물체 영역이 임의의 임계치보다 작은 경우와 가로 길이가 세로 길이와 같거나 클 경우에는 보행자로 검출하지 않았으며, 보행자가 겹쳐 발생하는 경우 사람의 형태 정보와 상이하여 실패하는 경우가 발생하였다. 또한 보행자 검출이 성공적인 경우의 추적은 대부분 성공적인 모습을 보였지만 움직임 정보가 유사하고 hue값이 유사한 경우의 다른 보행자에 대해 관심 보행자로 판단하여 추적하는 모습을 보였으며, 너무 빠른 이동을 보인 경우에는 추적에 실패하는 결과를 보였다.

IV. 결론

본 논문에서는 Pan-Tilt-Zoom 카메라로부터 영상을 획득하여 실시간으로 움직이는 다수 물체의 영역을 검출하고, 사람의 구조적 특징과 형태 정보를 이용하여 관심 보행자를 검출한 후, 움직임 정보와 색상 정보를 이용하여 관심 보행자만을 추적하는 방법을 제안하였다.

먼저, 차영상을 수행하고 자료의 양을 줄이고 빠른 처리를 위하여 이진화하였다. 이진화된 영상에 대해 형상이나 크기는 유지시키는 형태학적 처리인 닫힘 연산으로 물체의 형태를 명확하게 하였다. 다수 물체 영역은 이진 변환 밀도 누적값을 계산한 후 CMODE방법을 이용하여 검출한다. 검출된 다수 물체 영역에서 사람의 구조적

정보와 형태 정보를 이용하여 나무들의 흔들림으로 인한 영역이나 차량의 움직임 영역은 제거되고, 관심 보행자 영역만을 검출하였다. 검출된 보행자 영역에 대하여 영역의 무게중심 간의 거리 차를 이용한 움직임 정보로 동일성을 1차적으로 판별하고, 1차 판별된 보행자 영역에 대해 k -평균 알고리즘으로 세 점을 구하고 각 점 3×3 영역의 평균 hue값을 이용하여 동일성을 2차 판별한 후 추적하였다. 보행자 인식률을 높이기 위하여 배율을 조정하여 일정 크기로 확대하여 추적하도록 하였으며, 보행자가 화면의 임의 범위에 있지 않을 경우 카메라 시계를 조정하여 연속적으로 추적하도록 하였다.

실험 결과, 제안한 CMODE 방법이 라벨링 방법보다 평균 접근 횟수가 1/4배정도 덜 접근하였으며, 평균 검출시간도 3배정도 빠르게 검출됨을 확인할 수 있었다. 관심 보행자 검출은 나무의 흔들림으로 인한 영역이나 차량의 움직임 영역, 그림자 영역과 같이 복잡한 배경에서도 평균 96.5%의 높은 검출률을 보였다. 관심 보행자 추적은 위치 정보와 색상 정보를 이용하여 평균 95%의 높은 추적률을 보였으며, 관심 보행자는 카메라 시계와 배율을 조정함으로써 연속적으로 추적할 수 있었다.

향후 연구 방향은 보행자가 겹쳐 발생하는 경우에도 관심 보행자 영역만을 추출하는 연구와 빠르게 움직이는 경우의 추적을 위한 연구가 추가되어야 할 것이다.

참 고 문 헌

- [1] 김휘용, "화상감시를 위한 이동물체 추출 및 동영상 부호화 기법", 한국과학기술원 박사학위논문, 2004.
- [2] 임종석, "이동카메라 영상에서 컬러 정보를 이용한 다수 보행자 검출 및 추적", 영남대학교 박사학위논문, 2004.
- [3] Hampapur, A, L. Brown, J. Connell, A. Ekin, N. Haas, M. Lu, H. Merkl, S. Pankanti, A. Senior, C.F. Shu, and Y.L. Tian, "Smart Video Surveillance", IEEE Signal Processing Magazine, March, 2005, pp. 38-51.
- [4] 조주현, "배경 모자이크 평면을 이용한 물체 검출 및 동영상 부호화", 한국과학기술원 박사학위논문, 2006.
- [5] Hussein. M, W.A. Almageed, Y .Ran, L. Davis, "Real-Time Human Detection, Tracking, and Verification in Uncontrolled Camera Motion Environments", In Proc. of the Fourth IEEE International Conference on Computer Vision Systems, January, 2006, pp. 41-47.
- [6] Sotelo. M.A, I. Parra, D. Fernandez, E. Naranjo, "Pedestrian Detection using SVM and Multi-feature Combination", In Proc. of the 2006 IEEE Intelligent Transportation Systems Conference, September, 2006, pp. 103-108.
- [7] Zhang. Y, S.J. Kiselewich, A. Bauson, and R. Hammoud, "Robust Moving Object Detection at Distance in the Visible Spectrum and Beyond using a Moving Camera", In Proc. the 2006 Conference on Computer Vision and Pattern Recognition Workshop, June, 2006, pp. 131-138.
- [8] Fuentes. L.M, and S.A. Velastin, "People Tracking in Surveillance Application", In Proc. 2nd IEEE International Workshop on PETS, December, 2001.
- [9] Haga. T, K. Sumi and Y. Yagi, "Human detection in Outdoor Scene using Spatio-temporal Motion Analysis", In Proc. of the 17th International Conference on Pattern Recognition, Vol. 4, August, 2004, pp. 331-334.
- [10] Mckenna. S.J, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler, "Tracking Groups of People", Computer Vision and Image Understanding, 2000, pp. 42-56.
- [11] Wren. C, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Real-time tracking of the human body", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, 1997, pp. 780-785.
- [12] Zhou. J, and J. Hoang, "Real Time Robust Human Detection and Tracking System", In Proc. the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 3, June, 2005, pp. 149-156.
- [13] Haritaoglu. I, D. Harwood, L.S. Davis, " W4: Real-Time Surveillance of People and Their Activities", In Proc. of IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, August, 2000, pp. 809-830.