스테레오 카메라를 이용한 3차원 입력장치 움직임의 실시간 추출

윤상민*·**, 김익재*, 안상철*, 김형곤*, 고한석**
*한국과학기술연구원 영상미디어 연구센터, **고려대학교 전자공학과

Real Time Motion Extraction of 3-D Input Device Using Stereo Camera

Yoon, SangMin*,**, Ig-Jae Kim*, SangChul Ahn*, HyoungGon Kim*, HanSeok Ko**
* Imaging Media Research Center, KIST
** Intelligent Information and Signal Processing Lab., Dep. of Electronics Engineering, Korea University
{smyoon, kij, asc, hgk}@cherry.kist.re.kr, hsko@korea.ac.kr

요약

본 논문은 두 대의 카메라로 제안하는 물체의 색상, 움직임, 형태상의 특성을 이용하여 3차원 공간상의 움직임을 실시간으로 추출하는 것을 목적으로 한다. 본 논문에서 제안하는 물체는 구조상 물체 자체가 캘리브레이션 물체의 역할을 포함하여 캘리브레이션이 되지 않은 상황에서도 정확하게 물체의 3차원 정보를 추출할 수 있으므로 3차원 입력 디바이스로 이용할 수 있다. 3차원 움직임을 추출하기 위해 먼저 3차원 공간상의 물체와 좌우 영상의 상관관계를 구하고, 좌우 각 영상에서 원이 위치할 탐색영역은 MAWUPC 알고리즘을 이용하여 예측한다. 탐색영역 안에서 PCA를 사용하여 원의 정확한 위치를 찾으며 좌우 영상에서 얻은 원의 위치와 스테레오 카메라의 기하학적 구조를 종합하여 3차원 움직임을 추출한다. 추출한 3차원 움직임은 가상환경에서 가상 물체의 움직임을 제어하는데 응용할 수 있다.

1. 서 론

HCI (Human-Computer Interface) 분야 중에 카메라로부터 입력된 영상을 대상으로 컴퓨터 비전 방법을 이용하여 물체의 움직임을 추출하는 연구가 근래에 많은 주목을 받고 있다.[1] 이와 같이 카메라를 이용한 물체의 움직임 추출은 인간과 컴퓨터의 상호작용에서 가장 자연스럽고 직관적인 인터페이스로 사용될 수 있다. 입력 디바이스로는 현재 키보드,마우스,조이스틱이 가장 보편적이면서, 가장 인기 있지만 거리

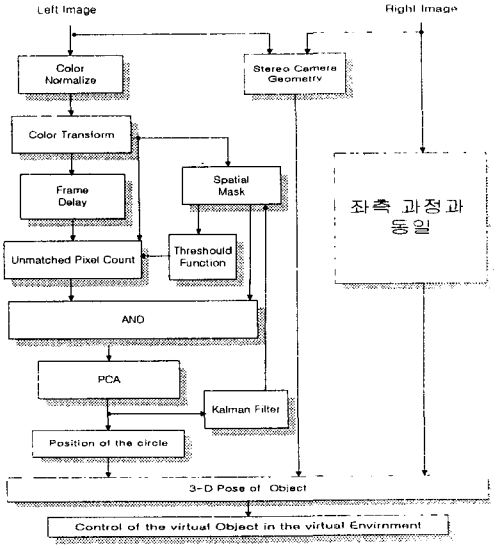


그림 1. 시스템 흐름도

와 간섭등에 의한 단점이 있으며 손과 얼굴을 입력 디바이스로 이용하는 방법은 3차원 움직임을 추출하는데 부정확하다는 단점이 있다. 하지만 본 논문에서 제안하는 물체는 3차원 공간에서 사용자 중심의 간단하면서도 정확한 3차원 입력 디바이스를 구현하기에 적합하다. 스테레오 카메라의 기하학적 구조와 물체의 색상, 움직임, 형태상의 특성을 효율적으로 이용하기 위해 그림 1과 같은 과정을 거쳐 정확한 3차원 정보를 얻는다.

2.스테레오 카메라의 기하학적 구조

스테레오 비전의 가장 큰 장점은 공간상에서 물체의 깊이 정보를 추출할 수 있다는 것이다. 이를 위해 먼저 좌우 카메라의 상관관계 및 물체와 영상과의 상관관계를 구해야 한다. 카메라 캘리브레이션 방법은 물체와 영상사이의 내부변수와 외부변수를 구하는 방법과 좌우 영상에서 생기는 에피폴의 기하학적 구조를 이용하는 방법이 있다.[2][3] 본 논문에서는 좌우 영상의 동일한 평면에 존재하지 않는 4개의 특징 점을 이용하여 perspective projection 행렬과 perspective distortion 행렬을 이용하여 스테레오 카메라의 상관관계를 구한다.[4] 기존에는 두 카메라의 상관관계를 구하기 위해서 별도의 캘리브레이션 물체가 필요하였지만 그림 2와 같이 제안하는 물체는 그 자체가 캘리브레이션 물체의 역할을 포함하며 3차원 공간상의 움직임 자체가 가상환경에서의 가상물체의 움직임을 제어할 수 있다.

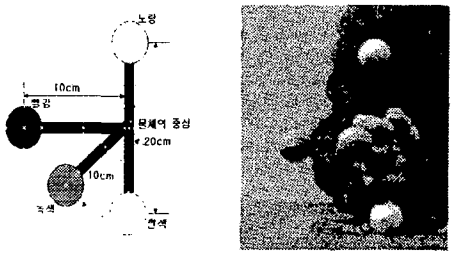


그림 2. 제안하는 물체

3차원 공간상의 물체와 영상과의 상관관계는 perspective projection 행렬과 perspective distortion 행렬을 통해서 식(1)과 같이 나타낼 수 있다.

$$\begin{aligned} \lambda_i[x_i, y_i, z_i, t_i]^T &= D[X_i, Y_i, Z_i, 1]^T \\ s[u_i, v_i, 1]^T &= P[x_i, y_i, z_i, t_i]^T \\ s'[u_i', v_i', 1]^T &= P'[x_i, y_i, z_i, t_i]^T \end{aligned} \quad (1)$$

3차원 공간상의 물체와 영상과의 상관관계를 나타내기 위한 perspective projection 행렬은 좌우 각각 P와 P'로 좌우 영상의 특징점의 매칭을 통해 3차원 구조를 알 수 있다. D는 perspective distortion 행렬로 $[x_i, y_i, z_i]$ 의 좌표축과 $[X_i, Y_i, Z_i]$ 의 좌표축을 각 카메라의 광학중심을 기준으로 동일선상에 있게 하기 위한 행렬이다. $[u_i, v_i]$ 와 $[u_i', v_i']$ 는 좌우 영상에서의 물체의 위치이며 $[X_i, Y_i, Z_i]$ 초기 위치의 물체의 중심을 기준으로 한 월드 좌표계에서의 물체의 3차원 공간상에서의 위치이다. i 는 i 번째 특징 점을 나타내고 s, s', λ_i 는 scaling factor이다. 이와 같은 방법을 통해 얻은 좌우의 평면 방정식을 이용하여 3차원 공간상의 물체의 위치를 얻을 수 있다. 이를 위해서는 먼저 좌우 영상에서 원의 정확한 위치를 구해야 한다.

3. 물체의 정확한 움직임 추출

3.1 색상변환

특정 색상을 가진 물체의 움직임을 영상에서 추출하기 위한 과정이다. 일반적으로 영상의 한 픽셀은 RGB 색상모델을 사용하여 표현되지만 RGB 색상모델은 그 픽셀 값 자체에 색상 성분과 밝기 성분을 포함하고 있다. 따라서 동일한 색상을 가진 물체라도 조명의 영향에 따라 RGB 값은 달라질 수 있다. 따라서 물체의 색상 성분만 남기고 밝기 성분을 제거하기 위해서 식(2)와 같이 RGB 색상 성분을 정규화한다.

$$\begin{aligned} I &= R + G + B, \\ r &= \frac{R}{I}, g = \frac{G}{I}, b = \frac{B}{I} \end{aligned} \quad (2)$$

정규화된 RGB 색상모델에서 물체의 색상 분포는 다음식과 같이 가우시안 분포로 모델링될 수 있다.

$$G(r, g) = N(m, \Sigma^2), \quad (3)$$

$$\bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i, \bar{g} = \frac{1}{n}\sum_{i=1}^{n} g_i, \quad (4)$$

$$\Sigma^2 = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_g^2 \end{bmatrix}. \quad (5)$$

위의 식과 같이 가우시안 분포를 역으로 식(6)처럼 색상변환 함수로 사용하면 입력되는 컬러영상의 색상이 모델링되어 대상 물체의 색상 분포에 가까울수록 큰 값을 가지는 그레이 영상 Z 로 변환할 수 있다

$$\begin{aligned} &G: R^2 \rightarrow R, \\ &Z(x, y) = G(r(x, y), g(x, y)). \end{aligned} \quad (6)$$

3.2 MAWUPC

MAWUPC (Motion Adaptive Weighted Unmatched Pixel Count)는 물체의 움직임 정보와 색상 정보를 이용하여 특정 색상의 움직이는 물체를 추출하고 칼만 필터를 결합하여 다음 프레임에서 물체가 위치할 부분을 예측하는 알고리즘이다. 먼저 물체의 변화 영역을 검출하기 위해서 AWUPC (Adaptive Weighted Unmatched Pixel Count)를 이용하고 있다.[5] AWUPC는 UPC에 색상 변화의 판단을 위해 입계값 함수에 적용하고, UPC 연산 결과에 가중치로 사용하는

것으로 구성된다.

$$AWUPC(x,y,t) = Z(x,y,t)\sum_{i=x-N}^{i=x+N}\sum_{j=y-N}^{j=y+N}U(i,j,t), \quad (7)$$

$$U(i,j,t) = \begin{cases} 1, & if\,|Z(i,j,t) - Z(i,j,t-1)| \rangle Th(Z) \\ 0, & otherwise \end{cases}, \quad (8)$$

$$Th(Z) = \frac{255}{1+\exp(\frac{Z(x,y,t)-255/2}{Q})}. \quad (9)$$

움직임 검출에 있어서 색상변환의 반영은 모델링된 색상 분포에 가까운 색상 영역을 강조할 뿐만 아니라, 이러한 영역의 움직임을 부각하게 된다. 임계값 함수는 시그모이드 함수를 이용한다. AWUPC에 매 프레임의 움직임 정보를 결합한 알고리즘이 MAWUPC이다.[6] 먼저 이전 프레임에서 얻은 위치로부터 현재 프레임의 물체의 위치를 예측하기 위해 칼만 필터를 사용하여 물체의 위치를 얻기 위한 탐색 영역을 설정하고 이 탐색 영역 안에서 2차원 가우시안 분포를 가지는 마스크를 적용하는 것이다. 가우시안 분포의 2차원 공간 마스크는 AWUPC에서 사용되는 색상 변환 값에 적용된다.

$$SM(x,y) = N(m_p, \Sigma_p^2), \quad (10)$$

$$Z'(x,y,t) = SM(x,y)Z(x,y,t), \quad (11)$$

$$MAWUPC(x,y,t) = Z'(x,y,t)\sum_{i=x-N}^{i=x+N}\sum_{j=y-N}^{j=y+N}U(i,j,t), \quad (12)$$

$$U(i,j,t) = \begin{cases} 1, & if\,|Z(i,j,t) - Z(i,j,t-1)| \rangle Th(Z') \\ 0, & otherwise \end{cases}, \quad (13)$$

3.3 PCA를 이용한 물체의 중심 추출

좌우 영상에서 찾은 4개의 구의 중심을 이용하여 물체의 3차원 정보를 얻는데 있어서 약간의 오차도 3차원 움직임을 복원하는데 많은 오차가 포함될 수 있다. 이러한 오차를 줄이기 위해서 구의 정확한 중심을 찾을 필요성이 있다. MAWUPC를 이용하여 현재 프레임에서 4개의 물체가 위치할 탐색영역을 설정하고 이 탐색영역 안에서 정확한 물체의 중심을 구하기 위해서 PCA(Principal Component Analysis)를 이용한다. 원의 중심을 구하는 것은 크게 3부분으로 나눌 수 있는데 먼저 다양한 조명에서 방향성 있는 조명의 세기와 방향을 변화시켜 색상별로 구의 영상을 훈련시켜 eigen-vector를 구한다. 두 번째는 탐색영역 안에서 색상변환 후의 영상에서 원의 지름 후보와 중심 후보를 결정한다. 마지막으로 입력 영상의 원의 지름과 원의 중심을 변화시켜가면서 구의 영상을 정규화하여 PCA 방법으로 재구성된 영상과의 유클리디언 거리를 비교하여 거리가 가장 작을 때의 원의 중심을 최종적인 원의 중심으로 결정한다. 원의 정확한 중심을 얻는 과정은 물체의 정확한 3차원 정보를 얻는데 도움을 줄 뿐만 아니라 물체의 보다 정확한 운동 특성상 정보를 얻도록 해주며, 결국은 칼만 필터의 보다 정확한 예측이 가능하도록 해준다. 그림 3는 PCA를 이용하여 정확하게 원의 중심을 구하기 위한 조명의 방향을 변화시켜 훈련 영상과 탐색영역 안에서의 원의 중심과 지름의 후보를 구하는 영상 그리고 원의 중심과 지름을 변화시키면서 재구성한 영상의 예이다.

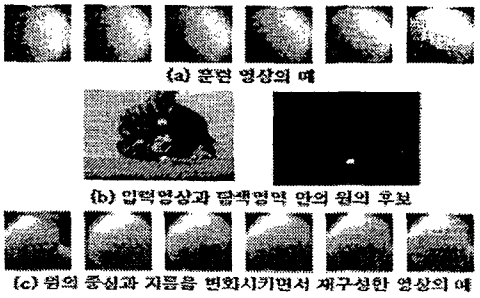


그림 3. PCA를 이용하여 정확한 구의 중심을 구하는 과정

4. 3차원 움직임 추출

본 논문에서 제안하는 물체의 3차원 움직임은 물체의 중심을 월드 좌표계의 중심으로 정의하고 X,Y,Z 축 방향으로의 이동과 각 축 방향으로의 회전을 틸트(tilt), 팬(pan), 요(yaw)의 회전으로 구분할 수 있다. 임의의 위치에서 빨강, 노랑, 흰색, 녹색, 물체의 중심을 $(X_1,Y_1,Z_1),(X_2,Y_2,Z_2),(X_3,Y_3,Z_3),(X_4,Y_4,Z_4),(X_0,Y_0,Z_0)$이라고 할 때 3차원 공간상에서의 회전은 다음 식과 같이 정의할 수 있다.

$$tilt = \tan^{-1}\frac{|Y_2-Y_0|}{|Z_2-Z_0|} = \tan^{-1}\frac{|Z_3-Z_0|}{|Y_3-Y_0|} = \tan^{-1}\frac{|Z_4-Z_0|}{|Y_4-Y_0|} \quad (14)$$

$$pan = \tan^{-1}\frac{|Z_1-Z_0|}{|X_1-X_0|} = \tan^{-1}\frac{|X_2-X_0|}{|Z_2-Z_0|} \quad (15)$$

$$yaw = \tan^{-1}\frac{|Y_1-Y_0|}{|X_1-X_0|} = \tan^{-1}\frac{|X_3-X_0|}{|Y_3-Y_0|} = \tan^{-1}\frac{|X_4-X_0|}{|Y_4-Y_0|} \quad (16)$$

5. 실험결과

본 논문에서 제안하는 알고리즘의 성능을 평가하기 위해서 SONY PC 100 캠코더 2대를 이용하여 초당 10프레임을 샘플링한 영상 시퀀스에 대해서 실험을 하였다. 실험에 사용된

영상의 크기는 360*240이며 일반적인 환경에서 실험하였다. 그림 4는 색상변환, MAWUPC 및 정확한 구의 중심을 구하는 과정과 좌우 영상에서 4개의 원의 중심을 추출한 결과 영상이다. 실시간으로 3차원 정보를 추출하기 위해서는 IEEE 1394 카메라를 이용하여 실험하였다. 영상 크기는 160*120이며 속도는 초당 5프레임이다. 정확도 측정은 그림 5과 같이 로봇암을 이용하였으며 표2에서 움직임을 비교하였다.

6. 결론 및 향후계획

본 논문에서는 스테레오 카메라와 제안하는 물체의 색상, 움직임 정보, PCA를 이용하여 연속되는 컬러 영상시퀀스에서 물체의 3차원 움직임 정보를 실시간으로 추출하였다. 이와 같은 방법을 이용하여 추출한 물체의 3차원 움직임은 실시간으로 가상환경에서의 가상 물체의 움직임을 제어하는데 응용할 수 있다. 앞으로 정확도 개선을 위해 서브픽셀로 개선하고 있다.

참고문헌

[1] D. M. Gavrila, "The Visual Analysis of Human Movement: A Survey, Computer Vision and Image Understanding," Vol. 73, No. 1, January, pp.82-98, 1999.

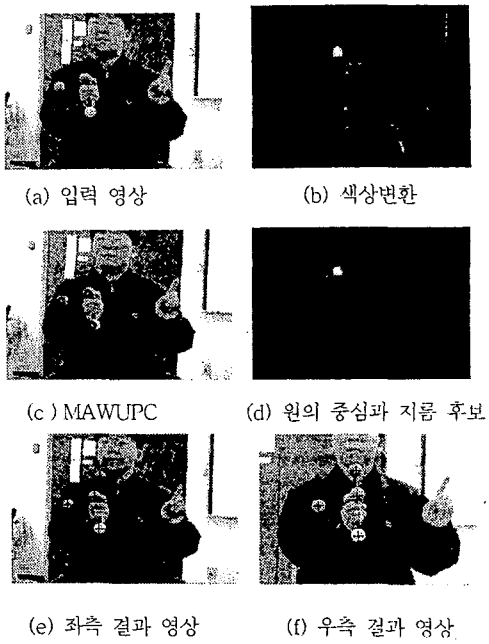
[2]R.T.Tsai. "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. In Proceedings CVPR '86, Miami Beach, Florida, pages 364-374. IEEE, June 1986.

[3]O.D. Faugeras, "What can be seen in three dimensions with an uncalibrated stereo rig." In proc. European Conference on Computer Vision, pages 563-578, 1992.

[4] Zhengyou Zhang, Oliver Faugeras, Rachid Deriche, "An Effecive Technique for Calibrating a Binocular Stereo Through Projective Reconstruction Using Both a Calibration Object and the Envirnmont" VIDERE1:1 pp58-68, 1997.

[5] H.G. Kim, N.H. Kim, S.C. Ahn, "Skin Region Extraction Using Moving Color Technique", The 4th IEEE International Workshop on Intelligent Signal Processing and Communication Systems (ISPACS '98), pp73-33, Nov. 1998, Australia..

[6] 이상환,"컴퓨터 비전을 이용한 사람의 3차원 정보 추출", 연세대학교 석사학위 논문, 2001.



(a) 입력 영상 (b) 색상변환

(c) MAWUPC (d) 원의 중심과 지름 후보

(e) 좌측 결과 영상 (f) 우측 결과 영상

그림 4 일반 환경에서의 실험 결과

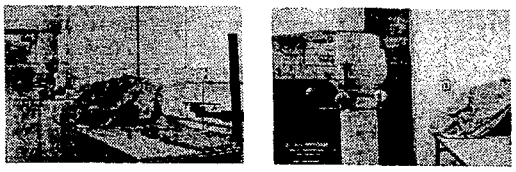


그림 5 로봇암에서의 구의 중심 추출 결과

Robot Arm			실험값		
X(cm)	Y(cm)	Z(cm)	X(cm)	Y(cm)	Z(cm)
2	0	0	1.9338	0.2486	0.3874
4	0	0	3.9196	0.1439	0.3139
6	0	0	6.2466	0.0180	-0.2290
0	-2	0	-0.0633	-1.8987	0.3962
0	-4	0	0.1191	-3.9668	0.2789
0	-6	0	0.2014	-6.1645	0.1624
0	0	2	-0.0738	0.3519	2.3691
0	0	4	-0.1752	0.3340	3.3854
0	0	6	-0.2138	0.3850	5.8377
Rotation (pan)		10도	Rotation (pan)		10.3651
		15도			14.5831

표 2. 로봇암의 움직임과 실험값과의 비교