

타원 모델링을 이용한 사람 머리 추적 시스템 구현

이명재*, 박동선*, 조재완**, 이용범**

*전북대학교 정보통신공학과 **한국원자력 연구소 로보트 기반기술 개발

E_mail : eagle@iceng.chonbuk.ac.kr

Human head Tracking system using the ellipse modeling

Myungjae Lee*, Dong-sun Park*, Jaiwan Cho** and Yong-bum Lee**

*Dep. of Info. and Comm. Eng., Chonbuk Nat'l Univ. **Advanced Robotics, KAERI

E_mail : eagle@iceng.chonbuk.ac.kr

Abstract

Recognizing a human part becomes very important for applications which are based on the interaction between computers and their users. In this paper, we design and implement a system which recognizes and tracks a human head using a sequence of images. Difference images are used to easily extract feature vectors from images with very complex backgrounds. A human head is represented with an ellipse and recognized by searching for a maximum value from preprocessed gradient images. The method is developed by considering the fact that the tracking system should be real-time. The designed system not only shows an excellent performance for the normal up-right position of the head, but also for the cases of 360° rotated head position, occluded images of heads, and tilted head positions.

적하였는데 확실적인 방법을 이용하여 추적을 수행하였으나 단순히 현재 위치만을 파악하였다.[2] Lee 등이 제시한 물체 추적 시스템은 적외선 영상에서 느리게 움직이는 진함을 추적하는 시스템으로 이 시스템은 다중 프레임 간출 개념을 이용하였다. Lee등이 제시한 시스템은 카메라의 추적 서보가 없으므로 능동적 의미의 추적 시스템의 범주에 들어가지 않으며 상관을 구하는데 추적장의 크기가 커지면 계산 소요 시간이 많아진다는 단점이 있다.[9] 미 공군에서는 표적으로 지정된 물체의 모양에 기반을 둔 추적을 수행하였는데 자기상관의 높이와 형태를 이용하였으며 예측은 하지 않았다.[2] 일본의 신일본 제철에서는 2차원 비전 센서를 이용한 고속 표적 추적 시스템을 개발하였는데 병렬처리를 수행하여 빠른 속도로 다양한 비전 작업들을 수행하였다.[2] 이 중에서 본 연구는 복잡한 배경에서도 손쉽게 원하는 목표물을 추출하기 위하여 차영상을 이용하였고, 정확한 edge영상을 얻기 위하여 움직임 검출한 후 검출된 영역만을 현재 영상 으로부터 그레이 영상을 가지고 edge영상을 얻었고, 움직임 검출이 안된 영상은 현재 영상을 가지고 edge를 구하였다. 또한 타원으로 모델링하여 사람의 머리를 인식하고 향상 화면의 중심에 사람의 머리를 위치 시키기 위한 능동적인 추적 시스템을 구현하였다. 사람 머리 부분의 360도 회전, occlusion, tilting에서도 좋은 성능을 발휘하였다.

I. 서론

현대 사회는 컴퓨터를 사용한 정보 치리가 일반화 되어가고 있다. 디지털 영상 처리를 이용하여 목표물을 인식하고 추적하는 시스템은 군사용 정찰, 도로상의 교통량 측정, 보안 장치에서의 무인 감시 시스템, 화상 통신 시스템 그리고 컴퓨터 비전 시스템에서 사람의 행상을 인식하고 검출하는 방법을 이용하여 인간과 컴퓨터 사이의 인터페이스를 구현하고자 하는 연구가 활발히 진행되고 있다.[2] 인간의 시각이 아닌 카메라를 통해서 영상을 입력받고 머리, 눈, 코, 입, 몸통, 손등이 어디에 위치해 있는지를 컴퓨터가 판단하고 인식한다는 것은 앞으로 여러 응용분야에서 중요한 역할을 담당할 것이다. Nick Cutania는 움직이는 비행기를 추

본 논문은 2장에서 특징 추출과 모델링, 3장에서는 알고리즘을 자세히 설명하고, 4장에서 Pan/Tilt구동에 대해서 설명하였다. 그리고 마지막 5장에서 실험 결과를 보여준다.

II. 특징 추출과 모델링

사람 머리를 추적 하기 위해서 유효한 특징 패턴

설정을 고려하였다. 사람의 머리는 타원의 형태를 갖고 있기 때문에 타원에 대한 모델을 설정하고 가장 근접하는 타원의 형태를 갖는 목표물을 사람의 머리로 인식하고 위치 검출을 하는 방법이다. 사람 머리의 실제 모습을 특징점으로 하기 위해 먼저 카메라에서 획득된 영상에서 잡음을 제거하고 에지 영상을 구하였다. 그리고 사람의 머리 크기도 사전에 미리 알고 있어야 한다는 제약 조건을 갖고있다.

타원의 단축과 장축의 비인 aspect ratio는 모든 사람의 영상을 실험적으로 조사한 결과 사람의 머리 비율 가장 잘 대표 할 수있는 1.2의 값을 사용한다[1].

타원의 크기는 25에서 35까지 선정하였는데, 상황에 따라 타원 모델의 크기는 변경될 수있다. 여기에서 타원의 크기는 타원의 단축의 크기를 의미한다.그림 1에서 타원 모델의 Outline을 보여주고 있다.



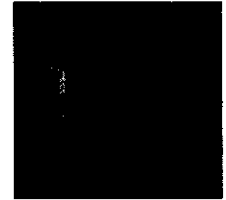
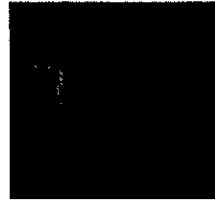
그림 1. 타원 모델의 Outline 모습

타원의 모델링 과정이 끝나면, 사람 머리의 특징벡터를 추출하기 위한 영상 전처리 과정이 필요하다. 복잡한 배경으로부터 손쉽게 목표물을 분리 하기위하여 차영상을 도입하였다. 차영상은 2개의 연속적인 명사도의 차이를 이용하는 방법이다.[4] 움직임이 검출되면 차영상을 통해 손쉽게 움직이는 물체의 영상을 얻을수 있다. 그러나 움직임이 거의 없다면 차영상을 통해서 제대로 영상을 얻을수 없게 된다.

본 논문은 정확한 gradient영상을 필요로 하기 때문에 차영상을 통해서 는 움직임을 검출하고 x,y축으로 projection을 구한 후 세그멘테이션을 수행하여 움직이는 물체의 영역을 검출한다. 그런 다음 현재 영상에서 검출된 영역만의 그레이 영상을 불러오고 나머지는 영값을 주므로써 복잡한 배경을 포함하고 있는 영상에서 손쉽게 움직이는 물체를 검출할 수 있고, 또한 정확한 gradient영상을 얻을 수도 있었다. 그림 2와 3은 이러한 과정을 보여주고 있다. 만약 움직임 검출이 없다면 움직이는 사람이 정지해 있다거나, 전혀 움직이는 물체가 없다는 것을 의미하므로 현재 영상을 Gaussian 필터를 수행한 후 gradient에지 영상을 얻는다. 이러한 경우 능동적인 추적을 수행하는데 있어서 Pan/Tilt는 움직임 필요가 없으며, 정확한 사람의 머리를 검출하기 위하여 그 전에 검출된 좌표값을 중심으로 탐색 영역을 설정하고 사람의 머리를 검출한다.

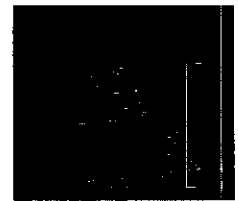
영상 획득은 흑백 512 × 480 영상을 CCD카메라로 획득하고 호스트로 전송시 256 × 240 영상으로 만든다. 사람 머리의 윤곽선을 제대로 나타내기 위해 잡음 제거를 위한 필터로 5 × 5 Gaussian 필터를 사용하였다. 수평쪽으로는 $[1 \ 4 \ 6 \ 4 \ 1]$ 마스크를 사용하여 Convolution을 수행하고,수직쪽으로 $[1 \ 4 \ 6 \ 4 \ 1]^T$

마스크를 사용하여 Convolution을 수행하였다. 그리고 gradient 에지 방법을 사용하였다.



(a) 이전 영상

(b) 현재 영상



(c) 차 영상과 x,y projection

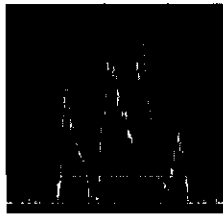


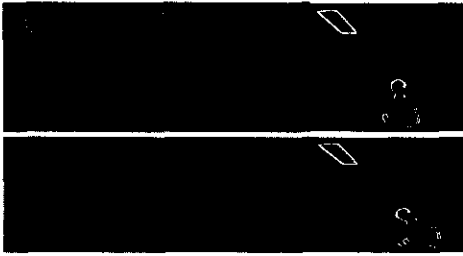
그림 2. 차 영상과 projection



그림 3. 움직임 추출된 영상과 edge 영상

수평쪽으로는 $[-1 \ 0 \ 1]$ 마스크를 사용하여 Convolution을 수행하고, 수직쪽으로는 $[-1 \ 0 \ 1]^T$ 마스크를 사용하여 Convolution을 수행 하였다.[5] 그림 4는 이러한 과정을 보여주고 있다.에지 정보로부터 사람 머리에 대한 정보를 찾아내므로 전처리 과정에서 사람 머리에 대한 정확한 정보를 제대로 나타내지 못하면 정확한 인식을 할수 없게 된다. 에지를 검출하는 과정에서는 조명에 더욱 신경을 써야 하며 threshold값을 선택하는 것도 특히 신경을 써야한다. Threshold값은 실

협에 의한 통계적인 값을 선택하였다.



(a)원 영상 (b)gaussian 영상 (c) edge 영상
그림 4. 획득된 영상과 전처리된 영상

또한 정확한 에지를 얻기 위해서는 잡음 제거와 에지를 구하는데 사용하는 필터의 크기를 크게 함으로써 좋은 결과를 얻을 수 있다.

III. 사람 머리 인식 및 위치 검출

사람의 머리는 수직축으로 고정된 타원으로 모델링을 하였다. 새로운 영상이 들어올 때마다, 타원의 상태 $s = (x, y, \sigma)$ (여기에서 (x, y) 는 영상 안에서의 타원의 중심 좌표값이며, σ 값은 단축의 길이에 대한 값을 나타낸다.) 는 타원의 주변을 따라 계산되는 gradient 크기의 정규화한 합의 최대값[1]을 얻기 위하여 local 탐색을 수행함으로써 얻어진다.

$$s^* = \arg \max_{s \in S} \left\{ \frac{1}{N_\sigma} \sum_{i=0}^N |g_i| \right\} \quad \text{수식(1)}$$

g_i = 타원 주변의 i Pixel안에서의 gradient값

N_σ = 타원 주변에 있는 Pixel들의 숫자값

사람의 머리를 찾는 탐색 과정에서 시간을 줄이기 위한 방법으로 다음과 같은 탐색 방법을 도입하였다. 탐색 영역은 세그멘테이션을 수행하여 얻어진 탐색 영역의 범위를 갖는다. 탐색 영역 안에서 x,y축 방향으로 2 pixels씩 이동하면서 탐색 영역 안에서의 gradient의 정규화한 합의 최대값을 찾는다. 그런 다음 gradient의 정규화한 합이 최대가 되는 점을 중심으로 새로운 탐색 영역을 설정한다. 탐색 영역은 x,y축 방향으로 ± 5 의 탐색 영역을 갖는다. 그런 다음 x,y축 방향으로 1 Pixel씩 이동하면서 세밀하게 탐색을 한다. 탐색을 할 때 gradient값이 있을 때만 탐색을 수행한다. 그래서 gradient의 정규화한 합이 최대가 되는 점을 사람의 머리의 위치로 결정하고 x,y축 좌표값을 Pan/Tilt로 보내 능동적인 추적을 가능케 한다.

IV. Pan/Tilt

위치 검출된 x,y축의 상대적인 거리를 이용하여 PC23으로 제어에 필요한 명령을 전송한다. PC23은 명령에 따라 X,Y축 방향으로 Pan/Tilt 제어를 하게 된다. PC23에 전달된 가속도,속도, 이동거리 등의 정보는 바로 STEP 모터를 움직이게 한다. 움직임이 끝날 때까지 기다린후 다음 영상을 획득하여 반복적인 수행을 한다. 본 논문에서는 Zoom기능은 사용하지 않았다. 그림 5는 PC에 연결된 PC23과 STEP모터를 이용한 Pan/Tilt의 모습을 보여준다.

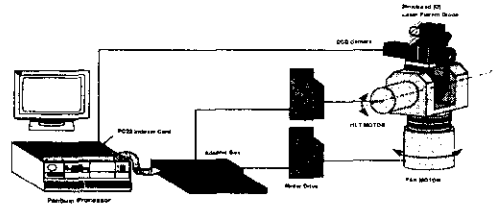


그림 5. PC23을 이용한 Pan/Tilt 제어 구성

V. 실험 결과

실험은 pentium (200MHz)에서 DT3851보드와 흑백 CCD카메라를 이용하였다. 흑백CCD카메라로부터 512×480 의 영상을 획득하여 호스트 컴퓨터로 256×240 의 영상을 전송한다. 실험 방법은 움직임이 적은 영상과 비교적 움직임이 큰 영상에서의 사람 머리 인식 및 추적 시스템을 비교한다. 먼저 움직임이 적은 영상에서 Occlusion, 360도 회전, Tilting에 대한 실험에서 정확히 사람 머리에 대한 위치를 검출하였다.

- (1) Occlusion : 사람의 정면에서 한쪽 손으로 사람의 얼굴 부분을 가로 질러가면서 얼굴 부분을 가린다. 그래도 추적 시스템은 정확히 사람의 머리 부분에 위치하게 된다.
- (2) 360 도 회전: 추적 시스템은 360도 회전에서도 별 영향을 받지 않는다. 사람 얼굴의 내부적인 정보가 아닌 사람 머리의 윤곽선을 이용하기 때문에 360도 회전에서도 정확히 사람 머리를 인식한다.
- (3) Tilting : 타원의 방향은 수직축으로 고정되어 있지만, 사람의 머리가 옆으로 기울어 질 때도 정확히 사람 머리를 인식한다.



그림 6. Occlusion한 경우의 그림

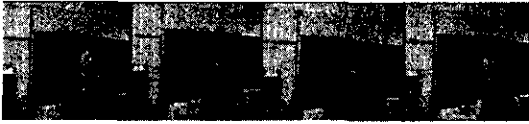


그림 7. 360도 회전하는 모습의 그림



그림 8. 좌우로 Tilt하는 모습의 그림

다음 그림 9, 10은 움직임이 큰 영상에서의 사람 머리에 대한 위치 검출과 추적하는 결과를 보여주고 있다. 움직이는 사람의 속도는 일정한 속도를 유지하면서 움직이도록 하였으며, 속도는 83mm/s 정도이다.

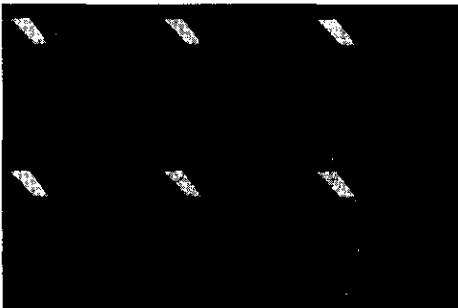


그림 9. Occlusion 과 360도 회전

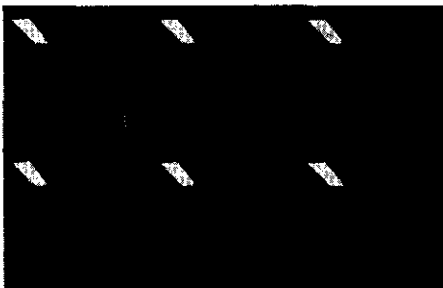


그림 10. Up-Down, Tilting

VI. 결론

본 논문에서는 관심있는 영역을 사람의 머리에 초점을 맞추고 미리 타원 모델링을 설정한 후 영상에서 gradient의 정규화한 합의 최대값을 찾음으로써 사람의 머리를 인식하는 간단한 사람 머리 추적 시스템을 구현하였다. 또한 복잡한 배경으로부터 원하는 영상을 얻기 위하여 차영상과 세그멘테이션 기법을 사용하였다. 이 시스템은 사람의 머리 부분의 360도 회전, Occlusion, Tilting 등의 문제점들을 극복하였다. 그러나 본 논문에서 구현한 시스템은 사람 머리의 윤곽선이 intensity gradients에 매우 의존도가 크다는 단점을 갖는다. 그러므로 보다 강력하고, 일반적인 추적을 제공하기 위해서는 사람 머리의 특징을 보다 잘 대표할 수 있고, 또한 실시간 적인 처리를 가능케하는 기술들이 추가되는 연구가 진행되고 있다.

참고 문헌

- [1] Stan Birchfield, "An Elliptical Head Tracker", Computer Science Department Stanford University CVTR'97
- [2] 백성현, 박동선, "효율적인 로봇 추적을 위한 신경 회로망 구조", 대한전자공학회 하계 종합학술대회 논문집 Vol 20, No 1, pp. 875 - 878, 1997
- [3] A.V.Nefian, M.khosravi, M.H. Hayes III "Real-time detection of human faces in uncontrolled environments", SPIE'97, Vol.3024, pp.211-219.
- [4] Rafael C.Gonzalez Richard E.Woods, Digital Image Processing, Addison Wesley
- [5] Ramesh Jain, Rangachar Kasturi, Brian G. Schunck, Machine Vision
- [6] Shoichiro Iwasawa, Kazuyuki Ebihara, Jun Ohya, shigeo Morishima, "Real-Time Estimation of Human Body Posture from Monocular Thermal Images
- [7] Jung-Hyun Hwang, "A study on Motion Tracking by An Active Camera", Department of Electrical Engineering Keio University, Yokohama, Japan
- [8] Lars-Peter Bala, Kay Talmi and Jin Liu "Automatic Detection and Tracking of Faces and Facial Features in Video Sequences"
- [9] H.J.Lee, L.F.Luang and Z.Cheng, "Multi-frame Ship Detection and Tracking in an Infrared Image Sequence", Pattern Recognition, Vol.23, pp.785-798, 1990