

신경망을 이용한 동작분석과 원격 응급상황 검출 시스템

Human Behavior Analysis and Remote Emergency Detection System Using the Neural Network

이동규*, 이기정**, 황보택근*, 임혁규**
한북대학교 컴퓨터공학과*, 경원대학교 소프트웨어대학**

Dong-Gyu Lee(dglee@hanbuk.ac.kr)*, Ki-Jung Lee(jcm5758@empas.com)**,
Taeg-Keun Whangbo(tkwhangbo@kyungwon.ac.kr)**, Hyuk-Kyu Lim(hklimgo@daum.net)**

요약

본 논문에서는 신경망을 이용한 동작분석 기법을 통한 자동화 영상감시시스템의 구현과 응급상황 검출에의 응용을 제안한다. 카메라로부터 입력된 영상은 통계적 배경 모델에 의한 배경 감산법에 의해 객체 영역이 분리되고, 분리된 객체영역의 특징을 표현할 수 있는 특징벡터의 형태로 변형된다. 특징벡터를 이용한 동작분석을 위해 신경망을 사용하였고 간단한 연산에 의해 동작을 구분할 수 있도록 하였다. 본 논문에서는 실험을 위해 stand, faint, squat 등 3가지의 동작 상태를 분류할 수 있도록 하였고, 실험 결과 응급상황을 검출하기위한 알고리즘으로 유용함을 보였다.

■ 중심어 : | 영상감시시스템 | 응급상황 | 신경망 | 동작분석 |

Abstract

This paper proposes an automatic video monitoring system and its application to emergency detection by analyzing human behavior using neural network. The object area is identified by subtracting the statistically constructed background image from the input image. The identified object area then is transformed to the feature vector. Neural network has been adapted for analyzing the human behavior using the feature vector, and is designed to classify the behavior in rather simple numerical calculation. The system proposed in this paper is able to classify the three human behavior: stand, faint, and squat. Experiment results shows that the proposed algorithm is very efficient and useful in detecting the emergency situation.

■ keyword : | Video Monitoring System | Emergency Situation | Neural Network | Human Behavior Analysis |

1. 서론

산업사회의 발전과 의료기술 발달로 고령 노인인구는 지속적 증가하고 핵가족화에 따라 혼자 지내는 노

인이 급증하고 있다. 이러한 현상으로 노인성 만성질환자의 비중이 증가하고 86%가 만성질환, 32.8%가 치매 증상으로 고생, 43%가 일상생활에 타인의 도움이 필요한 상태이다(경기발전5개년계획, 2003). 의료 서비스

* 본 연구는 2005년도 경원대학교 RIS 사업단 연구비에 의해 수행되었습니다.

접수번호 : #060713-003

접수일자 : 2006년 07월 13일

심사완료일 : 2006년 08월 31일

교신저자 : 황보택근, e-mail : tkwhangbo@kyungwon.ac.kr

측면에서 볼 때 독거노인은 증가하는 반면 의료서비스는 체계가 갖추어져 있지 않아 응급상황 발생 시 신속한 의료 서비스를 받지 못하고 있는 실정이다. 따라서 원격 모니터링을 통한 화상인식 시스템개발을 통해 일일이 찾아가지 않고 돌연사의 위험이나 응급상황을 미연에 방지하고자 이러한 시스템을 개발하게 되었다.

이러한 기술들은 유비쿼터스 시스템이나 지능형 센서 시스템, 원격제어, 원격 건강관리, 비디오 감시 장치 등과 연관되어 있다. 이러한 시스템들의 목적은 인간의 동작분석을 통한 사고방지를 목적으로 하며 주변에 설치된 카메라로부터 얻을 수 있는 영상 데이터를 처리하여 간접적인 방법으로 잠재적인 위험상황을 감지하는데 있다[1][2]. 본 논문의 주요 목적은 실내에 설치된 감시 카메라로부터 수집된 영상 데이터를 통해 일련의 영상처리 방법을 사용하여 객체영역을 추출하고, 정의되어 있는 인간의 동작 특성과의 비교를 통해 특정상황의 동작을 구분하고 인지함으로써 응급상황에 대처할 수 방법을 제시하고자 한다.

영상데이터를 통한 동작분석을 위해 여러 가지방법이 제시되었다[4-7]. 그러나 기존의 방법들은 처리할 데이터의 양이 많고 복잡한 분석방법에 의해 처리시간이 많이 걸리게 된다. 그러므로 연속적인 동영상으로부터 프레임별로 분석결과를 추정하기 위해서는 간단한 추출방법에 의해 우리가 원하는 동작의 특성만을 실시간으로 분석해 낼 수 있는 방법을 필요로 한다.

본 논문에서는 사람의 팔, 다리 등 세부적인 신체구조의 분석보다는 전체적인 응급상황의 인지만을 필요로 하므로 단순한 데이터의 처리를 통해, 실시간처리가 가능하고 복잡한 동작의 분석보다는 특정한 응급상황의 판별을 위주로 하는 동작 분석 방법을 제안한다. 특정한 동작 상태를 분석하기위해서 객체부분의 특징을 추출하고 이로부터 동작 상태를 분석하기위해 단순한 연산에 의해 가장 유사한 패턴을 분류할 수 있는 신경망(neural network)을 사용한다.

II. 관련연구

영상데이터를 이용한 동작분석을 위해서 카메라로부터 입력된 영상신호에서 움직임이 있는 객체영역을 분리해내기 위한 객체 분리 방법과 분리된 객체로부터 인간의 동작을 분석해내기 위한 동작 분류(classification)방법을 필요로 한다.

1. 객체분리

연속된 영상신호에서 움직임이 있는 객체의 분리를 위해 연속된 두 프레임간의 차를 구하여 움직임이 있는 부분만을 검출하는 방법과 저장된 배경화면과 입력 화면의 비교를 통해 객체부분을 분리하는 방법이 사용된다. 프레임의 차를 이용하는 방법은 움직임이 큰 부분에서만 객체부분을 분리할 수 있기 때문에 객체가 정지해 있는 경우 분리해낼 수가 없고 조명의 변화나 잡음의 영향을 많이 받게 된다. 그러므로 정확한 객체의 분리를 위해서는 일반적으로 배경감산(background subtraction)에 의한 객체분리 알고리즘이 사용되며 [그림 1]과 같이 저장되어 있는 배경화면(배경모델)과 현재의 화면을 감산함으로써 배경화면에 추가된 객체영역을 분리해 낸다[1][5][6]. 조명의 변화나 시간의 흐름에 따른 점진적인 배경모델의 변화를 반영하기위해 배경화면을 갱신해 나간다.

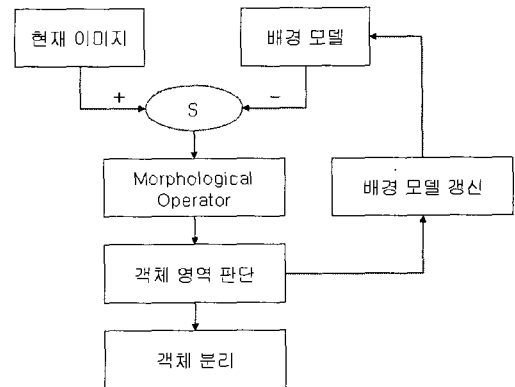


그림 1. 배경감산에 의한 객체분리

2. 동작분석(human behavior analysis)

영상 데이터를 통한 동작분석은 측정하려는 대상과 응용분야에 따라 인간의 신체를 팔과 다리, 몸통, 머리 등과 같이 여러 기능적인 부분으로 분해하여 각 부분의 연결 상태를 분석하는 신체구조분석(body structure analysis)[2][3], 신체의 분리 없이 인간의 움직임을 추적하는 움직임 추적(motion tracking)[5][8], 연속적인 동영상으로부터 인간의 활동 상태를 추정하는 움직임 인식(human activity recognition)[4-7] 등으로 구분할 수 있다.

신체구조분석에서는 인간의 신체 구조를 여러 가지 형태로 모델링하고 입력영상과의 비교를 통해 해당 신체부분의 동작 상태를 분석하는 방법이다. 움직임 추적은 신체의 각 부분의 움직임보다는 몸 전체의 움직임을 추적하기 위한 방법으로 객체의 위치나 속도, 모양, 색상의 변화와 같이 객체와 관련된 여러 정보의 변화를 분석하기 위한 방법이다. 움직임 인식은 각 신체 부분의 연속적인 움직임이나 전체적인 움직임의 특징을 분석하여 이러한 움직임이 어떠한 동작에 의해 발생했는지를 추정하고 인식하는 분야이다.

III. 영상분석을 통한 응급상황 검출시스템

1. 응급상황검출을 위한 시스템 구성

카메라를 통한 인간의 동작분석을 통해 응급상황을 검출해내고 그 결과를 전송함으로써 자동진단의 기능과 감시 장치의 역할을 수행하도록 한다. 실제로 영상 데이터만을 사용하여 응급상황을 판단하는 데는 여러 가지 제약이 있으나 검출된 결과를 각종 센서나 진단도구를 사용한 응급상황검출의 사전자료로서 사용한다. 영상 데이터를 분석하여 응급상황이 발생했을 가능성이 큰 상황을 정의하고 이러한 동작들을 검출함으로써 원격진단을 필요로 하는 대상을 선별하기 위한 방법을 제시하고자 한다. 응급상황에 대한 동작 분석결과가 인지되면 이를 서버에 전송하여 응급상황을 알리고 입력된 영상 데이터를 전송한다. 메인서버에서는 수신된 영상 데이터와 각종센서의 상태를 종합하여 응급

상황의 검증을 통해 필요한 조치를 취하도록 구성되어 있다. [그림 2]는 응급상황을 판별하기위한 동작 분석과정과 전체 시스템의 구성을 나타내고 있다.

카메라의 위치는 원하는 작업 공간 안에서 가능하면 전신의 모습이 포함되도록 설치하여야 한다. 기절이나 기타 응급상황에 의해 발생하는 비정상 동작의 검출을 목적으로 하므로 화면분석이 행해지는 공간은 서있거나 앉아있는 등 일반적인 작업공간이나 활동영역으로 제한한다. 이는 누워있거나 수면상태의 경우 현실적으로 화면만을 통해 대상의 응급상황을 구분할 수 없고 부가적인 맥박이나 혈압 등의 측정을 통해 알 수가 있다.

본 논문에서는 신체의 각 부분을 분리하여 다양한 움직임을 분석하는 대신 세 가지의 동작상태만을 가정하여 서있는 상태(stand), 쓰러진 상태(faint), 앉아있는 상태(squat)만을 구분하는 것을 목적으로 한다.

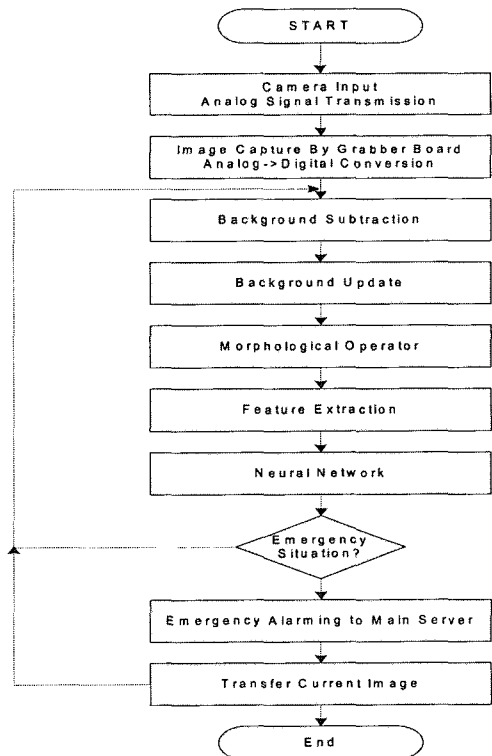


그림 2. 시스템 구성

이러한 동작 상태를 구분하기 위해서 카메라에 입력된 영상을 취득하기 위한 영상입력장치를 필요로 하고 입력된 영상을 프레임별로 저장한다. 저장된 영상을 배경화면과 비교하여 객체부분을 분리해낸다. 배경모델과 현재의 입력화면을 감산(subtraction)하여 객체영역을 추정하고 배경모델을 갱신한다. 객체의 분리를 위해 입력영상의 에지를 강조하고 형태학적 연산자(morph-ological operator)를 통해 객체 부분만을 분리해낸다. 분리된 객체의 영역으로부터 동작을 구분해 낼 수 있도록 특징벡터를 생성한 후 생성된 특징벡터를 이용해서 동작을 분석하기 위해 신경망을 사용한다. 대표적인 동작의 특징벡터를 생성한 후 이를 바탕으로 신경망을 학습시킨다. 학습된 신경망의 가중치를 이용해서 임의의 입력에 대한 특징벡터로부터 결과를 구해내고 이로부터 동작을 구분해 낸다.

동작분석을 위한 전체 알고리즘은 배경감산에 의한 객체분리, 수평과 수직 사영에 의한 특징벡터의 추출, 신경망에 의한 동작분석의 세 단계로 구성된다.

2. 배경감산(background subtraction)에 의한 객체 분리

본 논문에서는 배경화면과 현재 입력된 화면의 차이를 구하여 차이 값이 임계치 이상일 때 객체(object)영역으로 판단하는 기법을 사용한다. 목표로 하는 객체나 환경에 대한 사전 지식이 없는 상황에서 고정된 카메라를 사용한 움직임 검출에 가장 널리 사용되는 방법이 배경 감산법(background subtraction)이다. 연속된 장면에서 움직이는 객체는 현재의 입력 프레임과 객체가 포함되지 않은 배경화면(배경모델)의 차를 구함으로써 검출할 수 있다[8-11]. 그러나 배경모델도 시간이 지남에 따라서 밝기나 그림자의 위치 등이 변하게 되므로 이러한 배경화면의 변화를 반영하기 위해 입력된 영상으로부터 배경영역을 추정하고 갱신해 나가는 과정을 거치게 된다.

대부분의 배경감산을 이용한 객체분리는 화소단위의 갱신과정을 포함한다[8-10]. 입력된 영상으로부터 각 화소가 배경에 포함되는지 혹은 움직이는 객체의 일부 인지를 판단하여 배경모델의 갱신을 하거나 객체영역

으로 판별한다. 연속된 프레임에서 움직이는 객체는 현재의 프레임과 배경모델의 차를 구함으로써 검출할 수 있다. 이때 점 p에서 입력영상의 화소 값이 $I(p)$, 배경모델의 화소 값이 $B(p)$ 이고 배경감산에 의한 배경차를 $DB(p)$ 라하면

$$DB(p) = d(I(p), B(p)) \quad (1)$$

이고 이때 Distance는 RGB 칼라공간에서 다음과 같은 L-inf 거리를 사용한다.

$$d(x_i, x_j) = \max(x_i, c - x_j, d), \quad c = R, G, B \quad (2)$$

배경모델의 갱신은 시간에 따른 배경화면의 갱신을 위해 반드시 필요하며 갱신을 위해 사용되는 화면의 수와 갱신 규칙에 따라 성능이 달라진다. 어떤 시간 t에서 과거에 입력된 n개의 영상으로부터 해당화소의 갱신여부를 결정하기 위해 어떤 점 p가 객체에 포함되지 않는다면 다음과 같은 화소 공간 S에서 갱신될 화소의 값을 결정한다.

$$S = \{I^{t-n\Delta t}(p), I^{t-(n-1)\Delta t}(p), \dots, I^t(p)\} \cup \{B^t(p)\} \quad (3)$$

이것은 n개의 과거 화소 값과 현재의 배경모델의 화소 값으로부터 갱신될 화소 값이 결정됨을 의미한다. 이때 갱신될 다음 화소 값은 식4와 같이 S 공간 안에서 메디안 중간 값에 의해 결정한다.

$$B^{t+\Delta t}(p) = \arg \min_{i=1, \dots, k} \sum_{j=1}^k d(x_i, x_j), \quad \text{with } x_i, x_j \in S \quad (4)$$

배경감산에 의한 배경차 영상은 임계치와 비교하여 이치화 한다. 객체의 복장이나 색깔의 영향을 줄이기 위해 이치화된 영상을 이용하여 객체영역을 분리한다.

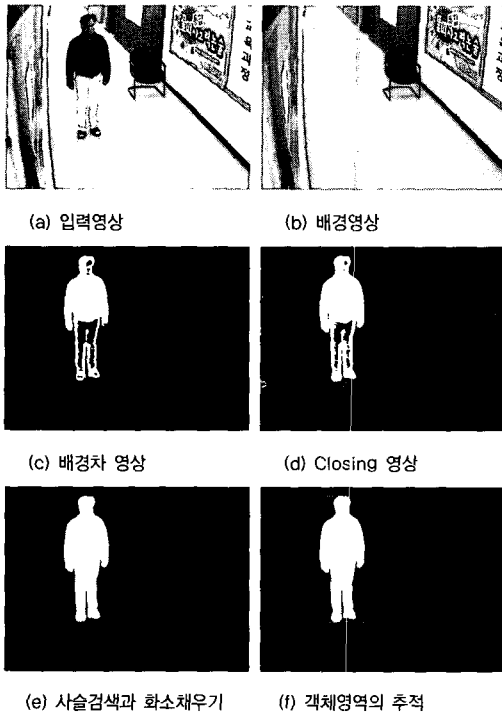


그림 3. 객체 분리과정

그러나 이치화 된 결과는 카메라의 미세한 환경변화에도 민감하게 반응하므로 객체영역의 인식 시 노이즈가 많이 생기게 된다. 이러한 문제점을 해결하기 위하여 형태학적 연산자인 closing 함수를 사용하여 끊어진 화소들을 연결하고 외곽 부분의 화소에 대해 사슬탐색(chain search) 알고리즘을 사용하여 정해진 크기 이하의 영역을 삭제한다. 삭제되고 남은 영역에 화소 값을 채워 객체영역을 설정한다. [그림 3]은 처리과정을 나타내고 있다.

3. 수평수직 투영에 의한 특징추출

검출된 객체의 영역으로부터 사람의 동작을 나타낼 수 있는 특징벡터로 변환을 한다. 본 논문에서는 신경망을 통해 응급상황을 판단하는 것을 목적으로 하므로 이를 구분할 수 있는 간단한 특징벡터를 선택하였다. 세 가지의 기본적인 동작 상태를 정의하고 이를 분리하기위해서 객체영역의 화소를 수평과 수직으로 투영(projection)하여 화소의 배치상태를 정의하였다. 수직

수평 투영에 입력화면의 각 화소를 사용하면 객체의 복장이나 색상에 영향을 많이 받으므로 객체영역과 배경영역을 분리하여 이치화 시킨 후 객체영역에 포함되는 화소의 수만을 계수하여 투영(projection)한다. 이 경우 사람의 머리카락, 다리 등을 정확히 구분할 수는 없으나 팔, 다리가 겹쳐지거나 구부리는 등 대략적인 정보를 알 수 있으므로 투영된 객체의 화소수를 특징벡터로 사용한다. 또한 객체영역의 세로 폭과 가로 폭의 비율을 계산하여 특징벡터로 사용한다. [그림 4]는 분리된 객체영역과 수평, 수직 투영에 의한 객체의 화소수를 나타낸다.



그림 4. 수평투영과 수직투영

그러나 객체의 크기나 영역에 따라 화소의 개수가 달라지므로 정규화(normalization)를 통해 크기에 상관없이 모양에만 영향을 받고 특징벡터의 개수가 일정하도록 해 주어야 한다. 이를 위해서 수평 수직으로 투영된 값을 일정한 크기의 벡터로 스케일 해 주어야한다. 여기서는 수평, 수직 공통적으로 16개의 벡터로 정규화하여 처리하며 이때 실제 화소의 라인번호가 정규화된 특징벡터의 위치와 일치할 수가 없으므로 쌍선형보간법(bilinear interpolation)에 의해 16개의 특징벡터로 정규화 한다. 특징벡터의 각 요소는 0에서 1사이의 값을 가지도록 하기위해 계수된 화소수를 전체화소의 개수로 나누어 주어야한다. [그림 5]는 특징벡터의 정규화 과정이며 가로, 세로크기가 각각 43, 59일 경우의 예이다. 수평투영에 의해 나타나는 59개의 특징벡터는 정규화 과정을 통해 16개의 정규화 특징벡터로 변환하기 위해서 $3.7(59/16=3.7)$ 개의 특징벡터가 1개의 정규화 특징벡터로 변환되어야 한다. 이를 위해 [그림 5]와 같이

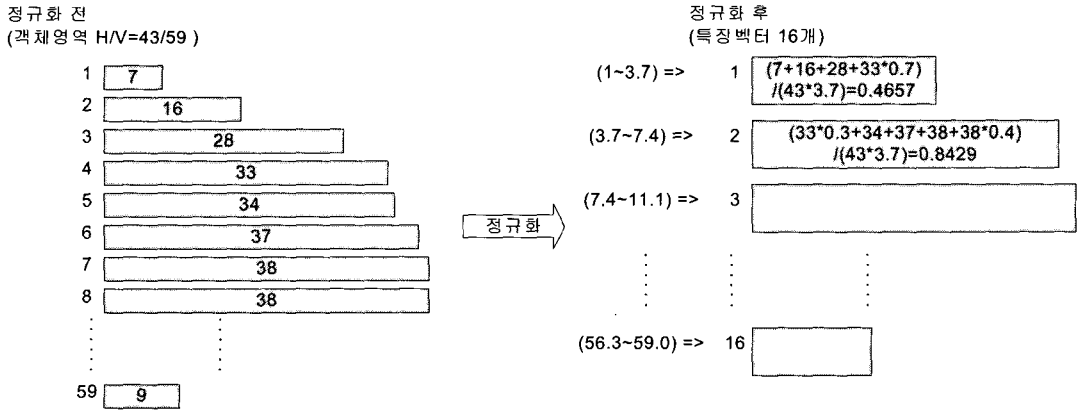


그림 5. 특징벡터의 정규화

3라인의 화소수를 더하고 소수점이하에 대응하는 라인 값에 0.7을 곱한 값을 더해준다. 또한 결과 값이 0과 1 사이의 값을 가지도록하기위해 나타날 수 있는 최대화 소수(43*3.7)로 나누어줌으로써 0에서 1사이 값을 가지는 정규화특징 벡터를 구할 수 있다.

4. 신경망을 이용한 동작분석

분리된 객체영역으로부터 동작을 분석하기위해 신경망을 사용한다. 신경망의 구성은 추출된 특징벡터를 입력으로 한다. 본 논문에서는 응급상황을 인지하기위한 시스템 구성을 목적으로 하기 때문에 분석을 위한 행동 양식을 제한하였다. 정상적인 동작상태, 기절이나 기타 응급상태로 인한 이상동작상태, 의자에 앉아있는 준 정지 상태 등으로 구분하였다. 이러한 3가지의 동작상태를 구분하기 위해서 본 논문에서 사용한 특징벡터는 수평방향의 투영벡터, 수직방향의 투영벡터, 객체영역의 세로/가로 비율 등을 사용하였다.

4.1 신경망의 구조

본 논문에서 사용한 특징벡터는 세로/가로 비율, 16개의 수평방향 투영벡터, 16개의 수직방향 투영벡터 등이다. 총 33개의 실수 값을 입력으로 하여 원하는 동작상태를 분석할 수 있어야 한다. 구분하기를 원하는 출력 상태는 정상 동작상태(stand)와 이상동작상태(faint), 준 정지상태(squat)등 3가지의 동작이므로 본

논문에서는 다층 퍼셉트론(multilayer perceptron-MLP)을 사용하여 표준 동작에 대한 특징 벡터를 정의하고 이를 학습시켰다. MLP의 입력층(input layer)는 10개의 노드, 은닉층(hidden layer)은 10개의 노드, 출력층(output layer)은 3개의 노드를 사용하여 구성하였고 출력의 3노드에 stand, faint, squat로 할당하여 가장 유사한 동작이 입력되면 해당 노드의 출력이 1에 가까운 출력이 나오도록 설정하였다. [그림 6]은 본 논문에서 사용한 MLP의 구조이다.

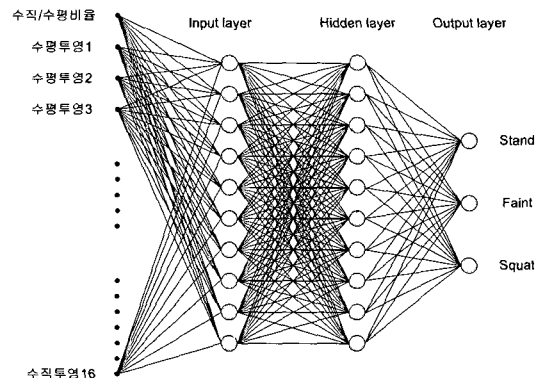


그림 6. 다층 퍼셉트론의 구조

4.2 신경망의 학습

신경망의 학습을 위해 대표적인 동작패턴을 정의하고 이에 대한 출력결과를 설정하여 학습시킨다. 동작의

종류별로 다수의 입력패턴을 추출한 후 신경망의 입력으로 사용하였으며 대응하는 출력을 지정하여 역전파(back-propagation) 알고리즘에 의해 노드사이의 연결강도를 나타내는 가중치(weight)를 갱신한다. 전체 입력패턴에 대한 신경망의 출력 결과와 원하는 출력결과사이의 오차 계산하여 임계치 이하가 될 때까지 학습을 반복하고 학습이 완료되면 레이어별 가중치를 저장한다. 학습과정에서 얻어진 가중치를 이용하여 다층 퍼셉트론의 구조를 구현하고 임의의 특징벡터가 입력되면 이를 입력으로 하여 신경망의 출력 값을 계산한다. 출력 값의 분포로부터 어떤 종류의 표준패턴에 가장 가까운지 결정함으로써 동작을 분류한다.

본 논문에서는 크기와 상태를 달리하는 표준패턴을 정의하기 위해 stand 에 해당하는 입력패턴 10개, faint 에 해당하는 입력 패턴 10개, squat에 해당하는 입력패턴 10개 등을 선정하고 이로부터 특징 벡터를 구하여 신경망을 학습시켰다. [그림 7]은 신경망의 학습을 위해 사용한 동작패턴의 예이다.

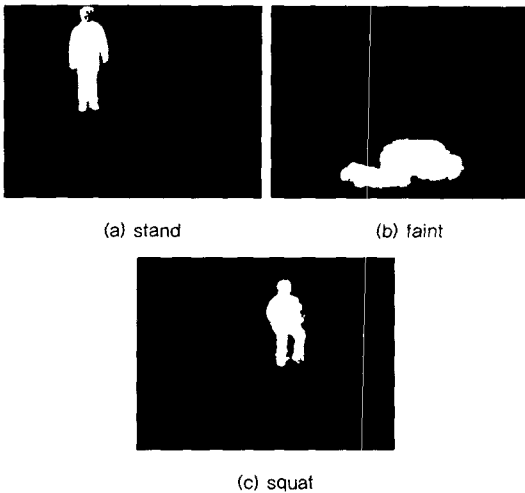


그림 7. 학습에 사용한 동작패턴

IV. 실험 및 고찰

본 논문에서 제안한 동작분석 알고리즘은 실내에 설치된 카메라로부터 입력된 영상을 이용하여 테스트하

였다. 감시영역 안에서 사람의 전신이 영상에 포함되도록 카메라를 설치하였고 초당 15프레임의 영상을 640*480의 해상도를 가지는 RGB 칼라의 영상을 입력으로 사용하였다. 초당1장의 프레임을 메모리에 저장하고 10장의 저장된 프레임을 사용하여 배경모델의 갱신에 사용하였다. 동작분석 알고리즘은 초당 3장의 영상에 대해서 처리를 하며 결과를 분석하고 응급상황이 발생했을 때 저장된 화면들을 전송한다. 즉 3장의 영상을 분석하여 결과를 출력하고 이 중 1장의 화면은 배경모델의 갱신을 위해 프레임 메모리에 저장한다. 프레임 메모리는 링 버퍼를 구성하여 FIFO 형태의 버퍼를 구현함으로써 최근 10초 동안의 영상을 1초 간격으로 저장하도록 한다.

동작 패턴은 서있는 자세(stand), 쓰러진 자세(faint), 앉아있는 상태(squat) 등 세 가지의 동작을 설정하고 신경망을 학습시켰다. 각 동작별 10가지의 영상을 선정하여 특징벡터를 추출하고 이 표준패턴을 입력으로 하여 신경망을 학습시켰다. faint 영상의 경우 좌우의 방향이 입력패턴에 영향을 미치므로 같은 영상을 좌우로 뒤집은 2가지의 영상을 만들어 총 20가지의 학습패턴을 사용하였다. [그림 8]은 특징벡터를 신경망의 입력으로 하여 출력 층에 나오는 세 노드의 출력 값의 변화를 프레임에 따라 나타낸 결과이다.

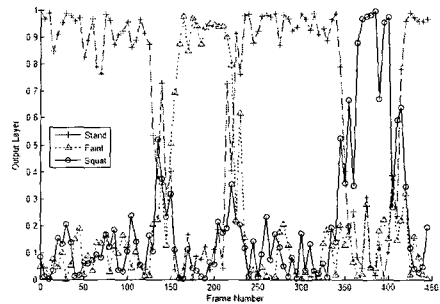


그림 8. 프레임에 따른 신경망의 출력

프레임의 앞부분에서는 서있거나 걸어서 이동하는 동안의 프레임에 따른 신경망의 출력을 나타내고 120 프레임 근처에서 faint 동작의 발생, 220 프레임 부근에서 다시 일어나 다른곳으로 이동함을 나타내고, 마지막

에는 의자에 앉을 때의 출력 값의 변화를 표시하고 있다. 3개의 출력 값 중 가장 크게 나오는 값을 기준으로 동작 상태를 결정해 나간다.

[표 1]은 동작 상태에 따라 구분된 정지영상을 입력으로 하여 출력되는 신경망의 출력을 바탕으로 인식된 결과를 나타내고 있다. 카메라에서 객체의 크기와 동작 상태를 달리한 여러 화면을 저장하고 이를 입력으로 하여 화상처리과정을 거쳐 객체영역을 분리한다. 신경망의 입력을 위해 특징벡터를 추출한 후 신경망에 입력한 결과를 표시하고 있다. [표 1]에서 오인식된 결과는 주로 동작의 중간상태 화면, 즉 서있는 상태에서 쓰러지기 직전의 상태를 입력으로 하거나 객체영역 분리과정에서 정확한 분리가 되지 않아서 일부 신체부위가 침식되어 객체분리가 정확히 되지 않은 경우에 발생했다. Stand와 Squat의 경우 앉거나 구부리고 있는 상태에 따라서 유사한 패턴으로 나타나는 경우가 있었으며 팔을 벌리거나 그림자에 의한 객체영역이 확장된 경우, 의자나 기타 부속물이 객체영역에 포함된 경우 오인식되는 경우가 있었다.

표 1. 동작상태에 따른 인식결과

동작상태	입력 프레임 수	분류결과			인식률 (%)
		Stand	Faint	Squat	
Stand	32	30	0	2	93.7
Faint	40	0	39	1	97.5
Squat	14	1	0	12	85.7

V. 결론

본 논문에서는 원격 응급상황 검출 및 처리를 위해 화상인식을 통한 동작분석 알고리즘을 제안하였다. 카메라로부터 입력된 영상으로부터 통계적 배경모델을 사용한 배경 감산법에 의해 객체 영역을 분리해내고, 분리된 객체로부터 동작분석을 위한 특징벡터를 추출한다. 응급상황 검출을 위해 검출하려는 동작을 제한하고 이를 위한 특징벡터의 설정과 신경망을 사용한 동

작 분류 알고리즘을 구현하였다. 실험결과 간단한 연산에 의해 동작 상태를 분류해 낼 수 있었고 효과적으로 응급상황과 기타 제한된 동작 상태를 자동 분류 및 처리할 수 있는 시스템을 구현하였다. 인식성능의 개선을 위해 배경부분과 객체부분을 정확히 분리함으로써 신체부분의 침식현상을 줄이고 동작분석의 성능을 개선할 수 있을 것으로 생각된다.

참고 문헌

- [1] J. K. Aggarwal and Q. Cai, "Human Motion Analysis: A Review," *Computer Vision and Image Understand*, Vol.73, No.3, pp.428-440, Mar. 1999.
- [2] C. Rigotti, P. Cerveri, G. Andreoni, A. Pedotti, and G. Ferrigno, "Modeling and Driving a Reduced Human Mannequin through Motion Captured Data: A Neural Network Approach," *IEEE Trans. on System, Man, and Cybernetics-part A*, Vol.31, No.3, pp.187-193, May 2001.
- [3] Z. Chen and H. J. Lee, "Knowledge-guided Visual Perception of 3D Human Gait from a Single Image Sequence," *IEEE Trans. on System, Man, and Cybernetics*, Vol.22, No.2, pp. 336-342, 2005.
- [4] R. Cucchiara, C. Grana, A. Parati, and R. Vezzani, "Probabilistic Posture Classification for Human-Behavior Analysis," *IEEE Trans. on System, Man, and Cybernetics-part A*, Vol.35, No.1, pp.42-54, Jan. 2005.
- [5] P. Spagnolo, M. Leo, G. Attolico, and A. Distanto, "Posture Recognition in Visual Surveillance of Archaeological Sites," *IEEE Proceedings of International Conference on Intelligent Robots and Systems*, Vol.2, pp.1542-1547, Oct. 2003.
- [6] Lim Hock Wyi Aloysiu, Guo Dong, Huang

Zhiyong, and Tele Tan, "Human posture recognition in visual sequence using pseudo 2-D hidden markov models," *IEEE Proceedings of International Conference on Control, Automation, Robotics and Vision*, Vol.1, pp.712-716, Dec. 2004.

[7] K. Takahashi and S. Sugakawa, "Remarks on Human Posture Classification Using Self-Organizing Map," *IEEE Proceedings of International Conference on System, Man and Cybernetics*, Vol.3, pp.2623-2628, Oct. 2004.

[8] R. Cucchiara, C. Grana, A. Parati, and R. Vezzani, "Computer vision system for in-house video surveillance," *IEEE Proc-Vis. Image Signal Process.*, Vol.152, No.2, pp.242-249, Apr. 2005.

[9] D. Gutchess, M. Trajkovic, E. Cohen-Solal, D. Lyons, and A. K. Jain, "A Background Model Initialization algorithm for Video Surveillance," *IEEE Proceedings of International Conference on Computer Vision*, Vol.1, pp.733-740, July 2001.

[10] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting Moving Objects, Ghosts and Shadows in Video Streams," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, No.10, pp.1337-1342, Oct. 2003.

[11] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4:Real-Time Surveillance of People and Their Activities," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.22, No.8, pp.809-830, Aug. 2000.

[12] M. K. Leung and Yee Hong, "First Sight: A Human Body Outline Labeling System," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.17, No.4, pp.359-377, Apr. 1995.

저자 소개

이 동 규(Dong-Kyu Lee)

정회원



- 1991년 2월 : 한양대학교 공과대학 학사
- 1993년 2월 : 한양대학교 전자공학 석사
- 2003년 8월 : 한양대학교 전자공학 박사

• 2004년 ~ 현재 : 한북대학교 컴퓨터공학과 전임강사
 <관심분야> : 디지털신호처리, 영상처리, 멀티미디어 통신

이 기 정(Ki-Jung Lee)

정회원



- 1999년 2월 : 서울시립대학교 국사학과 학사
- 2003년 2월 : 경원대학교 전자계산학과 석사
- 2004년 3월 ~ 현재 : 경원대학교 전자계산학과 박사과정

<관심분야> : 신경망, 멀티미디어 검색, 시맨틱 웹, 내용기반검색

황 보 택 근(Taeg-Keun Whangbo)

정회원



- 1983년 2월 : 고려대학교 공과대학 학사
- 1987년 8월 : CUNY 전산학과 석사
- 1995년 8월 : S.I.T. 전산학과 박사

• 1995년 ~ 1997년 : 삼성종합기술원 선임연구원
 • 1997년 ~ 현재 : 경원대학교 소프트웨어대학 부교수
 • 1983년 2월 : 고려대학교 공과대학 학사
 • 1987년 8월 : CUNY 전산학과 석사
 • 1995년 8월 : S.I.T. 전산학과 박사
 • 1995년 ~ 1997년 : 삼성종합기술원 선임연구원
 • 1997년 ~ 현재 : 경원대학교 소프트웨어대학 부교수
 <관심분야> : 영상처리, 데이터마이닝, 정보가시화, 컴퓨터그래픽스, 영상처리, 모바일

임 혁 규(Hyuk-Kyu Lim)

정회원



- 1991년 2월 : 한국방송대학 전자계산학과 학사
- 1994년 2월 : 한양대학교 산업대학원, 전자계산과 석사
- 1998년 4월~2000년 8월 : 한국건설기술연구원 선임연구원

- 2001년 6월~2004년 2월 : 브이앤아이(주) 대표이사
- 2004년 2월~현재 : 브이앤아이(주) 이사
- 2003년 3월~현재 : 경원대학교 소프트웨어대학 겸임교수

<관심분야> : 멀티미디어 처리, LBS, 영상처리