
장면 전환에서의 물체 추적을 통한 모델기반추적 방법 연구

The Model based Tracking using the Object Tracking method in the Sequence Scene

김세훈, Sehoon Kim*, 황중원, Jungwon Hwang**, 김기상 Ki-sang KIM**, 최형일, Hyungil Choi***

숭실대학교 미디어학과

요약 증강현실은 가상현실의 한 분야로 실제 환경에 가상 사물을 합성하여 원래의 환경에 존재하는 사물처럼 보이도록 하는 컴퓨터 그래픽 기법이다. 증강현실은 가상의 공간과 사물만을 대상으로 하는 기존의 가상현실과 달리 현실세계 기반위에 가상의 사물을 합성하여 현실세계 만으로는 얻기 어려운 부가적인 정보를 보강해 제공할 수 있는 특징을 가지고 있다. 실제 세계 기반위에 가상의 사물의 합성을 구현하는데 있어 중요하게 여겨지는 기반 기술인 레지스트레이션 방법이 있다. 레지스트레이션 방법은 실사영상과 3차원 그래픽 객체의 위치와 방향을 결정하는 방법으로서, 모델기반추적과 Move-Matching 방법이 사용되고 있다. 본 논문에서는 모델기반추적방법에 대하여 물체 추적을 통한 물체의 정보와 색상 분포를 이용하여 각 장면간의 물체 추적을 통하여 전환되는 장면에서의 모델을 통해 상대적 좌표계를 생성하는 방법에 대하여 연구하였다.

Abstract ~ Augmented Reality is a growing area in virtual reality research. The world environment around us provides a wealth of information that is difficult to duplicate in a computer. This evidenced by the worlds used in virtual environments, An augmented reality system generates a composite view for the user. It is a combination of the real scene viewed by the user and a virtual scene generated by the computer that augments the scene with addition information. The registration method represent to the user enhances that person's performance in and perception of the world. It decide the direction and location between real world and 3D graphic objects. The registration method devide two method, Model based tracking and Move-Matching. This paper researched at to generate a commerce correlation using a tracking object method, using at a color distribution and information, in the sequence scene.

핵심어: 증강현실, 모델기반추적, 물체추적, Registration, Calibration, Tracking, Mean Shift

1. 서론

증강현실은 가상현실의 한 분야로 실제 환경에 가상 사물을 합성하여 원래의 환경에 존재하는 사물처럼 보이도록 하는 컴퓨터 그래픽 기법이다. 증강현실은 가상의 공간과 사물만을 대상으로 하는 기존의 가상현실과 달리 현실세계 기반위에 가상의 사물을 합성하여 현실세계 만으로는 얻기 어려운 부가적인 정보를 보강해 제공할 수 있는 특징을 가지고 있다. R. T. Azuma[1]에 의하면 현실과 가상 세계가 결합되고 실시간처리로 상호대화형이며, 3차원 정합이 이루어져야만 증강현실에 범주로 포함된다고 보여지며, 사용자의 시각 및 현실과의 상호관계에 도움을 주는데 목적이 있다고 증강현실을 정의하고 있다.

증강 현실에 있어서 가장 기본적이고도 어려운 것은 실제 환경과 가상의 물체의 정확한 정합에 관한 문제이다. 정합이란 실제 세계 기반위에 가상의 사물의 합성을 구현하는데 있어 중요하게 여겨지는 기반 기술인 방법이다. 정합 방법은 실사 영상과 3차원 그래픽 객체의 위치와 방향을 결정하는 방법으로서, 가상현실에서도 영상과 운동 감각의 불일치 등으로 발생하는 정합 문제는 존재하지만 사용자가 보는 영상이 모두 가상의 것이므로 본질적인 문제는 아니다. 그러나 증강현실에서는 실제 환경과 가상의 물체가 시각적으로 결합되어, 가상 물체가 실제 환경에서의 물체처럼 움직이지 않는 것에 따른 오차 및 실제 환경과 일치 하지 않아 발생하는 어긋남 등이 발생하기 때문에 정합이 제대로 이루어 지지 않을 경우의 시각적인 불일치는 심각하게 나타난다.

이러한 시각적인 불일치를 제거하기 위해 실제 환경의 카메라의 특성을 파악하여 가상물체와 실제 환경의 3차원 좌표를 정확히 일치 시켜야 한다. 특히 실시간으로 사용자와 가상 물체 간에 상호 작용을 통해 사용자로 하여금 더욱 현실감을 느끼게 할 수 있게 되고, 실시간으로 입력되는 비 교정된 영상에서 가상 물체를 합성시키기 위해서는 실제 카메라의 파라미터를 알아내는 카메라 교정 작업이 필수적이다.

증강현실에서의 정합 문제를 해결하기 위해 모델기반추적과 Move-Matching방법이 사용되고 있다. 본 논문에서는 모델기반추적방법에 대하여 물체 추적을 통한 물체의 정보와 색상 분포를 이용하여 각 장면간의 물체 추적을 통하여 전환되는 장면에서의 모델을 통해 상대적 좌표계를 생성하는 방법에 대하여 연구하였다.

본 논문의 구성은 제 2절에서는 정합 문제 해결에 대한 기존 방안을 기술하고, 제 3절에서는 제안하는 방법에 대하여 기술하고 있으며, 제 4절에서는 실험 및 결과를 보이고, 마지막으로 제 5절에서는 결론 및 향후 연구에 관하여 기술한다.

2. 정합

증강 현실에 있어서 가장 기본적이고도 어려운 것은 실제 환경과 가상의 물체의 정확한 정합에 관한 문제이다. 가상현실에서도 영상과 운동 감각의 불일치 등 정합 문제는 존재하지만 사용자가 보는 영상이 모두 가상의 것이므로 본질적인 문제는 아니다. 그러나 증강현실에서는 실제 환경과 가상의 물체가 시각적으로 결합되어야 하므로 정합이 제대로 이루어 지지 않을 경우 시각적인 불일치는 심각하게 나타난다. 이러한 시각적인 불일치를 제거하기 위해 실제 환경의 카메라의 특성을 파악하여 가상물체와 실제 환경의 3차원 좌표를 정확히 일치 시켜야 한다. 특히 실시간으로 사용자와 가상 물체 간에 상호 작용을 통해 사용자로 하여금 더욱 현실감을 느끼게 할 수 있게 되고, 실시간으로 입력되는 비 교정된 영상에서 가상 물체를 합성시키기 위해서는 실제 카메라의 파라미터를 알아내는 카메라 교정 방법으로서 모델기반추적과 Move-Matching방법이 사용되고 있다.

2.1 Model-based Tracking

대부분의 기술들이 실제계의 한 구조물을 추적하고 실제계의 좌표를 추적하는 방식을 사용하고 있다. 그 예로 수평선[2]이나 카메라의 위치를 이용하여 포즈를 판단하는 기술[3,4], 점을 기반으로 추적하는 방식[5], 평면을 추적하는 방식[6]들이 사용되고 있다.

2.2 Move-matching

일반적으로, 정확한 정합을 위하여 3D 구조물과 카메라 모션을 동시에 계산하는 방식으로 수행하고 있다. 이러한 방식을 Structure-and-motion 기반 추적이라고도 하며 Neumann에 의하여 개발된 기술[6] 등이 있다.

3. 제안하는 방법

본 논문은 기존 증강현실하의 정합 문제에 대하여 효과적으로 적용하기 위한 방법을 제안한다. 증강현실하의 환경에서는 고정된 물체와 이동하는 카메라 간의 관계를 기준으로 전환하는 장면하의 Markerless 추적을 통한 정합[5]을 하는 경우가 대부분이라 판단되어 전환되는 장면 하에서 물체 추적을 수행하였다.

특정 조건을 만족하는 일반적인 물체에 대하여 고유한 색상의 분포를 이용하여, 전환되는 장면하의 물체 추적을 통하여 각 장면에서의 물체의 좌표와 명암분포를 이용하여 정합하는 방법과 전환되는 장면 안에서의 물체를 통하여 그림 1과 같이 카메라의 좌표계 생성을 목적으로 한다.

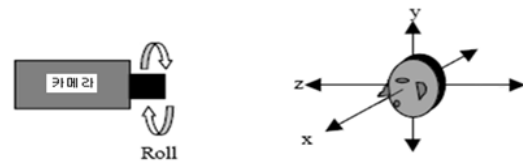


그림 1 좌표계 생성의 예

3.1 물체의 색상 분포 추출

전환되는 장면간의 물체 추적을 위해서는 색상의 분포를 이용하여 물체를 추적하는 방법을 사용 하였다. 효과적인 색상 분포 추출을 위해 HSV 색상계를 이용하여 물체의 HSV 각각의 1차원 히스토그램을 작성하여 색상분포 모델을 만들어 작성하였으며 식(1)과 같이 추적한다.

$$P((H, S, V)) = k * H \tag{1}$$

$$= \frac{H_{pixel}}{Max_{pixel}}$$

3.2 전환되는 장면간의 물체 추적

증강현실하의 환경에서는 고정된 물체와 이동하는 카메라 간의 관계를 기준으로 전환하는 장면하의 Markerless 추적을 통한 정합을 하는 경우가 대부분이라 판단되어 전환되는 장면 하에서 물체 추적을 하기 위하여 본 논문은 CAMSHIFT 알고리즘[7]을 이용, 카메라의 좌표를 검출하

고 카메라 좌표계를 생성하는 방법을 제안한다.

CAMSHIFT 알고리즘은 얼굴인식 분야에서 매우 인지도 있게 사용되고 실시간으로 처리가 가능한 알고리즘으로서 추적하는 물체에 대하여 고유한 색상의 분포를 이용하여, 전환되는 장면하의 물체 추적을 통하여 각 장면에서의 물체의 좌표를 이용하여 Self-calibration 방법을 이용한다.

전환되는 장면 하에 물체 추적을 위해서는 각 장면에 대하여 Mean Shift 알고리즘을 적용하여 영상에 대하여 탐색영역(W)을 정하고 크기(s)와 초기위치(P_k)를 결정하고, 다음과 같이 탐색영역의 mean의 위치와 그라디언트를 식(2)를 통하여 계산한다.

$$\begin{aligned} \overline{P}_k(W) &= \frac{1}{|W|} \sum_{j \in W} P_j \\ \overline{P}_k(W) - P_k &\approx \frac{f'(P_k)}{f''(P_k)} \end{aligned} \quad (2)$$

위치가 계산된 토대로 탐색영역의 중심($\overline{P}_k(W)$)을 계산 위치가 mean의 그라디언트(f'(p))가 수렴이 될 수 있도록 계산을 반복하여 물체를 탐색한다.

mean 탐색 영역의 중심을 계산하기 위하여 0번 장면의 모멘트를 식(3)과 같이 계산하고,

$$M_{00} = \sum_x \sum_y I(x, y) \quad (3)$$

식(4)를 이용하여, 첫 번째 모멘트에서 x와 y를 검출하게 된다.

$$M_{10} = \sum_x \sum_y x I(x, y); M_{01} = \sum_x \sum_y y I(x, y) \quad (4)$$

위의 계산을 이용하여 mean 탐색 영역의 중심을 식(5)와 같이 계산한다.

$$x_c = \frac{M_{10}}{M_{00}}; y_c = \frac{M_{01}}{M_{00}} \quad (5)$$

검출된 영역을 중심으로 다음 장면에 대하여 탐색영역의 위치를 (4)번 수식에서 나온 mean 탐색 영역 중심으로 이동 반복적으로 수행하여 각 장면에 대하여 탐색영역의 중심을 계속적으로 계산한다.

3.3 전환되는 장면간의 카메라 위치 계산

본 논문은 고정된 물체와 이동하는 카메라를 이용한 정합 과정에 대하여 제안하고 있다. 본 절에서 설명할 카메라의 위치 계산을 이용 정합과정에서 효과적인 정합문제 해결방안으로 제안한다.

먼저 카메라의 중심위치를 영상의 중앙의 위치라 가정하고, 식(6)과 같다하면,

$$C_x = I_{cx}; C_y = I_{cy} \quad (6)$$

앞 절에서 계산된 mean 탐색 영역의 중심 이용하여, 두 번째 모멘트를 식(7)과 같이 계산 한다.

$$M_{20} = \sum_x \sum_y x^2 I(x, y); M_{02} = \sum_x \sum_y y^2 I(x, y) \quad (7)$$

계산된 두 번째 모멘트를 이용, 카메라의 방위 정보를 식(8)과 같이 추출한다.

$$C_\theta = \frac{\arctan \left[\frac{2 \left(\frac{M_{11}}{M_{00}} - x_c y_c \right)}{\left(\frac{M_{20}}{M_{00}} - x_c^2 \right) - \left(\frac{M_{02}}{M_{00}} - y_c^2 \right)} \right]}{2} \quad (8)$$

추출된 카메라 방위 정보를 이용, 각 프레임간의 탐색영역의 크기정보를 이용하여 탐색된 물체의 길이와 넓이 정보를 식(9)와 같이 추출한다.

$$a = \frac{M_{20}}{M_{00}} - x_c^2, \quad (9)$$

$$b = 2 \left(\frac{M_{11}}{M_{00}} - x_c y_c \right),$$

$$c = \frac{M_{02}}{M_{00}} - y_c^2$$

$$l = \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}}$$

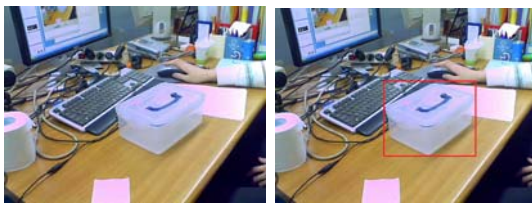
$$w = \sqrt{\frac{(a+c) - \sqrt{b^2 + (a-c)^2}}{2}}$$

이렇게 추출된 물체의 길이와 넓이 정보를 이용하여 각 장면간의 상대적이 깊이 정보를 추출한다.

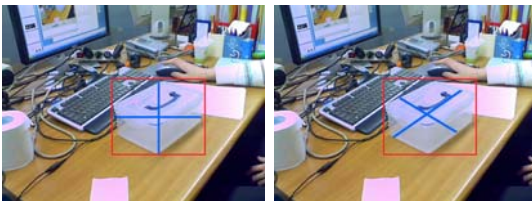
4. 실험 및 결과

본 실험을 위하여 사용된 시스템은 펜티엄 코어2듀오 프로세서와 1GB 램을 장착한 컴퓨터를, 운영체제는 Windows XP Home Edition, 컴파일러는 MS사의 Visual C++ 8.0을 사용하였다. 실험에 사용된 카메라는 로지텍 퀵캠을 이용하여 640x480의 영상을 획득하였다.

획득한 영상에 대하여 조건을 만족하는 물체의 추적(그림 2의 (b))을 통하여 좌표를 생성(그림 2의 (c))하고 카메라의 방위를 계산(그림2의 (d))하는 과정을 거쳐 결과를 산출한다.



(a) 원본영상 (b) 탐색영역생성



(c) 물체의 좌표계산 (d) 카메라의 방위계산

그림 2 전환되는 장면간의 차이를 이용한 위치 계산

일련의 과정을 통해 검출된 좌표를 이용해 가상의 물체를 실사영상에 보여주는 과정은 다음의 그림 3을 통하여 나타내었다.



(a) 원본 영상 (b) 합성된 영상

그림 3 가상의 물체를 실사영상에 합성

5. 결론

본 논문은 탐색영역을 이용하여 처리되는 영상의 범위를

줄이게 되어 기존의 방법을 통한 전체 영상의 처리에서 보다 더 적은 양의 계산을 통하여 빠른 처리가 가능하며, 카메라 1개를 통하여 시간의 흐름에 변화하는 물체의 위치에 대하여 명암분포와 색상분포를 이용하여 상대적 좌표 생성하는 방법을 제안하고 있다.

제안된 방법에 대하여 실험 및 결과를 통해 기존에 검증된 알고리즘을 이용하여, 증강현실하의 상황에서 효과적으로 적용하고 결과를 산출 할 수 있었으며, 단안 카메라 하의 환경에서 상대적인 위치 검출을 통하여 가상의 물체에 효과적으로 적용시킬 수 있음을 볼 수 있다.

향후 연구 방향으로서는 여러 물체를 동시에 추적하여 각 물체마다의 영상에 대한 상대 좌표 생성을 통해 가상의 물체를 실사영상에 합성하는 방법에 대하여 연구가 이루어진다면 보다 효과적인 결과를 검출할 수 있을 것이라 사료된다.

참고문헌

- [1] Ronald T. Azuma "A Survey of Augmented Reality", Teleoperators and Virtual Environments, Vol. 6, No. 4 pp.355-385, 1997
- [2] R. Behringer "Improving the precision of registration for augmented reality in an out door scenario by visual horizon silhouette matching" In First IEEE Workshop on AR(IWAR'98), 1998
- [3] D. Dementhon and L. Davis "Model based object Pose in 25 Lines of Code", IJCV, Vol.15, pp.123-141, 1995
- [4] M.A. Fishler and R.C.Bolles "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, Vol.24, No.6, pp 381-395, 1981
- [5] Gilles Simon, Andrew W.Fitzgibbon, Andrew Zisserman "Markerless Tracking using Planar Structures in the Scene", Proc. Int'l Symposium on AR, 2000
- [6] U. Neumann, S. You, Y.Cho, j.Lee, and J. Park "Natural feature tracking for augmented reality", IEEE Transaction on Multimedia, Vol. 1, No. 1, pp53-64, 1999
- [7] Gary R. Bradski, Microcomputer Research Lab, Santa Clara, CA, Intel Cooperation "Computer Vision Face Tracking For Use in a Perceptual User Interface"