

감시용 로봇의 시각을 위한 인공 신경망 기반 겹친 사람의 구분

Dividing Occluded Humans Based on an Artificial Neural Network for the Vision of a Surveillance Robot

도 용 태*
(Yongtae Do)

Abstract: In recent years the space where a robot works has been expanding to the human space unlike traditional industrial robots that work only at fixed positions apart from humans. A human in the recent situation may be the owner of a robot or the target in a robotic application. This paper deals with the latter case; when a robot vision system is employed to monitor humans for a surveillance application, each person in a scene needs to be identified. Humans, however, often move together, and occlusions between them occur frequently. Although this problem has not been seriously tackled in relevant literature, it brings difficulty into later image analysis steps such as tracking and scene understanding. In this paper, a probabilistic neural network is employed to learn the patterns of the best dividing position along the top pixels of an image region of partly occlude people. As this method uses only shape information from an image, it is simple and can be implemented in real time.

Keywords: video surveillance and monitoring (VSAM), security robot, occlusion, probabilistic neural network

I. 서론

근래에 와서 로봇이 종래의 산업용 매니퓰레이터(manipulator)와는 달리 이동성을 가지게 됨에 따라 실내 및 실외에서 자유로이 동작하게 되었다. 로봇의 동작 공간이 이처럼 확장되면서 주변 환경의 불확실성(uncertainty) 문제가 대두되고, 이는 시각 센서(vision sensor)의 중요성을 증대시키는 요인이 되고 있다. 특히 이제는 로봇의 동작 공간과 사람의 행동 공간이 더 이상 분리되지 않게 되었다.

사람과 로봇이 하나의 공간에 공존할 때, 사람은 로봇에 있어 명령을 내리는 주인이 될 수 있다. 특히 실내에서 활용되는 서비스 로봇은 권한을 가진 주인을 인식하고 이와 소통(interaction)하는 것이 필요하다. 그러나 실외에서 활용되는 로봇의 경우, 사람은 대체로 로봇에 주어진 임무의 대상이 되는 경향이 있다. 예를 들어 경비, 군사, 혹은 교통 통제용 로봇에 사용된 시각 센서에 있어서 사람은 주목하여야 할 중요한 대상이다. 그러나 비구조화된(unstructured) 일반 환경내의 사람을 시각 센서의 출력 영상에서 처리하는 데는 여러 가지 현실적인 난점이 있다. 이 같은 난점은 먼저 사람이 동적인 존재라는 특성에서 찾을 수 있다. 즉, 사람은 그 위치와 자세가 계속 바뀐다. 이러한 특성은 차량과 같은 이동체에서도 마찬가지이다. 그러나 사람은 차량 보다 작고, 특히 시각 장치가 넓은 영역을 탐지하도록 구성된 경우 영상 내에서 상당히 작게 나타난다. 또 사람은 그 이동 경로가 불규칙적이고 예측하기 어렵다. 뿐만 아니라 사람의 형태는 비정형적이어서 고정된 특징을 정의하여 활용하기 어렵다. 그리고 사람들은 다른 이들과 같이 다니는 경향이 있고, 이는 상호간에 겹침(occlusion)을 초래하여 영상 처리를 더욱 어렵게 한다[1].

본 논문에서는 실외와 같이 넓은 공간에서 사용된 감시용 로봇의 시각 영상 내에서 부분적으로 겹쳐진 사람들을 각 개

인별로 구분하는 기법을 제시한다. 이는 다수의 사람들이 영상에 나타나는 상황에서 각 개인별로 추적하거나 그들의 행동을 이해하기 위해서 기본적으로 실행되어야 하는 중요한 문제이지만, 실제로는 많은 주목을 받지 못하고 있다. 예를 들어 미국 CMU의 Robotics Institute에서는 DARPA의 지원 아래 '비디오 기반 감시 및 관측', 즉 VSAM (Video Surveillance And Monitoring) 연구를 광범위하게 수행한 바 있다[2]. 영상 내에서 검출되는 동적인 물체는 인공 신경망 (ANN: Artificial Neural Network)을 사용하여 사람, 사람 무리, 그리고 차량으로 분류된 후 각각 처리되었는데, 사람 무리 영상 내에 존재하는 겹침의 문제는 다루어지지 않았다. 또 Treptow 등[3]은 이동형 경비 로봇을 위하여 열화상 카메라에 의해 사람을 일단 검출하고, 이로부터 pan-tilt-zoom 카메라를 사용하여 사람의 얼굴 부분 영역을 추출하였다. 그러나 겹쳐짐은 여전히 다루어지지 않았다.

카메라 영상 내 복수의 사람들 사이에 존재하는 겹쳐짐을 다루는 가장 간단한 방법은 카메라를 사람들의 활동 영역 상부에 장착하고 수직으로 내려다 보게 함으로써 아예 겹쳐짐이 존재하지 않게 하는 것이다[4]. 그러나 이러한 방법은 로봇에 카메라가 탑재된 경우에는 적용하기 어렵고, 대상이 되는 사람의 형태적 특성에 대한 정보를 획득할 수 없는 단점도 있다. Sindhu와 Morris의 연구[5]에서는 비디오 영상에서 사람들의 속도에 기초하여 복수의 사람들이 동일한 위치에 겹쳐짐을 예측하였다. 그러나 부분적으로 겹쳐진 사람들을 구분하기 위한 시도는 하지 않았다. Lin 등[6]은 컬러(color)에 기반하여 겹쳐진 사람들을 구분하는 반복적 확률 갱신 모델을 제안하였다. Batista[7]도 컬러를 겹쳐진 사람들의 개인별 추적에 사용하였고, 특히 부가적으로 사람들의 키(height) 정보도 활용하고 다수의 카메라를 이용하여 공간적, 시간적 추정을 수행하였다. 컬러의 사용은 근래 급격히 진행중인 컬러 카메라의 저가화와 컴퓨터의 고성능화를 고려할 때 유용할 수 있으나, 자동 경비의 중요성이 보다 증대되는 야간의 경우에는 효과적이지 못하다. 또 주간이라도 체복과 같이 동일

* 책임저자(Corresponding Author)

논문접수: 2009. 1. 30., 채택확정: 2009. 2. 25.

도용태: 대구대학교 전자공학부(ytdo@daegu.ac.kr)

※ 이 논문은 2006학년도 대구대학교 학술연구비 지원에 의한 논문임.

한 색의 복장을 착용한 사람들 무리에 대해서는 적용하기 어려운 문제가 있다. 컬러의 사용 없이 목표의 형태적 특성만을 고려한 연구의 예로서, Heikkila와 Silven[8]는 교통 분석의 목적으로 카메라 영상 내 보행자를 자동으로 계수하는 연구를 수행하면서 겹쳐진 사람의 외곽 형태를 고려하였다. 그들이 사용한 영상에서 사람들은 비교적 큰 해상도로 나타났고, 수집된 영상을 추후 오프라인으로 한꺼번에 처리하였다.

본 논문에서는 카메라 영상 내 복수의 사람들 사이에 부분적인 겹쳐짐이 존재할 때, ANN을 사용하여 외곽 형태상에 존재하는 특성을 학습함으로써 각 개인을 실시간으로 구분하는(dividing)하는 방법에 대해 서술한다. 사용된 ANN은 확률 신경망(PNN: Probabilistic Neural Network)으로, PNN은 Bayesian 분류기의 신경망적 구현이다. 제안된 기법은 영상 내 목표의 형태 정보만을 이용하고 시간적 정보는 이용하지 않으므로 간단하여 계산 비용이 크지 않다. 또 학습에 의하므로 다양한 카메라와 설치 조건에서 유용할 수 있다.

II. 영상의 전처리와 신경망을 위한 특징추출

1. 영상 내 사람 영역을 추출하기 위한 전처리

사람을 대상으로 하는 로봇 시각 체계의 영상 처리 절차는 그림 1과 같다. 먼저 카메라 영상 내 동적인 대상을 추출한다. 이를 위해서는 배경 영상 차감법(background subtraction)[9]이 실행하기 간단하고, 목표물의 형태 특징을 대체로 간직하는 결과를 얻을 수 있어 널리 사용된다. 배경 영상 차감법은 카메라 관측 영역의 영상을 배경(background)과 전경(foreground)으로 구분한다. 기저장된 배경 영상을 기준으로 현재의 카메라 영상 중 변화가 뚜렷하고 그 변화가 어느 시간이상 지속되는 화소들로 영상 내 전경을 구성한다. 이 때 배경 영상은 관측 지점의 주변 밝기 등 환경 변화에 맞추어 지속적으로 갱신할 필요가 있다. 이 방법은 간단하지만, 일정 시간 동안 카메라의 이동이 없고 대상체만 움직일 경우 적용할 수 있다. 이동 로봇이 특별히 움직여야 할 경우가 아니면 전력 소모를 줄이기 위해 정지해 있을 수 있으며, 이 때 이상의 방법은 유용하다. 그러나 만약 로봇도 이동 중에 있으면 보다 복잡한 검출 방법이 필요하다[10].

본 논문에서는 사람만을 대상으로 하고 있으나, 관측 영역에 사람과 차량이 혼재할 경우에는 전경으로 추출된 객체가 사람인지 확인하여야 한다. 사람과 차량은 식 (1)과 같이 정의된 dispersedness를 적용하여 구분할 수 있다[2].

$$Dispersedness = \frac{perimeter^2}{area(pixels)} \quad (1)$$

일단 사람으로 판정된 전경 영역에는 사람의 그림자까지 포함되어 있을 수 있고, 그림자는 영역의 형태를 비예측적으로 변형시켜 차후의 처리를 어렵게 하는 요인이 된다. 우리는 참고 문헌 [11]에서 컬러 공간의 변환에 기초한 그림자 화소 판정법을 제안한 바 있다. 즉, [R,GB] 공간에서의 컬러 영상을 [H,S,V] 공간으로 변화시켰을 때, 그림자는 H와 S값은 유지하면서 V값을 낮게 하므로 이를 화소별로 모호화하여 그림자 화소를 추출할 수 있다. 이 방법은 컬러 정보를 활용한다는 문제가 있지만, 그림자는 밝은 광원이 있을 때만 생긴다는 점과 목표 사람 자체의 컬러가 아니라 그림자가 드리워진 배경의 컬러가 처리 대상이 된다는 점에서 앞 절에서

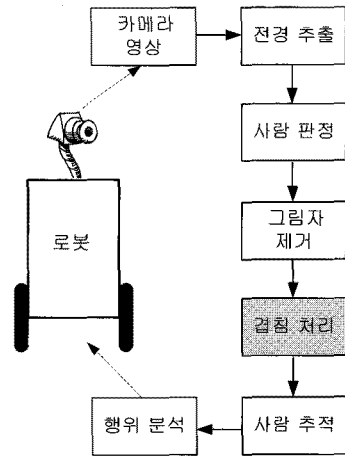


그림 1. 사람을 대상으로 하는 로봇 시각 체계.
Fig. 1. Robot vision system targeting humans.

[6,7]에 대해 논의한 컬러 기반 방법의 문제점이 부각되지는 않는다.

2. 신경망을 위한 겹쳐진 사람 영역의 특징 추출

부분적으로 겹쳐진 사람들의 영역이 앞 절에서 서술된 절차에 의하여 영상에서 추출되었으면, 이를 적절한 위치에서 구분함으로써 각 개인별 영역을 확정하여 추적 및 행위 분석과 같은 높은 단계의 영상 처리가 가능하게 된다. 본 논문에서는 ANN을 사용하여 겹쳐진 사람 영역을 개인별로 구분한다. 이를 위해서는 ANN에 입력될 적절한 특징이 영상에서 추출되어야 한다.

겹쳐진 사람 영역의 특징을 얻기 위하여 본 논문에서는 전경 영역의 최상위 화소들을 연결한 선을 얻었다. 이는 그림 2에 보인 바와 같다. 이와 같은 선을 정의하여 사용하는 것은 두 가지 이유를 가진다. 먼저, 전경 영역의 추출은 흔히 완벽하지 못하다. 즉, 추출된 영역 내부가 비거나, 혹은 떨어져야 할 부분이 연결되는 상황이 자주 생긴다. 특히 사람의 어깨 이하 부분이 그러한데, 이는 사람이 움직일 때 팔이나 다리가 많이 움직이기 때문이다[12]. 또 사람이 착용한 복장이 배경과 유사한 색, 혹은 밝기를 가질 때도 그렇다. 둘째로 특정 전경 영역에 대해서 최상위 화소들은 거의 연산없이 바로 주어진다. 이 선은 기존의 경계(boundary)와는 다르므로 경계 추적(boundary following)이 필요하지 않다. 일단 최상위 화소들을 연결한 선이 주어지면, 이 선의 국부적인 기울기 코드(slope code)을 얻고, 이를 ANN의 입력으로 사용하였다.

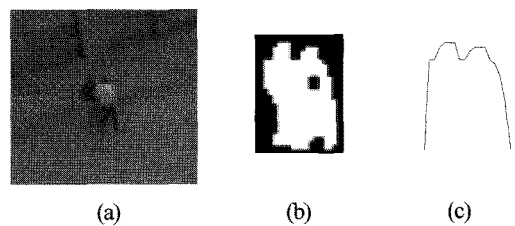


그림 2. 최상위 화소 연결선: (a) 원 영상, (b) 전경 영역, (c) 추출된 연결선.
Fig. 2. Line connecting the utmost pixels: (a) Original image, (b) Foreground image region, (c) Line extracted.

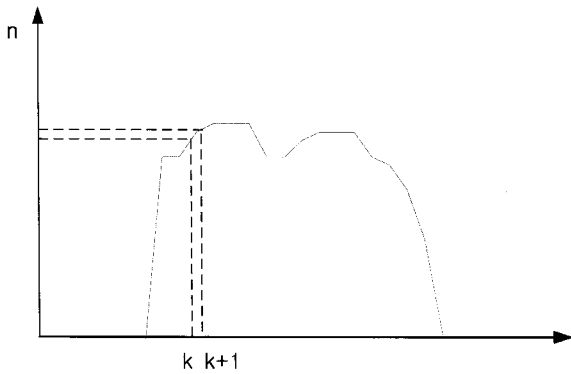


그림 3. 최상위 화소 연결선의 국부적 기울기.
Fig. 3. Partial slope of a line connecting the utmost pixels.

즉, 그림 2(c)과 같이 얻어진 연결선을 그림 3과 같이 다시 그릴 때, 임의의 k 번째 최상위 화소의 좌표에 대해, 다음과 같이 3가지 다른 코드를 얻는다:

- $n(k) < n(k+1)$ 인 경우 +1,
- $n(k) > n(k+1)$ 인 경우에는 -1,
- 나머지 경우에는 0.

III. 신경망을 이용한 겹친 사람의 구분

1. 확률 신경망

우리는 그림 2(b)와 같이 부분적으로 겹쳐져 하나의 전경 영역으로 추출된 복수의 사람을 각 개인별로 구분하기를 원한다. 사람들의 형태가 유동적이고, 겹침의 유형도 다양하므로 ANN의 학습에 의하여 적절한 구분 위치를 찾는 것을 제안한다.

확률 신경망(PNN)은 Bayesian 분류기를 ANN으로 구현하기 위해 제안되었다[13]. PNN은 한 번의 자료 제시만으로 학습이 완료되는 장점을 가지고 있다. 단점은 모든 학습용 자료가 신경망에 저장되어야 하므로 저장 공간과 계산 시간이 자료의 양에 따라 커진다는 것인데, 컴퓨터의 성능이 근래에 급격히 향상됨을 고려하면 큰 문제가 되지 않는다.

차원 m 을 가지는 임의의 유형 벡터(pattern vector) \mathbf{x} 가 두 개의 군집 K_1 과 K_2 중 하나에 속한다고 하자. 두 군집의 확률 밀도 함수(probability density function; pdf)를 각각 f_1 과 f_2 라고 할 때, Bayes의 규칙에 따르면 \mathbf{x} 는 다음의 조건에 따라 분류된다.

$$\mathbf{x} \text{ belongs } K_1 \text{ if } \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} > \frac{L_1 P_2}{L_2 P_1} \quad (2)$$

K_2 otherwise.

단, L_1 은 K_2 에 속하는 \mathbf{x} 를 K_1 으로 잘못 분류한 경우에 대한 손실 함수(loss function)이고, L_2 는 K_2 로 잘못 분류한 경우에 대한 손실 함수이다. 또 P_1 은 K_1 이 일어날 사전 확률(prior probability)이며, P_2 는 K_2 의 사전 확률이다. 손실 함수와 사전 확률은 쉽게 구할 수 있거나, 흔히 두 군집에 대해 동일하다고 간주한다. 그러면 문제는 pdf를 학습용 자료로부터 추정하는 일이 된다[14].

PNN에서는 비모수 추정(nonparametric estimation) 방법인

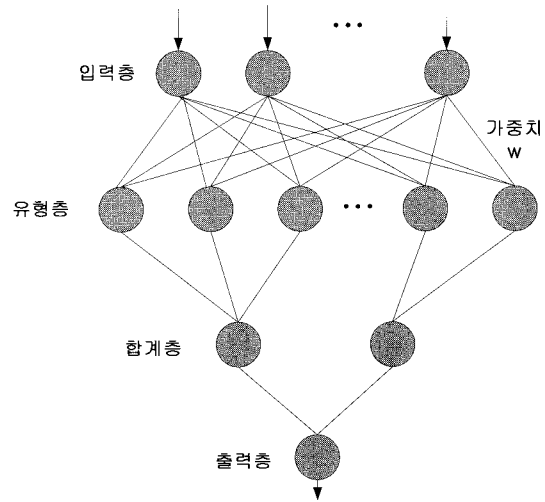


그림 4. 확률 신경망(PNN)의 구조.
Fig. 4. Structure of a probabilistic neural network(PNN).

Parzen 창[15]을 사용하여 분류할 군집에 의존적인 pdf를 식 (3)과 같이 얻는다. 이는 주어진 유형 벡터가 특정 군집에 속할 확률을 계산할 수 있게 한다.

$$f_j(x) = \frac{1}{(2\pi)^{m/2} \sigma^m P_j} \sum_{i=1}^{P_j} e^{-(x-x_i^j)^T (x-x_i^j)/2\sigma^2} \quad (3)$$

단, 식 (3)에서 m 은 입력 벡터의 차원, P_j 는 군집 j 에 속하는 학습 유형의 개수, x_i^j 는 j 번째 군집의 i 번째 학습용 입력 유형, 그리고 σ 는 평활화 인자(smoothing factor)이다. PNN의 학습, 즉 가중치 결정은 한 번의 학습용 자료 제시로 결정되므로, 실제의 신경망 학습은 이 평활화 인자를 적절하게 결정하는 것이 된다.

하나의 PNN은 그림 4와 같이 4개의 층으로 구성된다. 즉, 입력층(input layer), 유형층(pattern layer), 합계층(summation layer), 그리고 출력층(output layer)이 그것들이다. 입력층은 m 개의 입력 변수 $\{x_1, x_2, \dots, x_m\}$ 와 연결된다. 입력층의 뉴런들(neurons)은 유형층의 뉴런들과 완전 연결되어 있고, 유형층의 각 뉴런은 훈련 자료들의 특정 유형을 대표한다. 이 때 유형층 뉴런의 가중치(weight)는 훈련용 유형과 동일하게 설정된다. 합계층에서는 식 (3)의 지수함수 항이 계산된다. 유형층으로부터 합계층으로의 연결 가중치는 모두 1이며, 군집 별로 하나씩의 합계 뉴런을 갖는다. 출력층의 뉴런은 이진값을 출력하는데, 이는 식 (3)에 의한 가장 큰 pdf에 해당한다.

2. 신경망에 의한 겹친 사람의 구분

본 논문에서는 영상 영역의 구분(division) 문제를 유형 인식(pattern recognition)의 문제로 간주하는 것을 제안함으로써, 신경망을 활용할 수 있게 하였다. 즉, 주어진 영상 영역을 구분할 많은 위치들 중 납득할 수 있는 위치들로 하나의 유형 군집(pattern class)을 만들고, 다른 적당하지 않은 구분 위치들을 또 다른 유형 군집으로 만든다. 이들 두 군집을 각각 D (class of divisible positions)와 N (class of not divisible positions)이라고 하자.

본 논문에서는 PNN을 사용하여 II.2절에서 구한 최상위 화

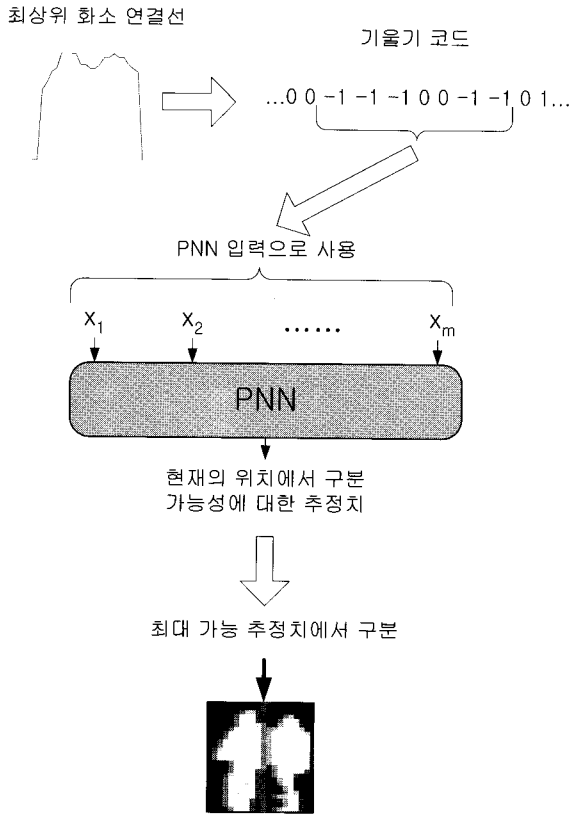


그림 5. 확률 신경망을 사용한 겹쳐진 사람의 구분 절차.
Fig. 5. Procedure of dividing occluded people using a PNN.

소 연결선을 따라 적절한 구분 위치를 찾는다. 부분적으로 겹쳐진 복수 사람의 다수 영상 영역을 육안으로 직접 최적 구분하고, 이를 신경망의 학습 자료로 사용한다. PNN을 위한 입력 벡터는 최상위 화소 연결선의 기울기 코드를 사용한다. 이는 PNN이 부분적으로 겹쳐진 두 사람을 구분할 수 있는 최적의 위치를 사람 신체의 가장 윗부분의 국부 정보에 기초하여 찾는다는 것을 의미한다. 겹쳐진 사람의 전경 영역의 너비가 c 개의 화소이면, $c-1$ 개의 기울기 코드가 얻어지고, 이들 중 m 개의 연속된 코드가 PNN의 입력으로 사용된다. 신경망의 출력은 입력된 코드들이 얻어진 위치의 중앙이 영역을 나누기에 적절한 곳인지를 추정하는 값이 된다. PNN은 Bayes의 분류기를 실현하는 것이므로 그 출력은 이진이다. 그러나 본 논문에서는 이진화된 출력을 사용하지 않고, 실수 값을 사용하여 최대의 값을 출력하는 위치를 영역 구분 위치로 결정한다. 그런데 실제 영상에서 최적의 구분 위치는 사람의 육안으로도 정확하게 찾기 어려운 경우가 많으므로, 육안으로 결정한 위치의 좌우 세 화소씩을 D 군집의 요소로 분류한다. 이는 실제적으로 주어진 영역에서 최적의 구분 위치는 단수인데 반해 나누기 적당하지 않은 위치는 나머지 다수이므로 N 의 크기가 D 에 비해 훨씬 큰 문제를 완화시킴으로써, 신경망의 학습 자료를 적절하게 하고자 하는 의도도 있다. 그림 5는 이상의 절차를 요약한 것이다.

IV. 실험 결과

그림 5와 같은 방식으로 PNN을 사용하여 겹쳐진 복수의

Person 1	Person 2	Person 3	Person 4	Person 5
34×14 [Pixels]	35×13 [Pixels]	32×13 [Pixels]	30×10 [Pixels]	29×13 [Pixels]

그림 6. 실험에 사용된 사람들의 영상.
Fig. 6. People images used for experiment.

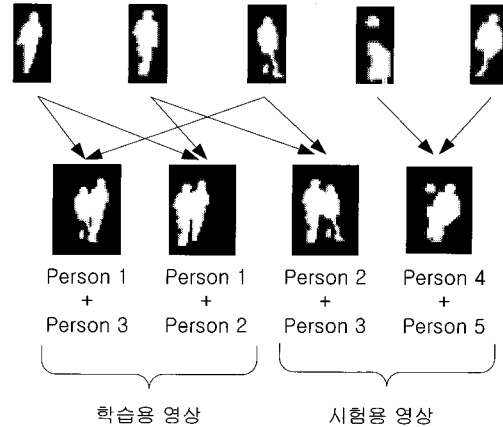


그림 7. 실험을 위해 사용한 합성 겹침 영상.
Fig. 7. Synthesized occluded images for experiment.

사람 영상을 구분하였다. 실제 카메라의 영상에서 얻어진 그림 6과 같은 다섯 명의 사람을 컴퓨터에서 다양하게 그림 7과 같이 겹쳤다. 이는 실제 영상에서는 다양한 겹침을 얻기가 극히 어렵기 때문이다. 그림 6에 나타난 실제 사람들의 몸 크기에는 상당한 차이가 있었지만, 각 사람들을 상자로 둘러쌀 때(bounding box), 그 크기의 차이는 의외로 크지 않았다. 이는 사용된 영상이 저해상도이기 때문인 것으로 보인다. 사용된 영상은 CMU의 VSAM 연구에서 획득된 것들이다[2].

총 60개의 겹쳐진 영상이 생성되어 PNN의 학습에 사용되었다. 학습용 자료를 얻기 위해 사용된 겹침 영상은 Person 1과 Person 2를 사용하여 30장, 그리고 Person 1과 Person 3를 사용하여 30장이다. 학습된 신경망은 또 다른 60장의 겹침 영상을 이용하여 시험되었다. 이 때 이용된 영상은 Person 2와 Person 3, 그리고 Person 4와 Person 5를 사용하여 만들어졌고, 각각 30장씩이다. Person 4와 Person 5는 신경망의 학습에 사용되지 않았고, 이진 전경 영역의 추출이 상당히 불완전하기 때문에 이에 대한 실험 결과는 흥미롭다.

실험에 사용된 임의의 전경 영상에 대해서 최상위 화소 연결선을 얻고, 이 선을 따라 7개의 연속적인 국부 기울기 코드를 PNN의 입력 벡터로 사용하였다. 즉, 그림 5에서 $m=7$ 로 하였다. 평활화 요소는 실험을 통하여 적당하다고 결론내려진 1.1을 사용하였다. 즉, 식 (3)에서 $\sigma=1.1$ 로 하였다.

이상과 같이 구성된 PNN은 60개의 학습 영상에 대해 56개의 구분 성공을 보였는데 이는 약 93%의 성공률이다. 학습된 PNN이 시험용 겹침 영상에 적용된 결과는 표 1과 같이 평균 85%의 구분 성공률을 보였다. 성공의 여부는 사람의 육안으로 설정된 최적의 구분 위치에 대해 III절에서 서술한 바와 같이 좌우로 3화소씩 허용하는 것으로 그림 8과 같이

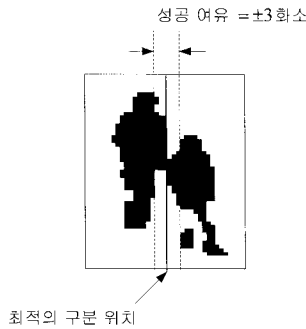


그림 8. 구분선의 성공 여유.

Fig. 8. Success margin for a dividing line.

표 1. 시험 결과 (입력벡터 크기=7, $\sigma = 1.1$).

Table 1. Test results (Input vector size=7, $\sigma = 1.1$).

겹침의 종류	성공적인 구분 수 (비율)
Person 2 + Person 3	24 (80%)
Person 4 + Person 5	27 (90%)

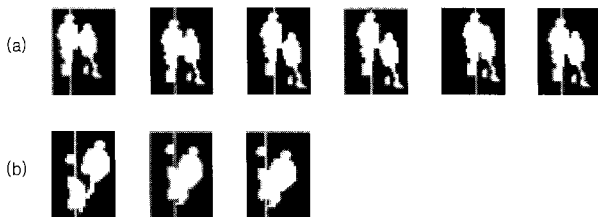


그림 9. 실험에서 구분 실패 판정된 경우들: (a) Person2+ Person3의 경우, (b) Person4+Person5의 경우.

Fig. 9. Cases determined as division failures in experiment: (a) Cases for Person2+ Person3, (b) Cases for Person4+Person5.

여유를 두어 판정하였다. 특히 Person 4와 Person 5를 사용한 겹침 영상에 대한 시험에서도 그 성공률이 떨어지지 않고, 오히려 좋은 결과를 보였다. 그림 9는 구분에 실패한 경우를 보여 준다. 대체로 비슷한 형태 유형에서 실패하였고, 실패한 구분선은 최적의 구분 위치와 크게 벗어나지는 않았다.

V. 결론

로봇이 카메라를 사용하여 보안 및 경비 등의 목적으로 넓은 공간에서 사람들을 관측할 때, 영상 내 사람들 사이의 겹쳐짐은 이들의 추적이나 행위 분석과 같은 높은 단계의 처리를 어렵게 하는 요인이 된다. 기존의 많은 관련 연구들에서 이러한 겹쳐짐의 문제는 고려되지 않았거나, 겹쳐진 사람들 전체를 한 개체로 처리하였다. 본 논문에서는 복수의 사람들이 저해상도 카메라의 영상에서 부분적으로 겹쳐져 있을 때, 이를 신경망을 사용하여 구분하는 방법을 제안하였다.

본 논문에서 제안된 기법의 특징은 겹쳐진 사람의 영상을 구분하는 문제를 유형 인식의 문제로 전환하여, 확률 신경망(PNN)을 사용하여 인식하는 것이다. 겹쳐진 사람들의 이진 영상 영역의 최상위 화소 연결선을 정의하고, 이 선을 따라 연속된 국부 기울기 코드를 신경망의 입력으로 사용하였다. PNN은 오류역전파에 의한 다층신경회로망과는 달리 한 번

의 자료 제시만으로 학습이 완료되는 장점이 있다. 제안된 방법은 사람의 머리 및 어깨와 같은 상측 형태 정보만을 이용하는 데, 이 부분들은 사람의 보행 중에도 안정적인 특성을 보이므로 유용하다. 또 시간적 정보를 활용하지 않아 처리가 간단하다는 장점도 있다. 제안된 기법은 다양한 겹침에 대한 실험에서 약 85%의 성공률을 보였다.

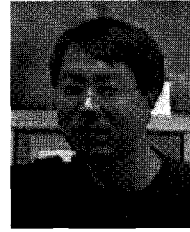
본 논문의 연구에서 목표 사람들은 배경 차감법에 의하여 검출하였으므로, 카메라는 정지된 상황에 있다는 가정이 사용되었다. 이는 보안을 위해 사용되는 고정된 카메라나 전력 손실을 최소화하기 위해 정지해 있는 로봇의 경우에는 적절하지만, 주행 중인 로봇에 장착된 카메라에는 적용상의 제약이 된다. 또 겹쳐짐의 정도가 심각하지 않은 경우를 주로 다루었는데, 이는 이진영상에 의한 목표 사람들의 윤곽만을 고려하였기 때문이다. 이러한 문제들은 향후 더욱 연구해야 할 과제로 남아 있다.

참고문헌

- [1] S. Khan and M. Shah, "Tracking people in presence of occlusion," *Proc. of Asian Conf. Computer Vision*, 2000.
- [2] T. Kanade, R. T. Collins, and A. J. Lipton, "Advances in cooperative multi-sensor video surveillance," *Proc. of Image Understanding Workshop*, pp. 3-24, 1998.
- [3] A. Treptow, G. Cielniak, and T. Duckett, "Active people recognition using thermal and grey images on a mobile security robot," *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, pp. 2103-2108, 2005.
- [4] J. Barandiaran, B. Murguia, and F. Boto, "Real-time people counting using multiple lines," *Proc. Int. Workshop on Image Analysis for Multimedia Interactive Services*, pp. 159-162, 2008.
- [5] A. J. Sindhu and T. Morris, "A region based approach to tracking people before, during, and after occlusions," *Proc. IASTED Int. Conf. Visualization, Imaging, and Image Processing*, 2004.
- [6] Z. Lin et al., "Simultaneous appearance modeling and segmentation for matching people under occlusion," *Proc. Asian Conference on Computer Vision*, pp. 404-413, 2007.
- [7] J. P. Batista, "Tracking pedestrians under occlusion using multiple cameras," *Lecture Notes in Computer Science*, vol. 3212, pp. 552-562, 2004.
- [8] J. Heikkila and O. Silven, "A real-time system for monitoring cyclists and pedestrians," *Proc. IEEE Workshop on Visual Surveillance*, pp. 74-81, 1999.
- [9] P. D. Z. Varcheie, M. Sills-Lavoie, and G-A. Bilodeau, "An efficient region-based background subtraction technique," *Proc. Canadian Conf. Computer and Robot Vision (CRV)*, pp. 71-78, 2008.
- [10] W. Abd-Almageed, B. J. Burns, and L. S. Davis, "Identifying and segmenting human-motion for mobile robot navigation using alignment errors," *Proc. Int. Conf. Advanced Robotics (ICAR)*, pp. 398-403, 2005.
- [11] 박수우, 김정대, 도용태, "비디오 카메라 영상 내 동적 물체의 그림자 화소 검출 기법," *멀티미디어학회논문지*, 제8권 제10호, pp. 1314-1321, 2005.
- [12] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4: Who? When? Where? What? A real time system for detecting and tracking people," *Proc. Int. Conf. Face and Gesture Recognition*, pp.

222-227, 1998.

- [13] D. E. Specht, "Probabilistic neural networks for classification, mapping, or associate memory," *Proc. IEEE Int. Conf. Neural Networks*, vol. 1, pp. 525-532, 1988.
- [14] A. T. C. Goh, "Probabilistic neural network for evaluating seismic liquefaction potential," *Canadian Geotechnical Journal*, vol. 39, no. 1, pp. 219-232, 2002.
- [15] E. Parzen, "On estimation of a probability density function and mode," *Annals of Mathematical Statistics*, vol. 33, pp. 1065-1076, 1962.



도 용 태

경북대학교 전자공학과(공학사). 서강대학교 전자공학과(공학석사). 영국 Univ of Hull 전자공학과(Ph.D). 1999년 미국 CMU Robotics Institute 객원부교수. 2004년 영국 Imperial College Vision & Image Processing Lab 객원교수. 2004년 뉴질랜드 RSNZ Linkage Fund 초청 과학자. 현재 대구대학교 전자공학부 교수. 관심분야는 로봇 시각, 인공 신경망, 검사 자동화.