

# Hough 변환을 이용한 비디오 영상내 사람의 계수

## People Counting Using Hough Transform in Video Images

서영교\*, 도용태\*\*

\* 대구대학교 대학원 정보통신공학과(전화:(053)850-4425, 팩스:(053)850-4425, E-mail : nerinbi@biho.daegu.ac.kr)

\*\*대구대학교 정보통신공학부(전화:(053)850-6625, 팩스:(053)850-6619, E-mail :ytdo@biho.daegu.ac.kr)

Abstract : Monitoring and analyzing people in video images are important and attract attention from many researchers in computer vision field. In this paper, we propose a new technique by which people overlapped one another in an image can be extracted and counted. After extracting moving people from video images as foreground, their heads are searched by Hough Transform inside the foreground. Since heads are comparatively stable in spite of motion in video images, semicircles along the head tops can be important features to be found in an edge map. In our experiment, the technique successfully separated people who existed at the same vertical positions, which was impossible by existing techniques. Meanwhile, it showed high dependency on edge information and false results were obtained when the edges were rather incomplete.

Keywords : Video images, VSAM, Hough Transform

### I. 서론

컴퓨터 및 기계시각의 발전에 따라, 그 활용역역이 급속히 확대되는 추세에 있으며, 비디오에 의한 감시 및 관측(Video Surveillance and Monitoring; VSAM)은 근래에 와서 많은 연구자들의 관심을 끌고 있는 분야가 되었다[1]. 차량과 사람은 VSAM에서 가장 중요한 두 가지 대상이며, 특히 사람의 경우 크기가 작고 동작이 다양하여 그 해석이 어려우므로 연구의 필요성이 증대되고 있다.

본 논문에서는 비디오 영상내 사람의 계수(counting)를 위한 방법에 대해 서술한다. 사람을 추적하고 감시하기 위해서는 사람을 개개인으로 분리하는 작업이 기본이 되며, 이는 계수작업과 직접적으로 연관되어 있다. 특히 VSAM관련 연구 중에서 계수에 관한 연구는 미미하므로, 그 연구의 의의가 크다. 기존의 연구결과로는 Quasi-Topological Codes(QTC)를 이용하는 기법[2]과 convex-hull/corner 탐색을 수행하고 수직 투영에서 국부 최소치를 찾아 분리하는 Hydra[3], Draping에 의한 우리의 선행연구[4]가 있다. 이들 기법들은 모두 대상이 되는 사람이나 사람군의 윤곽을 중심으로 작업하므로, 한 장의 특정 영상 내에서 겹쳐진 사람들에 대한 해법은 제시하지 못하는 문제가 있다.

본 논문에서 특히 중점적으로 다루고 있는 것은 두 명 이상이 대략적으로 동일한 수직선상에 위치하여 중첩됨으로서 기존의 기법으로는 확인이 불가능한 상황이다. 이때 각 개인을 분리 가능할 수 있다면, 이후 계

수나 추적 및 감시등의 작업이 효율적으로 이루어질 수 있다. 우리는 기존의 기법이 유용하지 못하는 상황에서 Hough 변환을 적용하여 사람의 머리에 해당하는 원호(circular arc)를 찾음으로서, 사람을 적절히 계수하는 기법을 제안한다.

논문의 구성은 우선 II장에서 비디오 영상으로부터 사람 부분을 전경(foreground)영역으로 검출하기 위한 전처리 단계를 설명한다. 이때 기존의 기법 중 Draping을 이용하여 부분적으로 겹쳐진 사람을 분리하는 방법에 대해서도 간단히 서술한다. III장에서는 Hough 변환을 이용해 수직방향으로 겹쳐진 사람을 분리하고 계수하는 기법을 제안한다. 이어 IV장에서는 실제 실험의 결과에 대해 설명하고, 마지막으로 V장에서 결론을 내린다.

### II. 비디오 영상에서 사람의 검출

연속적으로 입력되는 비디오 영상을 분석하면 관측 영역 내에서 움직이는 대상과 정지된 대상을 쉽게 구분할 수 있다. 동적인 대상 중 특히 관심을 갖게 되는 것은 차량과 사람이며, 본 논문에서는 동작의 유형이 불균일하고 크기가 작아 해석이 어려운 사람에 대해서만 다룬다. 비디오 영상에서 움직이는 사람을 전경영역으로 하여, 정적인 배경(background) 영역과 구분하기 위해 우리는 그림 1과 같은 절차를 사용하였다.

먼저 카메라로부터 직접, 혹은 녹화된 파일로부터 영상을 입력 받고, 적응 배경 차감법(Adaptive Backg-

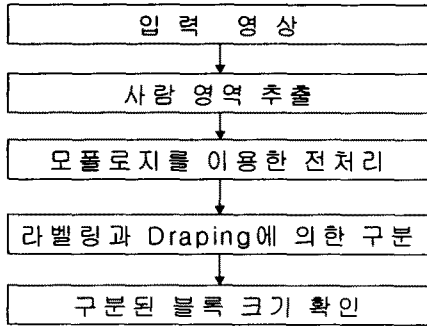


그림 1 비디오 영상에서 사람을 추출하기 위한 절차

Fig 1. People extraction steps in video images.

round Subtraction; ABS)을 이용하여 동적 부분을 분리한다. ABS의 기법을 사용한 예로는 Haritaoglu[5]가 W4에서 사용한 방법과 Fujiyoshi[6]가 VSAM을 위해 사용한 방법 등이 있다. W4에서는 다수의 배경 영상을 이용하여 화소 당 밝기 값의 최소와 최대 값을 저장한 후, 이 범위를 벗어나는 경우 전경으로 판정하는 기법을 사용하였다. Haritaoglu는 동적 물체가 없을 때 배경영상을 갱신하였는데, 영상에 동적 부분이 계속 나타날 경우 배경영상을 갱신할 수 없어 구름이나 태양의 조도 변화 등 환경의 변화에 연속적으로 적용할 수 없는 단점이 있다. 한편 Fujiyoshi는 실시간으로 배경영상을 갱신하면서, 전경을 분리하는 모델을 사용하였다. 그 모델은 아래 식(1)과 같다.

$$\frac{\overline{p}_{n+1}}{\sigma_{n+1}} = \frac{\alpha \overline{p}_{n+1} + (1-\alpha) \overline{p}_n}{\alpha \overline{p}_{n+1} - \overline{p}_{n+1} + (1-\alpha) \sigma_n} \quad (1)$$

여기서  $\overline{p}_n$ 은 n번째 입력 영상의 특정 화소 값이고,  $\overline{p}_n$ 은 그 특정 화소에 해당하는 배경 영상의 화소 값이며,  $\sigma_n$ 은 이의 표준 편차이다. 만약  $\overline{p}_n$ 이  $\overline{p}_n$ 에서  $2\sigma_n$ 보다 더 크거나 작다면, 배경영역이 통상 확률적으로 변화할 수 있는 양 이상 변화했다고 간주할 수 있으므로 동적인 개체가 그 부분에 존재한다고 가정할 수 있는 것이다. 이 방식은 W4보다 계산량이 많아진다는 단점을 가지고 있지만, 실시간으로 배경을 갱신할 수 있다는 장점을 지닌다. 본 논문에서는 이 방식을 사용하여 동적인 부분을 추출하였다.

일단 동적부분을 추출한 후, 3x3 화소 크기의 닫힘(closing)과 3x5 크기의 열림(opening) 모폴로지(morphology) 연산을 수행하여 작은 잡음과 구멍을 제거 하고, 라벨링(labeling)을 수행하여 영역을 구분하였다.

모폴로지와 라벨링을 통하여 얻어진 전경 영역에서 각 구분된 영역의 크기를 구하고, 일정 크기 이상인 영역에 대해서는 그 영역을 둘러싸는 사각형을 얻었다. 그리고 그 사각형 내에서 Draping 기법에 의한 분할을 행하였다. 이는 사각형 내에 부분적으로 겹쳐진 사람이 존재한다면 이를 분리하고 계수하기 위한 것이다.

Draping 기법에 대해 좀 더 상세히 서술하면, 그림 2(a)와 같은 영상에서 그림 (b)와 같이 동적인 대상을 전경 영역으로 추출하고, 그림 (c)에서와 같이 동적부분의 최상위 화소들로 draping 선을 구한다. 그리고 그림 (d)에서 보인 바와 같이 draping 선의 국부 최소에서 전경 영역을 수직으로 분리한다. 일단 분리된 영역들은 각각 그 높이와 크기가 일정 임계치 이하이면 무시함으로써, 그림자와 같이 관심의 대상이 되지 않는 부분은 버렸다. 그림에서 볼 수 있듯이 3명의 부분적으로 겹쳐진 사람들이 성공적으로 분리되었다.

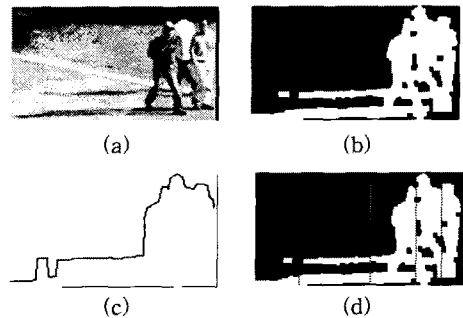


그림 2 Draping에 의한 사람의 구분: (a) 입력영상, (b) 전경영역의 추출, (c) Draping 선, (d) 수직 분할

Fig. 2. Dividing people by draping: (a) Input image, (b) Extracting foreground, (c) Line by draping, (d) Vertical division

### III. Hough 변환을 이용한 겹쳐진 사람의 분리

앞 절에서 서술한 바와 같이 비디오 영상에서 동적인 부분을 추출하는 방식은 잘 정립되어 있고, Draping은 부분적으로 겹쳐져 있는 사람군을 각각의 개인으로 나누는 간단한 방법이 될 수 있다. 그러나 사람이 수직으로 완전히 겹쳐져 있는 경우에는 Draping 선상에 국부 최소가 존재하지 않으므로, 분리할 수 없게 된다. 그림 3은 그러한 예를 보여 준다.

본 논문에서는 이 문제에 대한 한 가지 해결책으로 원호 Hough 변환(Circular Arc Hough Transform; CAHT)을 사용하여 사람의 머리를 찾고 계수하는 기법을 제안한다. CAHT가 유용할 수 있는 근거는 우선 사람의 머리부분이 다리나 팔등과 비교하여 이동 중 가장 안정된 변화를 보이는 부분이어서 관측하기 쉽

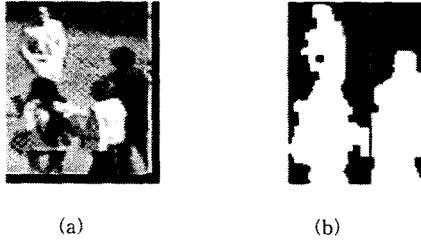


그림 3. 수직으로 겹쳐진 사람의 예: (a) 원 영상, (b) Draping에 기초한 분할의 결과  
 Fig. 3. Example of vertically overlapped people: (a) Original image, (b) Division result by draping

고, 특히 머리의 상단부 원호부분이 머리카락의 모양 등에 상관없이 반원의 형태인 경우가 많기 때문이다.

이산적인 화소로 구성된 영상에서 반원은 그 반경에 따른 독특한 특성을 가진다. 이러한 특성을 미리 알고 있으면 반원을 찾는 시간과 노력을 줄일 수 있다. 표 1은 몇 가지 반경에 따른 반원상의 화소수와 연결된 화소간의 각도의 진행을 보여준다.

반원을 찾기 위해 먼저 경계의 검출이 행하여 졌다. 경계는 최대한 가늘면서도 연결성이 좋도록 찾겨져야 하며, 우리는 Canny의 경계 검출법을 사용하였다. 각 경계점에서 Gerig의 기법[7]을 기반으로 그 경계점을 지나는 가장 가능성이 높은 반원을 찾았는데, 구체적인 절차는 다음과 같다:

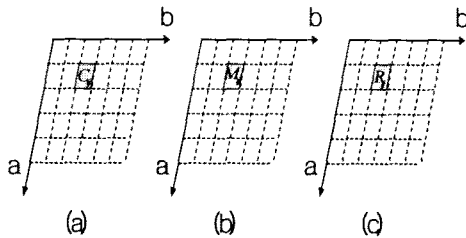


그림 4. 반원을 찾기 위해 필요한 3개의 배열: (a) 화소를 중심으로 하는 반원의 개수, (b) 화소를 중심으로 하는 반원의 최대 개수, (c) 최대의 반원 개수를 가질 때의 반원의 반경

Fig. 4. Three arrays needed for finding semicircles: (a) Number of semicircles centered at the pixel, (b) Maximum number of semicircles centered at the pixel, (c) Radius when the maximum number of semicircles are found

반경	반원상의 기준 화소수	반원상 화소간의 기준 각도 [degree]
5	14	12
6	16	11
7	18	10

표 1 이산 화소로 구성된 영상에서 반원들의 특성  
 Table 1. Characteristics of semicircles on an image consists of discrete pixels

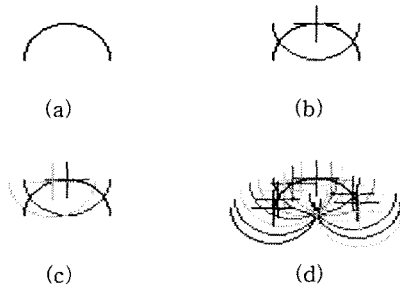


그림 5. CAHT에 의해 구하는 반원: (a) 경계 영상, (b) 한 경계점에서 그려진 원호, (c) 다른 경계점에서 그려진 원호의 중첩, (d) 다수의 경계점으로부터 결정되는 반원의 중심

Fig 5. Semicircle detection by CAHT: (a) Edge image, (b) Arc drawn from an edge point, (c) Overlap of another arc drawn from another edge point, (d) Center of semicircle determined from multiple edge points

(i) 3개의 2차원 배열을 그림 4와 같이 구성한다. 여기서 배열내 변수  $C_y$ 와  $M_y$ 의 값들은 모두 0으로 초기화 한다.

(ii) 특정 반경  $r$ 을 고려하고,  $r_y$ 를 이 값으로 초기화 한다.

(iii) 영상내 경계점에서,  $r$ 을 반경으로 가지는 반원을  $180^\circ \sim 360^\circ$ 각도 사이에 그린다. 그림 5는 그러한 과정을 보여주는데, 특히 (b)그림은 '+'로 표시된 한 경계점에서 그려진 아래쪽 반원을 보여준다.

(iv) 그려진 반원상에 있는 화소들에 대해서 해당  $C_y$ 의 값을 1증가 시킨다. 이러한 절차를 모든 경계점들에 대해 수행하면 그림 5의 (c)와 (d)같은 절차가 수행되며, 이 때  $C_y$ 에는 그 화소를 원점으로 하고 반경을  $r_y$ 로 하는 반원상의 경계점 수가 저장된다.

(v)  $C_y$ 와  $M_y$ 를 비교하여, 만약  $C_y$ 가 더 크면  $M_y$ 의 값은  $C_y$ 로 대체되고,  $r_y$ 도 현재의  $r$ 값으로 대체된다. 즉, (i,j)위치의 화소를 중심으로  $r_y$ 반경의 반원이 가장 많은 경계점을  $M_y$ 개 만큼 그 반경 상에 가진다. 다시

말하자면, 그 화소는 반경  $r_y$ 의 반원의 중심일 가능성이 가장 크다.

(vi) 만약 탐색하고자 하는 가장 큰 반원의 반경보다 반경  $r$ 이 아직 작거나 같으면,  $r$ 을 1만큼 증가 시키고 절차 (iii)으로 간다. 아니면 다음 단계로 간다.

(vii) 만약  $M_y > k \times (r_y \text{ 당 기준 경계점 수})$ 이면, (i,j)를 중심으로  $r_y$ 를 반경으로 하는 반원이 실재한다고 판정한다. 아니면 그 위치를 중심으로 하는 반원은 존재하지 않는다고 결론 내린다. 이 때 기준 경계점 수는 표 1에서 보인 바와 같고, 계수  $k$ 는 실험을 통해 0.65 정도가 최적으로 얻어졌다.

#### IV. 실험

실험은 Pentium-III 550MHz PC에서 VC++6.0을 사용하여 프로그램 되었다. 본 논문의 연구가 사람에 한정되어 있으므로, 사람들만 다니는 지점을 대상으로 카메라가 설치되었다. Meteor-II 프레임그래버로 실제 영상을 획득하여 이미지 파일로 저장하고, 그 이미지들 이용하여 분석 작업을 행하였다.

얻어진 영상에서 우선 II장에서 설명한 방식으로 동적인 영역을 전경영역으로 검출하였다. 전경영역에 draping이 행하여지고 수직 분할을 행하였다. 일단 분할된 영역에 대해 본 논문에서 제안된 방식과 같이 CAHT가 행하여 졌다. 그림 6은 결과의 예를 보여준다. 각 그림에서 좌측은 입력 영상 중 draping에 의해 수직 분할된 영역을 보여주고, 중앙은 경계검출을 한 결과이며, 우측은 CAHT에 의해 얻어진 머리의 상반원을 보여준다. 그림 (a),(b),(c)의 경우 겹쳐진 사람들을 적절히 찾고 있다. 그러나 (d)의 경우 가장 위쪽에 있는 사람을 찾지 못하고 있고, (e)의 경우 사람의 허리 부분에서 잘못 찾은 결과를 보여준다.

#### V. 결론

비디오 영상으로부터 차량이나 사람을 추적하고 그 행위를 분석하는 VSAM은 근래에 와서 컴퓨터 시각분야에서 많은 사람들의 관심을 끌고 있다. 본 논문에서는 특히 사람들 중심으로 기존의 연구들에서 해결책을 제시하지 못했던 수직방향으로 중첩된 사람의 분리와 이를 통한 계수 작업에 대하여 연구한 결과를 서술하였다. 인간의 신체부위 중 비디오 영상에서 비교적 안정적으로 모양을 유지하는 머리의 상측 반원을 Hough 변환을 이용하여 찾음으로서, 겹쳐진 사람들도 구분해 낼 수 있게 되었다. 특히 머리 반원의 유형을 분석하여 이용함으로써 변환의 시간을 상당히 줄일 수 있었다. 그러나 실제 실험에서 제시된 방법은 경계검출의 결과에 상당히 의존적이며, 이에 따라 오추출이나 미추출하는 경우가 발생하였는데, 이는 향후 보완 연구의 방향이 될 수 있을 것이다. 또 머리의 반원 외 다른 정보를 융합함으로써 보다 신뢰할 수 있는 결과에 도달할 수 있을 것으로 사료된다.

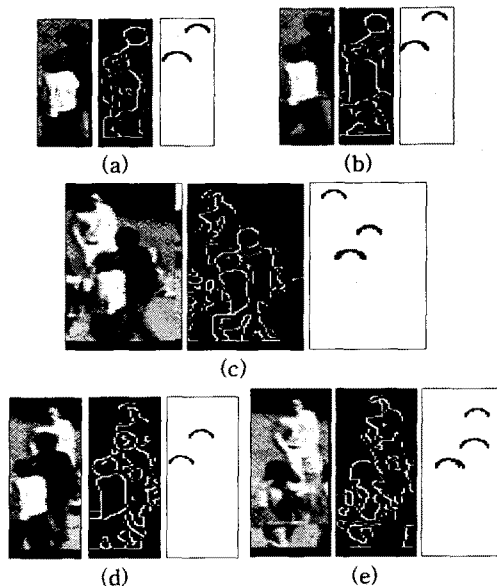


그림 6. CAHT에 의한 영상내 겹쳐진 사람의 검출예  
Fig. 6. Examples of detecting overlapped people on images by CAHT

#### 참고 문헌

- [1] T. Kanade et al., "Advances in Cooperative Multi-Sensor Video Surveillance", Proc. DARPA Image Understanding Workshop, pp. 3-24, 1998.
- [2] J. Heikkila, O. Silven, "A Real-Time System for Monitoring of Cyclists and Pedestrians", Proc. Second IEEE Workshop on Visual Surveillance, pp. 74-81, 1999.
- [3] I. Haritaoglu, D. Harwood, L. Davis, "Hydra: Multiple People Detection and Tracking Using Silhouettes", Proc. Second IEEE Workshop on Visual Surveillance, pp. 6-13, 1999.
- [4] Y. Do, T. Kanade, "Counting People from Image Sequences", Proc. CISST, pp. 185-190, 2000.
- [5] I. Haritaoglu, D. Harwood, L. Davis, "W4: Who? When? Where? What? A Real Time System for Detecting and Tracking People", Proc. IEEE International Conference on Automatic Face and Gesture Recognition, pp. 222-227, 1998
- [6] H. Fujiyoshi, A. Lipton, "Real-time Human Motion Analysis by Image Skeletonization", Proc. Fourth IEEE Workshop on Applications of Computer Vision, pp. 15-21, 1998.
- [7] G. Gerig, F. Klein, "Fast Contour Identification through Efficient Hough Transform and Simplified Interpretation Strategy", Proc. ICPR, pp. 498-500, 1986.