

---

# 스마트폰 적용을 위한 휴먼 바디 추적 방법에 대한 연구

김범영\* · 최유진\* · 장성욱\* · 김윤상\*

\*한국기술교육대학교

## A Study on Human Body Tracking Method for Application of Smartphones

Beom-yeong Kim\* · Yu-jin Choi\* · Seong-wook Jang\* · Yoon-sang Kim\*

\*Korea University of Technology and Education (KOREATECH)

E-mail : yoonsang@koreatech.ac.kr

### 요 약

본 논문에서는 스마트폰 적용을 위한 휴먼 바디 추적 방법을 제안한다. 기존 휴먼 바디 추적 방법은 크게 센서 기반 방법과 비전 기반 방법으로 구분된다. 센서 기반 방법은 위치 정보의 누적 오차로 인해 추적 성능이 떨어지는 단점이 있다. 비전 기반 방법은 누적 오차가 없지만 스마트폰 적용을 위한 연산량 감소가 요구되고 있다. 본 논문에서는 스마트폰 적용을 위한 휴먼 바디 추적 방법으로 개선된 HOG 알고리즘을 이용한다. 개선된 HOG 알고리즘은 다운샘플링과 프레임 샘플링을 통해 구현된다. 다운샘플링에는 가우시안 피라미드가 적용되고, 프레임 샘플링에는 uniform sampling이 적용된다. 제안한 알고리즘을 2개 기기, 4개 해상도, 4개 프레임에서 측정하였고, 실시간으로 적용이 가능한 다운샘플링과 프레임 샘플링 파라미터 중에서 가장 검출률이 좋은 값을 도출하였다.

### ABSTRACT

In this paper we propose a human body tracking method for application of smartphones. The conventional human body tracking method is divided into a sensor-based method and a vision-based method. The sensor-based methods have a weakness in that tracking accuracy is low due to cumulative error of position information. The vision-based method has no cumulative error, but it requires reduction of the computational complexity for application of smartphone. In this paper we use the improved HOG algorithm as a human body tracking method for application of smartphone. The improved HOG algorithm is implemented through downsampling and frame sampling. Gaussian pyramid is applied for downsampling, and uniform sampling is applied for frame sampling. We measured the proposed algorithm on two devices, four resolutions, and four frame sampling intervals. We derive the best detection rate among downsampling and frame sampling parameters that can be applied in realtime.

### 키워드

Smartphone, Human Body Tracking, HOG, Augmented Reality

### 1. 서 론

최근, VR/AR 어플리케이션을 위한 스마트폰 기반의 휴먼 바디 추적 연구가 활발하게 진행되고 있다. 휴먼 바디 추적 방법은 크게 센서 기반 방법[1,2]과 비전 기반 방법[3,4]으로 구분된다. Tan은 센서 기반 방법의 휴먼 바디 추적을 통해 학습에 적용되는 mobile AR을 개발하였다[1]. Pryss는 센서 기반의 휴먼 바디 추적을 통해

mobile AR을 위한 프레임워크를 개발하였다[2]. Wanger는 mobile AR적용을 위해 비전 기반 방법의 휴먼 바디 추적에서 사용 가능한 새로운 마커 패턴을 제안하고, 기존 마커를 이용하여 구현한 mobile AR 결과와의 성능을 비교하였다[3]. Ziegler는 비마커 기반 mobile AR에 적합한 비전 기반 방법의 휴먼 바디 추적 기술들을 평가하였다[4].

센서 기반 방법은 GPS, 디지털 나침반, 가속

도 센서, 자이로 센서 등을 이용하여 사물의 위치와 움직임, 속도, 방향 등을 정밀하게 추적하지만, 위치정보의 누적오차로 인해 추적 성능이 떨어지는 단점이 있다[5-6]. 비전 기반 방법은 스마트폰 화면을 통해 보이는 실세계 대상의 3차원 위치를 2차원 위치로 변환하여 추적하고, 센서 기반 방법의 단점인 누적 오차가 없는 장점이 있으나[5-6], 스마트폰 적용을 위해서는 연산량의 감소가 요구되고 있다[7].

비전 기반 방법 중에서 휴먼 바디를 추적하는 대표적인 방법으로는 HOG(Histogram of Oriented Gradient)를 이용한 방법, Harr 웨이블릿(Wavelet)을 이용한 방법, LRF(Local Receptive Field)를 이용한 방법 등이 있으며, 이 중 가장 뛰어난 성능을 가지는 알고리즘은 HOG를 이용한 방법이다[8].

따라서, 본 논문에서는 스마트폰 적용을 위한 휴먼 바디 추적 방법으로 개선된 HOG 알고리즘을 제안한다. 제안된 휴먼 바디 추적 방법은 스마트폰 환경에서 요구되는 연산량의 감소를 위해, 다운샘플링과 프레임 샘플링을 HOG 알고리즘에 적용하여 개선하였다. 개선된 HOG 알고리즘은 다양한 스마트폰 환경(2개 기기/4개 해상도/4개 프레임)에서 측정된 성능(연산시간/검출률)을 통해 검토된다.

## II. 스마트폰 적용을 위한 휴먼 바디 추적 방법

### 2.1 관련연구 : HOG 알고리즘

HOG 알고리즘은 보행자 검출을 위해 일반적으로 널리 이용되고 있는 비전 기반 방법으로[9], 관심영역 검출을 위해 색상, 밝기, 방향 성분을 특징점을 이용한다[10]. HOG 알고리즘은 다양한 비전 기반 방법들 중에서 높은 추적 성능을 갖는 장점이 있다[8]. 따라서 HOG 알고리즘을 이용한 많은 휴먼 바디 추적 연구가 진행되고 있다[11-13]. Penny Chong은 HOG 알고리즘과 칼만 필터를 함께 적용함으로써 휴먼 바디 추적을 위한 연산시간을 감소시켰다[12]. Tudor Barbu는 HOG 기반 템플릿 매칭 과정을 이용한 강인한 휴먼 바디 추적 방식을 제안하였다[13]. Antonio Prioletti는 운전 보조를 위해 HOG 파트 기반 접근 방식을 이용하여 그림 1과 같이 휴먼 바디를 추적하였다[11].



그림 1. HOG 알고리즘을 이용한 운전 보조화면

### 2.2 제안된 휴먼 바디 추적 방법 : HOG 알고리즘 개선 방법

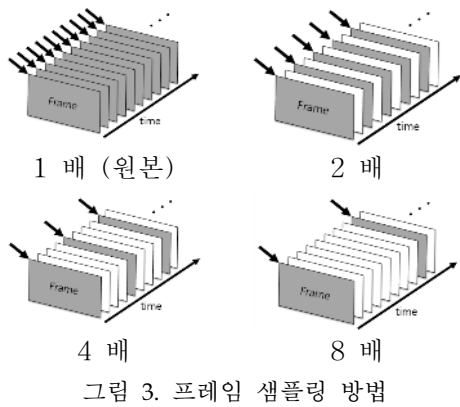
HOG 알고리즘을 이용한 방법은 연산량이 많은 단점이 있기 때문에[14] 동영상의 모든 프레임과 전체 영역을 처리하면 많은 연산량을 갖는 문제가 있다. 이러한 문제를 해결하기 위해, 제안된 휴먼 바디 추적 방법은 가우시안 피라미드를 이용한 다운샘플링으로 영상의 해상도를 줄이고, 프레임 샘플링을 하여 전체 프레임에 대해 연산을 수행하지 않고 일부 프레임만 연산하여 HOG 알고리즘의 연산량을 감소시킨다.

가우시안 피라미드는 다중 해상도 이미지를 나타내기 위한 구조이다. 원본 이미지에 대해 단계적으로 평활화(smoothing)와 다운샘플링을 수행하여 해상도와 크기가 점진적으로 감소하는 일련의 이미지를 얻는다. 가우시안 피라미드는 해상도가 감소된 각각의 가우시안 이미지들로 이루어져 있으며 원본 이미지의 너비가  $w$ , 높이가  $h$  라면  $i$ 번째 가우시안 이미지의 너비와 높이는 각각  $\frac{w}{2^i}$ ,  $\frac{h}{2^i}$ 로 감소한다[15]. 그림 2는 원본 이미지와 가우시안 피라미드를 이용하여 다운샘플링된 이미지를 보여 준다[16].



그림 2. 다운샘플링 방법

프레임 샘플링에는 그림 3과 같이 전체 동영상에서 균등하게 선택하는 uniform sampling과, 동영상의 시작 부분에서 연속된 프레임을 선택하는 continuous sampling이 있다[17]. 본 논문에서 제안하는 방법에는 전체 동영상에 대해 연산을 수행해야하기 때문에 uniform sampling을 적용하였다.



제안된 방법은 입력 이미지에 대해 프레임 샘플링을 수행한 후 복제 이미지를 생성한다. 복제된 이미지는 다운샘플링 과정을 거친 후 이진화되어 HOG연산에 사용된다. HOG연산의 결과로 나온 휴먼 바디의 위치는 입력 원본 이미지 위에 증강하여 표시된다. 그림 4는 제안된 휴먼 바디 추적의 흐름도를 보여 준다.

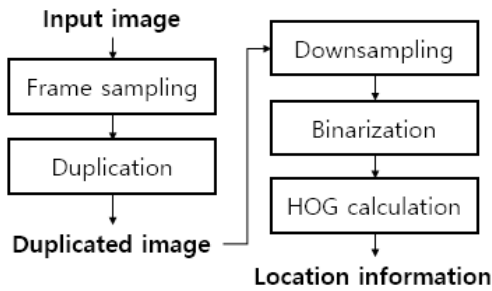


그림 4. 제안된 휴먼 바디 추적 방법 흐름도

### III. 실험 및 결과

제안된 휴먼 바디 추적 방법의 스마트폰 적용을 위해, Android 환경에서 OpenCV 3.2를 이용하여 개선된 HOG 알고리즘을 그림 5와 같이 구현하였다. 개선된 HOG 알고리즘은 다양한 스마트폰 환경에서 성능(연산시간/검출률)이 측정되었다. 다양한 스마트폰 환경으로 2개 기기, 4개 다운샘플링 파라미터, 4개 프레임 샘플링 파라미터가 고려되었다. 2개 기기는 저사양 스마트폰 (1.6GHz CPU, Galaxy S4)과 고사양 스마트폰 (2.1GHz CPU, Galaxy S6)이 사용되었다. 다운샘플링 파라미터는 1, 2, 4, 8이 있으며, 각각 가우시안 피라미드로 1, 2, 4, 8배 다운샘플링한 해상도를 의미한다. 프레임 샘플링 파라미터는 1, 2, 4, 8이 있으며, 각각 1, 2, 4, 8 프레임마다 uniform sampling을 적용한 것을 의미한다. 다운샘플링과 프레임 샘플링 파라미터를 하나씩 선택하여 총 16가지(4×4) 파라미터 순서쌍에 대해 실험

를 진행하였다. 입력 데이터로는 1920×1080 해상도의 FHD(Full High Definition) 동영상이 사용되었다.



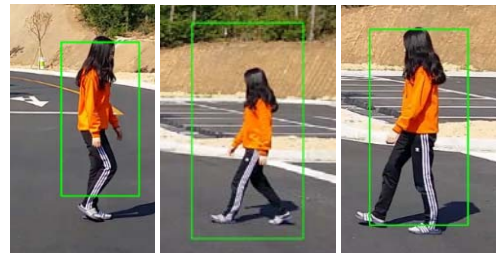
그림 5. 개선된 HOG 알고리즘 구현 결과

#### 3.1 평가 방법

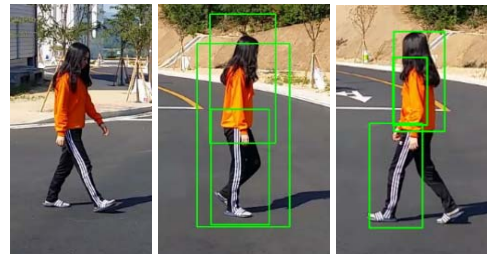
다운샘플링과 프레임 샘플링 파라미터에 따른 개선된 HOG 알고리즘의 성능(연산시간/검출률)을 측정하는 방법은 다음과 같다. 연산시간은 1개 프레임 이미지를 처리하는 시간의 평균값이며, 검출률은 식 3.1을 통해 측정한다.

$$recall\ rate = \frac{\text{추적이 성공한 프레임 수}}{\text{휴먼 바디가 존재하는 프레임 수}} \quad (3.1)$$

그림 6은 휴먼 바디 추적 성공 케이스, 그림 7은 휴먼 바디 추적 실패 케이스를 보여준다. 그림 7(a)와 같이 보행자가 인식되지 않거나, 그림 7(b), 7(c)와 같이 하나의 추적 대상에 대해 검출 영역이 중복되어 나타난 경우 추적 실패로 판단하였다.



(a) (b) (c)  
그림 6. 추적 성공 화면



(a) (b) (c)  
그림 7. 추적 실패 화면

3.2 평가 결과

저사양 스마트폰에서 개선된 HOG 알고리즘의 성능 평가 결과는 표 1, 표 2와 같다. 다운샘플링 파라미터 4개와 프레임 샘플링 파라미터 4개에 따른 개선된 HOG 알고리즘의 성능을 평가하기 위해, 총 16개의 측정 결과를 비교하였다.

저사양 스마트폰 환경에서의 성능 평가 결과, 프레임 샘플링 파라미터의 값이 커질수록 연산시간은 감소하였고, 검출률은 큰 차이가 나타나지 않았다. 다운샘플링 파라미터의 값이 커질수록 연산시간은 240×135 영상을 제외하고 평균 26%로 감소하였고, 검출률은 960×540영상과 480×270영상에서 각각 87%와 88%로 가장 높았다. 1920×1080 영상에서는 그림 6-(b)와 같이 하나의 추적 대상에 대해 검출 결과가 중복되어 나타나는 경우가 발생하여 상대적으로 검출률이 낮았고, 240×135영상에서는 정보의 손실이 심하여 그림 6-(a)와 같이 대상을 인식하지 못하여 검출률이 가장 낮게 나타났다.

고사양 스마트폰 환경에서의 성능 평가 결과는 표 3, 표 4와 같다. 저사양 스마트폰 환경에서의 성능 평가 결과, 프레임 샘플링 파라미터의 값이 커질수록 연산시간은 감소하였고, 검출률은 큰 차이가 나타나지 않았다. 다운샘플링 파라미터의 값이 커질수록 연산시간은 240×135 영상을 제외하고 평균 28%씩 감소하였다. 저사양 및 고사양 스마트폰 환경에서 성능을 비교한 결과, 검출률은 차이가 나타나지 않았으며 연산시간은 3배 차이가 나타났다.

평가 결과로부터, 스마트폰 환경에서 개선된 HOG 알고리즘을 실시간(15 fps[18])으로 적용하기 위한 파라미터를 도출하였다. 저사양 기기에서는 240×135 해상도에서 2 프레임마다 샘플링을 적용한 경우, 고사양 기기에서는 480×270 해상도에서 8 프레임마다 샘플링을 적용한 경우가 검출률이 가장 높았다.

표 1. 저사양 스마트폰에서 개선된 HOG 알고리즘의 연산시간 평가 결과

|                   |     | 다운샘플링     |         |         |         |
|-------------------|-----|-----------|---------|---------|---------|
|                   |     | 1920×1080 | 960×540 | 480×270 | 240×135 |
| 프레임 샘플링 (1/frame) | 1/1 | 12088     | 3275    | 902     | 97      |
|                   | 1/2 | 6851      | 1692    | 451     | 47      |
|                   | 1/4 | 3613      | 782     | 231     | 39      |
|                   | 1/8 | 1974      | 397     | 126     | 24      |

표 2. 저사양 스마트폰에서 개선된 HOG 알고리즘의 검출률 평가 결과

| 검출률(%)            |     | 다운샘플링     |         |         |         |
|-------------------|-----|-----------|---------|---------|---------|
|                   |     | 1920×1080 | 960×540 | 480×270 | 240×135 |
| 프레임 샘플링 (1/frame) | 1/1 | 62.81     | 88.23   | 87.89   | 31.49   |
|                   | 1/2 | 60.55     | 87.54   | 88.23   | 31.49   |
|                   | 1/4 | 61.40     | 87.54   | 87.54   | 30.80   |
|                   | 1/8 | 58.48     | 86.16   | 88.58   | 28.03   |

표 3. 고사양 스마트폰에서 개선된 HOG 알고리즘의 연산시간 평가 결과

| 연산시간(ms)          |     | 다운샘플링     |         |         |         |
|-------------------|-----|-----------|---------|---------|---------|
|                   |     | 1920×1080 | 960×540 | 480×270 | 240×135 |
| 프레임 샘플링 (1/frame) | 1/1 | 4199      | 988     | 309     | 54      |
|                   | 1/2 | 2122      | 515     | 158     | 24      |
|                   | 1/4 | 1124      | 255     | 86      | 16      |
|                   | 1/8 | 547       | 128     | 42      | 8       |

표 4. 고사양 스마트폰에서 개선된 HOG 알고리즘의 검출률 평가 결과

| 검출률(%)            |     | 다운샘플링     |         |         |         |
|-------------------|-----|-----------|---------|---------|---------|
|                   |     | 1920×1080 | 960×540 | 480×270 | 240×135 |
| 프레임 샘플링 (1/frame) | 1/1 | 62.81     | 88.23   | 87.89   | 31.49   |
|                   | 1/2 | 60.55     | 87.54   | 88.23   | 31.49   |
|                   | 1/4 | 61.40     | 87.54   | 87.54   | 30.80   |
|                   | 1/8 | 58.48     | 86.16   | 88.58   | 28.03   |

IV. 결 론

본 논문에서는 스마트폰 환경에서 휴먼 바디 추적을 위해 개선된 HOG 알고리즘을 제안하였다. 제안된 방법은 스마트폰 환경에서 요구되는 연산량의 감소를 위해, 다운샘플링과 프레임 샘플링을 이용하였다. 구현을 위해 Android 환경에서 OpenCV 3.2가 사용되었고, 입력 영상으로는 1920×1080 해상도의 FHD 영상이 사용되었다.

제안한 알고리즘을 2개 기기, 4개 해상도, 4개 프레임에서 측정하여 연산시간과 검출률에 대한 결과를 얻었다. 실험 결과, 알고리즘을 실시간으로 적용이 가능한 15fps 이상의 성능을 만족하는 다운샘플링과 프레임 샘플링 파라미터 중에서 가장 검출률이 좋은 값을 도출하였다.

도출한 파라미터를 통해 복잡한 과정 없이 연산량을 감소시킴으로써 스마트폰 환경에서 휴먼 바디 추적이 실시간으로 가능한 해상도와 프레임 샘플링 값을 결정할 수 있을 것이다.

## 사사표기

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2015년도 문화기술 연구개발 지원사업으로 수행되었음

## 참고문헌

- [1] Tan, Q., & Chang, W., "Location-Based Augmented Reality for Mobile Learning: Algorithm, System, and Implementation.", *Electronic Journal of e-Learning*, 13(2), pp. 138-148, Feb. 2015.
- [2] Pryss, R., Geiger, P., Schickler, M., Schobel, J., & Reichert, M., "Advanced Algorithms for Location-Based Smart Mobile Augmented Reality Applications." *Procedia Computer Science*, 94, pp. 97-104, 2016.
- [3] Wagner, D., Langlotz, T., & Schmalstieg, D., "Robust and unobtrusive marker tracking on mobile phones.", In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 121-124, Sep. 2008.
- [4] Ziegler, E., "Real-time markerless tracking of objects on mobile devices", Doctoral dissertation, B. Sc. Thesis, 2009.
- [5] Höllerer, T., Hallaway, D., Tinna, N., & Feiner, S., "Steps toward accommodating variable position tracking accuracy in a mobile augmented reality system.", In *2nd International Workshop on Artificial Intelligence in Mobile Systems (AIMS'01)*, pp. 31-37, Aug. 2001.
- [6] Papandrea, M., "A smartphone-based energy efficient and intelligent multi-technology system for localization and movement prediction.", *2012 IEEE International Conference on*, pp. 554-555, Mar. 2012.
- [7] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H., "Mobilenets: Efficient convolutional neural networks for mobile vision applications.", *arXiv preprint arXiv:1704.04861*, Apr. 2017.
- [8] Pyeonggang Kim, "Design and Evaluation of Improved Pedestrian Recognition Algorithm Using HOG-UDP", Master degree Gachon University, Feb. 2015.
- [9] Dalal, N., & Triggs, B., "Histograms of oriented gradients for human detection.", In *Computer Vision and Pattern Recognition 2005, CVPR 2005, Vol. 1*, pp. 886-893, Jun. 2005.
- [10] Seo, C. J., & Ji, H. I., "Pedestrian Detection using HOG Feature and Multi-Frame Operation.", *The Transactions of the Korean Institute of Electrical Engineers P*, 64(3), pp. 193-198, Sep. 2015.
- [11] Prioletti, A., Møgelmoose, A., Grisleri, P., Trivedi, M. M., Broggi, A., & Moeslund, T. B., "Part-based pedestrian detection and feature-based tracking for driver assistance: real-time, robust algorithms, and evaluation.", *IEEE Transactions on Intelligent Transportation Systems*, 14(3), pp. 1346-1359, May. 2013.
- [12] Chong, P., & Tay, Y. H., "A novel pedestrian detection and tracking with boosted HOG classifiers and Kalman filter.", In *Research and Development (SCOREd), 2016 IEEE Student Conference on*, pp. 1-5, Dec. 2016.
- [13] Barbu, T., "Pedestrian detection and tracking using temporal differencing and HOG features.", *Computers & Electrical Engineering*, 40(4), pp. 1072-1079, May. 2014.
- [14] Nabila, M., Yousra, B. J., & Eric, W., "Optimized HOG Descriptor for on Road Cars Detection.", In *Proceedings of the 10th International Conference on Distributed Smart Camera*, pp. 166-171, Sep. 2016.
- [15] Yin, J., & Cheng, G., "Splicing and display of massive rock thin section images.", In *Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2016 IEEE*, pp. 1882-1886, Oct. 2016.
- [16] Burt, P., & Adelson, E., "The Laplacian pyramid as a compact image code.", *IEEE Transactions on communications*, 31(4), pp. 532-540, Apr. 1983.
- [17] Jiang, Y. G., Dai, Q., Mei, T., Rui, Y., & Chang, S. F., "Super fast event recognition in internet videos.", *IEEE Transactions on Multimedia*, 17(8), pp. 1174-1186, Aug. 2015.
- [18] Sang-min Seo, Jonhno Kwon, & Yoon-Ho Choi, "Development of Android App for Supporting Smooth Multimedia Streaming Service Using Frame Buffer.", *Journal of Internet Computing and Services(JICS)*, 17(1), pp. 55-64, Feb. 2016.