

광각렌즈 장착 카메라 및 팬-틸트-줌 카메라를 이용한 이동 물체의 실시간 추적 및 제어

¹김도환^o ²최경주 ¹이일병
¹연세대학교 컴퓨터 과학과
²(주)LG CNS, 연구개발센터
¹{zelda64^o, yblee}@csai.yonsei.ac.kr
²{choikj}@lgcns.com

Real Time Tracking and Control of Moving Objects using the Camera with Wide Lens and Pan-Tilt-Zoom Camera

¹Do-hwan Kim^o ²Kyungjoo Cheoi ¹Yillbyung Lee
¹Dept. of Computer Science, Yonsei University
²R&D Center, LG CNS Co., Ltd.

요 약

본 논문에서는 광각렌즈를 장착한 CCD 카메라를 통하여 실시간으로 입력되는 영상 데이터로부터 블록을 기반으로 한 영상처리 방법을 사용하여 움직임 정보를 추출한 후, 팬틸트줌 기능을 갖고 있는 카메라를 이용하여 이동하는 물체를 추적해 가는 실시간 이동물체 추적 시스템을 제안한다. 본 시스템은 입력되는 전체 영상 화면을 여러 개의 블록으로 나누어 처리함으로써 보다 빠른 연산 속도를 보이면서도 잡영에도 강하다는 특성을 가진다.

1. 서 론

본 논문에서는 광각렌즈(Wide Lens)를 장착한 카메라와 팬틸트줌(Pan, Tilt, Zoom)기능이 되는 카메라를 이용해서 순간적이면서 예측하기 어려운 움직임을 보이는 물체를 빠르고 정확하게 추적하는 시스템을 제안하고자 한다. 이러한 시스템을 개발하기 위해서는 시스템의 고속화가 필수적인데, 본 논문에서는 영상을 입력받는 방법과 탐지하고자 하는 물체와 배경과의 경계를 찾는 컨투어링(contouring) 방법을 개선함으로써 시스템 전체의 속도를 향상시켰으며 동시에 영상처리에 방해가 되는 불필요한 영상 신호인 잡영 제거에도 효과적인 알고리즘을 제안하였다.

2. 연구 배경

CPU의 고성능화와 소프트웨어 발전에 따라 최근의 컴퓨터 영상처리 능력은 CCD(Charge Coupled Device)를 컴퓨터의 눈과 같은 존재로 급부상시켰다. 이미 선진국에서는 군수업체와 대학 연구소를 선두로 활발하게 컴퓨터 시각 연구가 이루어지고 있지만 아직 우리나라는 이 분야에 대한 이해와 관심이 그 중요성에 비해 크게 떨어져 있는 실정이다. 특히, 시각에 의한 제어기술의 기본이라고 할 수 있는 이동물체 추적에 있어서 속도와 정확도에 대한 문제는 앞으로도 더 해결해야 될 주요 과제로 남아있다. 이에 본 논문에서는 이동물체 추적 문제에 대해 보다 근본적인 부분부터 다시 시작함으로써 기존의 방법에 비해 좀 더 효율적인 추적 시스템을 제안하고자 하였다. 본 논문에서 제안하는 시스템은 광각렌즈를 장

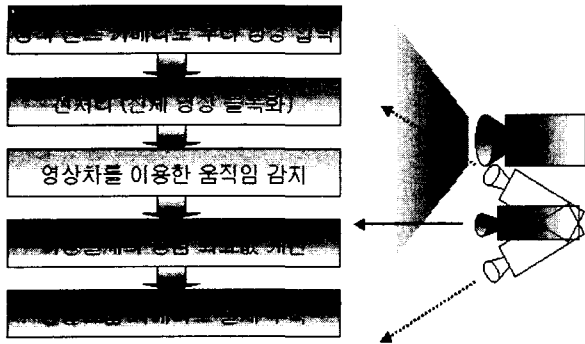
착한 카메라를 통해 보다 넓은 영역에 걸쳐서 움직임을 감지하며, 이렇게 감지된 움직임 감지 영역에서 해당 물체를 팬틸트줌 카메라로 정확하게 포착함으로써 컴퓨터를 이용한 영상처리방법을 사용하여 움직이는 물체를 추적할 수 있으며, 기계적으로도 카메라를 제어할 수 있음을 보여준다. 본 논문에서 사용된 개선된 블록기반 알고리즘은 전체 영상을 수백 개의 블록으로 나누고 블록(block)을 집합으로, 화소(pixel)를 원소라 가정함으로써 각 화소를 일정한 크기들의 블록들에 포함시켜 그들의 대표 값으로 나머지를 처리하도록 했다. 이렇게 하면 수 만개의 화소를 컨투어링하는 대신에 그 1/100에 달하는 수백 개의 블록만을 처리하므로 보다 빠르고 정확하게 목표물을 추적할 수 있게 도움을 준다. 또한 블록 방식의 특성을 심분 활용한 일체화된 잡영 제거 연산은 잡영을 제거하기 위한 영상처리를 추가적으로 할 필요가 없게 하므로 전체 시스템에 속도와 더불어 잡영에도 매우 강하다는 장점을 부여해준다.

3. 관련 연구

선진국에서 움직임 추적에 관한 연구들은 오랜 시간동안 체계적으로 진행되어 왔으며 국내에서는 최근에 들어 몇몇 대학 연구소를 중심으로 이루어지고 있다. 국내 연구로는 블록정합 알고리즘과 삼각형 기반 방법의 결점을 해소하기 연구[1]등이 진행되고 있고, 외국의 경우에는 다수의 카메라를 이용한 색상과 움직임 정보를 통한 얼굴 추적[2], 다수의 비디오 센서를 이용한 감시 시스템[3] 등이 연구되고 있다.

4. 제안하는 이동물체 추적 시스템의 전체 구조

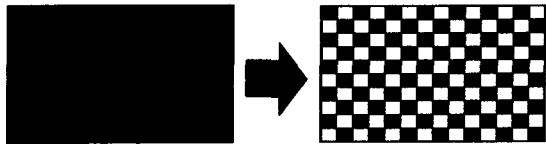
[그림 1]은 본 논문에서 제안하는 블록 기반 이동물체 추적 시스템 전체 구조를 보여준다. 먼저, 광각렌즈를 장착한 고정된 카메라로부터 연속적으로 입력되는 영상들을 일정크기의 블록으로 나누는 블록화 전처리가 수행되고, 이렇게 블록화 된 영상 시퀀스에 블록기반 연산을 수행하여 움직임이 감지 된 블록들을 추출한다. 그 후에 추출된 블록들의 인접 여부를 따져서 블록들을 해당 물체별로 분류시킨다. 그림 각 단계별로 자세한 설명과 함께 블록개념으로 연계 되는 이점을 보도록 하겠다.



[그림 1] 시스템 구조 [그림 2] 카메라 구성

1) 카메라를 통한 영상 입력 : [그림 2]에서와 같이 화각이 표준렌즈에 비해 2배 이상 넓은 광각렌즈를 장착한 1번 카메라를 고정시키고 펜틸트 기능으로 자유롭게 시야를 바꿀 수 있으며 줌인, 줌아웃(Zoom In, Zoom Out)이 가능한 2번 카메라와 조합시키면 넓은 시야를 갖는 시각 시스템이 구성된다. 실세계 시각 정보는 일차적으로 1번 카메라를 통해서 입력받는다.

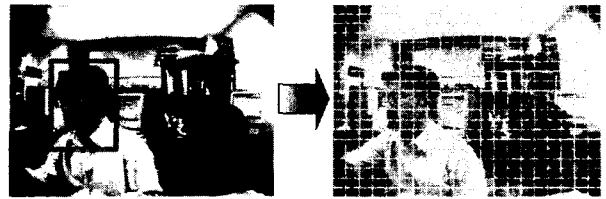
2) 영상 시퀀스의 블록화 : 1번 카메라를 통해 들어온 영상을 같은 크기의 블록으로 나누고, 이렇게 나뉘어진 영상을 토대로 움직임을 추적한다. 본 논문에서는 320×240 크기의 입력 영상을 사용할 경우, 16×12 크기의 블록으로 나누면 총 400개의 블록이 생성된다. 이렇게 생성된 블록들은 주변의 화소들을 담당하고 있는 하나의 집합으로 볼 수 있는데 이는 영상 정보에 대한 개념을 간략화하여 컨투어링을 빠르게 하는 역할을 한다.



[그림 3] 비교화소 감소에 의한 블록내 비교 영역 절감

[그림 3]은 12×10 크기의 하나의 블록을 상징하며 내부에 존재하는 각각의 작은 사각형들은 화소들을 나타낸다. 이렇게 좌로부터 홀수 또는 짝수에 해당되는 화소만

입력받아도 움직임 감지에 성능저하를 가져오지 않으면서 동시에 영상의 입력과 처리에 두 배 이상의 연산을 절약할 수 있다. [그림 4]의 좌측 그림은 카메라에서 전처리 없이 받아들인 실 영상이고 우측 그림은 블록화 전처리가 수행된 후 비교영역이 감소된 컴퓨터 메모리 내에 상주하고 있는 영상정보를 나타낸다. 우측 화면이 뿌옇게 보이는 이유는 [그림 3]처럼 각 블록내에서 짝수나 홀수에 해당되는 화소들만을 입력받았기 때문이다.



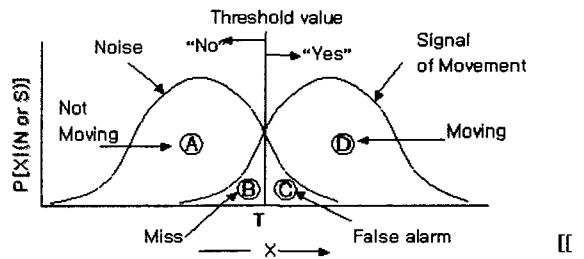
[그림 4] 비교 영역 절감

3) 블록 기반 영상 처리 방법에 의한 움직임 추출 : 식(1)에서 $I_0(x,y)$ 는 이전 프레임의 (x,y) 에 위치한 이전 프레임의 화소값이고 $I_1(x,y)$ 는 같은 위치의 다음 프레임의 화소값이며 I_d 는 그 차이의 절대 값이다. 식(2)는 크기가 $a \times b$ 인 블록에서 식(1)에서 나온 각각의 화소에 해당되는 결과인 I_d 를 블록 사이즈인 $a \times b$ 로 나누어 이를 모두 합한 것이다. 식(3)을 통해 식(2)의 결과가 일정한 임계치 값 T 보다 크면 움직임이 발견된 것으로 판단하게 된다. 이와 더불어 배경 전반에 발생하는 잡영이 소멸되는 효과가 생긴다.

$$I_d = |I_0(x, y) - I_1(x, y)| \dots (1)$$

$$X = \sum_x^a \sum_y^b I_d(x, y) / (a \times b) \dots (2)$$

$$X > T \dots (3)$$

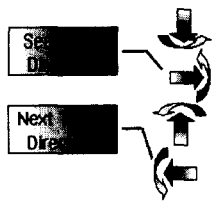


[그림 5] Bayes' theorem of signal

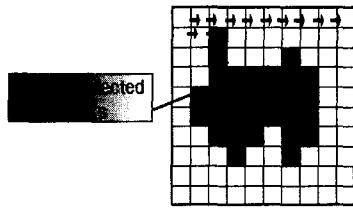
[그림 5]는 임계값 T 에 따른 잡영과 움직임 정보에 대한 판별 여부를 나타낸 것이다. 영상은 언제나 올바른 정보와 잡영이 공존하기 마련이다. 좌측의 가우시안 곡선은 영상에서 노이즈 신호를 나타내고 우측의 곡선은 움직임 신호를 나타낸다.

4) 선택된 이동물체의 추적 : 일단 움직임 영역을 감지하면 먼저 인접한 화소들만을 하나의 영역으로 보고 전체 움직임 영역을 물체별로 구분 지어줄 필요가 있다. 여기

서 이동물체의 외곽을 찾는 컨투어링 작업은 전체연산에서 상대적으로 가장 시간을 소요하며 또한 오류도 많이 발생한다. 이는 화소기반처리에서 발생하는 필연적인 결과라고 할 수 있다. 하지만 여러 개의 화소를 블록으로 묶어주면 많은 연산을 절약할 수 있게 되고 오류 또한 줄어들게 된다. 본 논문에서는 최대 76,800(320×240)개의 화소들로 처리해야 할 컨투어링을 400(20×20)개의 블록들로 처리함으로써 이 부분에 있어 많은 계산 절감을 볼 수 있었다.



[그림 6] 탐색 방법



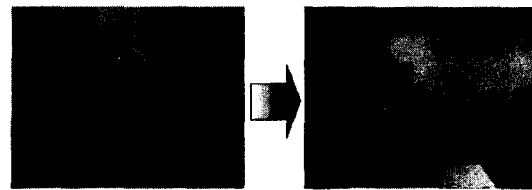
[그림 7] 객체에 대한 컨투어링 작업

[그림 6]은 본 논문에서 사용된 움직임 블록 탐색 알고리즘이다. 가운데 끝은 화살표는 지나온 방향을 나타내고 그것을 둘러싸고 있는 휘어진 화살표는 탐색해야 할 방향을 나타낸다. [그림 7]은 컨투어링 결과를 나타내는 것으로서 처음에 탐색이 시작될 때는 입력된 영상(320×240)의 맨 위의 좌측에서부터 우측으로 탐색을 하게 되며 우측까지 다 끝났으면 위에서 한 블록 내려와서 다시 좌측으로부터 탐색이 시작된다. 일단 움직임 블록이 감지가 되면 지나온 방향의 +90도인 방향을 제1순위로 탐색하게 되고 그곳에 움직임 블록이 없으면 지나온 방향과 일치한 방향인 0도 방향을 탐색하게 되고 그곳에도 움직임 블록이 없으면 -90도 방향을 탐색하게 되며 만일 그곳에도 없으면 유평하게 된다. 언뜻 보면 '연결요소(connected component)'와 유사하지만 화소값이 아닌 블록값으로 처리함으로써 속도가 매우 빠르다는 점과 대각 방향 탐색을 하지 않음으로써 불필요한 연산이 없고 또한 이해하기 쉬어 적용이 용이하다는 점이 특징이다.

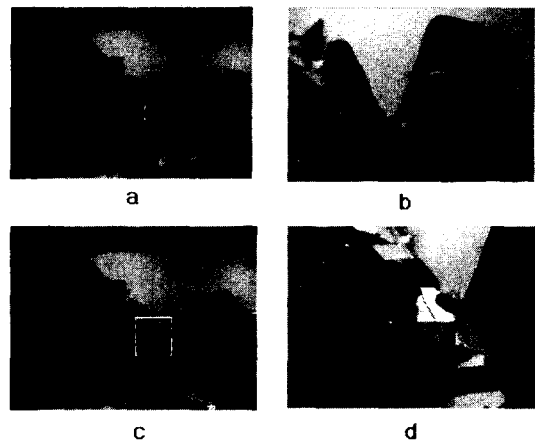
5. 실험결과

실험 장비는 두 대의 SONY EVI-D30 카메라와 View Cast Osprey-100 영상캡처보드를 사용하였으며 약 1/20초의 프레임 속도로 320×240 크기의 입력 영상을 받아 들었다. [그림 8]은 일반렌즈를 착용할 때(좌측)와 광각렌즈를 착용할 때(우측)의 카메라의 화각(시야의 각도)차이를 잘 보여주고 있다. 동일한 상태의 같은 카메라에 광각 렌즈를 장착한 사진인데 좌측 사진에 비해 우측 사진의 시야가 훨씬 넓은 것을 알 수 있다. [그림 9]는 추적 시스템으로 본인을 촬영한 것으로 추적 실험을 위해 다른 움직임 중에서도 비교한 움직임 값이 큰 곳을 추적하도록 설정해 놓았다. [그림 9.a]는 추적 시스템 앞에서 손을 흔들고 있는 본인을 광각렌즈를 장착한 1번 카메라가 움직임이 큰 부분인 손을 감지하고 [그림 9.b]처럼 렌즈

틸트줌 기능이 있는 2번 카메라에게 추적신호를 보내어 2번 카메라가 적절하게 해당 물체를 포착하고 있는 사진이다. [그림 9.c] 또한 손의 움직임을 감지해서 [그림 9.d]에서 보는 바와 같이 2번 카메라에게 제어신호를 보내 손을 자동촬영 하도록 한 결과이다. 제어신호는 RS232 Serial 방식으로 전송해주었다. 비교적 빠르고 예측하기 어려운 움직임에도 신속하고 정확하게 반응했다. 여기에 색상, 무게중심, 벡터값 등을 부여한다면 멀티 추적 기능도 가능할 것으로 기대한다.



[그림 8] 일반 렌즈와 광각렌즈의 차이



[그림 9] 추적 실험 결과

6. 참고문헌

[1] 김학수, 임원택, 이재철, 이원규, 박규택, 다중선형회귀모형을 이용한 움직임 추정방법, 대한전자공학회 논문지, 제34권, S편, 제10호, pp.98-103, 1997.10.
 [2] Hongo, H.; Ohya, M.; Yasumoto, M.; Niwa, Y.; Yamamoto, K. Focus of attention for face and hand gesture recognition using multiple cameras Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on, Page(s): 156 -161, 2000.
 [3] Kanade, Collins, Lipton, Anandan, Burt, Wikson, Cooperative Multi-Sensor Video Surveillance, DARPA Image Understanding Workshop(IUW), New Orleans LA, pp 3-10, May. 1997.