

## Edge Pixels의 밀도에 의한 혼잡도 측정

양준철<sup>0</sup> 김희승

서울시립대학교 컴퓨터통계학과  
Soulbane7@Hotmail.com, hskim@Uos.ac.kr

### Measuring the degree of congestion by the density of edge pixels

Jun Chul Yang<sup>0</sup> Hee Sung Kim

Department of Computer Science and Statistics, University of Seoul

#### 요 약

컴퓨터 비전 연구에서의 주요 관심은 객체의 특징을 이용하여 객체를 분간하거나 또는 계수하는데 있어 왔다. 최근 대단위의 사람들이 온집하는 공공장소에서의 사고에 대비한 대책의 기준으로 혼잡도라는 정보의 중요성이 대두 되고 있다. 본 실험에서는 객체들이 존재하는 전경(Foreground) 영역을 객체들이 없는 배경 영역(back ground)으로부터 분리한 후 전경 영역에서의 edge pixel 들의 수를 계수하여 혼잡도의 정도를 구한다. 전경 영역과 배경 영역은 소 영역별로 RGB에 대한 표준편차와 평균을 비교 분석해서 구분하고 배경 영역을 삭제한다. 전경 영역에서 edge detection 방법을 이용하여 환경에 일맞은 edge pixels수를 계수하고 pixel수와 혼잡도 사이의 관계를 구한다. 이러한 측정 방법의 장점은 다양한 환경에서도 혼잡도라는 기본 특징정보를 추출할 수 있다는 것이다.

#### 1. 서론

컴퓨터 비전 연구는 그 동안 하나의 객체를 인식하는데 초점이 맞춰져 왔다. 하지만 객체의 특성이나 패턴 등의 인식을 통한 통계적 값을 추출하는 것도 다양한 분야에 이용될 수 있고 매우 가치 있는 일이다. 그 중 혼잡도라는 정보는 다양한 방면에 이용될 수 있다. 마라톤과 같은 대단위 사람들이 참여하는 경기 또는 운동장이나 체육관과 같은 대규모 장소의 공연, 등에서 혼잡도 측정을 통한 군중의 이동을 제어할 수 있고 대형매장에서 고객의 흐름과 시간대별 고객 성향 파악, 공공장소에서의 출입 통제 관리 시스템 등 다수의 사람이 모이는 곳에서 여러 가지 응용 분야에 응용할 수 있다. 특히 지하철의 경우 특정 객차에 사람이 몰리는 현상을 방지해 전량의 객차에 골고루 승객이 분포될 수 있도록 시스템적으로 유도할 수 있다. 이 정보를 통하여 특정 사고상황에 유연하게 승객들을 안내할 수 있는 방법으로도 이용할 수 있다.

최근 대단위의 사람이 이용하는 건물이 급증하고 시위나 특정 경기응원을 위하여 동원되는 사람이 수만에서 수 십만을 오가는 현재의 상황에서 혼잡도라는 정보의 필요성은 더욱 절실하다 하겠다. 공간은 한정되어 있고 이용하려는 사람은 늘어나게 됐을 때 시스템적으로 설정된 공간관리 프로그램은 특정부분에 사람이 몰리는 것을 예방해주고 너무 많은 사람이 몰렸을 경우에 사고를 방지하기 위하여 잠시 사람의 유입을 막아 효율적인 공간 또는 시설의 이용을 유도할 수 있다. 이러한 연유로 혼잡도라는 정보는 그 중요성이 크게 부각된다. 본 논문에서는 전경 영역에서의 edge pixels 수를 이용하여 혼잡도와 edge pixel 수 사이의 상관관계를 보여주는 데 중점을 두고 실험을 진행한다. 실험에 쓰인 영상은 640\*480 8Bit영상을 사용했다.

#### 2. 기존의 혼잡도 측정 방법

공공장소 또는 지하철 같은 대중교통 시설에 대한 혼잡도(이용도)에 대한 연구는 조금씩 있어 왔다. 대부분의 논문에서는 혼잡도 자체의 연구보다는 Tracking[1] 또는 Counting 분야에 초점이 맞추어져 왔다. 또 하나의 카메라가 아니라 여러 개의 카메라를 이용한 방법[2][4]도 사용되었다. 인간 객체의 계수를 위해서 머리(Head)부분에 초점을 맞춘 논문[3]도 있었고 움직임은 인간의 인식을 위해 다수의 객체가 아닌 단 한 사람의 운동량을 Sampling 하는 방법도 연구되었다. 칼라영상을 이용한 방법 중 아파트 통로 또는 지하도 통로에 설치된 여러 대의 카메라를 이용하여 움직임 영역을 구하는 방법도 있었다. 이를 위하여 각 카메라 영상 별로 얻은 정보를 비교 분석하는 HSI 히스토그램이 사용되기도 하였다.[5]

위의 논문들의 공통점은 다수의 객체 계수 보다 특정 환경과 상황에서 소수의 객체 수를 인식하는 방법이라는데 있다. 지하철의

경우 MRF 를 이용한 방법으로 플랫폼에서의 객체 거치를 시도한 연구도 있다.[6] 이는 공공장소에서의 안정성과 보안성을 증가시키기 위한 목적으로 비디오 감시를 이용한 방법이다. 전처리 과정으로 전경과 배경의 분리를 시도하였다. 비디오 카메라의 영상연속성에 기반을 둔 것으로 MRF 불연속점을 발견하여 전경의 분리를 시도한 방법이다. 분리된 전경의 플랫폼 영역의 혼잡도는 기하학적 모듈과의 결합을 통해서 백분율로 표시하였다.

#### 3. 영상처리

##### 3.1 영역분할 및 전처리

실험을 위하여 사용하는 카메라가 항상 고정되어 있다면 단순한 차영상(Subtract)를 이용해서 배경(Background)의 영상을 전경(Foreground)의 영상으로부터 분리되는 실험을 쉽게 할 수 있을 것이다. 하지만 지하철이나 공공장소에서의 촬영 카메라는 약간의 흔들림을 고려해야 함으로 주어진 영상을 20X20 개의 분할된 소영역으로 만들었다.

도합 400 개의 분할영상을 넘버링하고 각각의 영상 별로 전경과 배경의 RGB 값을 비교[식 1]하여 임의의 문턱치(Threshold)를 넘지 않는 부분영상은 제거한다. 즉 RGB 값에 의해서 전경과 배경의 영상이 같은 이미지를 담고 있는지 그렇지 않음을 판단하는 것이다. 여기서 다른 이미지정보를 이용하지 않고 RGB 정보를 사용하는 이유는 HSI 값이나 단순한 Gray 정보로는 같이 넘버링 된 전경과 배경의 분할영역이 서로 유사한지를 판별할 수 있는 정보를 제공하기 힘들다고 판단하기 때문이다. 또 400 개의 분할영역으로 구분했을 때 HSI 나 Gray 값의 의존도는 RGB 에 비해 더욱 떨어지기 때문에 가장 확연히 차이가 들어나는 RGB 값을 본 논문에서는 판별의 기준정보로 삼아서 실험을 하였다. 영역을 분할하고 전처리 하는 방법을 표현하면 다음과 같다. [식 1]

[식 1]

$$R_g R_{ij} = \sum_{k,l=1}^{k,l=m,n} red_{kl} / m \times n$$

$red_{kl}$  : ij 번째 분할영상이 M x N 픽셀일때 kl 번째 픽셀의 Red 값 (Fg : 전경, Bg : 배경)

마찬가지의 방법으로  $R_g R_{ij}, B_g G_{ij}, F_g G_{ij}, B_g B_{ij}, F_g B_{ij}$  값을 구한다. 이것이 의미하는 바는 Red, Green, Blue 에 따른 평균값이다. 구해진 값을 각 분할영역의 RGB 평균값이라 하면 각각의 칼라 채널 별로의 값이 아니라 전체 RGB 값을 구하고자 함으로 다음과 같이 나타낼 수 있다. [식 2]

[식 2]

$$\alpha = \frac{|b_g R_{ij} - F_x R_{ij}| + |b_g G_{ij} - F_x G_{ij}| + |b_g B_{ij} - F_x B_{ij}|}{3}$$

$$\beta = \sigma_{ij}$$

$b_g R_{ij}$  : ij 번째 배경영상의 적색(Red)평균값

$F_x G_{ij}$  : ij 번째 전경영상의 초록색(Green)평균값

$\sigma_{ij}$  : ij 번째 영상의 표준편차(Standard deviation)

RGB 의 값을 통합하는 것보다 영상 각각의 칼라 채널 별로 그 값을 비교하면 특정 칼라채널에 의해 부분영상의 RGB 정보가 결정되는 현상을 줄일 수 있다. 그리고 표준편차 값의 추가를 통해서 영역의 구별을 더욱 명확하게 진행할 수 있다.

구해진  $\alpha$ ,  $\beta$  값이  $\epsilon$  값을 넘지 않는 범위에서 영상을 비교 분석한다. (a, b 는 상수이고 a + b=1 인 확률 값이다. 환경의 변화가 심하지 않은 경우 표준편차의 의존도가 심해지고 주변환경이 복잡한 경우 평균값의 의존도가 심해진다.) [식 3]

[식 3]

$If : a\alpha + b\beta < \epsilon$  전경과 배경의 ij 번째 영상은 같은 영상

$If : a\alpha + b\beta > \epsilon$  전경과 배경이 ij 번째 영상은 다른 영상

$\epsilon$  : Threshold 값 (실험에 의한 값)

Threshold 값을 넘지 않는 분할영역을 RGB 값에 의해서 같은 영상으로 정의한다. 전경과 배경의 오차 값이 작다는 말은 전경의 영상에 어떤 다른 이미지가 포함되지 않고 배경으로 찍었던 영상의 이미지를 그대로 보관하고 있다는 뜻이 됨으로 우리는 그 부분영역의 영상에는 우리가 관심이 있는 특정객체(사람 또는 대상)의 간섭이 없다는 것으로 생각할 수 있다.

아래의 [그림 1]은 전 처리 작업이 끝난 전경의 영상으로 이후에 있을 Edge 처리에 사용될 영상이다.

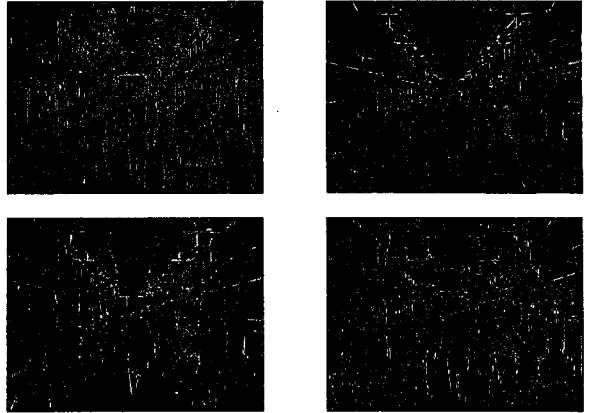
[그림 1]



### 3.2 Edge 처리

어떤 Edge 검출기 방법이 가장 적합한지를 위해서 각각의 Edge 검출기를 모두 사용하여 처리된 영상[그림 3.2]을 얻어보았다. Robert, Laplacian, Log, Prewit, Canny, Sobel 이상의 6 가지 방법중 Canny 검출기가 가장 확실한 영상의 Edge 를 표현해 주었다. 여타 검출기도 나름대로 영상의 영역을 구분해 주었지만 Sobel, Prewit, Robet 검출기는 영상의 Edge 가 너무 생략되어 표현되었고 Laplacian 검출기는 Edge 뿐만 아니라 그 외의 부분까지 표현되었다. 따라서 픽셀 비교방법에 가장 적합한 Canny 방법으로 Edge 처리 방법으로 사용되었다. [그림 2]

[그림 2]



좌상:Canny    우상:Laplacian    좌하:Sobel    우하:Log

### 3.3 전경과 배경의 픽셀

공공 장소에서의 혼잡도 측정을 위한 개체 계수는 인물의 다양한 모습과 크기 때문에 쉽지 않다. 그러므로 Canny 검출기의 특성을 살려 전경과 배경의 전 처리된 영상을 가지고 두 영상의 픽셀 수의 총합을 수학적 모습으로 나타내 보았다. 먼저 Edge 영상의 흑백 영상도에 따라 각각의 픽셀들이 0~255 사이의 값을 가지므로 픽셀 수의 계산에 있어서 정확한 데이터를 제공하지 않기 때문에 전처리된 영상을 Binary 영상으로 전환한다. [식 4]

[식 4]

$$If : P_{ij} > 0 \quad Then \quad P_{ij} = 1 \quad Else \quad P_{ij} = 0$$

$P_{ij}$  : Edge 처리후의 ij 번째 픽셀의 값

Binary 영상으로 전환한 후에 row 와 column 별로 픽셀 수를 계수해서 그 총합을 구하고 결과값을 혼잡도로 표시한다. [식 5]

[식 5]

$$v = \frac{\mu}{\omega} \geq \delta$$

$\mu$  : 전처리된 전경영상의 Pixels

$\omega$  : 한 객체의 Pixels

$v$  값과  $\delta$  의 비교를 통해 혼잡도를 표시할 수 있게 된다. 여기서 ij 번째 분할 영역 중 각각의 영역 별로 Threshold 값에 의하여 삭제된 영상에 차이가 있음을 유념해서 보정치(Correction Value)를 준다.

### 3.4 영역 보정

동일한 혼잡도를 갖는 영상이라 할지라도 객체의 지역 밀집 여부에 따라서 삭제되는 영역 수의 차이가 존재한다. 객체가 밀집되어있다면 더욱 많은 영역이 삭제될 것이고 그렇지 않다면 더욱 적은 영역이 삭제될 것이다. 때문에 많이 삭제된 영상에 대해서 보정치를 더해주어서 혼잡도는 같지만 삭제된 영역의 수가 틀린 영상들이 균일한 값을 가질 수 있도록 만든다. 따라서 혼잡도를 위한 최종수학적 표현은 다음과 같다. [식 6]

[식 6]

$$v = \frac{\mu + \frac{\phi \times \mu}{400}}{2 \times \omega} \geq \delta$$

$\phi$  : 삭제영역의 수

삭제영역의 수에 대한 확률표시로서 전체영역수인 400 으로 나누고 거기에 전경픽셀의 값을 곱해 줌으로써 영역 수 차이에 대한 보정을

수행할 수 있다. 여기서 전경 픽셀 수에 보정 픽셀 수가 더해졌기 때문에 우리가 원하는 혼잡도를 나타내기 위하여 2로 나누었다.

4. 실험

640\*480 8bit 영상을 획득하여 실험하였고 지하철내부, 신호대기 차량, 공, 호두, 글 등의 복잡한 객체에서부터 단순한 객체까지 다양한 실험을 진행하였다. 각각 취득한 영상은 100, 장으로 한정하였고 식에 대입하여 나온 혼잡도의 값을 그래프로 그려보았다. 한 객체의 픽셀 수는 무작위적으로 뽑은 영상을 바탕으로 통계적인 값을 사용했다.

4.1 환경이 객체에 미치는 영향이 큰 실험 (지하철, 신호대기차량)

실험에서는 지하철 1호선 서울역에서 청량리방향 또는 용산에서 정동북부 방향까지의 영상을 사용했다. 최대한 객체 안의 영상정보를 유지하기 위하여 사람의 이동이 가장 적은 맨 앞쪽 객차 또는 맨 뒤쪽객차의 영상을 찍었고 몇몇 객차의 경우 노약자 우대석이 장애인전용으로 의자 설치가 되어 있지 않아서 약간의 차이는 있었지만 무시할만한 수준으로 생각된다. 인위적인 계산에 의해서 지하철 차량 별 혼잡한 정도를 몸으로 느끼고 거기에 따라 총 6 단계의 혼잡도를 두었다. 혼잡도 6의 경우 승객들의 밀착상태가 상당하므로 비교 대상에서 제외시켰다. 카메라에 지지대를 대고 1m 60cm 정도의 높이에서 얻은 사진으로 작업을 하였다.

건널목에 신호대기로 멈춰있는 차량의 영상을 획득한 후 실험을 진행하였다. 객관적 조건을 고정시키기 위해 오후에 획득한 사진만을 대상으로 하였고 승용차, 버스, 트럭 등 다양한 형태의 차량이 포함된 사진을 바탕으로 혼잡도를 계산하였다.

4.2 환경이 객체에 미치는 영향이 작은 실험 (공, 호두, 글)

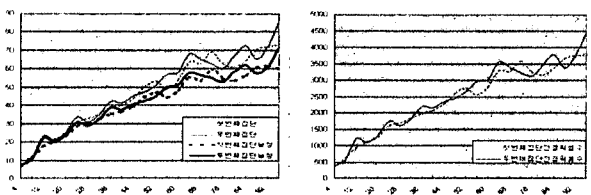
총 25 가지 종류의 객체 수를 갖는 영상으로 실험을 진행하였고 카메라의 위치는 90°의 각도를 유지한 채로 영상을 획득하였다. 동일한 수의 객체를 갖는 영상이라 할지라도 객체의 분포여부에 따라 다른 형태의 영상을 취득하였다.

5. 결과

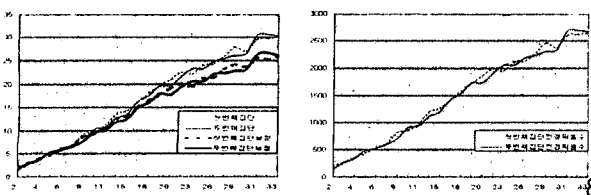
지하철 실험의 경우 주변환경의 영향이 실험에 상당히 많은 혼란을 주었다. 특히 카메라의 위치가 원래 생각했던 곳이 아닌 지하철 1m 60cm 정도의 높이밖에 되지 않고 그 시점이 좌상에서 우 하로 향하는 사선모양의 시선이 아니라 좌중에서 우중으로 향하는 평행선 모양의 시선을 담아내기 때문에 환경의 영향이 더욱 심했다. 여기서 말하는 지하철 자체의 환경영향은 400 개의 부분 영역 중 손잡이, 창문, 지지대 등의 영상이 집약되어 있는 영상으로 문턱 값 통과 여부에 따라 픽셀 값의 변화폭이 커지는 모습을 보였다.

지하철과 차량대기 실험의 공통점은 영상을 얻는 카메라의 위치로 인한 시점의 한계 때문에 혼잡도의 그래프 증가형태는 뿔아낼 수 있었지만 특유의 대표 값을 뿔아내기가 힘들었다. 특히나 앞의 사람에게 의하여 가리워지는 경우(지하철), 앞 버스에 의해 뒤의 승용차가 가리워 지는 경우(차량)로 인해 객체의 수가 다름에도 불구하고 비슷한 혼잡도의 값을 나타내는 경우도 있었다. 하지만 실험의 조건이 갖춰진 공, 호두, 글 등의 실험에서는 픽셀 수 비교를 통하여 혼잡도 값을 얻어 낼 수 있었다. [그림 3.1~3.6]

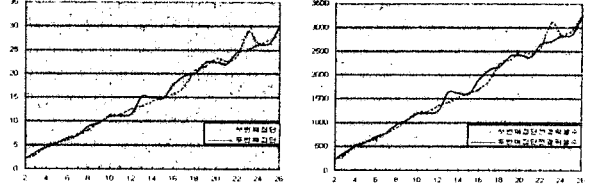
[그림 3.1. 3.2] X 축:공의 실개체수, Y 축:혼잡도와 Pixels 수



[그림 3.3. 3.4] X 축:글의 실개체수, Y 축:혼잡도와 Pixels 수



[그림 3.5. 3.6] X 축:호두의 실개체수, Y 축:혼잡도와 Pixels 수



[그림 3.1~3.6]의 그래프를 참고하면 전 처리된 전경 Pixels 수의 그래프와 한 객체의 픽셀 수에 의하여 나타난 혼잡도의 값이 동일한 형태의 그래프를 보이는 것을 알 수 있다. [그림 3.3~3.4]를 보면 혼잡도와 실개체수 사이의 관계가 선형함수 형태의 그래프를 보이는데 이것은 픽셀 수와 개체 혼잡도의 상관관계를 보여주는 결과이다. 이 결과를 바탕으로 동일한 실험을 반복함으로써 특정 개체 영역에서의 혼잡도의 대표 값을 얻어 낼 수 있었다. 이는 다양한 형태의 혼잡이 나타나는 실생활에서 빠른 혼잡도 측정을 가능하게 한다. 보정이 적용된 그래프도 비슷한 형태를 보임으로써 영역 수 차이에 의하여 나타나는 불일치성도 극복할 수 있다. 이는 영역 전처리 과정에 있어서 신뢰성을 부여하는 결과가 된다.

6. 향후 방향

정면에서 찍은 영상을 얻지 못하는 경우에 400 개 부분영역의 원근각에 따른 변화를 고려하는 방법이 필요하다. 가까운 곳의 픽셀 수와 먼 곳을 찍은 영상의 픽셀 수가 의미하는 바는 서로 차이가 있으므로 향후 이러한 두 영상 간의 의미 격차를 줄일 수 있는 방법을 찾아야 할 것이다. 전경과 배경의 픽셀 수의 상관관계를 통한 수학적 모델은 통계적인 데이터만 충분하다면 다양한 분야에 이용할 수 있다. 즉 영상의 영역이 제한되어 있고 영상에 미치는 주변환경이 미미한 경우 이 방법은 단순하면서도 정확한 정보를 제공할 것이다. 계수의 대상이 되는 객체가 포함된 영상에 특이 할만한 이미지가 포함되어 있지 않다면 픽셀 수의 상관관계 만큼 확실한 데이터는 없다. 이것은 대상이 되는 객체가 포함되지 않는 영역의 영상 복잡도는 이 실험에서 고려대상이 되지 않는다는 의미이다.

7. 참고문헌

- [1] M.Rossi, A.Bozzoli, "Tracking and Counting Moving People", IEEE Trans. 0-8166-6950-0/94, IRST-Istituto per la Ricerca Scientifica e Tecnologica I-38050 Povo, Trento, Italy
- [2] Kazuhiko Hashimoto, Katsuya Morinaka, Nobuyuki Yoshiike, Chihiro Kawaguchi and Satoshi Matsueda, "People count System Using Multi-Sensing Application", IEEE Trans. 0-7803-3829-4/97, Kyushu Matsushita Electric Co. Ltd
- [3] Xiaowei Zhang and Graham Sexton, "Automatic human head location for pedestrian of counting", IEEE Conference Publication No.433, CompressionLabs.Inc.USA, University of Northumbria at Newcastle UK.
- [4] Vera Kettner, Ramin Zabih, "Counting people from Multiple Cameras", IEEE Trans. 0-7695-0253-9/99, Cornell University Ithaca, NY 14853
- [5] Thou-HO Chen, Che-Wei Hsu, "An Automatic Bi-Directional Passing-People Counting Method Based on Color Image Processing", IEEE Trans. 0-7803-7882-2/03, Department of Electronic Engineering
- [6] Nikos Paragios, Visvanathan Ramesh, "A MRF-based approach for real-time subway monitoring", IEEE Trans. 0-7695-1272-0301, Image and Visualization department
- [7] 김희승, "영상인식", 생능출판사, 1993.
- [8] Rafael C.Gonzalez, Richard E.Woods, Steven L.Eddins "Digital Image Processing Using Matlab", "Prentice-Hall, Inc.", 2004.