

비디오 기반의 독거노인 위급 상황 탐지를 위한 행동 분석

Motion analysis for Home Surveillance of the Aged who Lives Alone based on Video Images

이 상 용*, 김 영 백

(Sang-Yong Rhee and Young-Baek Kim)

Abstract : In this paper, motion analysis algorithm is presented for home surveillance of the aged who lives alone. For the first step, we acquire images from a camera. To enhance the image, we use median filtering and binarize it to reduce processing time. And then morphological operations are performed to remove small blobs and small holes. At the fourth step, blobs are analysed to extract foreground region. Then, motions are predicted from these images by using optical flow technique, and the predicted motion data are refined by comparing our cardboard models so as to judge behavior pattern.

Keywords : motion analysis, card board model

I. 서론

경제 성장, 의료기술의 발달, 그리고, 생활수준향상으로 사람의 평균수명이 증가하고, 핵가족화가 진행되고 있다. 이 결과 단독 노인세대가 급증하고 노인이 사회적으로도 소외되고 있다. 한국의 노인 인구는 2000년에 전체 인구의 7.4%이며, 2019년에는 14.4%로 증가될 전망이다[1]. 현재, 24시간 보호가 필요한 노인 환자의 수는 매년 증가하고 있는 것에 비해, 노인 환자를 위한 종합적인 관리체계가 미흡하고, 전문적인 병원 및 요양소, 보호시설도 전국적으로 매우 부족한 상황이다.

노인 환자 간병은 대부분 가족들이 담당하고 있으며, 하루 24시간 계속되는 노인 환자의 관리는 가족들의 사회적, 경제적 활동에 매우 심각한 장애 요인이 되고 있다. 따라서 부족한 의료 시설 및 요양소 시설을 대신하여 노인 환자가 있는 가정의 원활한 사회적, 경제적 활동을 지원할 수 있도록 노인 환자 관리 시스템이 요구되고 있다.

본 논문에서는 이러한 노인 환자 관리 시스템들이 필요로 하는 기능 중에 하나인 노인의 행동을 분석하는 알고리즘을 개발하고 구현하는데 목적이 있다.

II. 관련 연구

컴퓨터를 사용하여 인간의 행동을 분석한다는 것은 인체 각 부위가 시간의 흐름에 따라 어떤 모습으로 변화하는가를 자동으로 분석하고 그 변화를 추상적인 의미로 해석하는 것을 의미한다[8]. 즉, 동영상으로부터 신체 영역을 추출한 다음 특정 부분을 식별하고 각 부분들이 하나의 의미를 갖기 위해 어떤 변화를 거치는지를 알아내는 것이다. 인간 동작 분석을 위한 가장 이상적인 방법은 분석 대상에 마커를 붙이지 않고 자연스럽게 사람이 눈으로 대상 인간을 관

찰하듯 비디오카메라로 입력된 영상만을 분석하여 그 의미를 알아내는 것이다. 일반 비디오 영상으로부터 행동 분석하는 방법은 크게 3가지로 나눌 수 있다.

첫 번째 방법은 신체의 특징점을 추출하여 그 특징점의 운동을 시간에 대한 패턴으로 구별하여 인식하는 방법이다. 이 경우 대개 영상이 주어지면 전처리 과정을 거쳐 특징점을 추출하고 이 특징점을 추적하여 자세를 추정하고 이 자세의 시간적인 변화를 동작의 의미로 해석하는 방법이다[5,9,10]. 이 방법은 복장이나 카메라의 시선방향에 따라 특징점이 달라지고, 복잡한 움직임의 경우 특징점들이 서로 겹치기 때문에 특징점 추출이 어려워 고도의 인식은 불가능하다. 따라서 머리, 양 손 끝, 양발의 위치 등 가장 안정적으로 추출할 수 있는 특징점들을 이용하여 제스처를 추정하는 경우가 많다[5,10]. 두 번째 방법은 인체의 형상 모델을 이용하는 방법이다[11]. 미리 일반화된 인체의 2차원 혹은 3차원 모델을 만들어 놓고 그 모델을 적당한 가설에 의해 변형시켜 2차원 영상 공간으로 투영한 다음 입력된 영상과 생성된 영상을 비교하여 인체의 자세를 인식하는 방법이다. 이 방법은 다양한 형상 모델 때문에 탐색공간이 커져 계산량이 많아지는 단점이 있다. 따라서 원통과 2차원 곡면과 같은 단순 기하구조를 이용하여 계산량을 줄이는 방법이 많이 연구되고 있다[11]. 세 번째 방법은 영상에 나타나는 인체의 전체적인 모습(whole bodily appearance)을 여러 가지 매개변수(parameter)를 이용하여 표현하고 그 변화를 해석하여 동작을 인식하는 방법이 있다[12,13]. 이 방법은 특징점을 추출하지 않고 대개의 경우 인체 부분을 이진 영상으로 변환하여 의복에 의한 영향을 제거한 후, 그 시계열 영상이 갖는 통계적 특성이나 기하학적 정보를 분석하여 동작을 인식하는 방법이다. 인체의 복잡한 특징점을 추출하지 않고 전체의 외관을 하나의 분석 대상으로 삼기 때문에 잡음과 환경변화에 영향을 받지 않아 구현이 쉽고 인식결과가 안정하다는 장점이 있으나 상세한 행동 분석이 안된다는 단점이 있다.

* 책임저자(Corresponding Author)

논문접수 : 2007. 1. 25., 채택확정 : 2007. 2. 17.

이상용, 김영백 : 경남대학교 컴퓨터공학부

(syrrhee@kyungnam.ac.kr/baroaleum@nate.com)

III. 광류

광류(optical flow)는 사람의 망막에 맺히는 명암의 흐름 정보에 의해 움직이는 물체를 판단한다는 생각에 착안하여 이론화한 것으로서 관측자의 움직임이나 물체의 움직임 또는 관측자와 물체의 움직임에 의해, 연속된 영상에서 발생하는 명암 패턴의 명확한 움직임을 뜻한다. 즉, 영상에서 픽셀의 운동방향과 속도를 벡터로 표현한 것이다[3]. 광류는 이동하는 물체와 관측자의 상대적인 움직임만을 나타내므로 움직임이 있는 영상의 해석에 유용하고 추적 시스템을 구현하는데 적합하다. 따라서 관측자의 이동으로 생기는 배경의 변화를 배제하면서 이동하는 물체의 속도성분을 추출함으로써 이동하는 관측자가 이동물체를 추적할 수 있다.

광류 추정 방법은 물체의 각 점에서 관측된 밝기는 시간에 관계없이 항상 일정하다는 가정에 근거를 두고 있다. 따라서 어떤 화소에서 발생하는 밝기의 변화는 전적으로 운동의 영향 때문이라 가정한다[4].

시간 t에서 (x,y)좌표에 존재하는 한 점의 밝기 값을 I(x,y,t)라 하고, 짧은 시간 간격 δt 후에 이 점이 수평 및 수직 방향으로 δx 및 δy 만큼 이동하였다면, 이동된 점의 밝기는 (1)과 같다.

$$I(x+\delta x, y+\delta y, t+\delta t) \tag{1}$$

위 식의 우변을 Taylor 시리즈로 전개하면 (2)와 같다.

$$I(x+\delta x, y+\delta y, t+\delta t) = I(x,y,t) + \delta x \frac{\delta I}{\delta x} + \delta y \frac{\delta I}{\delta y} + \delta t \frac{\delta I}{\delta t} + h.o.t \tag{2}$$

여기서, h.o.t는 Taylor 시리즈에서 생기는 고차항이다. 만약 δx , δy , 및 δt 가 매우 작은 값이라면 위 식의 고차항은 제거될 수 있으므로 (3)과 같이 된다.

$$\delta x \frac{\delta I}{\delta x} + \delta y \frac{\delta I}{\delta y} + \delta t \frac{\delta I}{\delta t} = 0 \tag{3}$$

다시 양변을 δt 로 나누면 (4)와 같이 된다.

$$\frac{\delta x}{\delta t} \frac{\delta I}{\delta x} + \frac{\delta y}{\delta t} \frac{\delta I}{\delta y} + \frac{\delta I}{\delta t} = 0 \tag{4}$$

이때 $\frac{\delta x}{\delta t} = u$, $\frac{\delta y}{\delta t} = v$ 로 놓으면 (5)와 같이 된다.

$$\frac{\delta I}{\delta x} u + \frac{\delta I}{\delta y} v + \frac{\delta I}{\delta t} = 0 \tag{5}$$

따라서 (5)는 (6)과 같다.

$$I_x u + I_y v + I_t = 0 \tag{6}$$

(6)을 다르게 표현하면 (7)과 같다.

$$\nabla I(x,t) \cdot v + I_t(x,t) = 0 \tag{7}$$

이를 광류 제약 방정식이라 한다[4]. 여기서 $v = (u, v)^T$ 는 속도, $\nabla I(x,t) = (I_x(x,t), I_y(x,t))^T$ 는 공간적인 밝기 변화를, 그리고 $I_t(x,t)$ 는 시간적인 밝기 변화율을 나타낸다.

하지만 (7) 하나만으로 두 개의 변수를 갖는 속도 v를 구한다면, 부정해만 구할 수 있다는 문제점이 있다. 따라서 광류를 추정하기 위해서는 부가적인 제약조건이 필요하다.

Horn 등[4]은 속도장이 전역적으로 평탄하다는 제약조건을 부가하여 광류를 추정하였고, Lucas 등[5]은 속도장이 국부적으로 평탄하다는 제약조건을 부가하여 광류를 추정하였다.

IV. 노인의 행동 분석 알고리즘

행동 분석 알고리즘의 전체적인 흐름은 그림 1과 같은데 고정된 카메라로부터 영상을 입력받고 입력되는 영상에 대하여 노이즈제거를 위한 전처리를 수행하고, 모델링되어 있는 배경과 비교연산을 통하여 이진화를 수행한다. 이렇게 이진화된 영상에서 다시 잡영을 제거하기 위하여 형태론적 연산(morphological operation)인 열림(opening), 닫힘(closing) 연산을 수행하고, 얼룩(blob)을 추출하고 이 얼룩들을 분석하여 노인영역을 검출한다. 만약 노인이 없으면 모델링된 배경을 업데이트 시키고 노인이 검출되었다면 노인영역에 대하여 부위별 검출을 시도하고 부위별로 검출된 데이터를 이용하여 행동을 분석한다.

1. 노이즈 제거를 위한 영상처리

방안에서 생기는 자체 노이즈, 형광등에서 방출하는 열 전자에 의한 노이즈, 카메라 케이블에 의한 노이즈는 대부분 고, 저주파 성분의 값을 가지므로 미디언 필터를 사용하여 필터링을 한다.

2. 배경 모델링

영상의 각 픽셀당 세가지 값으로 배경을 모델링한다. 모델링되는 값은 학습에 사용된 모든 영상에 대한 최소, 최대 밝기값과 각 학습 영상의 인접한 두 프레임간의 최대 차이값이다.

배경 모델링은 지속적으로 업데이트 되는데 입력된 영상에서 어떠한 오브젝트도 발견되지 않은 경우 원형 큐에 배경 데이터가 입력된다. 따라서 기존의 데이터는 제거되고 가장 최신의 데이터만 큐에 남게 되어 배경에 조명상태 등의 변화가 오더라도 바로 바로 적용이 가능하다.

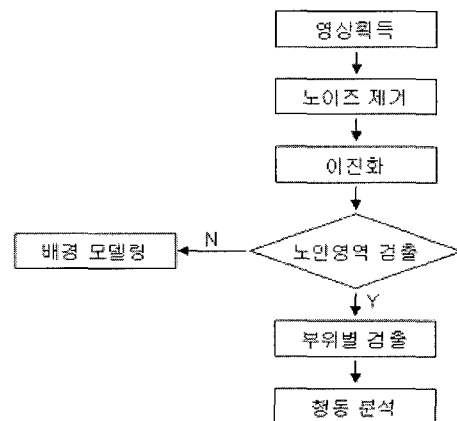


그림 1. 흐름도.

Fig. 1. Flow chart.

3. 노인 검출

본 연구에서는 배경 영역과 노인 영역을 분할하기 위하여 이진화, 형태학적 연산, 얼룩 영역과 배경 영역 분할, 노인 추적 이 네 가지 단계를 거친다.

3.1 이진화

입력된 영상(I)의 각 픽셀(x)과 모델링 된 배경 영상의 각 픽셀(x)에서 최소 밝기값(M), 최대 밝기값(N), 인접한 두 프레임 사이의 차이값(D)이 주어지면 (8)에 의하여 값이 0 과 1로만 이루어지는 이진화된 영상을 얻는다.

$$|M(x) - I(x)| > D(x) \text{ or } |N(x) - I(x)| > D(x) \quad (8)$$

3.2 형태학적 연산

이진화된 영상의 영역을 분할하기 위해 열림(opening)과 닫힘(closing)의 형태학적 연산을 수행한다.

3.3 얼룩 영역과 배경 영역 분할

유사한 성질을 가지고 모여 있는 화소들을 얼룩(blob)이라고 부르는데[6], 이진화된 영상에서 존재 하는 얼룩들을 검출하고 그 얼룩들의 위치와 크기 정보를 획득한다.

3.4 노인 추적

입력된 영상에서 노인 검출하는데 사용되는 데이터는 얼룩의 위치와 크기이다. 지속적으로 얻어지는 이 얼룩을 추적함으로써 노인에 대한 오검출을 막는다. 노인의 추적은 얼룩을 둘러싸는 박스를 치고 이 박스가 갑자기 나타났는지 이전에 있던 박스와 매치되는지 하나였던 박스가 여러 개로 나뉘었는지에 따라서 다른 처리과정을 거친다.

첫 번째로 이전에 없던 박스가 나타나면 이것이 노인인지 잠시 나타난 노이즈인지 알 수 없으므로 이 박스에 대한 정보를 저장하고 있다가 이 박스와 매치되는 박스가 연속해서 나오면 노인 영역으로 간주하고 그렇지 않으면 노이즈로 간주한다.

두 번째로 기존에 있던 박스와 현재의 박스가 매치되면 노인 영역으로 간주한다. 마지막으로 하나였던 박스가 여러 개의 박스와 매치되면 바로 이 얼룩들을 각각의 오브젝트로 인식하는 것이 아니라 기존의 박스정보를 저장하고 있다가 이 박스들이 다시 합쳐지면 이것이 노이즈로 인하여 하나였던 얼룩이 여러 개로 갈라진 것이라고 판단하고 그렇지 않은 경우는 새로운 노인이 영상에 진입하였다고 간주한다.

4. 노인의 부위별 검출

본 연구에서는 노인의 행동 패턴(걷다, 앉다, 서다, 눕다, 쓰러지다)을 인식하기 위하여 사람의 각 부위(머리, 몸, 넓적다리, 종아리)를 나누어서 검출하고 각 부위들의 위치정보를 행동 패턴을 인식하는 특징벡터로 사용한다. 노인의 각 부위를 검출하는 것은 Shanon X. Ju 등의 Cardboard Model[7]을 이용하여 구현하였다.

Cardboard Model은 사람의 형상을 평면의 박스 조각으로 표현 가능하다는 것을 가정한다. 그림 2는 Shanon X. Ju 등이 사람의 형상을 모델링한 것이다.

모델링 데이터는 다음과 같은 방법으로 얻어진다. 우선, 동작 하나 하나에 대해서 카드보드 모델을 만든다. 그리고 그것을 둘러싸는 박스를 지정한 후, 그 박스의 넓이와 높이

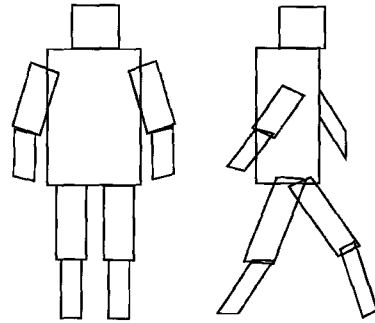


그림 2. 사람의 카드보드 모델.

Fig. 2. The cardboard person model.

를 기준으로 각각의 부위가 박스의 어디에 위치하는지를 저장함으로써 얻어진다.

다음으로 모델링된 데이터와 검출된 노인 영역을 비교하기 위하여 모델링된 데이터에 대하여 크기 정규화를 수행한다. 크기 정규화는 검출된 노인 영역을 둘러싸는 박스를 구하고, 그 박스와 카드보드 모델을 둘러싸는 박스를 넓이와 높이에 대하여 각각 1:1로 맞춤으로써 이루어진다. 여기서 노인영역에 포함되어있는 그림자가 문제가 될 수 있는데, 입력된 영상은 문제가 될 수 있는 그림자들을 미리 제거하였다.

마지막으로 정규화를 마친 모델링 데이터와 검출된 노인 영역을 2차원 영상공간으로 투영하여 비교하고 임계치 이상으로 매칭되는 모델링 데이터 가운데 가장 높은 매칭율을 보이는 모델링 데이터를 이용하여 노인의 부위별 데이터를 획득한다.

5. 움직임 추정

검출된 노인영역에서 각 부위를 검출을 하기위해 검출된 영역을 모델링된 모든 데이터와 순차적으로 비교해나가는 것은 매우 비효율적이고 시간이 너무 오래 걸린다. 따라서 본 연구에서는 이러한 비효율성을 제거하기 위하여 입력된 영상에 대하여 Horn 등[4]의 방법을 이용하여 광류를 계산하고 광류의 방향과 가속도의 분석을 통하여 다음 동작을 예측함으로써 모델링 데이터와 비교하는 경우의 수를 줄인다.

6. 행동 분석

걷다, 앉다, 서다, 눕다, 쓰러지다와 같은 행동들은 하나의 동작으로 이루어지는 것이 아니라 일련의 행동 패턴을 가진다. 따라서 연속된 5프레임에서 4프레임 이상 같은 행동 패턴의 동작이 나타나면 노인이 어떤 행동을 취했는지 결정한다.

V. 실험 및 결과분석

1. 실험환경

본 연구를 테스트한 환경은 마이크로 소프트사의 Window 2000 Professional 운영체제를 사용하고, 컴파일러는 Microsoft Visual C++ 6.0을 사용했다. 그리고 영상처리 라이브러리는 Matrox사의 Matrox Imaging Library 7.1을 이용하여 환경을 구축하였다. 하드웨어 사양은 Pentium IV 1.8

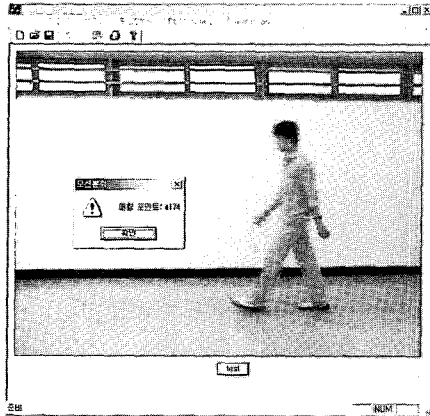


그림 3. 테스트 프로그램 전체 화면.

Fig. 3. Test scene of the algorithm.

GHz, RAM 512MB에서 프레임 그래버는 Matrox사의 Meteor II MC/4를 사용했다.

실험을 위해 사용된 영상은 Color Progressive Scan Camera 1기를 이용하여 27세의 남자 영상을 촬영하고 초당 5프레임의 영상을 320X240 해상도를 가지는 JPG 파일로 저장하여 사용 하였다. 그림 3은 알고리즘을 테스트하기 위해서 작성된 프로그램의 전체 화면이다.

2. 실험결과 분석

실험 결과 제안된 방법은 광류의 계산과 모델링 데이터와 검출 영역을 비교하는데 평균 0.75초가 걸려 실시간으로 노인의 행동을 분석하는 데는 실패 하였다. 그리고 모델링 되지 않은 돌발행동이 학습된 행동 패턴에 끼어들게 되면 대부분 인식에 실패하였다. 하지만 모델링된 각 행동 패턴에 대하여 10회씩 테스트한 결과 95% 이상 노인의 행동 분석에 성공하였다.

VI. 결론 및 향후 연구과제

본 논문에서는 비디오 영상에서 독거노인의 위급 상황 탐지를 위한 행동 분석을 위한 연구를 수행하고 알고리즘을 구현하였다.

카메라로부터 영상을 입력받고 노이즈 제거를 위한 필터링을 거친 후 이진화를 수행하여 영상에서 배경영역을 제외시킨 전경 영역을 얻고 형태론적 연산(열림과 닫힘)을 통해 분리된 점과 높은 피크들, 불필요한 구멍을 제거하였다. 입력된 영상에 대하여 광류를 추정하여 움직임을 예측하고 예측된 데이터를 기반으로 카드보드 모델 1과 비교하여 동작 정보를 얻어낸 후 연속된 프레임 정보를 이용한 일련의 행동 패턴을 판단하여 표시하였다.

구현된 알고리즘은 광류 추정을 위한 계산과 많은 모델링 데이터와의 비교연산을 위한 시간이 너무 길어 실시간 구현이 어려웠다. 따라서 보다 개선된 광류 추정 기법과 정확한 움직임 추정으로 모델링 데이터와의 비교연산을 줄인다면 실시간으로도 구현이 가능할 것으로 기대된다. 본 논문은 점차 노령화 추세의 사회에서 노인 복지를 위한 기술

로서 영상에서 인간의 행동 분석 알고리즘을 제시하는데 그 의미가 크다 하겠다.

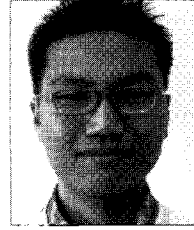
참고 문헌

- [1] 국립창원대학교 이동가족학과 노인문제 연구실: http://www.changwon.ac.kr/~goldage/sub3_menu03.html
- [2] A. Azarbayejani, C. Wren, and A. Pentland, "Real-time 3-D tracking of the human body," MIT Media laboratory Perceptual Computing Section Technical Report No. 374.
- [3] 박성현, "이동 카메라에 의한 이동 물체의 추적기법에 관한 연구," 석사학위논문, 부산대학교, 2000.
- [4] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, no. 1-3, pp. 185-204, Aug. 1981.
- [5] I. Haritaoglu, D. Harwood and L. S. Davis, "W4: Who? When? Where? What? A real time system for detection and tracking people," *International Conference On Face and Gesture Recognition*. 14-16, April 1998.
- [6] 김갑기, 김광인, 김항준, "신경망을 이용한 자동차 번호판 추출," *한국 정보과학회 논문집 제 26 권 제 2 호*, pp. 476-478, 10, 1999.
- [7] S. Ju, M. Black, Y. Yacoob. "Cardboard people: A parameterized model of articulated image motion," *International conference on Face and Gesture Analysis*. 1996.
- [8] V. I. Pavolovic, R. Sharma, and T. S. Hunag, "Visual interpretation of hand gestures for human-computer interaction: A review," *IEEE Transaction on PAMI*, vol. 19, no. 7, p. 677, July 1997.
- [9] I. Haritaoglu, D. Harwood and L. S. Davis, "Ghost: A human body part labeling system using shilhouettes," *International Conference on Pattern Recognition*, 1998.
- [10] J. Sherrah and S. Gong, "VIGOUR: A system for tracking and recognition of multiple people and their activities," *Proc. ICPR, Barcelona Spain*, vol. 1, pp. 179- 182, 2000.
- [11] D. M. Gavrila and L. S. Davis, "Towards 3D modelbased tracking and recognition of human movement: a multi-view approach," *Int. Workshop on Face and Gesture Recognition*, vol. 0000162479, pp. 272-277, 1995.
- [12] T. Watanabe and M. Yachida, "Real time recognition of gesture and gesture degree information using multi input image sequence," *ICPR*, vol. 1, p. 1855, 1998.
- [13] J. W. Davis, A. F. Bobic, "The representation and recognition of action using temporal templates," MIT Media Lab Technical Report 402, 1997.



이 상 용

1982년 고려대 산업공학과 졸업. 1984년 고려대 대학원 산업공학과(공학석사). 1992년 포항공대 대학원 산업공학과(공학박사). 1992년~현재 경남대학교 컴퓨터공학교수. 관심분야는 컴퓨터 비전, 뉴로-퍼지, 지능로봇, 생체인식.



김 영 백

2005년 경남대학교 컴퓨터공학과 학사. 2005년~현재 경남대학교 대학원 컴퓨터공학과 석사과정. 관심분야는 생체인식, 상황인식, 패턴분석.