

# 비교정 디지털 동영상의 향상된 AR 비디오 시스템 설계

김혜경<sup>o</sup>, 오해석

송실대학교 컴퓨터학과 멀티미디어연구실

hkkim@multi.soongsil.ac.kr, oh@computing.soongsil.ac.kr

## Design of Improved AR Video System Using Uncalibrated Video Sequences

Hye-Kyung Kim<sup>o</sup>, Hae-Seok Oh

Multimedia Lab., Dept. of Computer Science, Soongsil University

### 요약

본 논문에서는 향상된 AR 비디오 시스템을 위하여 장면내의 객체에 대한 그림자 모양 안쪽의 광휘 분포로부터 실제 장면의 조명 분포를 추정하기 위한 새로운 메소드를 설명한다. 장면의 조명 분포는 확장된 빛의 근원에 대한 이산 샘플링에 의해 근사화되었고, 장면에 대한 조명 분포는 그림자 모양 안쪽의 광휘 분포로부터 장면내의 다른 객체 위에 알려진 모양의 객체에 의해 추정되었다. 그림자 안쪽 표면에 대한 반사 특성 대신에, 장면에 대한 조명 분포와 표면에 대한 반사 특성을 동시에 추정하는 반복적인 최적화 구조를 기반으로 한다. 또한, 장면에 대한 조명 분포의 적응적 샘플링 방법을 소개한다. 전체적인 조명 분포에 대한 균일한 분리를 사용하기보다는 이전의 반복 지점에서 조명 분포에 대한 샘플링 방향을 적응적으로 증가한다. 적응적인 샘플링 구조를 사용한 추정방법은 보다 소수의 샘플링을 사용하여 전체적인 조명을 보다 효율적으로 추정할 수 있었다. 제안된 메소드는 복잡한 조명 환경아래라도 조명 분포를 추정하는 데 매우 효율적이다.

### 1. 서론

AR의 목적은 가상정보로 실제 환경을 부각시키기 위하여 실제와 가상을 병합하는 것이다. 초기에는 간단한 텍스트 주석 달기로부터 증가한 범위는 실제 환경으로 삽입된 실제 오브젝트에 대한 가상의 모조품을 실제 오브젝트에 부착하는 연구까지 진행되었다. 후자의 경우에서 궁극적인 목적은 현실과 가상 객체들간의 식별이 거의 불가능하게 만드는 것이다. 그러므로, 실제 환경에 대한 가상 객체의 견고한 registration, 현실과 가상 객체들의 공통의 occlusion 문제 그리고 이런 조명 모델을 가진 가상 객체를 만들기 위하여 실제 환경에 대한 조명 분포의 추출 문제가 반드시 극복되어야 한다.

본 논문에서는 Augmented Reality Video System을 구현하기 위해 제한했던 방법을 설명하고자 한다. 일부 움직이는 부분을 구성하기 위해 실제 환경으로부터 전체적으로 교정되지 않은 비디오 장면 안으로 가상 객체를 합성하기 위해 사용한 방법과 조명 모델을 적용한 가상의 객체 영상을 실제 비디오 영상에 삽입하기 위해 적용한 방법을 설명한다.

### 2. AR 비디오 시스템의 움직임과 구조회복 방법

AR 비디오 시스템의 구조와 움직임 회복 알고리즘에 대한 적합성을 설명한다. 주요한 목적이 비디오 장면 내내 카메라의 움직임 회복이라 할지라도, 시스템은 또한 실제 환경의 가공하지 않은 3D 구조를 회복한다. 이것은 occlusion 해결과 실제 환경에 대한 조명 분포를 추출하는 문제를 다루는데 매우 유용하다. 또한, 회복된 움직임 파라미터의 사용과 실제 환경 내부에 등록된 가상 객체에 대한 3D 구조를 설명한다.

### 2.1 움직임과 구조회복 알고리즘

#### 1) 영상 특성 선택과 매칭

컴퓨터 비전내의 움직임 회복은 거의 항상 영상들의 처음부터 끝까지 특징들에 대한 트래킹을 기반으로 하고, 실제 환경에 대한 카메라의 움직임 파라미터를 결정하도록 사용한다. 특징은 points, lines, curves 이나 영역과 같은 모든 특징들이 들어온다. 우리가 사용한 특징들은 입력 비디오 장면에 대한 각각의 영상을 적용한 Harris Corner Detector 알고리즘의 결과이다. 영상내의 점들과 코너들로 이루어진 결과는 영상의 세기를 두 직교 방향에서 의미적으로 변경하여 위치를 결정한다. 가능한 매칭 코너들의 초기 집합은 다른 영상들에서 코너를 찾고 있는 각 코너 둘레의 작은 검색 영역을 사용하여 구성하고 정밀한 조사를 통하여 코너를 가진 매우 큰 표준화된 강도의 교차 상관관계를 가진다. 대응하거나 일치하는 코너들은 각각의 다른 에피폴라 라인에 놓여있는 에피폴라 기하를 통하여 강제적이다. 이런 제약은 (1)로부터 코너가 매치 되기를 바라는 두 영상들간의 선형방정식으로 표현될 수 있다.

$$x_1^T F_{12} x_2 = 0 \quad (1)$$

$x_1 = (u_1, v_1, 1)^T$  와  $x_2 = (u_2, v_2, 1)^T$  는 첫 번째와 두 번째 영상에서 매칭 코너들에 대한 동종의 영상 좌표를 나타낸다.  $F_{12}$  는  $3 \times 3$ 의 단일한 행렬이고, 두 영상들간의 에피폴라 기하를 나타낸다.

실제 환경안에 움직임 부분이 비디오 장면안에서 작은 동안 outliers 와 같은 움직임 부분에 속하는 코너를 취급하여, 행렬  $F_{12}$  와 매칭 코너들에 대한 결정에서 주로 버려질 것이다.

2) 움직임과 구조회복 초기화

영상위의 3D 구조점과 프로젝션 사이의 관계는 유사좌표의 선형적인 관계에 의해 (2)와 같이 표현된다.

$$m_k \sim P_k M \quad (2)$$

여기서  $M=(X, Y, Z, 1)$  와  $m_k = (x_k, y_k, 1)^T$  는 각각 3D 구조점의 유사 좌표와 영상  $k$  에 대한 프로젝션이다.  $P_k$  는  $3 \times 4$  행렬이고, 이것은 프로젝션 작업을 설명하고 ‘ $\sim$ ’ 는 이런 균등이 규모인자에 이르기까지 효과가 있음을 나타낸다.

장면에 대한 두 개의 초기 영상들은 관련 프레임을 결정하고, 세계 프레임은 첫 번째 영상의 카메라로 정렬되었다. 두 번째 카메라가 선택되었고 그 결과 epipolar geometry 는 검색된  $F_{12}$  에 대응한다.

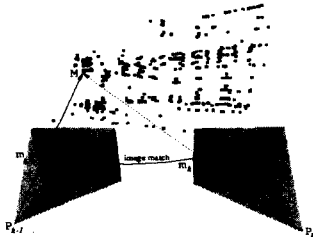
$$P_1 = [ \quad \quad \quad I_{3 \times 3} \quad | \quad 0_3 ] \quad (3)$$

$$P_2 = [ [ e_{12} ]_x \quad F_{12} + e_{12} \pi^T \quad | \quad \sigma e_{12} ]$$

여기서  $[ e_{12} ]_x$  는  $e_{12}$  에 대한 벡터의 내적을 나타낸다. 등식 (3)는 에피폴라 기하(epipolar geometry)에 의해 완벽하게 결정되는 것이 아니고(i.e.  $F_{12}$  and  $e_{12}$ ), 4 이상의 자유도를 갖는 것이다(i.e.  $\pi$  and  $\sigma$  ).  $\pi$  는 참조 평면의 위치를 결정하고,  $\sigma$  는 복원에 대한 전역적인 단계를 결정한다.

3) 움직임과 구조회복 갱신

처음으로 임의의 대응하는 행렬  $P$  가 계산되어진 영상을 가져와 그 영상의 코너들과 복원된 3D 구조점들 사이에서 2D/3D의 매치를 검색한다. 다음으로 가져온 영상은 행렬  $F$  와 두영상들간의 코너 매치를 계산한다. 다음 영상에 대한  $P$  를 결정하도록 이전의 복원된 3D 구조점들을 사용하여 새로운 행렬  $P$  로 초기화한다.



[그림 1] 코너 매치

4) Metric 구조와 움직임

비교정 카메라에 대한 내부 파라미터 위의 일부 제약은 이용 가능하다. 카메라 셋팅이 변경되지 않으면 내부 파라미터들은 장면 위에 고정되고, 영상의 기울기도 없고, 주요한 포인트는 영상 중심에 인접한 상태로 항상 비율은 항상 고정되어 있다. Metric 캘리브레이션에 대한  $P$  행렬들의 인수분해는 이런 제약을 만족하는 내부 파라미터를 야기한다. 그러므로 Self-calibration 이 가능한  $P$  행렬을 찾는데, 절대적인 원추곡선의 개념을 기반으로 한다.

무한면과 절대적인 원추곡선은 각각 affine 과 metric 특성의 공간을 부호화한다. 이것은  $\Pi_\infty$  의 위치를

투영 구조에서 알고있는 상태를 의미하고, affine 불변식은 평가의 기준이 될 수 있다. 절대적인 원추곡선이 Euclidean 변환에서 불변일 때 그 영상은 단지 내부 카메라 파라미터(초점 길이, ... )에 의존하고 외부 카메라 파라미터들 (카메라 pose) 에는 의존하지 않는다. 아래의 (4)번 등식은 절대적인 원추 곡선의 두 영상에 대해 적용한다.

$$\omega_k^* \propto K_k K_k^T \quad (4)$$

여기서  $K_k$  는 영상  $k$ 에 대한 카메라의 내부를 구성하는 삼각 기동 행렬의 상부이다. 등식 (4) 는 내부 카메라 파라미터 위의 제약이 절대적인 원추 곡선에 대한 두 영상의 제약으로 손쉽게 변환되었음을 보인다.

$$\omega_k^* \propto P_k \Omega^* P_k^T \quad (5)$$

등식(5)에서,  $\Omega^*$  는 두 부분으로 된 절대적인 2차 곡면으로 절대적인 원추 곡선(absolute conic)과 무한면을 부호화한다.  $\omega_k^*$  위의 제약은 그 결과 이 등식을 통하여 back-projected 될 수 있다. 결과는 절대적인 원추 곡선(그리고 무한 평면)에 대한 위치에서 제약의 집합이다. 제안된 시스템은 처음으로 선형적인 매소드를 개략적인 캘리브레이션을 얻기위해 사용한다. 이러한 캘리브레이션은 두 번째 단계에서 비 선형 최적화를 통하여 더욱 정제된다.

3. 향상된 AR 비디오 시스템

향상된 AR 비디오 시스템을 위한 알고리즘의 순서를 개략화하면 다음과 같다.

- 1단계 : 초기화 단계로, 비디오 영상으로부터 가져온 두 영상의 움직임과 구조 양쪽의 어느쪽에서든 재구성될 수 있는 projective 프레임을 초기화한다.
- 2단계 : 마지막 영상을 가져와 처리하고 registering 이 필요한 더 나은 비디오 장면으로 또 다른 영상을 가져온다.
- 3단계 : 이미 처리된 영상들로부터 이들 영상들과 2D-3D 코너 매치를 registration 하는 동안 구축하여 사용한다.
- 4단계 : 2D-3D 매치를 사용하여 행렬  $P$ 를 결정한다.
- 5단계 : 행렬  $P$ 를 사용한 새로운 3D 구조점들을 복원한다.
- 6단계 : 비디오 장면의 마지막이던 단계 2로 돌아간다.
- 7단계 : 각 남아있는 프레임에 대하여 2D-3D 매치를 획득하여 사용한다.
- 8단계 : 단계 4와 유사하게, 이런 프레임에 대한 행렬  $P$ 를 계산한다.
- 9단계 : 처음의 가상 객체들은 이전 단계에서 얻어진 있는 그대로의 3D 복원을 사용하여 실제 환경안에 개략적으로 배치되었다.
- 10단계 : 가상 객체의 위치에 대한 미조정은 일부 선택된 영상위에서 결과를 보여주고 만족스러운 배치가 얻어질 때까지 가상객체를 조정한다.

4. 조명 모델

4.1 제안된 매소드의 기본적인 단계

- 1단계 : 그림자 표면의 반사 파라미터를 초기화한다. 전형적으로, 그림자 표면을 Lambertian 으로 가정하고, 확산한(흩어진) 파라미터  $K_d$  는 그림자 표면위에서 가장 밝은 포인트의 픽셀 값으로 된 집합이다.

반사하는 파라미터들은 제로로 ( $K_s=0, \sigma=0$ ) 구성 된 집합이다.

2단계 : 가공 방향의 빛의 근원에 대한 광휘 값  $L(\theta_i, \phi_i)$  을 추정하여 실제 장면에 대한 조명 분포 모델로 이용한다. 그림자 영상 내의 그림자 안쪽에 있는 반사 파라미터 ( $K_d, K_s, \sigma$ ) 와 이미지 밝기를 사용하여 광휘 분포  $L(\theta_i, \phi_i)$ 을 계산한다.

3단계 : 그림자 표면의 반사 파라미터 ( $K_d, K_s, \sigma$ ) 를 최적화 기법을 사용하여 장면  $L(\theta_i, \phi_i)$ 에 대하여 얻어진 광휘 분포로부터 추정한다.

4단계 : 장면  $L(\theta_i, \phi_i)$ 의 광휘 분포를 획득된 반사 파라미터 ( $K_d, K_s, \sigma$ )로부터 추정한다.

5단계 : 추정된 값들인  $L(\theta_i, \phi_i), K_d, K_s, \sigma$  내 에서 어떤 주요한 변화도 없다면 다음 단계로 진행한다. 그렇지 않다면, 3단계로 돌아가라. 장면의 광휘 분포와 그림자 표면의 반사 파라미터들 두 가지를 반복적으로 측정하여, 장면의 조명 광휘 분포에 대한 샘플링 방향의 주어진 집합에 대하여 최고의 추정값을 얻는다.

6단계 : 얻어진 조명 광휘 분포가 충분한 정확성을 가지고 실제 광휘 분포로 근사화 되었다면 추정 과정을 종료한다. 그렇지 않다면, 다음 단계로 진행해 나간다.

7단계 : 얻어진 조명 광휘 분포  $L(\theta_i, \phi_i)$ 을 기초로 조명 분포의 샘플링 방향을 적응적으로 증가한다. 끝 이어 단계 2로 돌아가라.

#### 4.2 광휘 분포의 추정

주어진 반사 파라미터 ( $K_d, K_s, \sigma$ )의 집합에 대한 장면  $L(\theta_i, \phi_i)$ 의 광휘 분포를 추정하는 방법은 다음 식을 이용한다.

$$a_1 L_1 + a_2 L_2 + a_3 L_3 + \dots + a_{1n} L_n = P$$

$L_i (i=1, 2, \dots, n)$  는 장면에 대한 광휘 분포의 n개 샘플링 방향에 의해 묘사된 n 개의 알려지지 않은 조명 광휘이다. 계수  $a_i (i=1, 2, \dots, n)$ 는 ( $K_d \cos \theta_i + K_s \frac{1}{\cos \theta_r} e^{-\frac{\gamma(\theta_i, \phi_i)^2}{2\sigma^2}}$ )  $S(\theta_i, \phi_i) (i=1, 2, \dots, n)$  을 표현한다. 즉, 이미지 픽셀, 조명 방향, 그리고 occluding 객체의 모양에 대응하고 있는 표면 포인트의 3D 기하학을 기반으로  $\theta_i, \theta_r, \gamma(\theta_i, \phi_i)$  그리고,  $S(\theta_i, \phi_i)$ 를 계산한다. 위에서 사용한 P는 이미지 픽셀  $P(\theta_e, \phi_e)$  의 값이다.

#### 4.3 반사 파라미터 추정

그림자 표면에 대한 반사 파라미터를 장면  $L_i$  에 대하여 추정한 광휘 분포 ( $K_d, K_s, \sigma$ )를 사용하여 추정하는 방법은 그림자 표면의 반사 파라미터를 그림자 이미지에서 관찰된 픽셀 강도와 대응하는 표면 포인트의 픽셀 값들 간의 차이를 제공한 것의 합계를 최소화함에 의해 추정한다. 즉,

$$f = \sum_{j=0}^n (P_j - P_j)^2 \quad (6)$$

#### 4.4 광휘 분포에 대한 적응적 샘플링

광휘 값  $L_1, L_2, L_3$  를 가진 세 개의 샘플링 방향이 [그림2]에서 설명된 것처럼 삼각형  $M_1$ 의 형태로 놓여있다고 가정하면, 새로운 샘플링 방향이  $L_1$ 과  $L_2$  사이에 추가되는 것이 필요한지 필요하지 않음을 결정하기 위하여 사용하는 비용 함수는 다음 등식 (7)과 같이 구한다.

$$U(L_1, L_2) = \text{diff}(L_1, L_2) + \alpha \min(L_1, L_2) \text{angle}(L_1, L_2) \quad (7)$$

$\text{diff}(L_1, L_2)$ 는  $L_1$ 과  $L_2$  간의 광휘의 차이이고,  $\min(L_1, L_2)$ 는  $L_1$ 과  $L_2$  에 대하여 더 작은 광휘를,  $\text{angle}(L_1, L_2)$ 는  $L_1$ 과  $L_2$  방향들간의 각도이고,  $\alpha$  는 두 가지 요소들의 상대적인 무게를 결정하기 위하여 손으로 만든 명확한 파라미터이다.



[그림 2] 적응적 샘플링

#### 5. 실험 결과

장면에 대한 광휘 분포와 표면의 반사 특성을 이용한 가상객체의 명암과 부드러운 그림자 모양을 가진 영상을 생성한 실험 결과는 다음과 같다.



#### 6. 결론 및 향후 연구 과제

본 논문에서는 올바른 음영을 가진 가상의 객체를 실제 장면의 비디오 영상위에 추가하기 위한 새로운 방법과 함께 장면에 대한 조명분포와 표면의 반사 특성을 동시에 추정하는 방법을 제안하였다. 또한 전체적인 비디오 영상이 교정되지 않은 영상을 대상으로 가상의 객체를 추가하기 위한 Self-calibration 방법 또한 제안하였다. 향후연구과제는 실제 영상을 제거한 위치에 가상의 객체 영상을 합성하는 방법과 동일한 가상의 객체 영상을 다양한 형태로 합성하는 방법을 연구할 것이다.

#### 7. 참고 문헌

[1]A. Fournier, A.Gunawan, "Common Illumination between Real and Computer Generated Scene," Proc. Graphics Interface '93, pp.254-262, 1993.  
[2]Y. Sato, "Object shape and reflectance modeling from observation," Proc. of SIGGRAPH 96, pp. 429-438, August, 1996.