

강화학습을 사용한 실시간 이동 물체 추적에 관한 연구

A Study of Real Time Object Tracking using Reinforcement Learning

김상헌, 이동명, 정재영, 윤학수, 박민욱, 김관형
동명정보대학교 컴퓨터공학과

Sang-heon Kim, Dong-myong Lee, Jae-young Joung,
Hak-su Youn, Min-yook Park, Kwan-hyung Kim
Dept. of Computer Science and Engineering,
TONGMYONG University of Infomation Technology
E-mail : honey1321@dreamwiz.com

요 약

과거의 이동로봇 시스템은 완전한 자율주행이 주된 목표였으며 그때의 영상정보는 단지 모니터링을 하는 보조적인 수단으로 사용되었다. 그러나 지금은 이동 물체의 추적, 대상 물체의 인식과 판별, 특징 추출과 같은 다양한 응용분야에서 영상정보를 이용하는 연구가 활발히 진행되고 있다. 또한 제어 측면에서는 전통적인 제어기법으로는 해결하기 힘들었던 여러 가지 비선형적인 제어를 지능제어 방법을 통하여 많이 해결하곤 하였다. 그러한 지능제어에서 신경망을 많이 사용하기도 한다. 최근에는 신경망의 학습에 많이 사용하는 방법 중 강화학습이 많이 사용되고 있다. 강화학습이란 동적인 제어평면에서 시행착오를 통해, 목적을 이루기 위해 각 상황에서 행동을 학습하는 방법이다. 그러므로 이러한 강화학습은 수많은 시행착오를 거쳐 그 대응 관계를 학습하게 된다. 제어에 사용되는 제어 파라메타는 어떠한 상태에 처할 수 있는 상태와 행동들, 그리고 상태의 변화, 또한 최적의 해를 구할 수 있는 포상알고리즘에 대해 다양하게 연구되고 있다.

본 논문에서 연구한 시스템은 비전시스템과 Strong Arm 보드를 이용하여 대상물체의 색상과 형태를 파악한 후 실시간으로 물체를 추적할 수 있게 구성하였으며, 또한 물체 이동의 비선형적인 경향성을 강화학습을 통하여 물체이동의 비선형성을 보다 유연하게 대처하여 보다 안정하고 빠르며 정확하게 물체를 추적하는 방법을 실험을 통하여 제안하였다.

keyword : 강화학습, 영상처리, 물체추적

1. 서론

현대 산업 사회에서는 하드웨어 기술과 컴퓨터를 이용한 비전기술 및 제어의 발전에 따라 이를 이용한 이동로봇에 대한 연구가 활발히 진행되고 있다. 근래에는 센서(sensor) 기술의 발달로 인간의 시각에 해당하는 카메라(camera)의 성능과 처리 기술이 고도로 발달되어 있으며, 이를 제어하는 제어기술이 컴퓨터와 멀티미디어

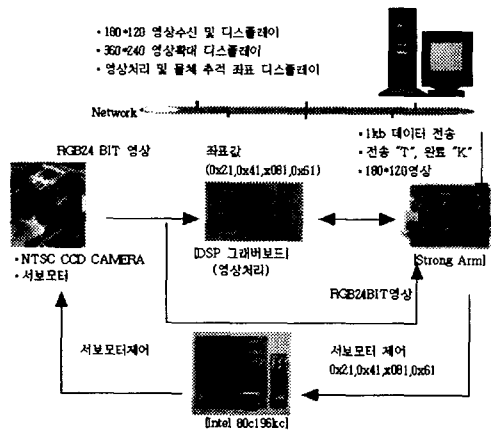
(multimedia) 그리고 신호처리 등의 기술의 급격한 발달로 인하여 다양한 형태의 영상신호처리 방법들이 제안되고 발달되고 있다. 오늘날 산업 현장에서도 머신 비전 시스템(machine vision system)과 고성능의 마이크로 프로세서(micro processor)를 도입하여 영상응용과 멀티미디어 응용분야로 생산자동화에 주력하고 있는 실정이다. 이동 로봇으로 '대상 인식', '목표 추적'과 같

은 행동을 구현할 때 초음파 센서가 일반적으로 사용된다. 그러나 초음파 센서는 주위 환경에서 거리 정보를 얻기에는 용이하지만 목표 대상을 인식하는 데에는 어려움이 있다. 예컨대, 초음파 센서만을 장착한 이동 로봇으로는 자율 주행시 로봇이 목표지점에 도착하였는지의 여부를 확인할 수 없다는 결점을 가지고 있다. 따라서 본 논문에서는 CCD 카메라를 이용하여 대상의 색상과 형태를 파악한 후 강화학습을 통한 목표 대상의 위치를 실시간적으로 추적하는 알고리즘의 방법을 제안한다.

2. 본론

2.1 시스템 모델

본 논문에서 제시하고자 하는 시스템은 [그림 1]과 같다.



[그림 1] 전체 시스템 구조

[그림 1]은 CCD 카메라로부터 받아진 영상을 이용하여 그래버보드에서 영상처리하여 타겟물체의 중심값을 ARM보드로 전송한다. CCD 카메라는 한프레임을 그래버보드로 전송하고 다음 프레임에서는 ARM보드로 전송을 반복한다. ARM보드는 그래버보드로부터 받은 중심좌표값과 CCD 카메라로부터 받은 영상 데이터를 조합하여 80C196KC에 서버모터 제어값을 넘기고 호스트 PC와 TCP/IP를 통한 데이터를 전송한다.

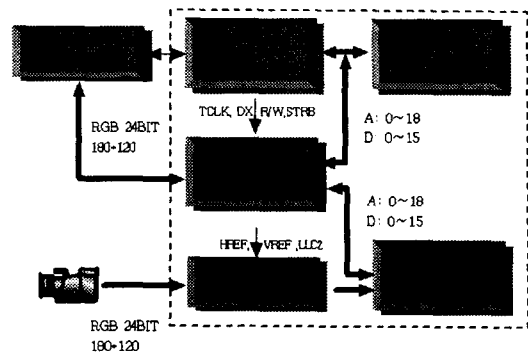
호스트PC는 ARM보드로부터 받은 영상데이터를 실시간적으로 모니터링한다.

2.2 비전시스템

CCD 카메라는 매초 33프레임 간격으로 180*120 사이즈의 RGB 24BIT 포맷을 구성하여 화소 단위로 그래버보드에 이미지 데이터를 전송한다. CCD 카메라로부터 전송되는 영상을 TI社의 TMS320C32-60 칩을 사용하여 전처리 과정을 통해 물체의 중심좌표값을 찾고 ARM보드로

전송하는 과정을 거친다.

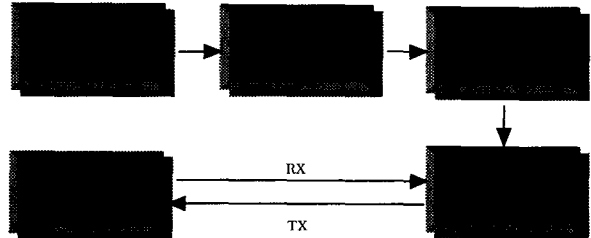
[그림 2]는 그래버보드 구조를 나타낸 것이다.



[그림 2] 그래버보드 시스템 구조

대상물체의 색(Blue) 범위를 결정하여 해당 물체의 색(Blue)을 구별해 낼 수 있다.

비전처리의 전체적인 과정은 다음과 같다.



[그림 3] 비전시스템 구조

[그림 3]은 카메라로부터 받아진 영상을 이용하여 파란색 물체에 대해 중심좌표를 나타내는 전체적인 흐름도이다.

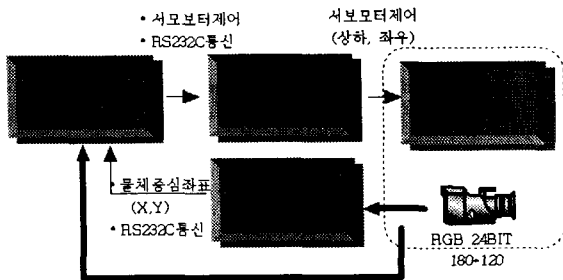
카메라의 해상도나 조명의 불균일성 등에 의해, 대상물체 영상에 잡음이 생길수 있다. 이러한 잡음은 보통 불규칙적으로 발생하므로 잡음에 해당하는 연결성분들의 화소의 수를 임의로 정하여, 화소 수 미만인 연결성분들을 제거하였다. 이 방법은 잡음에 의해 생기는 화소들을 제거하기 위한 단순하면서도 아주 효율적인 방법이다.

물체의 중심좌표를 찾기 위해서는 물체의 크기와 위치를 알아야 한다. 2진화상에서 물체의 중심 위치는 그 물체의 면적 중심과 같으므로 면적 중심법을 이용해서 물체의 위치(\bar{x} , \bar{y})를 구할수 있다. 면적중심법을 이용하여 ARM보드로 타겟물체의 중심좌표를 전송한다..

2.3 물체 추적 시스템 구조

물체가 중앙에 위치하지 않거나 위치한 영상에 대해 ARM보드로 실시간 제어한다.

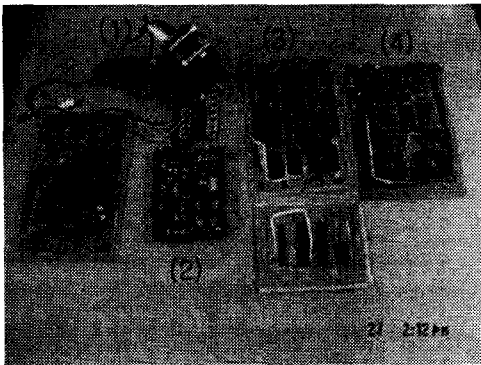
[그림 4]는 물체 추적 시스템 작동 구조를 나타낸 것이다. RS232(시리얼) 제어용 포트를 사용하여 실시간적으로 물체의 중심좌표를 확인하고 80C196KC 보드에 제어 신호를 보낸다.



[그림 4] 물체 추적 시스템 구조

80C196KC는 받은 제어값으로 서보모터의 상하좌우 카메라를 이동시키게 된다.

이동된 카메라에 의해 입력영상에서 보여지는 물체의 위치가 달라지게 되고, 물체의 위치가 카메라의 중앙에 위치할때까지 계속적으로 카메라가 이동된다.



[그림 5] 본 모듈의 실제 사진

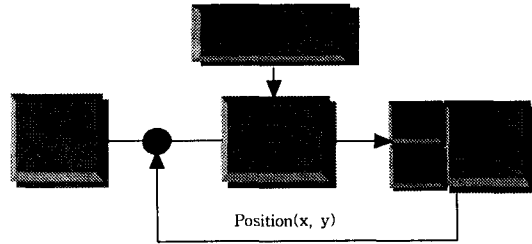
[그림 5]는 본 논문에서 사용한 시스템 모듈로서 (1)NTSC CCD CAMERA, (2)ARM 보드, (3)그래버보드(TMS320C32-60),(4)영상신호처리(SA A7111A), 80C196KC로 구성되어 있다.

3. 강화학습을 사용한 실시간 물체추적

3.1 강화 학습 모델

Q-Learning은 가장 널리 사용되는 강화학습 방법들 중 하나로 이 학습법은 현재 상태에서의 행위로부터 얻게 되는 총 보답을 예측하는 행위값에 대응시키는 행위함수를 학습하는 알고리즘이다. 그러나 Q-Learning은 빠른 실시간 성능을 가지더라도, 에이전트가 학습을 할 경우 학습 속도가 느리다. 이는 과거에 학습했던 지식을 사용하지 않기 때문이다. 에이전트가 같은 환경 내에서 여러 문제를 해결해야 할 경우, 과거에 문제들을 해결하면서 얻어온 환경에 대한 지식을 활용할 수 있다면 학습시간을 줄일 수 있는데 이러한 연구들이 활발히 진행되고 있다.

[그림 6]은 강화학습의 모델로써 환경변수값이 제어기를 통해 타겟물체의 이동제어값을 산출한다.



[그림 6] 강화학습의 모델

3.2 환경 지식을 이용하는 강화학습

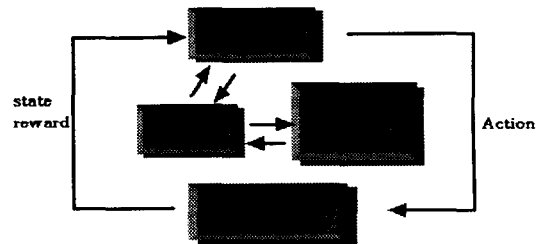
Q-learning은 대표적인 off-policy 강화학습으로 행위 함수의 갱신 $Q(st,at)$ 과 책략 $\pi(st)$ 은 다음과 같이 수행된다.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t))$$

$$\pi(s_t) = \text{argmax}_{a_t} Q(s_t, a_t)$$

α : 학습률

에이전트가 같은 환경 내에서 여러 문제를 해결해야 할 경우, 과거에 문제들을 해결하면서 얻은 환경에 대한 지식을 활용할 수 있다면 학습시간을 줄일 수 있다. 환경지식을 이용한 Q-Learning은 Local state feature, Example Set 그리고 Classifier를 가지고 과거의 지식을 이용하여 학습을 하게 된다.



[그림 7] 환경 지식을 이용하는 강화학습

3.2.1 Local state feature

Q-learning에서 과거에 학습한 행위 함수를 새로운 문제에 그대로 적용할 수 없는 이유는 그 행위함수가 과거의 문제에 국한되어 학습이 되어 있기 때문이다. 그래서 주어진 문제와는 독립적인 특성을 갖는 Local state feature라는 저장공간을 두어 학습에 이용하면 학습을 더욱 효과적으로 할 수 있을 것이다. 즉, Local state feature란 에이전트 주위의 상태를 저장하는 공간이다.

3.2.2 Example Set

에이전트가 주어진 문제에 대해 행위함수를 학습한 후 환경 지식을 추출하기 위해 Local state feature와 각 상태에서 취한 행위, 그리고 그 행위에 대한 평가가 이루어져야 한다. 이 세 가지의 정보를 갖는 저장 공간을 Example이라고 한다.

3.2.3 Classifier

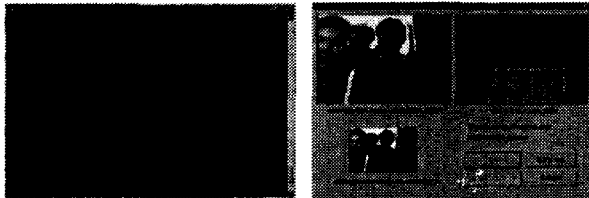
최종적으로 에이전트는 여러 문제들을 통하여 환경 지식을 추출할 때 각 문제를 해결하면서 생성한 Example Set을 이용한다.

3.3 강화학습을 사용한 실시간 물체추적

CCD 카메라를 제어하기 위한 제어 시스템은 [그림 8]과 [그림 9]와 같다. 비선형적인 환경요소와 타겟물체의 (x, y)좌표를 이용하여 제어기를 통한 값에 의해 두 개의 모터를 제어한다. CCD 카메라가 이동하게 되면 변화된 타겟물체의 위치와 이동각에 따라 기준선과의 오차를 최소화하기 위하여 제어기에서 좌우 모터의 값을 제어하게 된다.

4. 실험 및 결과

[그림 8]과 [그림 9]는 실험에 사용한 사용자 인터페이스 화면이다.



[그림 8] ARM 실행화면 [그림 9] 사용자 인터페이스

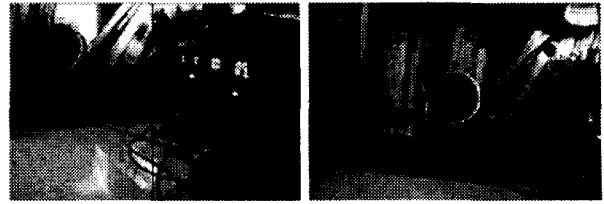
[그림 8]은 ARM보드에서 물체중심을 확인하고 PC에 데이터를 전송하는 화면이다.

[그림 9]는 호스트 PC의 화면으로 원영상 180*120의 크기와 물체의 중심 좌표를 화면에 모니터링 하는 화면이다. 좌상단 화면은 180*120의 원영상을 기하학적 변환으로 360*180사이즈로 확대 하여 디스플레이 하였다. 최근접이웃보간법을 사용하여 블록현상을 제거하였다. 우상단 화면은 원영상을 기반으로 그래버보드에서 물체의 중심 좌표를 찾기 위해 비전처리 하는 과정을 호스트 PC에서도 구현하여 CCD 카메라의 주변환경(배경 및 조도)등에 의해 색상 인식 및 형태 인식에 사용하는 대상물체 색의 각 성분(R, G, B)의 범위를 결정하여 해당 물체의 색을 구별해 낼 수 있도록 하였다.

[그림 10]과 [그림 11]의 (a) 영상은 공이 아래 쪽으로 잡힌 영상이고, (b) 영상은 카메라가 그 방향으로 이동하여 움직여 캡처된 영상이다.



(a) (b) [그림 10] 실험 I



(a) (b) [그림 11] 실험 II

카메라는 물체를 항상 중심점에서 ± 10 pixel 지점의 오차범위안에서 중앙쪽으로 위치하도록 움직였다. 여기서 보인 [그림 10]과 [그림 11]은 추적(tracking)중에 한 순간의 캡처영상이다. 처리속도는 카메라의 2자유도(Pitch, Yaw)속도인 최대 80도/초의 속도를 충분히 낼 수 있을 정도였다.

5. 결론 및 향후과제

본 논문에서는 에이전트가 새로운 문제를 학습할 때 과거에 학습했던 환경지식을 이용하는 알고리즘을 적용하여 실시간 물체 추적을 하였다.

CCD 카메라의 경우 주변 환경(배경 및 조도)등에 매우 민감하기 때문에 실제 상황에 적용하기 위해서는 주변 환경에 적응하여 색상 인식 및 형태 인식에 사용하는 파라미터들을 가변 시킬 수 있는 알고리즘의 개발이 필요하다.

본 논문에서는 영상을 일종의 Sensor처럼 사용하여 실시간 물체를 추적하였다. 이 실험을 통해 3차원 이동물체를 대상으로 이동카메라(좌표 및 앵글의 변화가 존재)의 시스템 설계로 확장할 것이다. 또한 이동물체의 공간상 좌표 및 예정경로를 추정할 수 있게 할 것이다. 실험에서는 카메라의 상하좌우(2자유도(Pitch, Yaw)) 움직임만을 보였지만, 로봇팔과도 직결하여 일종의 케어 로봇이나 골기퍼와 같은 모델을 만들 수 있다. 또한, 2족 보행로봇이나 탐사로봇등 움직이는 로봇에 탑재하여 고성능 지능시스템의 하나의 핵심 요소가 될 수 있다.

5. 참고문헌

- [1] 정재은, 유닉스시스템 프로그래밍, 한빛미디어, 2003
- [2] Rafael C. Gonzalez, Richard E. Woods "Digital Image Processing", 그린출판사, 1998
- [3] 이상엽, "Visual C++ programming Bible", 영진출판사, 1999
- [4] Randy Crane, 영상처리 이론과 실제, 홍릉과학출판사, 1997
- [6] 김종환, "로봇 축구 시스템", 대영사, 2000