라이브 마커를 이용한 대화형 증강 현실

장문석 박한훈 박종일
한양대학교 전자통신전파공학과
{fireduck, hanuni}@mr.hanyang.ac.kr, jipark@hanyang.ac.kr

Interactive Augmented Reality Using Live Markers

Jang, Moon-Seok Park, Hanhoon Park, Jong-Il
Division of Electrical and Computer Engineering, Hanyang University

요약

본 논문에서는 대화형 증강 현실을 위한 간단하고 유용한 방법을 제안한다. 대화형 증강 현실을 구현하기 위한 가장 일반적인 방법은 ARToolkit 라이브러리를 사용하는 것이다. 본 논문에서는 ARToolkit 라이브러리의 기능을 그대로 사용하지만 ARToolkit 라이브러리에서 사용하는 그래픽 마커를 라이브 마커로 대체한다. 따라서, 그래픽 마커를 사용할 때보다 직관적이면서도 편리한 사용자 인터페이스를 제공해 줄 수 있다. 단, 라이브 마커를 사용하기 때문에 ARToolkit 라이브러리의 마커 인식(Identification)을 위한 패턴 매칭 기술은 사용될 수 없으므로, 보다 정교한 객체 인식(object recognition) 기술이 요구된다. 본 논문에서는 원근 변환(perspective transformation)에 불변한 특성을 가지는 ART 기술자(Angular Radial Transform descriptor)를 사용한다. ART 기술자에 대한 간단한 성능 검증을 통해 라이브 마커에 적용 가능성을 확인하였고, 다양한 실험을 통해 라이브 마커의 효용성을 확인하였다. 본 논문에서 제안한 라이브 마커는 기존의 ARToolkit 사용자들에게 발전된 인터페이스 개념을 제공함으로써, 다양한 증강 현실의 구현에 응용될 수 있으리라 기대된다.

1. 서론

증강 현실은 컴퓨터에 의해 만들어진 그래픽을 현실 세계에 중첩함으로써, 사용자에게 그래픽이 마치 현실 세계에 존재하는 것처럼 느끼도록 만들어 주는 기술을 말한다[1]. 또한, 사용자에게 현실 세계와 관련된 다양한 부가적인 정보를 제공해 줄 수 있다. 현재 증강 현실을 구현하기 위한 다양한 방법이 제안되고 있으며, 그 중에 ARToolkit [2] 라이브러리는 간단하고 편리하게 증강 현실을 구현하기 위해 제공되는 공개형 라이브러리로 가장 널리 사용되고 있다. ARToolkit 라이브러리는 마커 기반 증강현실을 위한 것으로, 미리 제작된 다양한 그래픽 마커들을 카메라를 통해 인식(identification)하여 정해진 그래픽을 카메라 움직임에 대응하여 마커 위에 중첩시켜 준다.

ARToolkit 라이브러리에서 사용되는 그래픽 마커는 쉽고 강건하게 인식되지만, 마커 위에 중첩되는 그래픽과는 직접적인 연관성이 없다. 즉, 사용자에 의해 'Hiro'라는 마커는 아무런 연관성이 없는 정육면체를 중첩하도록 미리 설정된다. 물론, 마커와 중첩되는 그래픽의 연관성을 위해 마커를 중첩될 그래픽과 연관되도록 제작할 수 있다. 그러나, 그래픽과 연관되도록 마커를 제작해야 한다는 것은 사용자에게 큰 불편함을 제공할 것이다. 본 논문에서는 사용자에게 보다 직관적인 인터페이스를 제공하면서 마커 제작의 불편함을 해소하기 위해 그래픽 마커 대신 현실 세계에 존재하는 사물 자체를 마커로 사용하는 방법을 제안한다. 즉, 사용자가 자동차 그래픽을 중첩하고자 한다면, 미리 설정된 그래픽 마커를 사용하는 것이 아니라, 실제 자동차를 마커로 사용할 수 있다.

라이브 마커는 그림 1과 같이 ARToolkit 라이브러리에서 사용되는 그래픽 마커의 사각 프레임 만을 가진 것으로, 사각 프레임을 통해 관찰되는 사물과 함께 하나의 마커 역할을 한다. 라이브 마커를 인식하기 위해서는 객체 인식(object recognition) 기술이 필요하다. 객체 인식을 위해 목적 및 방법에 따라 다양한 기술들이 제안되어 왔다[3]. 본 논문에서는 원근 변환(perspective transformation)에 불변한 특성을 가지는 ART 기술자(Angular Radial Transform descriptor)를 사용한다. ART 기술자는 영역기반 기술자(region-based descriptor)로서, 상이한 모양을 가지는 객체간의 인식에는 우수한 인식률을 가지지만, 개와 늑대 같은 유사한 형태의 객체들은 구별하지 못한다[3]. 또한, 객체의 3차원 움직임에 의해 상이한 모양의 객체들이 유사한 모양으로 바뀌기도 하고, 유사한 모양의 객체들이 상이해지기도 한다. ART 기술자는 이러한 변화에 대응할 수 없다.[1] 사실, 본 논문에서 제안하는 라이브 마커는 객체 인식 기술의 성능에 크게 좌우된다. 그러나, 객체 인식 기술은 본 논문의 범주가 아니며, 단지 향상된 객체 인식 기술의 사용은 라이브 마커의 응용성을 확장시켜 줄 것이다.

본 논문은 다음과 같이 구성된다. 2 장에서는 라이브 마커에 대한 전반적인 개념과 마커 인식을 위해 사용된 ART 기술자에 대해 설명한다. 3 장에서는 ART 알고리즘에 대한 간단한 성능실험 및 라이브 마커를 이용한 예비 실험 결과를 보여준다. 4 장에서는 라이브 마커의 활용 예를 보여준다. 5 장에서는 결론 및 추후 과제를 보여준다.

2. 라이브 마커를 이용한 증강 현실

1) 자세한 설명은 3장을 참조하기 바란다.

라이브 마커는 그림 1에서 보는 것처럼 사각 프레임을 손에 들고 실제 세계의 물체를 프레임 안에 위치시킨 후 카메라를 이용하여 획득된다. 획득된 마커의 인식을 위해서 원근 변환에 불변한 ART 알고리즘을 이용하여 물체의 모양을 기술하는 계수들을 추출하고, 계수간의 유클리디안 거리를 계산한다. 카메라의 포즈(회전 및 변이)를 계산하기 위해서 사각 프레임 상의 네 점의 3차원 실제 좌표와 2차원 영상좌표를 이용한다. 본 논문에서는 ARToolkit 라이브러리의 함수를 사용한다. 라이브 마커로부터 획득된 카메라의 포즈를 이용하여 라이브 마커로부터 인식된 물체와 관련된 그래픽을 라이브 마커 위에 중첩시켜 준다. 이 과정 역시 ARToolkit 라이브러리의 함수를 그대로 사용한다. 그래픽이 중첩된 라이브 마커는 다른 물체가 인식되기 전까지는 라이브 마커에 머물러 있기 때문에, 사용자는 라이브 마커를 자유롭게 움직여 봄으로써 중첩된 그래픽 영상과 쉽게 인터랙션할 수 있다.

본 논문에서는 실험의 편의를 위해 인식의 대상이 되는 객체는 흰색 바탕에 있다고 가정한다. 간단한 영상처리를 통해 배경과 물체를 명확하게 구분할 수 있으며, 불필요한 배경으로 인해 발생할 수 있는 인식오류를 줄일 수 있다.

가. 영상으로부터 라이브 마커 획득

라이브 마커의 획득방법은 그림 1과 같다. 물체를 라이브 마커의 프레임 안에 위치시키고 카메라로 프레임을 촬영하면 쉽게 마커를 얻을 수 있다. 객체 추출 및 인식 기술이 뒷받침 된다면, 이론적으로 현실세계에 존재하는 모든 물체가 마커가 될 수 있다.

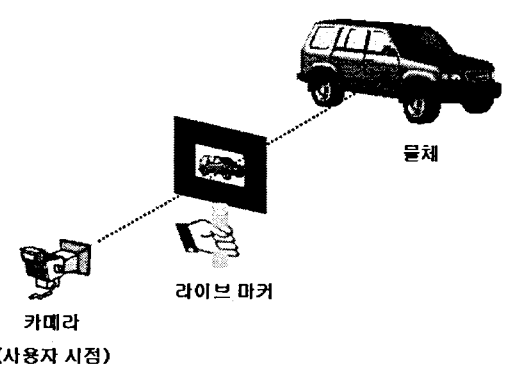


그림 1. 라이브마커 획득방법.

나. 라이브 마커 인식

라이브 마커의 인식을 위해서 ART 알고리즘을 사용한다. ART 알고리즘은 회전, 스케일, 원근 변환 등에 불변한 특성을 가지므로 객체 인식에 좋은 성능을 가진다[3]. 영상에 ART 알고리즘을 적용하기 전에 우선 획득된 영상을 이진화해야 한다. 본 논문에서는 실험의 편의를 위해 모든 객체는 흰색 바탕 위에 놓여 있기 때문에, 간단한 저역통과 필터 통과 후 문턱값 처리를 통해 쉽게 객체 영역은 1, 배경 영역은 0인 이진 영상을 획득할 수 있다.

ART는 객체의 모양을 기술하기 위해 다양한 주파수 성분을 담고 있는 기저 영상들의 합으로 이진 영상을 표현한다. 여기서, 각 계수들은 다음과 같은 식으로부터 얻을 수 있다.

$$F_{nm} = \langle V_{nm}(\rho,\theta), f(\rho,\theta) \rangle$$

$$= \int_0^{2\pi} \int_0^1 V_{nm}^*(\rho,\theta), f(\rho,\theta) \rho d\rho d\theta$$

여기서, F_{nm}은 n, m 차 ART 계수이고, $f(\rho,\theta)$는 극좌표계에서의 영상 함수다. 그리고 , $V_{nm}(\rho,\theta)$는 각도(angular) 와 방사(radial) 방향과 관련된 기저함수로서, 다음과 같이 분리될 수 있다.

$$V_{nm}(\rho,\theta) = A_m(\theta) R_n(\rho).$$

각도 기저함수는 회전 불변값을 얻기 위해 다음과 같은 지수함수가 사용된다.

$$A_m(\theta) = \frac{1}{2\pi} exp(jm\theta).$$

방사 기저함수는 다음과 같이 두개의 다른 변환으로 정의된다.

ART-C (코사인 방사 기저함수):

$$R_n^C(\rho) = \begin{cases} 1 & n = 0 \\ 2cos(\pi n\rho) & n \neq 0 \end{cases}$$

ART-S (사인 방사 기저함수):

$$R_n^s(\rho) = \begin{cases} 1 & n = 0 \\ 2sin(\pi n\rho) & n \neq 0 \end{cases}$$

이 식으로부터 얻어진 계수들을 차수별로 각각 비교하여 그 차의 합계를 통해 물체들의 모양간의 유사성을 판별한다.

$$\text{Dissimilarity} = \sum_i \mid M_d[i] - M_q[i] \mid$$

본 논문에서는 실험의 편의를 위해 비교적 간단한 객체만을 이용하기 때문에, 저주파 성분과 관련된 적은 수의 계수들만을 이용하더라도 높은 인식률을 가질 수 있다. 복잡한 물체의 인식은 고주파 성분들과 관련된 계수들을 포함하여 많은 수의 계수들을 비교함으로써 해결 할 수 있다. 단, 잡음에 의한 영향으로 계수의 개수를 늘린다고 인식률이 좋아지는 것은 아니므로, 잡음에 강건한 유사성 판별 함수를 사용하거나[4], 보다 정교한 인식 기술을 도입해야 한다.

3. 실험 방법 및 결과

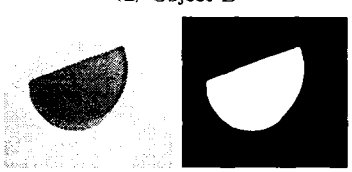
실험에 앞서 라이브 마커의 인식을 위해 사용한 ART의 간단한 성능테스트를 했다. 그림 2에서와 같이 총 네 종류(삼각형 모양, 반원 모양, 정사각형 모양, 직사각형 모양) 의 블록을 사용하여 상호 간의 유사성에 대한 판별 결과를 표 1에 나타냈다. 실험의 편의를 위해 비교적 간단한 객체만을 이용했기 때문에, 저주파 성분과 관련된 적은 수의 계수들만을 이용하더라도 높은 인식률을 가질 수 있었다.

(a) Object *A*

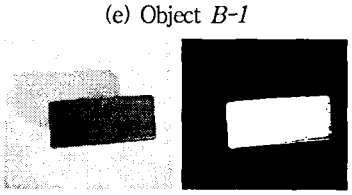
(b) Object *A-1*

(c) Object *A-2*

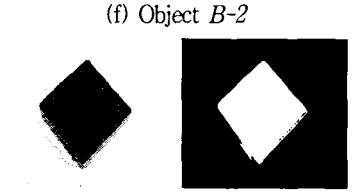
(d) Object *B*



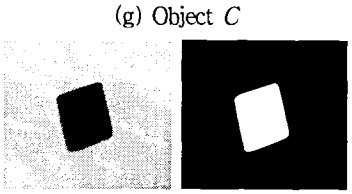
(e) Object *B-1*



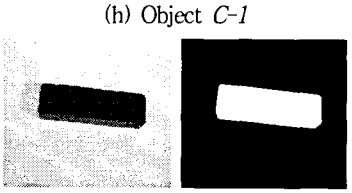
(f) Object *B-2*



(g) Object *C*



(h) Object *C-1*



(i) Object *D*

그림 2. 실험에 사용된 블록 영상 및 이진화 영상. 각각을 *A*, *B*, *C*, *D*로 구분. (b), (c), (e), (f), (h)는 같은 모양의 블록을 다른 방향에서 바라본 영상을 나타냄.

표 1. ART 계수를 이용한 블록 간의 유사성 비교

	A	*A-1*	*A-2*	*B*	*B-1*	*B-2*	*C*	*C-1*	*D*
A		**0.81**	1.94	1.03	1.22	2.26	1.83	1.99	2.85
A-1	**0.81**		1.96	1.31	1.56	2.26	2	2.07	2.8
A-2	1.94	1.96		1.4	1.37	0.5	1.52	1.66	1.2
B	1.03	1.31	1.4		**0.66**	1.66	1.78	2.05	2.25
B-1	1.22	1.56	1.37	**0.66**		1.76	1.67	1.93	2.4
B-2	2.26	2.26	0.5	1.66	1.76		1.67	1.69	0.74
C	1.83	2	1.52	1.78	1.67	1.67		**0.4**	2.14
C-1	1.99	2.07	1.66	2.05	1.93	1.69	**0.4**		2
D	2.85	2.8	1.2	2.25	2.4	0.74	2.14	2	

표의 대각 방향은 같은 물체에 대해 정확히 인식을 하게 되므로 빈칸은 0(일치)을 나타내며, 다른 물체 간의 비교에서 숫자가 작을수록 그 유사도는 높다는 것을 나타낸다. 표에서 보는 것처럼 대부분의 경우, 다른 방향에서 관찰하더라도 같은 모양을 가지는 블록 간의 유사도가 큰 것으로 나타났다. 그러나, 일부 서로 다른 모양을 가지는 블록(예, *A-2* 와 *B-2*) 간의 유사도 큰 경우가 있는데, 이는 보는 각도에 따라 다른 모양을 가졌던 블록의 이진영상의 크기와 모양이 비슷해 질 수 있기 때문에 발생하는 문제다. 이러한 문제는 모양 정보만을 이용하는 ART 기술자의 경우 해결할 수 없으므로, 컬러나 텍스쳐와 같은 부가 정보를 이용하거나 신경망을 이용한 학습 기법 등이 도입되어야 한다. 이와 관련된 내용은 본 논문의 범위를 벗어나므로 생략한다.

앞에서 서술한 것처럼 사용자가 관심있는 모든 물체는 라이브 마커를 이용한 증강현실의 대상이 될 수 있다. 그러나, 본 논문에서는 실험의 편의를 위해 큐브 모양을 가지는 블록을 이용하여 간단한 실험 결과를 보여준다. 주어진 실험 결과는 라이브 마커의 개념 및 효용성을 이해하는 데 충분하다고 생각된다. 그림 4와 같이 큐브 모양의 블록을 라이브 마커의 프레임을 통하여 촬영하여 라이브 마커 영상을 획득한 후, 잡음 제거를 위해 저역 통과 필터를 적용하고, 문턱값 처리를 이용하여 이진화 하였다(그림 3 참조). ART 알고리즘을 이용하여 이진 영상의 ART 계수를 추출해서 데이터베이스에 저장된 ART 계수와 비교한다. 그림 4에서 보는 것처럼, 이진 영상의 물체와 관련된 그래픽을 사각 프레임의 네 점으로부터 계산된 카메라 포즈를 이용하여 원래 영상에 중첩시켜 준다.

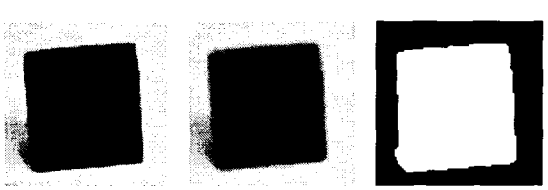
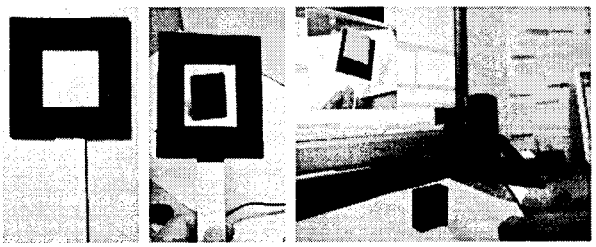


그림 3. 라이브 마커 영상의 이진화. 저역통과 필터와 문턱값 처리 수행.

(a) ARToolkit 라이브러리의 그래픽 마커를 이용한 증강 현실



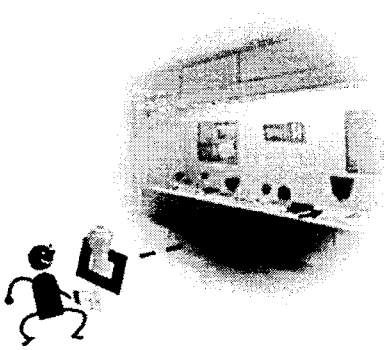
(b) 라이브 마커를 이용한 증강 현실

그림. 4 그래픽 마커와 라이브 마커를 이용한 증강 현실. 라이브 마커를 사용하는 것이 보다 직관적임을 알 수 있다.

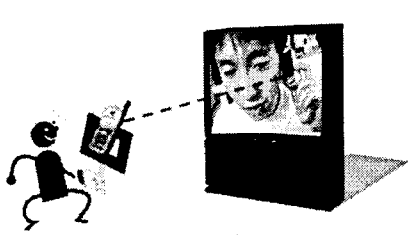
그림 4의 (a)와 (b)의 가운데 영상들은 똑같이 큐브를 나타내는 마커지만 (a)와 같이 큐브와 상관없는 문자열인 'Hiro'를 사용하는 것보다 (b)와 같이 라이브 마커를 사용하여 큐브 그래픽을 중첩시켜 주는 것이 사용자에게 보다 직관적이고 편리한 인터페이스를 제공해 줄 수 있다.

4. 라이브 마커의 활용 예

라이브 마커는 다양한 분야에서 다양한 방법으로 활용될 수 있다. 그 예로 그림 5에서 보는 것처럼, 박물관에서 전시물을 관람하거나 TV를 시청하는 경우, 라이브 마커 위에 사용자가 관심을 가지는 물체와 관련된 3차원 그래픽 모델을 보여줌으로써, 사용자는 전시물을 좀 더 가까이에서 자세히 볼 수 있을 뿐 아니라 자유롭게 3차원 모델과 인터랙션 할 수 있다. 이 모든 것이 사용자의 손 위에서 이루어지게 되는데, 이는 tangible 마커로서의 라이브 마커의 효용성을 보여준다.



(a) 박물관 및 전시회



(b) TV 영상

그림. 5. 라이브 마커의 활용 예.

5. 결론

본 논문에서는 ARToolkit 라이브러리에 라이브 마커의 개념을 도입하여 실제 물체를 마커로서 활용하는 새로운 증강 현실 인터페이스를 제안하고, 그 유효성을 검증하였다. 라이브 마커는 전통적으로 사용된 그래픽 마커의 개념과는 다른, 직관적인 인터페이스를 제공한다는 관점에서 접근하였다. 기존의 마커 시스템을 이용한 많은 유용한 증강 현실 시스템과 더불어 이 논문에서 제안한 라이브 마커를 이용하면 다양한 증강 현실 시스템의 구현이 가능하리라 전망되며, 앞으로 라이브 마커 인식에 대한 연구와 활용 분야에 대한 지속적인 연구가 필요하다.

6. 감사의 글

본 연구는 보건복지부 보건의료기술진흥사업의 지원에 의하여 이루어진 것임. (02-PJ3-PG6-EV04-0003)

7. 참고 문헌

[1] R. Azuma, "A survey of augmented reality," PRESENCE: Teleoperators and Virtual Environments, vol.6, no.4, pp.355-385, Aug. 1997.

[2] H. Kato and M. Billinghurst, "Marker tracking and hmd calibration for a video-based augmented reality conferencing system," Proc. of IWAR, Oct. 1999.

[3] 김용성, 내용기반 이미지 검색을 위한 형태 기술자, 박사학위 논문, 한양대학교 대학원, 2000.

[4] D.A. Forsyth and J. Ponce, Computer Vision: A Modern Approach, Prentice Hall, 2003.