스마트폰 환경하의 실시간 처리를 위한 가변 탐색영역을 이용한 마커 추적 방법

Realtime Processing for Marker Tracking in Smart-Phone Environment Using Deformable Searching Area

김세훈, Sehoon Kim*, 임성준, SungJun Lim**, 이민호 Minho Lee**, 김계영 Gye-yuong KIM**, 최형일, Hyungil Choi***

숭실대학교 미디어학과

요약 모바일상의 혼합현실 기반의 소프트웨어 기술은 현재 국내외적으로 시작 단계에 있다. 하지만 급속한 모바일 인터넷 인프라의 확산, 모바일 기기의 성능 향상 및 개방형 모바일 OS의 등장으로 인해 모바일 혼합현실과 같은 새로운 융합형 IT기술이 사용자의 요구에 부흥하여, 수요가 급신장 되고 있다. 모바일 혼합현실 기반 기술 중 하나인 마커의 검출과 추적은 혼합현실의 가장 기본이며 매우 중요한 과정이다. 본 논문에서는 낮은 CPU를 이용하는 모바일 환경에 최적화된 실시간 마커 추적 방법에 대하여 제안한다. 낮은 CPU의 처리속도를 보완하기 위하여 처리영역을 축소하여 가변 탐색영역내의 공간에서 처리를 수행하고, 좌표변환정보를 이용하여 추적을 수행하는 캘리브레이션 과정을 대신하여 캠 시프트 알고리즘을 이용하여 효과적인 추적 방법을 제안한다.

Abstract This paper introduces a Mixed-Reality based Software technology in Smart-Phone Environment. The field of Mixed-Reality in mobile environment is relatively young. but Cause to develop Mobile infra and improvement of mobile device, open-platform mobile OS, the request extended This paper suggest the method for Marker Detection and Marker Tracking method. This method is the one of some kind of a base-technology in Mixed Reality. this method is to effect to location and registration.
This paper suggest the method in low CPU computing power. Using a deformable searching area, the method improve computing power. and Using a Cam-shift algorithm, we suggest a calibration free method.

핵심어: 스마트폰, 가변탐색영역, 캠시프트, 혼합현실
Keywords: Smartphone, Deformable Searching Area, Cam Shift, Mixed-Reality

본 논문은 2009년 숭실대학교 교내 연구 학술비 지원에 의하여 연구되었음.
본 연구는 문화체육관광부 및 한국문화콘텐츠진흥원의 문화콘텐츠기술연구소육성사업의 연구결과로 수행되었음.
*주저자 : 숭실대학교 미디어학과 e-mail: krhoonse@hotmail.com
**공동저자 : 숭실대학교 미디어학과 e-mail: Lim0101@naver.com **공동저자 : 숭실대학교 미디어학과 e-mail: mino0507@naver.com
**공동저자 : 숭실대학교 컴퓨터학과 교수 e-mail : gykim11@ssu.ac.kr
***교신저자 : 숭실대학교 미디어학과 교수; e-mail: hic@ssu.ac.kr

1. 서론

모바일상의 혼합현실 기반의 소프트웨어 기술은 현재 국내외적으로 시작 단계에 있다. 하지만 급속한 모바일 인터넷 인프라의 확산, 모바일 기기의 성능 향상 및 개방형 모바일 OS의 등장으로 인해 모바일 혼합현실과 같은 새로운 융합형 IT기술이 사용자의 요구에 부흥하여, 수요가 급신장 되고 있다. 또한 GPS, RFID/USN, 고성능 웨어러블 및 모바일 기기에 대한 광범위한 보급으로 인해 혼합현실 기술의 실질 적용 가능성이 높아지고 있다[1].

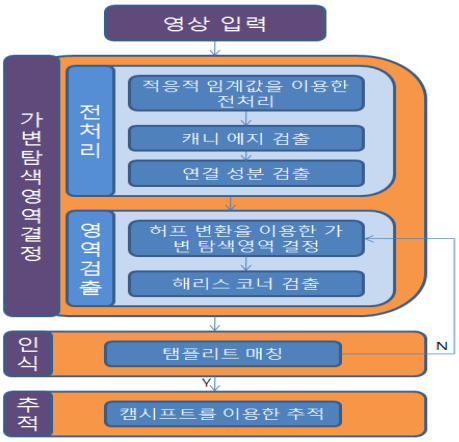
노키아의 "심바안", 구글의 "안드로이드", 애플의 "애플폰", MS의 "윈도우 모바일" 등을 통해 글로벌 IT기업은 차세대 모바일 OS를 선점하기위하여 치열한 경쟁을 펼치고 있고, 노키아의 "MARA 프로젝트", HIT Lab의 "ARToolKIT Plus", 지멘스의 모바일 단말용 혼합현실 게임 "Virtual Mosquito Hunt"등의 모바일 콘텐츠 기술 개발을 진행하고 있다.

윈도우 모바일(Windows Mobile)은 포켓 PC나 스마트 폰과 같은 모바일 단말기를 위한 윈도우CE 기반의 운영체제이다. PC의 운영체제 휴대전화나 휴대용 멀티미디어 플레이어 같은 모바일 기기를 위한 운영체제로서, 각 기능들이 모듈화되어 있어 개발하고자 하는 타깃 디바이스(H/W)의 용도에 따라 선택적인 기능 추가가 가능하고, 임베디드 시스템과 같은 제한적인 하드웨어 환경에 적합하도록 구성되어 있다.

혼합현실은 카메라로부터 획득한 실제 영상으로부터 3차원 공간정보를 획득하여, 부가적인 정보를 담고 있는 2차원 또는 3차원 가상객체를 실제 영상에 정합하여 보여주는 기술이다. 이러한 기술을 모바일 단말기 상에서 구현하기 위하여, 모바일 기기의 마커 검출, 추적, 위치 인식, 영상 정합 등의 모바일 혼합현실 기반 기술이 필요하다.

모바일 혼합현실 기반 기술 중 하나인 마커의 검출과 추적은 혼합현실의 가장 기본이며 매우 중요한 과정으로, 입력되는 영상으로부터 마커를 인식하고, 마커를 추적함으로서 인식하고 추출된 정보를 이용하여 위치정보를 얻거나 영상의 정합의 위치를 결정하게 되는 모바일 혼합현실에서 가장 기본이 되는 매우 중요한 과정이다. 카메라로 입력받은 영상에 대하여, 전처리 과정을 거친 후 연결성분 검출과 허프 변환을 이용하여 마커를 영역을 검출하고, 탬플리트 매칭을 이용하여 마커를 인식하는 방법[2], 사전에 정의된 특징(3개의 코너스톤과 2개의 직교 가이드 바)을 마커에 부여하여 마커의 정보를 추출하는 방법[3] 등의 검출 방법과, 호모그라피 연산을 이용하여, 카메라와 화면의 상관관계를 이용하여 위치를 정의 하는 방법[4], 평면적인 정보를 인식하고 추적하는 방법[5], 호모그라피 연산과 코너정보 예측, 추출, 갱신을 이용하여 추적하는 방법[6][7] 등이 있다. 하지만 기존의 제안된 방법은 캘리브레이션이라는 위치인식 기술이 전제된다는 하에 위치정보를 추출 및 추적하고 있고, 실제 모바일 환경하의 단안 카메라 정보를 이용하여 캘리브레이션을 수행하여 적용하기엔 많은 어려움이 존재한다. 또한 PC에 비해 상대적으로 매우 낮은 CPU를 이용하는 모바일 환경에 대하여 실시간으로 검출가능하고 적용 가능한 최적화된 알고리즘이 필요하다.

본 논문에서는 낮은 CPU를 이용하는 모바일 환경에 최적화된 실시간 마커 추적 방법에 대하여 제안한다. 제안하는 방법의 개요는 다음의 [그림1]과 같다.



[그림 1] 제안하는 방법의 개요

낮은 CPU의 처리속도를 보완하기 위하여 처리영역을 축소하여 가변 탐색영역내의 공간에서 처리를 수행하고, 좌표변환정보를 이용하여 추적을 수행하는 캘리브레이션 과정을 대신하여 캠시프트알고리즘[8]을 이용하여 효과적인 추적방법을 제안한다.

본 논문의 구성은 다음과 같다. 제 2절에서는 스마트폰 환경에서 영상에 대한 전처리 방법을, 제 3절에서는 마커를 추적하고 인식하는 방법의 개요를 기술하며, 제 4절에서는 캠시프트알고리즘을 이용한 추적방법에 대하여 기술하고, 제 5절에서는 실험 및 결과, 제 6절에서 결론에 관하여 기술한다.

2. 스마트폰 환경에서의 영상 처리의 특징

스마트폰의 특징은 카메라가 내장이 되어있고, PC에 비해 상대적으로 낮은 CPU(220~ 520Mhz)가 탑재 되어 있으며, 마이크로 소프트 운영체제인 윈도우 모바일이 사용된다.

2.1 기기의 특징

스마트폰에서 사용되어지는 CPU는 일반적으로 ARM계열 프로세서가 사용되며, PC에 비해 상대적으로 낮은 클럭(

220~520Mhz)과 낮은용량의 메모리를 가지고 있다.

2.2 운영체제의 특징

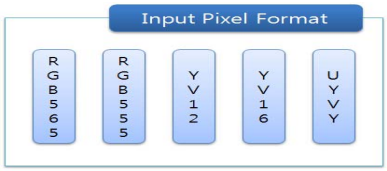
윈도우 모바일(Windows Mobile)은 포켓 PC나 스마트 폰과 같은 모바일 단말기를 위한 윈도우CE 기반의 운영체제이다. PC의 운영체제 휴대전화나 휴대용 멀티미디어 플레이어 같은 모바일 기기를 위한 운영체제로서, 각 기능들이 모듈화되고 각 모듈에 대한 API를 제공함으로서 응용프로그램 개발에 도움이 되고 있다. 또한 모듈화된 기능을 이용하여 개발하고자 하는 타깃 디바이스(H/W)의 용도에 따라 선택적인 기능 추가가 가능하고, 임베디드 시스템과 같은 제한적인 하드웨어 환경에 적합하도록 구성되어 있다.

2.3 카메라 특징

스마트폰에서 카메라를 이용하여 영상을 입력받기 위해, 윈도우 모바일의 주요 API인 Camera API를 이용하여 영상을 입력받는다. 획득된 영상의 픽셀 특징은 다음의 [그림 2]와 같다.

[그림 2] 카메라를 이용한 입력 영상의 픽셀 형식

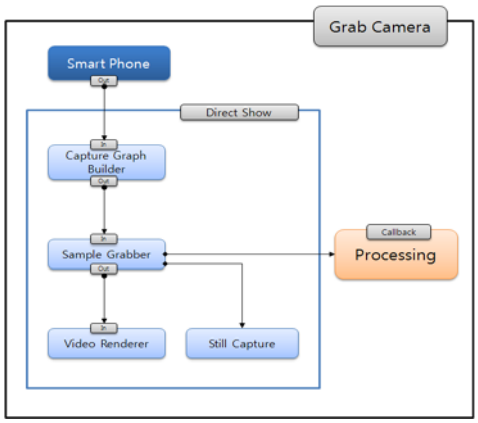
A는 Alpha, B는 파란픽셀, G는 녹색픽셀, R은 빨간색픽셀을 나타내며, RGB565는 R과 G는 5byte, G는 6byte의 자료 값을 나타낸다. 하지만 DirectX를 이용한 영상획득의 경우에는 다음의 [그림 3]과 같은 픽셀 형식을 가진다.



[그림 3] Direct X 픽셀형식

다이렉트 엑스를 이용한 영상획득의 경우 대표적으로 16bit RGB 형식과 YUV형식을 입력받는다. 16bit RGB의 하나인 RGB565는 R이 5bit, G가 6bit, B가 5bit의 형태이며, RGB555는 R이 5bit, G가 5bit, B가 5bit이며, 1bit의 널데이터를 가진다. 또한 YUV의 경우에는 32비트의 형식으로 존재하며, YV12, YV16같은 경우 영상에서 정보를 표현하는 방법에 따라 분류되며, UYVY같은 경우 중복된 값이 아닌 비트가 분할된 값을 영상으로 획득하게 된다.

본 논문에서는 영상을 획득하고, 획득된 데이터를 가공하기 위해 DirectX를 이용한 영상처리 방법을 이용한다. DirectX의 경우 Windows Mobile 5.0이후에 포함된 API의 일종이다. DirectX의 API중, 미디어를 조정하거나 열고 닫기를 하는 API를 DirectShow라 하며, 일종의 인터페이스로 표현한다. 하지만 Windows Mobile의 DirectShow는 영상을 받아오거나 출력하는 기능만 가능한 문제가 있다. 이러한 문제를 해결하기 위하여, DirectShow의 커스팀 필터를 생성하여 연결하는 방법으로 모바일에서 영상을 획득 가공 하는 과정을 수행할 수 있다. 다음의 [그림 4]는 DirectShow에서 영상을 획득하는 과정에 대한 도표이다.



[그림 4] 카메라의 영상획득 및 가공

[그림 4]를 보면, 스마트폰에서 CaptureGraphBuilder를 이용하여 영상을 획득하고, Sample Grabber를 이용하여 영상의 처리를 하게 된다. 바로 이 Sample Grabber가 Windows mobile에서 지원하지 않는 부분인데, 커스팀 필터 제작을 통하여 PC에서와 같은 인터페이스를 구성하여 연결할 수 있다. Sample Grabber를 이용하여 가공한 영상을 최종적으로 Video Renderer를 이용하여 화면에 출력 할수 있다.

3. 가변 탐색영역의 결정

스마트폰 자체의 낮은 처리속도를 보완하기 위하여 본 논문에서는 가변탐색영역을 이용한 처리 방법을 제안한다.

가변탐색영역이란 검출된 마커에 일정한 크기의 탐색영역을 결정하고 그 탐색영역 내에서만 연산이 이루어지게 결정하는 것으로서, 처리영역의 제한을 통하여 적은 처리영역으로 효과적인 결과를 검출하기 위한 방법이다.

가변 탐색영역을 결정하려면 획득한 영상에서 마커의 정보가 필요하다. 마커의 정보를 검출하기 위하여 전처리 과정을 수행하고, 탐색영역을 결정하게 된다.

3.1 전처리

입력된 영상에서 유효한 정보를 처리하기 위하여 전처리 과정을 수행한다. 전처리 과정에서는 쓸모없는 정보를 없애기 위하여 적응적 임계값을 이용한 이진화방법[4]과 케니에지 연산자[4]를 이용하여 이진화한 정보에 연결성분을 검출(레이블링)하는 방법을 이용 획득된 카메라 영상에서 잡음이 제거된 영상과 유효한 정보를 획득한다.

3.2 영역검출

획득된 유효한 정보를 이용하여, 영역을 정의하기 위하여, 허프 변환 알고리즘[9]을 이용한다. 허프 변환 알고리즘을 이용하여 마커의 외각성분을 조사하고, 결정된 위치에 가변 탐색영역의 크기와 위치를 설정한다.

탐색영역의 유효성을 확인하기 위하여 가변탐색영역내의 영상에 대하여, 해리스 코너 검출[4] 알고리즘을 이용하여 가변탐색영역내의 특징점을 추출, 탬플리트 정합을 이용하여 가변탐색영역의 유효성을 판단한다.

4. 캠시프트 알고리즘을 이용한 적응적 마커 추적

스마트폰을 이용하여 카메라를 이용하여 영상을 획득할 때, 고정된 마커와 이동하는 카메라 간의 관계를 이용하여 본 논문은 캠시프트 알고리즘[8]을 이용하여 효과적인 추적을 수행하는 방법을 제안한다.

캠시프트 알고리즘은 전환되는 장면 하에 물체 추적에 대하여 반복적인 민 시프트 알고리즘(Mean-Shift Algorithm)을 적용하는 방법이다. 가변탐색영역(W)에 대하여 크기(a)와 초기 위치(Pk)를 결정하고, 다음의 수식 1을 이용하여 탐색영역의 중심위치와 그라디언트를 계산한다.

$$\overline{P_k}(W) = \frac{1}{|W|}\sum_{j \in W} P_j \qquad (1)$$
$$\overline{P_k}(W) - P_k \approx \frac{f'(P_k)}{f'(P_k)}$$

가변탐색영역의 중심($\overline{P_k}(W)$)의 계산된 위치가 중심위치의 그라디언트(f'(p))가 수렴할수 있도록 계산을 반복하여 물체를 탐색한다.

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

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

식(4)를 이용하여, 첫 번째 모멘트에서 x와 y를 검출하게 된다.

$$M_{10} = \sum_x \sum_y xI(x,y); M_{01} = \sum_x \sum_y yI(x,y) \qquad (4)$$

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

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

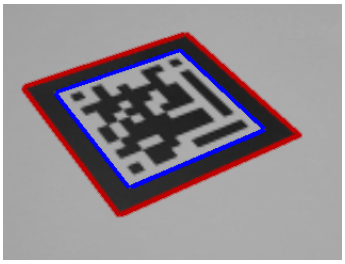
검출된 영역을 중심으로 다음 장면에 대하여 탐색영역의 위치를 (4)번 수식에서 나온 mean탐색 영역 중심으로 이동 반복적으로 수행하여 각 장면에 대하여 탐색영역의 중심을 반복적으로 계산하여 추적을 수행한다.

5. 실험 및 결과

본 실험을 수행하기 위하여 사용된 환경은 ARM계열 CPU와 Windows Mobile 5.0 이상의 운영체제를 이용하여 수행하였으며, 영상의 획득, 가공을 위해 제 2절에서 제안한 방법인 Direct X를 이용한 영상처리의 결과는 다음의 [그림 5]와 같다.

[그림 5] Direct X를 이용한 영상의 획득

탐색영역 검출의 결과를 다음의 [그림 6]과 같다.



[그림 6] 탐색영역의 검출의 예

다음의 표1은 추적과정의 결과를 나타낸다.

	탐색영역 검출	추적
정확도(%)	90	85
처리속도(f/s)	8	1

6. 결론

본 논문에서는 스마트폰이라는 제한된 환경하에서의 증강현실의 방법을 제안한다. 스마트폰이 가지는 특징으로는 ARM 계열 CPU와 적은용량의 RAM을 가지며, Windows Mobile이라는 운영체재를 탑재하고 있다. 이러한 환경에 적용하여 증강현실을 수행하기에는 PC와 다른 적용이나 접근이 필요하다. 하지만 Windows Mobile이라는 운영체제를 이용하여, 개발하는 경우 몇 가지 문제점만 주의하면 PC와 비슷한 개념으로 알고리즘을 적용 할수 있다. 스마트폰에 탑재된 CPU는 ARM계열의 CPU로서 PC의 CPU와는 다르게 부동소수점 연산을 제공하지 않는다. 따라서 부동소수점 연산 사용 시 매우 큰 속도의 저하가 발생하게 된다. 이러할 경우 비트연산을 이용하여 소수점연산을 하게 되면 속도의 보완이 가능하다. 또한, Windows Mobile에서 Windows Mobile 5.0이후에서 지원하는 Direct X를 이용한 카메라 영상획득에는 운영체제가 지원하는 API에서 ISampleGrabber라는 인터페이스가 존재하지 않아 영상을 가공하기 힘든 문제점이 있다. 이 문제를 해결하기 위해 본 논문은 2장에서 제안한 방법과 같이 커스텀 필터를 제작하여 인터페이스로서의 연결로 획득된 영상의 가공이 가능하게 제안하였다.

본논문에서 제안한 방법인 추적의 방법은 캠시프트알고리즘을 이용하여 효과적으로 단안카메라에서 추적하는 방법을 제안하였다. 검출된 탐색영역에 대하여, 매 프레임에 민시프트알고리즘을 적용, 이미 정의된 탐색영역 주변의 추적과정을 이용하여 고정된 물체의 상대적인 좌표계 및 움직임 정보를 추출할 수 있다.

참고문헌

[1] ETRI "모바일 혼합현실 원천핵심기술 개발 “, 연구기획 보고서, 2008.5

[2] Ivan Poupyrev, Desney S. Tan, Mark Billinghurst, Hirokazu Kato, Holger Regenbrecht, Nobuji Tetsutani, "Developing a generic Augmented-Reality Interface", IEEE, pp44~50, 2002

[3] Shahzad Malik, Gerhard Roth, Chris McDonald, "Robust 2D Tracking for Real-Time Augmented Reality",

[4] Xu, Qiang, "Visual Code Marker Detection",

[5] Gilles Simon, Andrew W.Fitzgibbon, Andrew Zisserman, "Markerless Tracking using Planar Structures in the Scene",

[6] Hirokazu Kato and Mark Billinghurst, "Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System

[7] Jun Rekimoto, "Matrix: A Realtime Object Identification and Registration Method for Augmented Reality"

[8] 김세훈, 황중원, 김기상, 최형일, "장면 전환에서의 물체 추적을 통한 모델기반추적 방법 연구 “, HCI 2008, vol.3 , 2008