

한 대의 USB port 카메라와 자바를 이용한 3차원 정보 추출

지창호*(부산대 대학원 지능기계공학과), 이동엽(부산대 대학원 메카트로닉스협동과정)
이만형(부산대 기계공학부)

3-D Position Analysis of an Object using a Monocular USB port Camera through JAVA

C. H. Ji(Mech. Eng. Dept., PNU), D. Y. Lee(Mech. Eng. Dept., PNU), M. H. Lee(Mech. Eng. Dept., PNU)

ABSTRACT

This paper's purpose is to obtain 3-Dimension information by using a monocular camera. This system embodies to obtain the height of object by using trigonometry method between a reference point of circumstance and an object. It is possible to build up system regardless of operating system, and then set it up. An comfortable USB port camera is used everywhere without the capture board. The internet can be used by using the applet and JMF everywhere. We regard the camera as a fixed. And we have developed a Real-Time JPEG/RTP Network Camera system using UDP/IP on Ethernet.

Key Words trigonometry(삼각법), java(자바), JMF(자바 미디어 프레임워크), applet(애플릿), USB port camera(USB 포트 카메라)

1. 서론

예전부터 사람의 오감과 비슷한 측정장치를 만들기 위해 많은 연구를 해왔다. 현재는 인간의 감각보다 뛰어난 센서들이 실제 산업현장에서 사용되고 있다. 그러나 다른 센서들과 비교해 보면 시각에 대한 연구는 최근에 이르러서야 많은 연구가 되고 있다.

카메라를 이용한 감시 시스템들 또한 많은 연구가 되고 있다. 현재 사용되고 있는 감시 카메라는 영상을 보여주는 시스템으로 되어있어 사람이 카메라로부터 들어오는 영상을 항상 감시해야 한다. 특정한 물체의 움직임에 대한 인식을 할 수 없으며 또한 물체에 대한 정보도 얻을 수 없다.

대부분 감시카메라는 특정 지역에 고정되어 있는 경우가 대부분이며 특별히 제작된 것은 상하 좌우를 움직일 수 있는 시스템으로 구성되어 있다. 본 시스템은 움직임은 하지 않고 고정된 상태에서 들어오는 영상정보를 처리한다.

본 시스템의 특징은 카메라가 고정된 상태에서 받아들이는 한정된 영상정보에서 모서리와 기둥의 특징점을 사용하여 카메라가 설치된 지역의 물체 절대 위치를 계산해 낼 수 있다.

본 논문은 USB port 카메라를 이용하여 쉽게 시스템구현이 가능하도록 하였으며 현재 사용중인 사무실의 컴퓨터와 건물에 구축된 네트워크 장비를 이용하여 감시시스템을 간단히 구성할 수 있으며 또한 영상처리 기법을 적용하여 단지 보여주는 시스템이 아닌 알림기능을 부여하는데 있다.

이를 구현하기 위해 사용된 프로그램인 자바를 이용하여 운영체제에 관계없는 시스템구축을 가능하게 하며 애플릿을 이용하여 인터넷을 사용할 수 있는 것이면 어디서든 이용가능한 시스템을 구현한다.

2. 삼각관계의 추출

2.1 카메라 설치장소의 상황

주로 감시 카메라의 설치 위치는 진입로나 복도 등의 교차지점 부분에 설치되므로 아래 그림과 같은 상황을 설정하였다.

그림 2에서 물체의 높이를 측정하기 위해서 필요한 간단한 삼각 관계식을 추출해 내교자 한다. 여기서 기준높이는 바닥에서부터 수직으로 서있는 기둥이나 벽의 모서리를 의미하며 이 기준높이가 바로 물체높이 측정을 위한 기준으로 이용된다.

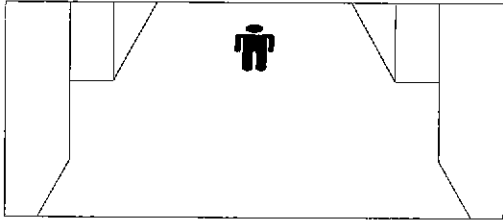


Fig 1 Position of a camera

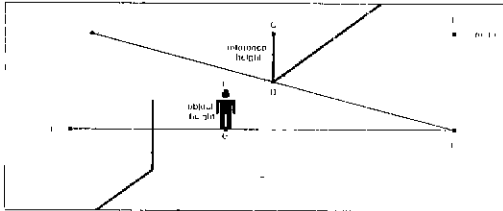


Fig 2 Height of an object and reference height

2.2 삼각관계식 유도

그림 2에서 물체의 높이 H 를 복도로부터 수직인 위치에 있다고 가정하고 그림에서 복도위의 직선 FD 와 AD 를 밑변으로 하고 카메라의 높이 DE 를 공통 높이로 하는 두 개의 삼각형 $\triangle EDF$ 와 $\triangle EDA$ 이 존재함을 볼 수 있다

이 두 삼각형이 복도와 수직인 것을 이용하여 두 도형의 관계를 살펴보면 각각의 삼각형에 대한 닮은 비에 의해 아래와 같은 식을 얻을 수 있다

$$\frac{\overline{FG}}{\overline{FD}} = \frac{\overline{GH}}{\overline{DE}} \quad (1)$$

$$\frac{\overline{AB}}{\overline{AD}} = \frac{\overline{BC}}{\overline{DE}} \quad (2)$$

식 (1)과 (2)의 공통분모 DL 에 관하여 정리하면 식(3)과 같이 된다

$$\overline{GH} = \frac{\overline{FG} \overline{AD}}{\overline{FD} \overline{AB}} \overline{BC} \quad (3)$$

따라서 물체의 높이 GH 는 기준높이(BC)의 상대적인 높이임을 알 수 있다.

이 기준 물체가 수직에 가까울수록 측정된 물체의 높이도 정확해진다. 따라서 복도의 벽면 모서리를 기준높이로 사용하여 물체의 높이를 측정하였다

2.3 절대 좌표계

위에서 물체의 높이를 계산하기 위해 삼각기본식

(3)을 유도했다. 이 식을 실제 영상에 적용하기 위해서는 각 점에 대한 실제 복도에서의 절대좌표가 필요하다. 이것은 입력 영상 위치와 실제 복도 위치와의 관계를 나타내는 좌표변환함수를 통해 가능해진다. 이 변환관계를 구하기 위해서는 입력영상과 복도 모델간의 매칭점이 필요하고 4개 이상의 매칭점에 의해 정의된다

3. 영상-모델 변환

입력 영상의 위치를 실제 복도상의 위치로 변환하기 위해서 필요한 일반적인 평면-평면 매핑(plane-to-plane mapping)인 평면투영변환(plane projective transformation)에 대해서 알아본다.

3.1 평면투영변환

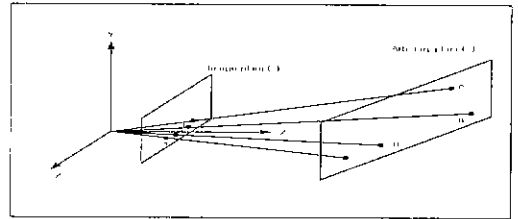


Fig 3 Image of a 3-D arbitrary plane according to Camera projection

그림3 은 카메라투영에 의해서 3차원 공간상에 존재하는 임의의 평면 S 위에 존재하는 점들이 영상평면(image plane) s 에 맺히는 것을 보여준다. 이 매핑을 일반적으로 평면-평면 매핑(plane-to-plane mapping)이라 부른다.

카메라투영에 의해서 평면이 평면으로 매핑되는 관계는 투영기하학 관점에서 평면투영이 성립한다. 이 평면투영은 평면 $S(A, B, C, D)$ 와 평면 $s(a, b, c, d)$ 사이의 최소 4개의 대응점(corresponding point)에 의해서 정의된다. 이렇게 정의된 변환을 평면투영변환이라 한다

이 평면투영변환은 동차 좌표계를 이용하면 행렬을 사용한 선형식으로 표현할 수 있다. 이 동차 좌표계를 이용하면 평면투영변환은 3×3 정방 행렬(square matrix)로 표시되어지고 평면 S 와 평면 s 의 관계는 다음과 같은 선형적인 행렬식으로 표현된다

$$X_2 = TX_1$$

또는,

(여기서 λ 는 스케일(scale)을 의미)

$$\lambda \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \quad (4)$$

으로 표현된다. 여기에서 $X_2=(x_2, y_2, 1)$ 은 평면 S 좌표계(S coordinate system)내에서의 주어진 점의 위치를, $X_1=(x_1, y_1, 1)$ 는 평면 s좌표계(s coordinate system)내에서의 주어진 점의 위치를 나타낸다 이 두 좌표의 관계를 나타내는 3*3정방행렬 T를 평면투영좌표계라 한다 변환 T은 9개의 자유도(freedom)에 의존하게 되는데 스케일 항인 λ 을 고려하면 실제로 아래와 같은 8개의 자유도에 의존하게 된다.

$$\lambda \begin{bmatrix} x_{2i} \\ y_{2i} \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix} \begin{bmatrix} x_{1i} \\ y_{1i} \\ 1 \end{bmatrix} \quad (5)$$

여기에서 x_{1i} 와 x_{2i} 는 두평면간의 4개 이상의 대응을 나타낸다 식 (5)에서 $a_{33}=1$ 이 된다.

식 (5)로부터 다음과 같은 식을 얻는다

$$\lambda y_{2i} = a_{21}x_{1i} + a_{22}y_{1i} + a_{23} \quad (6)$$

$$\lambda x_{2i} = a_{11}x_{1i} + a_{12}y_{1i} + a_{13} \quad (7)$$

$$\lambda = a_{31}x_{1i} + a_{32}y_{1i} + 1 \quad (8)$$

위의 식으로부터 다음과 같은 두 개의 독립된 식이 얻어진다

$$x_{2i} = \frac{a_{11}x_{1i} + a_{12}y_{1i} + a_{13}}{a_{31}x_{1i} + a_{32}y_{1i} + 1} \quad (9)$$

$$y_{2i} = \frac{a_{21}x_{1i} + a_{22}y_{1i} + a_{23}}{a_{31}x_{1i} + a_{32}y_{1i} + 1} \quad (10)$$

따라서 하나의 대응점으로부터 2개의 독립된 식을 얻을 수 있으므로 4개의 대응점만 있으면 평면투영변환 T(8개의 자유도를 가짐)가 구해짐을 알 수 있다.

식(9), (10)을 정리하면

$$x_{1i}a_{11} + y_{1i}a_{12} + a_{13} - x_{1i}x_{2i}a_{31} - y_{1i}x_{2i}a_{32} = x_{2i} \quad (11)$$

$$x_{1i}a_{21} + y_{1i}a_{22} + a_{23} - x_{1i}x_{2i}a_{31} - y_{1i}x_{2i}a_{32} = y_{2i} \quad (12)$$

으로 정리되고 이 두 식을 4개의 대응점 (N=4)에 대해서 행렬 형태로 나타내면,

$$Aq = z$$

또는

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x_{2i} & -y_1x_{2i} \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y_{2i} & -y_1y_{2i} \\ x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x_{2i} & -y_1x_{2i} \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y_{2i} & -y_1y_{2i} \\ x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x_{2i} & -y_1x_{2i} \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y_{2i} & -y_1y_{2i} \\ x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x_{2i} & -y_1x_{2i} \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y_{2i} & -y_1y_{2i} \end{bmatrix} \begin{bmatrix} \lambda \\ a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \\ a_{13} \\ a_{23} \\ a_{31} \\ a_{32} \end{bmatrix} = \begin{bmatrix} x_{2i} \\ y_{2i} \\ x_{2i} \\ y_{2i} \\ x_{2i} \\ y_{2i} \\ x_{2i} \\ y_{2i} \end{bmatrix} \quad (13)$$

으로 나타낸다 여기서 벡터 q는 아래와 같이 구할 수 있다

$$q = A^{-1}z \quad (14)$$

벡터 q의 각각의 항으로부터 변환 T를 얻을 수 있다 여기서 행렬 A의 역(inverse)이 존재하기 위해서는 A의 각 열(row)들이 독립적이어야 한다 이것이 만족되기 위해서는 대응점으로 쓰인 4개의 점 중에서 적어도 3개의 점이 일직선상에 존재하지 않아야 (non-collinear)한다 그렇지 않으면 4개의 점간의 종속성(dependency)으로 인해 행렬A의 열들간에 종속성이 생기게 되므로 역이 존재하지 않게 된다 이렇게 구해진 T를 이용하면 한 평면에 있는 나머지 점들에 대해서도 식 (4)에 의해 임의의 평면에 있는 대응점의 위치가 구해짐을 알 수 있다.

3.3 복도의 모델과 입력영상 좌표

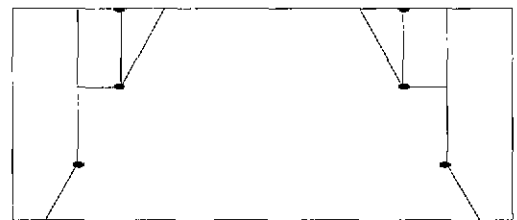


Fig 4 Model of a corridor

입력영상에서 주어진 위치가 실제 복도에서 어느 위치에 해당하는지를 알아내기 위해서 실제 복도에 대한 모델이 필요하다

그림4는 실제로 카메라가 설치된 복도의 규격을 측하여 만든 복도의 모델을 보여준다 이 모델을 기반으로 영상 좌표를 모델기준좌표로 변환한다. 그림 4에 복도의 존재하는 특징점들이 표시되어 있다. 복도 모델을 평면 S라하고 입력 영상을 평면 s라 하면 입력 영상에서 보이는 모델의 메칭점과 입력 영상간

의 4개의 매칭점들을 이용하여 모델과 입력영상간의 변환T를 구해낼 수 있다

4. 실험 결과

본 실험은 복도의 모델과 유사한 지역에 카메라를 부착하여 기준점을 잡고 그 기준점으로 부터의 물체의 크기를 측정할 수 있도록 하였다

그림 5는 복도의 실제 모습을 나타내고 있다. 자바로 프로그램을 구성하여 익스플로러에서 바로 실행이 가능하다

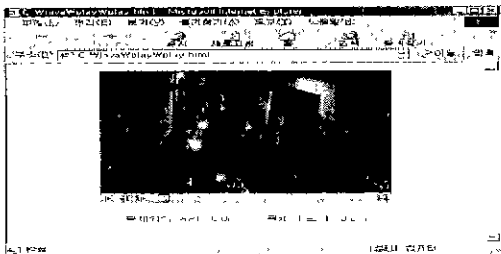


Fig. 5 Actual figure of a corridor

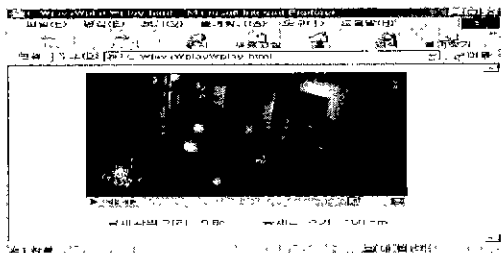


Fig. 6 Measurement of the distance from the camera and the height of an object

그림 6은 복도에 나타난 물체에 대한 거리와 크기를 측정한 값을 나타내고 있다

한편, 데이터 전송실험에서는 RTP방식을 이용하였다. 그림 7는 RTP 과정을 보여주고 있다

File format은 raw video/audio format으로 한다. 이를 JPEG 형식으로 인코딩한 다음 패킷으로 묶어 RTP방식으로 데이터를 전송한다

그림 8은 data의 전송량을 나타내고 있다



Fig. 7 RTP process

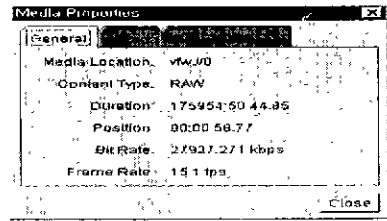


Fig. 8 Data transfer rate

5. 결론

본 실험에서 물체의 거리와 크기를 측정하는데는 성공했지만 많은 문제점을 발견하였다. 갑작스런 조명의 변화는 카메라의 측정이 불가능해지고, rtp통신시 많은 지연문제가 생김을 볼 수 있었다.

향후 카메라의 조명에 대한 연구가 더욱 필요하며, rtp통신시 지연문제를 해결하기 위해서 많은 data 양을 줄일 수 있는 압축의 기술이 필요하다

참고문헌

- 1 Johnson I Agbinya and Daviud Rees, "Multi-Object Tracking in Video." CSIRO Telecommunication and Industrial Physics. Image & Signal Processing Group. PO Box 76.Epping, NSW 2121, Australia,1999
- 2 Seo, Y., Choi, S., Kim, H and Hong, K-S, "Where are the ball and player? Soccer game analysis with colorbased tracking and image mosaic." Proc ICIAP,1997.
- 3 Cavallaro, R. "The Foxtrax hockey puck tracking system," IEEE Computer Graphics Applications 17 : 6 - 12 , 1997.
- 4 JAVA Media Framework API Guide, Sun Microsystems, Inc. 1998-99
- 5 Universal Serial Bus Specification, Compaq, Intel, Microsoft, NEC, 1998
- 6 Th Markousis, D Tsirikos, M. Vazirgiannis, Y. Stavarakas, "WWW-enabled delivery of interactive multimedia documents" Computer Communications 23 242-252, 2000.
- 7 Christoph Bregler, "Learning and Recognizing Human Dynamics in Video Sequences," Proc. IEEE Conf. Comp. Vision and Pattern Recognition, San Juan, Puerto Rico, 1997