

Color calibration과 Stereo Information을 이용한 프로젝션 화면 내의 전경물체 검출

홍광진⁰ 정기철 강현
송실대학교, 정보과학대학, 미디어학과
hongmsz@hci.ssu.ac.kr kcjung@ssu.ac.kr

Foreground object detection using color calibration and stereo information In projection display

Kwangjin Hong⁰ Keechul Jung, Hyun Kang
School of Media, College of Information Science, Soongsil University.

요 약

프로젝션 화면(projection display) 상에 보여지는 가상의 물체를 사용자가 직접 조작할 수 있는 인터페이스를 제공하기 위해서는 전경 물체를 검출해내는 과정이 필수적이다. 이전의 색상 정보만을 이용하는 방법은 몇 가지 제약 조건을 가지고 있었다. 본 논문은 색상 보정(color calibration)과 스테레오 정보(stereo information)를 이용하여 프로젝트 화면 내의 전경물체를 검출하는 방법을 제안한다. 실험에서는 프로젝터를 통해 책상 표면에 투사되는 영상과 일반 캠코더를 통해 얻어진 영상 사이의 왜곡을 기하 왜곡과 색상 왜곡으로 정의하여 모델링 하였고, 스테레오 정보를 이용하여 얻어진 최종 결과를 통해 제안된 방법의 실효성을 입증할 수 있었다.

1. 서 론

디지털데스크(DigitalDesk), 오그먼트데스크(Augmented Desk), 인헨스드데스크(Enhanced Desk) 등의 시스템들로 소개되는, 프로젝트 화면에서의 상호 작용에 관한 최근의 연구들은, 유비쿼터스 컴퓨팅(Ubiquitous Computing)과 증강 현실(Augmented Reality)을 구현하기 위한 좋은 기회를 제공하고 있다. 이러한 프로젝트 화면을 사용하는 시스템에서는 전경물체의 검출이 반드시 필요하다. 기존에는 전경물체 검출을 위해 별도의 장치나 영상처리 방법을 사용했다. 오그먼트데스크를 제안한 Yoichi Sato는 시스템에 적외선 카메라를 설치하여 전경물체를 검출하는 방법을 사용했다 [1]. 다른 예로는 Rekimoto가 제안한 스마트스킨(Smartskin) 센서를 이용하는 방법 [2], 데이터글러브(DataGlove)와 HMD(Head Mounted Display)를 사용하는 방법 등이 있다. 또한, 영상처리 방법을 사용하는 예로는 색상을 이용하거나, 동영상의 프레임 차이를 이용하여 전경물체를 검출하는 것 등이 있다. 그러나 이러한 방법들은 한계를 가지고 있다. 프레임의 차이를 이용하는 방법은 움직이지 않는 물체를, 적외선 카메라를 이용하는 방법은 무생물물, 스마트스킨 센서를 이용하는 방법은 제스처 등의 다양한 동작을 검출하지 못한다.

프로젝션 화면에서의 상호 작용은 프로젝터를 통해 화면에 투사되는 가상물체와 손이나 종이와 같은 실제물체 사이에서 이루어진다. 전경물체는 프로젝터의 입력 영상(프레임 버퍼 영상)과 캠코더를 통해 얻어지는 영상(카메

라 영상)을 사용하여 검출할 수 있다[3]. 그러나 두 영상의 색상 정보만 이용해서 전경물체를 검출하기 위해서는 몇 가지 제약 조건을 가지는 환경을 필요로 한다. 본 논문은 스테레오 정보를 이용하여 좀더 일반적인 상황에서도 전경물체의 검출이 가능한 방법을 제안한다.

2. 전경물체 검출

프로젝션 화면 내에서 색상 보정과 스테레오 정보를 사용하여 전경물체를 검출하기 위해 다음의 방법을 사용한다. 첫 번째로 카메라 보정에 사용되는 방법으로 기하 왜곡 문제를 해결하고, 두 번째로 색상 밝기 추정 방법을 사용하여 색상 왜곡을 해결한다. 끝으로 스테레오 정보를 가지고 있는 영상(스테레오 영상)을 사용하여 전경물체를 검출한다.

가상물체의 색상정보는 프레임 버퍼 영상으로부터 쉽게 얻을 수 있고, 실제물체의 색상 정보는 카메라 영상으로부터 얻을 수 있다. 하지만 프레임 버퍼 영상과 달리 카메라 영상은 많은 왜곡을 가지고 있다. 본 논문에서는 카메라 영상이 가지는 왜곡으로 기하 왜곡과 색상 왜곡을 사용한다. 기하 왜곡은 카메라 영상과 프레임 버퍼 영상 사이에 서로 대응하는 픽셀의 위치 차이로 정의하고, 색상 왜곡은 카메라 영상과 프레임 버퍼 영상 사이에 서로 대응하는 픽셀의 색상 차이로 정의한다. 또한 기존의 색상정보만을 이용한 검출 방법으로는 배경이 어두울 경우 전경물체의 검출이 어렵다는 문제점을 수정하기 위해 스테레오 정보를 사용하기로 한다.

2.1 기하 보정 (Geometry Calibration)

Ashdown[4]와 Sukthankar[5]은 서로 다른 좌표계를 일치시키기 위한 방법으로 투영 변환을 사용하였다. 본 논문에서도 투영 변환 방법을 사용하여 기하 보정을 한다. 투영 변환을 사용하면 이러한 대응하는 픽셀들의 관계를 설명할 수 있다. 기하 왜곡을 투영 변환 관계에 있는 두 픽셀의 위치로 모델링 해보면, 식(1)과 같이 표현된다.

$$(x, y, w) = H \begin{pmatrix} u \\ v \\ z \end{pmatrix}, \quad H = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} u_0 & v_0 & 1 & 0 & 0 & 0 & -u_0x_0 & -v_0x_0 & -x_0 \\ u_1 & v_1 & 1 & 0 & 0 & 0 & -u_1x_1 & -v_1x_1 & -x_1 \\ u_2 & v_2 & 1 & 0 & 0 & 0 & -u_2x_2 & -v_2x_2 & -x_2 \\ u_3 & v_3 & 1 & 0 & 0 & 0 & -u_3x_3 & -v_3x_3 & -x_3 \\ 0 & 0 & 0 & u_0 & v_0 & 1 & -u_0y_0 & -v_0y_0 & -y_0 \\ 0 & 0 & 0 & u_1 & v_1 & 1 & -u_1y_1 & -v_1y_1 & -y_1 \\ 0 & 0 & 0 & u_2 & v_2 & 1 & -u_2y_2 & -v_2y_2 & -y_2 \\ 0 & 0 & 0 & u_3 & v_3 & 1 & -u_3y_3 & -v_3y_3 & -y_3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \end{bmatrix} = 0 \quad (2)$$

9개의 상수로 이루어진 투영변환 행렬은 4개 이상의 프레임 버퍼 영상과 카메라 영상의 대응하는 픽셀들에 의해서 계산된다. 식 (2)에서 행렬을 **A**라고 하고 **A**는 4개 이상의 대응하는 픽셀들에 의해서 만들어진다. 투영 행렬 H의 상수 a에서 i까지는 **A^TA**의 최소 고유 값과 고유 벡터에 의해 결정된다 [7].

2.2 색상 보정 (Color Calibration)

본 논문에서는 색상 정보의 밝기 값을 사용한다. 한 픽셀(R, G, B)의 밝기 정보(Y)는 아래 식 (3)에 의해 결정된다.

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (3)$$

프로젝션 화면상에서 밝기 값의 왜곡은 시간과 공간에 따라 다양하게 변한다. 본 논문에서는 공간 위치에 따라 화면의 동일 밝기 픽셀의 변화는 전경물체에 의한 변화 값 보다 극히 작다고 가정하여, 동일 밝기 픽셀은 공간 위치에 따라 변화가 없다고 가정한다. 반면에 시간에 따른 밝기 값의 변화는 좀 더 복잡하다. 조명의 변화나 프로젝트 화면내의 전체 밝기에 따라 밝기의 시간적 변화가 있다.

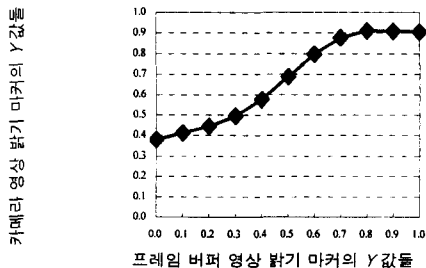


그림 1. 밝기 마커의 프레임 버퍼 영상과 카메라 영상간의 관계 도표.

이런 경우에도 시간에 따른 변화를 측정하기 위해 프로젝트션 화면의 일부에 **밝기 마커(brightness color marker)**를 만든다. 밝기 마커는 10단계의 회색 값을 가지는 10개의 블록으로 이루어져 있다. 이는 프로젝트션 화면의 상단에 항상 보이도록 설치된다. 그림 1은 특정 시간에서의 프레임 버퍼 안의 밝기 마커의 Y값들과 카메라 영상내의 밝기 마커의 Y값들의 관계 곡선을 보여준다. 프로젝트, 카메라의 세팅 또는 조명의 상태에 따라 전체 그래프가 아래위로 움직이지만 전체의 형태는 변하지 않는다.

2.3 스테레오 정보 이용

기하보정과 색상보정을 통해 얻어진 수정된 카메라 영상과 프레임 버퍼 영상의 색상 차이를 이용하여 현재 픽셀이 전경물체에 속하는지 여부를 결정할 수 있다. 하지만 색상정보, 특히 밝기 값만을 이용할 경우 몇 가지 제약 조건을 가지게 된다. 프로젝트는 화면을 투사할 때 R, G, B의 세가지 빛을 조합하기 때문에 검은색에 가까울수록 빛의 세기가 0에 가깝게 된다. 따라서 배경 화면이 너무 어두울 경우, 보정 과정을 거쳐 수정된 카메라 영상은 배경과 전경물체의 밝기가 모두 0으로 보정이 되기 때문에 전경물체를 검출해 내지 못하게 된다.

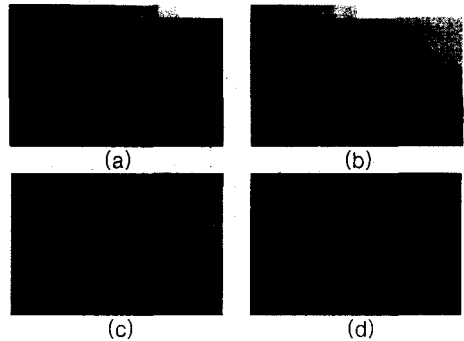


그림 2. 검은색 바탕에서의 전경물체 검출 결과.

그림 2는 검은색 바탕에서의 전경물체 검출 결과를 보여준다. (a)는 프레임 버퍼 영상이고, (b)는 기하 보정된 카메라 영상, (c)는 (a)와 (b)의 픽셀의 차를 이용해 얻어진 영상(차영상), (d)는 강현, 정기철 등에 의해 제안된 color calibration을 사용하여 얻어진 전경물체 검출 결과이다 [3].

본 논문에서는 색상 정보와 스테레오 영상을 사용하여 전경물체를 검출한다. 스테레오 영상은 일반적인 영상이 x, y의 2차원 정보를 가지는 것과 달리, x, y, z의 3차원 정보를 가지고 있다. 스테레오 카메라는 두 개 이상의 카메라를 통해 들어오는 영상들의 disparity를 계산하고, 이를 이용해서 영상 내의 depth를 측정한다.

그림 3는 두 개의 영상을 이용한 depth를 측정하는 것을 보여준다. L과 R은 점 P를 바라보는 두 대의 카메라에서 상이 맺히는 면을 나타내고 있다. L에 맺힌 P의 상을 P_L, R에 맺힌 P의 상을 P_R이라 하고, 각각의 카메라의 중심에서부터 떨어져있는 거리를 x_L, x_R이라고 할 때,

카메라에서 점 P까지의 거리 z는 다음의 식 (4)와 같이 구할 수 있다 [6].

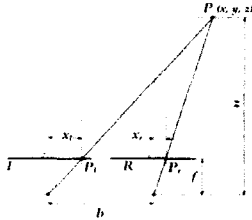


그림 3. 두 영상을 사용한 depth 검출.

$$\frac{x}{z} = \frac{x_1}{f}, \quad \frac{x-b}{z} = \frac{x_2}{f} \quad \therefore z = \frac{bf}{(x_1 - x_2)} \quad (4)$$

따라서 스테레오 영상을 이용하면 바탕 화면이 투사되는 바닥과 손이나 실제 문서 등의 전경물체를 검출해낼 수 있다. 이 때, 보다 확실한 검출을 위해 스테레오 영상을 가장 어두운 값을 0(검은색), 가장 밝은 값을 255(흰색)로 변환하여 사용한다.



그림 4. 스테레오 영상 (a)변환 전, (b)변환 후.

우선 수정된 카메라 영상과 프레임 버퍼 영상의 차영상을 만든다. 그리고, 이 영상과 스테레오 영상의 서로 대응하는 픽셀들의 밝기의 기하평균 값을 이용하여 전경물체를 검출한다. 기하평균을 이용하는 이유는 스테레오 영상에서 0의 값을 가지는 부분이 배경을 나타내므로 차영상에서 배경 부분을 구분해내기 쉽기 때문이다.

3. 실험 및 결과

구현된 디지털데스크 시스템은 프로젝터, 카메라, 일반 책상과 2.66GHz의 일반 IBM PC장비로 구성되어있다. 프로젝터는 벤큐의 HD2100 프로젝터이며, 카메라는 소니의 DCR-PC5 캠코더와 Pointgrey의 Digiclops Vision System을 사용하였다. 시스템은 Visual C++ 6.0과 DirectShow 9.0 SDK를 사용하여 720X480 크기의 컬러 영상을 사용한다. 그림 5은 스테레오 정보를 이용한 전경물체 검출의 결과이다. (c)까지의 영상은 그림 2와 동일하고, (d)는 변환된 스테레오 영상, (e)는 차영상과 스테레오 영상의 기하평균을 이용해 얻어진 결과 영상이다. (f)가 제안된 방법을 통해 얻어진 결과 영상이다. 전경물체는 손을 사용하였고, 배경화면은 검은색(R=0, G=0, B=0)의 윈도우 바탕화면을 사용하였다.

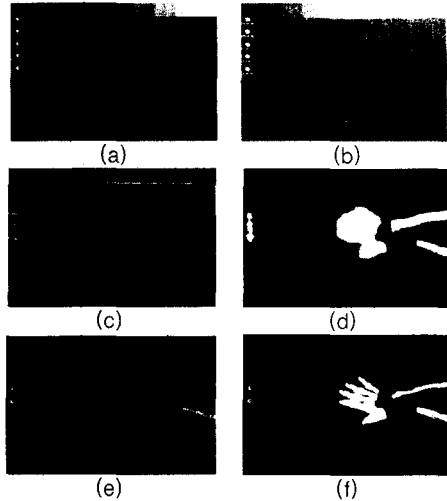


그림 5. 스테레오 정보를 이용한 전경물체 검출 결과.

4. 결론

본 논문은 프로젝션 화면 상에서 색상 보정과 스테레오 정보를 이용해 전경물체를 검출하는 방법을 제안하였다. 프로젝션 화면에서 발생하는 대표적인 왜곡 두 가지를 사용하였고, 투영 변환 기법을 사용하여 기하 왜곡을, 밝기 값을 사용하여 색상 왜곡을 보정하였으며, 여기에 스테레오 영상을 이용한 계산을 통해 전경물체를 검출하였다. 그 결과 배경이 어두운 환경에서도 만족스러운 성능을 얻을 수 있었다.

현재의 시스템은 색상 보정을 위해 카메라 영상의 밝기 정보만을 이용한다. 색상과 채도의 변화, 스테레오 정보를 고려하여 다수의 전경물체 검출 및 구분 방법을 제안할 것이다.

5. 참고문헌

- [1] Yoichi Sato, Yoshinori Kobayashi, Hideki Koike, "Fast Tracking of Hands and Fingertips in Infrared Images for Augmented Desk Interface," Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 462-467, 2000.
- [2] Rekimoto, J., "SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces," Proceedings of CHI, ACM, 2002, pp. 113-120, 2002.
- [3] 강현, 정기철 등, "색상 정보를 이용한 디지털데스크 상의 전경물체 검출," Proceedings of the 30th KISS Fall Conference, Vol. 30, No. 2, pp. 520-522, 2003.
- [4] Ashdown, M., Robinson, P., "The Escritore: A Personal Projected Display," Journal of WSCG, Vol.11, No. 1, pp. 33-40, 2003.
- [5] Sukthankar, R., Stockton, R.G., Mullin, M. D., "Smarter Presentations: Exploiting Homography in Camera-Projector Systems," proceedings on ICCV, 247-253, 2001.
- [6] Ramesh Jain, Rangachar Kasturi, Brian G. Schunk, "Machine Vision," McGraw-Hill, Inc., pp.289-293, 1995.
- [7] Hartley, R., Zisserman, A., "Multiple View Geometry in Computer Vision," Cambridge University Press, 2000.