

점유센서를 위한 유사성 메트릭 기반 입출입 사람 매칭

정재준 김만배

강원대학교 컴퓨터정보통신공학과

jjj602@kangwon.ac.kr manbae@kangwon.ac.kr

Incoming and Outgoing Human Matching Using Similarity Metrics for Occupancy Sensor

Jaejune Jung and Manbae Kim

Computer & Communications Engineering, Kangwon National University

요약

기존의 사람간의 유사성 측정 시스템은 적외선 빔이나 열 감지 영상 장치를 통해 측정하였다. 하지만 이와 같은 방법으로 측정하면 2명 이상의 객체를 분류해내는 기술은 제공하지 않는다. 이에 본 논문은 고정된 카메라를 이용하여 각 사람의 피부색과 옷차림 등의 RGB 정보를 이용한 사람 유사성 측정 기법을 제안한다. RGB카메라 영상을 통하여 객체의 RGB 히스토그램을 얻은 후 각 객체에 대해 Bhattacharyya metric, Cosine similarity, Jensen difference, Euclidean distance로 histogram similarity를 계산하여 객체 추적 및 유사성 측정을 통해 객체를 분류한다. 제안된 시스템은 C/C++를 기반으로 구현하여, 유사성 측정 성능을 평가하였다.

1. 서론

오늘날 CCTV와 같은 감시 시스템의 사용량이 크게 늘어났고 이에 따른 영상의 수도 늘어나게 되었다. 영상의 증가에 따라 영상을 자동으로 분석해주는 기술 또한 필요하게 되었다. 주요 분야는 사람 분류(human classification) 기술로, 실내의 공간에 입출입하는 사람들을 분류하는 것이다. 주요 응용분야는 비전 기반 점유센서(occupancy sensor)로 실내에 사람이 한명이라도 있으면 점등(light on)하고, 반대로 사람이 존재하지 않으면 소등(light off)하게 된다.

주로 사용되는 센서로는 PIR(pyroelectric infrared) 센서가 가장 많이 활용되어 왔고, 이외에도 ultrasonic 센서, microwave 센서, magnetic 센서, radio-frequency 센서, acoustic 센서 등도 개발되어 왔다. PIR은 사람의 신체에서 발생하는 열을 감지하고, 열의 변화량을 계산하여 움직임 측정한다.

최근에는 비전 센서를 이용하여 PIR의 한계를 넘어 보다 지능화된 감시 시스템을 구축하려는 추세이다. 본 논문은 고정된 RGB 카메라를 이용하여 입실, 퇴실을 하는 사람을 분류하는 점유센서를 활용하는 사람 계수 측정 기법을 제안한다. 사람이 가지고 있는 피부색, 옷의 색을 저장하여, 한 사람과 여러 명이 입, 퇴실할 때 각각 RGB성분을 저장하고, 유사성을 비교하여, 어떠한 사람이 입, 퇴실 했는지를 분류하는 방법을 제안한다.

2. 제안 방법

먼저, 현재영상과 배경영상의 차분화를 이용해 움직이는 객체를 검출한다. 이러한 기법을 이용하기 위해서는 첫째로 background model을 만들어야 한다. background 버퍼의 초기 값은 입력 영상의 grayscale 버퍼의 값이며, t번째 grayscale 버퍼를 Y_t 라고 하고, background 버퍼를 B_t 라고 가정할 때 B_t 를 구하는 연산은 다음과 같다.

$$B_t = \begin{cases} B_t + 1 & \text{if } Y_t > B_t \\ B_t - 1 & \text{otherwise} \end{cases} \quad (1)$$

객체를 찾기 위해서 background 영상을 차분하여 foreground mask를 얻는다. t번째 foreground mask를 F_t 라고 할 때 F_t 를 구하는 연산은 다음과 같다.

$$F_t = \begin{cases} 255 & \text{if } |B_t - Y_t| > T \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

이때, 임계값 T로 12를 사용했다.

목표 객체인 사람의 RGB 히스토그램을 얻기 위해서 foreground 버퍼에서 객체의 bounding box를 구한다. 먼저 영상의 가로크기 만큼 배열을 생성하고 foreground 버퍼에서 255값이 나오는 횟수의 누적 합을 배열에 저장한다. 그 배열에서 가장 큰 값에 해당하는 index를 기준으로 index를 감소, 증가 시켜 배열의 값이 0이 나오는 부분이 bounding box의 left, right line이 된다. 즉 객체의 중심을 기준으로 bounding box의 좌, 우 boundary를 구한다. 같은 방식으로 영상의 세로크기만큼의 배열을 이용하여 top, bottom line이 구해진다.

왼쪽에서 오른쪽으로 들어오는 incoming 영상과 오른쪽에서 왼쪽으로 나가는 outgoing 영상에서 bounding box내의 히스토그램을 통해 유사성을 구한다. RGB 히스토그램을 구할 때 픽셀의 RGB값을 각각 r, g, b 라고 한다면 r, g, b를 $16 \times 16 \times 16$ 으로 양자화 한 후, bin = 4096 사이즈의 히스토그램 p 배열을 저장한다. 그리고 각 픽셀에 대해 배열의 $[(r \times (16 \times 16)) + g \times 16 + b]$ index에 1을 더한다. 이후 bounding box내의 픽셀 개수만큼 나누어주면 bounding box에 대한 RGB 히스토그램 p가 완성된다. 움직임이 있는 프레임마다 신뢰도 향상을 위해 히스토그램을 업데이트 해주어야 한다. 히스토그램의 각 index에 대해 업데이트 연산은 다음과 같다.

$$p_i = 0.8 \times p_i + 0.2 \times q_i \quad (3)$$

여기서 q는 이전 프레임의 히스토그램이다.

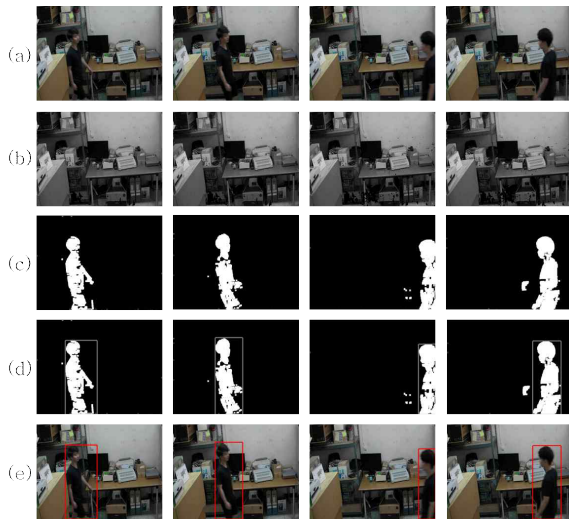


그림 1. 2장의 연속 incoming, outgoing 프레임 (a) 입력영상, (b) background 영상, (c) foreground mask, (d) foreground mask 영상의 bounding box, (e) 입력영상의 bounding box

동일한 조건에서 2명이 동시에 등장하는 incoming, outgoing 영상으로 위와 같은 작업을 수행한다. bounding box를 구하는 과정에서만 이전 방법과 달리 foreground 영상에서 목표객체의 좌표를 직접 구해서 bounding box를 얻었다.

4명의 incoming, outgoing 히스토그램을 나누고 incoming, outgoing 간의 유사성을 4개의 metric을 통해서 가장 성능이 좋은 metric을 찾는다. metric은 Bhattacharyya metric(BM), Cosine similarity(CS), Jensen Difference(JD), Euclidean distance(ED) 총 4가지 metric을 이용했다. 첫 번째로 Bhattacharyya metric을 이용하는데 연산은 식 (4)와 같다. p는 incoming histogram, q는 outgoing histogram 이다.

$$d(y) = \sqrt{1 - \rho[\hat{p}(y), \hat{q}]}$$

$$\hat{\rho}(y) \equiv \rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u}$$
 (4)

cosine similarity의 연산은 식(5)과 같다.

$$d = \frac{\sum_{i=1}^{bin} (p_i q_i)}{\sqrt{\sum_{i=1}^{bin} p_i^2} \sqrt{\sum_{i=1}^{bin} q_i^2}}$$
 (5)

Jensen Difference의 연산은 식 (6)과 같다. H는 entropy이다.

$$d = H\left(\frac{p+q}{2}\right) - \frac{H(p) + H(q)}{2}$$

$$H = \sum_{i=1}^{bin} p_i \log_2 p_i$$
 (6)

Euclidean distance의 연산은 식 (7)과 같다.

$$d = \sqrt{\sum_{i=1}^{bin} (p_i - q_i)^2}$$
 (7)

3. 실험 결과

제안 방법의 성능을 평가하기 위해서, 연구실에서 고정된 카메라를 통해 입, 퇴실 하는 4명의 사람들을 촬영했다. 영상의 해상도는 640x480 이고 RGB입력이다. 실험의 입력영상은 모든 샘플 1,300장중에 객체가 등장하는 incoming, outgoing 영상 각각 100장씩 사용했다.

표 1은 한 비디오 샘플에서 수행되는 한 사람과 두 사람의 모든 경우의 Similarity를 비교한 값을 나타낸다. A, B, C, D는 사람이고, 좌측은 incoming이고, 표 상단은 1명 outgoing, 2명 outgoing을 보여준다. A+B는 사람 A, B가 동시에 퇴실하는 경우이다.

표1에서 보듯이. 단일 목표객체들 사이의 similarity 대해서는 모든 metric에 대해서 100% 매칭된다. 2명 outgoing에서는 가장 우수한 score를 가지는 2개는 청색으로 표시된다. 적색은 이 2개에 포함 안 된 사람이다. Bhattacharyya metric(BM)을 보면 outgoing C+B에서 incoming C와 B를 포함하여 정확히 2명의 사람이 누구인지를 판별한다. A+D, A+D, A+B도 정확히 판별된다. 그러나 C+A에서는 C를 검출하지 않아, 오류가 발생한다, C+A에서는 C가, B+D에서는 B가 검출이 안 된다. 전체적으로 두 명이 퇴실해도 이 두 명의 컬러 히스토그램은 각 사람의 히스토그램의 정보를 포함하고 있는 것을 알 수 있다.

두 명이 outgoing 사이의 Bhattacharyya metric(BM), Cosine similarity(CS), Jensen Difference(JD), Euclidean distance(ED)를 이용한 similarity 측정 결과, 총 6가지 경우에 대해서 4가지 경우에서 100% 매칭, 2가지 경우에서 50% 매칭이 되었고, ED를 사용하였을 때 3가지 경우에서 100% matching이 되었고 나머지 3가지 경우에서 50% matching이 되는 것을 확인 할 수 있다. 실험결과에 의하면 CS, JD가 상대적으로 BM, ED보다 성능이 좋음을 알 수 있다.

4. 결론

본 논문은 영상인식을 통한 객체 구별을 위해 객체가 가지고 있는 RGB컬러 기반의 Similarity 측정 기법을 제안하였다. 그 기법으로 Bhattacharyya metric, Cosine metric, Jensen difference metric, Euclidean distance metric을 사용한 histogram similarity 비교를 통한 유사성 비교 방법을 제시하였다. 실험결과에 의하면 JD, CS가 더 우수한 매칭 메트릭으로 판단된다.

감사의 글

이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No.2017R1D1A3 B03028806). 본 연구는 과학기술정보통신부 및 정보통신 기술진흥센터의 대학ICT연구센터 지원사업의 연구결과임 (IITP- 2018-0-01433).

참고문헌

- [1] Z. Zhang et al. "A parking occupancy sensor detection algorithm based on AMR sensor", IEEE Sensors Journal, 15(2), Feb. 2015.
- [2] B. Iyer et al. "Dual-input dual output RF sensors for indoor human occupancy and position monitoring", IEEE Sensors Journal, 15(7), Feb. 2015.
- [3] P. Liu et al. "Occupancy inference using pyroelectric infrared sensors through hidden markov model", IEEE Sensors Journal, 16(4), Feb. 2016.
- [4] J. Trogh et al. "Enhanced indoor location tracking through body shadowing compensation", IEEE Sensors Journal, 16(7), Feb. 2016.
- [5] Z. Hua et al. "Track section occupancy detection model based on infrared ray sensors and time-series change rate matching," IEEE Sensors Journal, 16(4), Feb. 2016.

표 1. 4명의 incoming 및 outgoing p값. A, B, C, D는 사람이고, C+B는 두 사람 C, B가 같이 퇴실하는 것이다. 단일객체에서 Bold는 가장 우수한 score를 가리킨다. 우측의 청색은 가장 우수한 2개를 표시하고, red는 오류를 표시한다.

		1명 outgoing				2명 outgoing						
		A	B	C	D	C+B	A+D	D+C	A+B	C+A	B+D	
incoming	BM	A	0.301671	0.466993	0.476775	0.370435	0.379196	0.245251	0.385837	0.318056	0.275478	0.276825
		B	0.467220	0.312301	0.512388	0.500437	0.314221	0.512787	0.491455	0.291895	0.438857	0.416570
		C	0.558254	0.458322	0.222687	0.520882	0.333886	0.533108	0.274730	0.426764	0.358160	0.446867
		D	0.379620	0.547349	0.492826	0.339275	0.465520	0.231985	0.392154	0.421491	0.326322	0.308672
	CS	A	0.870430	0.562533	0.434098	0.820559	0.638752	0.927387	0.632348	0.851407	0.815624	0.897507
		B	0.600987	0.880340	0.577690	0.493523	0.798737	0.476014	0.591572	0.820610	0.653418	0.628718
		C	0.256267	0.691061	0.978256	0.363918	0.879959	0.314864	0.863896	0.599620	0.717261	0.519855
		D	0.699931	0.493124	0.477954	0.959243	0.551931	0.940740	0.739008	0.709114	0.823252	0.906767
	JD	A	0.121651	0.280859	0.297150	0.178000	0.190348	0.081453	0.200493	0.131812	0.103923	0.101264
		B	0.286026	0.123222	0.335865	0.326794	0.132985	0.345710	0.315082	0.114836	0.254342	0.232404
		C	0.398183	0.269860	0.064633	0.348336	0.148744	0.369724	0.103053	0.241882	0.176587	0.267635
		D	0.193930	0.382896	0.310254	0.146790	0.284515	0.072123	0.199856	0.233607	0.141920	0.124459
	ED	A	0.195577	0.328278	0.448611	0.227251	0.309203	0.150836	0.336532	0.222493	0.232108	0.179025
		B	0.284310	0.119294	0.361728	0.289523	0.181633	0.352671	0.313289	0.153269	0.270858	0.253251
		C	0.540761	0.390194	0.119831	0.492355	0.283931	0.537870	0.262532	0.411066	0.355752	0.439296
		D	0.357474	0.435940	0.484621	0.218253	0.418124	0.186137	0.337657	0.365516	0.289408	0.250396