휴대폰 상에서의 손동작 기반 증강현실 인터페이스 구현

최 준 영^{a)}, 박 한 훈^{b)}, 박 정 식^{a)}, 박 종 일^{c)†}

Implementation of Hand-Gesture-Based Augmented Reality Interface on Mobile Phone

Junyeong Choi^{a)}, Hanhoon Park^{b)}, Jungsik Park^{a)}, and Jong-Il Park^{c)†}

요 약

최근 휴대폰의 발전과 더불어 휴대폰 환경에 적합한 인터페이스에 대한 연구가 많이 이루어지고 있다. 본 논문에서는 특히 사람에게 친숙한 손동작을 이용한 상호작용에 주목하여, 휴대폰 상에서 손동작을 이용한 비전 기반 인터페이스를 구현하였다. 사용자가 한 손에 휴대폰을 들고 휴대폰의 카메라를 통해서 다른 손의 손바닥을 바라보면, 사용자의 손바닥 위에 가상 객체가 증강되고 사용자는 자신의 손과 손가락의 움직임을 통해서 가상 객체와 상호작용 할 수 있다. 구현된 인터페이스는 사람에게 친숙한 손동작을 기반으로 동작하기 때문에, 사용자는 별도의 훈련 과정 없이 친숙하게 사용가능하다. 또한, 추가적인 센서나 마커를 요구하지 않기 때문에, 사용자가 원하는 시간과 장소에서 자유롭게 가상 객체와 상호작용 할 수 있다. 일부 시간이 많이 걸리는 처리 과정을 휴대폰 환경에 적합하도록 최적화, 고속화함으로써, 구현된 인터페이스는 상용화된 휴대폰(Galaxy S2)에서 5 fps로 동작한다.

Abstract

With the recent advance in the performance of mobile phones, many effective interfaces for them have been proposed. This paper implements a hand-gesture-and-vision-based interface on a mobile phone. This paper assumes natural interaction scenario when user holds a mobile phone in a hand and sees the other hand's palm through mobile phone's camera. Then, a virtual object is rendered on his/her palm and reacts to hand and finger movements. Since the implemented interface is based on hand familiar to humans and does not require any additional sensors or markers, user freely interacts with the virtual object anytime and anywhere without any training. The implemented interface worked at 5 fps on mobile phone (Galaxy S2 having a dual-core processor).

Keyword : Augmented reality, interface, gesture recognition, hand shape recognition

a) 한양대학교 전자컴퓨터통신공학과

Hanyang University Department of Electronics and Computer Engineering

b) NHK 방송기술연구소

NHK Science & Technology Research Laboratories

c) 한양대학교 컴퓨터공학부

Hanyang University Department of Computer Science and Engineering

† 교신저자 : 박종일 (jipark@hanyang.ac.kr)

※ 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2011년도 콘텐츠산업기
술지원사업의 연구결과로 수행되었음

· 접수일(2011년9월6일), 수정일(2011년10월25일), 게재확정일(2011년11월2일)

1. 서 론

최근 휴대폰은 빠른 속도로 발전, 보급되고 있으며, 또한 휴대폰의 휴대성 때문에 많은 사람들은 휴대폰을 다양한 어플리케이션을 위한 강력한 플랫폼으로 여기고 있다. 그러나 대부분의 휴대폰은 단순한 인터페이스(버튼, 터치스크린, 음성 인식 등)만을 채택하고 있기 때문에, 사용자에게

친숙하고 휴대폰 환경에 적합한 인터페이스의 개발이 요구되어 진다.

Myron Krueger는 Artificial Reality란 책에서 자연스러운 상호작용이란 음성과 동작을 뜻한다고 서술하였다^[1]. 또한 손은 사람과 사람 사이의 상호작용에서 가장 중요한 상호작용 도구이기 때문에, 본 논문은 손과 동작, 두 가지 핵심어를 포함하는 손동작에 초점을 맞추고, 휴대폰 환경에 적합한 손동작 기반 인터페이스를 구현하였다.

구현된 인터페이스는 자연스러운 상호작용 시나리오를 가정한다. 사용자는 한 손에 휴대폰을 들고, 휴대폰의 카메라를 통해서 다른 손의 손바닥을 바라보면, 손바닥 포즈 계산을 통해서 손바닥 위에 가상 객체가 증강된다. 또한, 구현된 인터페이스는 손가락 동작을 인식할 수 있기 때문에, 사용자는 손가락을 움직여서 가상 객체와 상호작용할 수 있다.

비슷한 인터페이스는 그동안 꾸준히 연구되어 왔다. 이와 유사한 인터페이스로 산타바바라 대학에서는 새로운 개념의 손 기반 인터페이스를 제안하였다^[8]. 이 인터페이스는 착용식 컴퓨팅(wearable computing) 환경에서 사용자의 몸에 장착된 카메라를 통해서 손을 바라보는 상황을 가정한다. 이 인터페이스는 손바닥 포즈를 계산함으로써 손바닥 위의 가상 객체와의 상호작용을 제공한다. 그러나 이 인터페이스는 손바닥 포즈를 계산하기 위해서 손가락 끝 점의 좌표를 필요로 하기 때문에, 사용자는 상호작용을 하는 동

안 자신의 손가락을 움직일 수 없다. 결국, 상호작용의 범위나 종류가 극히 제한된다.

최근 울트라 모바일(ultra-mobile) PC 환경에서 더욱 발전된 비전 기반의 손을 이용한 인터페이스가 제안되었다. 이 인터페이스는 사용자가 한 손에 모바일 기기를 들고 있고, 모바일 기기의 카메라를 통해서 다른 손을 바라보는 자연스러운 상호작용을 가정한다. 사용자의 손바닥 포즈뿐만 아니라 손가락의 움직임까지 인식할 수 있기 때문에, 사용자의 손바닥 위에 있는 가상 객체와의 더욱 다양하고 세련된 상호작용을 제공할 수 있다^[9]. 그러나 이 인터페이스는 손목 위치 정보와 엄지와 검지 사이의 골 위치 정보를 이용해서 손바닥 포즈를 계산하는데, 이 정보들은 손바닥 포즈를 계산하기에는 불충분한 정보이기 때문에 손바닥 포즈는 한 쪽 방향으로만 변할 수 있다고 제한하였다. 즉, 이 인터페이스는 손의 피치(pitch)와 요(yaw) 방향으로 $0^{\circ} \sim 60^{\circ}$ 는 인식할 수 있지만 $0^{\circ} \sim 60^{\circ}$ 는 인식할 수 없다. 또한, 이 인터페이스는 널리 사용되지 않은 기기인 울트라 모바일 PC에 구현되었기 때문에, 스마트폰과 같이 사용자들에게 친숙한 환경에서의 개발이 필요하며, 아직까지 그 예를 찾아보기 힘들다.

본 논문에서는 모바일 폰 환경에서 더욱 세련된 상호작용을 제공하는 비전 기반 인터페이스를 구현하였다. 구현된 인터페이스는 가능한 모든 손바닥 포즈(피치 방향으로 $-60^{\circ} \sim 60^{\circ}$, 요 방향으로 $-45^{\circ} \sim 60^{\circ}$, 롤(roll) 방향으로 -180

그림 1. 모바일 기반 손동작을 이용한 증강현실 인터페이스. 사용자가 손가락을 펴면, 손바닥 위의 가상객체가 춤을 춘다.

Fig. 1. Hand-gesture-based augmented reality interface on mobile phone. The android character is rendered on user's palm nad dances when user closes his/her hand.

°~180°를 인식)를 더욱 정확히 인식할 수 있다. 또한, 착용식 컴퓨팅 환경에 구현된 인터페이스^[8]는 가상 객체와의 손가락 기반 상호작용을 제공하지 않으며, 울트라 모바일 PC 환경에 구현된 인터페이스^[9]는 손바닥을 펴거나 주먹을 쥐는 등 단순하고 제한된 상호작용만을 제공했지만, 본 논문에서 모바일 폰에 구현한 인터페이스는 형상 분해 기반의 손가락 검출 방법을 채택함으로써 손가락을 이용한 메뉴 선택과 같은 손가락 기반의 정교한 상호작용도 제공할 수 있다. 이 인터페이스를 모바일 환경에 구현하기 위하여, 기존에 제안된 손 형상 인식 방법^[10]을 모바일 환경에 적합하도록 변형하고, 휴대폰에 이식하였다. 최종 결과물로 구현된 인터페이스는 상용화중인 휴대폰에서 상호작용 가능한 속도로 동작하며, 그림 1은 구현된 인터페이스를 이용한 간단한 상호작용의 예를 보여준다.

II. 관련 연구

휴대폰 인터페이스에 관한 연구는 크게 하드웨어 기반이나 비전 기반으로 분류 될 수 있다. 그러나 하드웨어 기반의 인터페이스는 일반적으로 추가적인 장비나 물리적 센서를 요구하기 때문에, 현재 상용화된 휴대폰에는 적용하기 어려울 수 있으며, 이러한 장비나 센서는 사용자들에게 불편할 수 있다^{[2][3]}. 따라서 비전 기반 인터페이스에 관한 관심이 자연스럽게 높아지고 있다.

휴대폰 환경에서의 비전 기반 인터페이스의 예로, 카메라 영상에서의 광류(optical flow)로부터 휴대폰의 움직임을 간접적으로 측정하여 이를 상호작용 입력으로 사용하는 방법이 있다^[4]. 그리고 화면 위쪽에 스테레오 카메라를 장착하고 사용자가 바라보는 곳을 추적함으로써 사용자의 시점을 인터페이스 입력으로 사용한 인터페이스도 있다^[5]. 또한, 손동작을 이용한 인터페이스들도 제안이 되었는데, 휴대폰을 들고 있는 손의 검지 끝을 검출, 추적하여, 마우스의 기능을 휴대폰에 적용한 인터페이스가 있다^[6]. 또한, 살색을 효과적으로 표현할 수 있는 색상 공간으로 카메라 영상을 변환한 후, 변환된 영상의 그래디언트(gradient)를 분석하여 손의 위치 및 모양을 추정하고 상호작용 입력으로 사

용하는 포인팅(pointing) 시스템이 제안되기도 했다^[7]. 그러나 이러한 현존하는 비전 기반 인터페이스는 단순하고 제한된 상호작용만을 제공한다. 따라서 휴대폰 환경에서 세련된 상호작용을 제공하는 방법에 대한 연구는 중요한 이슈 중 하나이며, 본 논문에서는 휴대폰 환경에서 손과 손가락 움직임을 이용한 세련된 상호작용을 제공하는 효과적인 인터페이스를 구현하였다.

III. 손바닥 포즈 및 손동작 인식

구현된 인터페이스의 전체적인 흐름은 그림 2와 같으며, 크게 세 단계(손 영상 검출, 손바닥 포즈 계산, 손가락 동작 인식)로 구성된다. 구현된 인터페이스는 YCrCb 좌표계 상의 문턱값 처리(thresholding)와 거리 변환을 이용해서 입력 영상으로부터 손을 검출한다. 그 뒤, 데이터베이스 안의 모든 손바닥 포즈에 대한 주성분 분석법(PCA)과 히스토그램 기반의 비교를 통해서 손바닥 포즈를 인식한다. 그리고 형상 분해 기반의 방법을 이용해서 손동작을 인식한다. 이때, 더욱 정확한 인식을 하기 위해서, 손동작 인식 전에 손바닥 포즈 정보를 이용해서 손 영상을 손 정면 영상으로 변환한다. 구현된 인터페이스에서 가상 객체는 손바닥 포즈 정보를 기반으로 사용자의 손바닥 위에 증강되고, 사용자의 손가락 움직임에 맞춰 반응한다.

1. 손 영상 검출

입력 영상으로부터 손 영상을 검출하기 위해서, 우선 입력 영상에 YCrCb 색상 좌표계 상에서의 문턱값 처리를 적용해서 살색 영역을 검출한다. 그러나 검출된 살색 영역은 손뿐만 아니라 살색을 가진 배경을 포함하고 있기 때문에, 거리 변환을 이용해서 살색 배경을 제거하고 그림 2-(b)와 같은 손-팔 영역을 찾는다. 손-팔 영역은 손가락 끝부터 손목까지는 폭이 변화하다가, 손목 이후부터는 폭이 일정하다는 특징이 있기 때문에, 손-팔 영역의 폭이 일정해지기 시작하는 부분을 찾아서 이를 손목이라고 정의한다. 그리고 손목의 폭과 위치를 기반으로 그림 2-(c)와 같이 손을

포함하는 최소 영역을 찾고, 이 영역을 일정한 크기로 변환한다(본 논문에서는 100×100 , 그림 2-(d))^[10].

2. 손바닥 포즈 계산

검출된 손 영상은 손바닥 포즈 계산을 방해할 수 있는 배경 정보를 포함하고 있기 때문에, 그림 2-(e)와 같이 YCrCb 색상 좌표계상의 문턱값 처리를 통해서 배경 정보를 제거한 뒤, 거리 변환을 적용한다. 거리 변환된 손 영상은 그림 2-(f)에서 보이는 것과 같이 손가락 부분의 비중이 낮아지며, 손바닥의 기울어진 방향과 크기를 잘 표현하고 있다.

손바닥 포즈를 계산하기 위해서 거리 변환된 손 영상을 두 종류의 데이터로 변환한다. 하나는 주성분 분석법 공간상의 좌표 값이며, 다른 하나는 그래디언트 히스토그램(gradient histogram)^[11]이다. 첫 번째 데이터를 구하기 위해서, 다양한 손바닥 포즈를 가지고 있는 많은 수의 손 영상으로 데이터베이스를 구성하고, 이 데이터베이스의 모든 손 영상에 거리 변환을 적용한 뒤 주성분 분석법 공간을 만든다. 이렇게 만든 주성분 분석법 공간상에 입력 영상으로부터 얻은 거리 변환된 손 영상을 투영하고, 좌표 값을 구한다. 그리고 데이터베이스의 모든 손 영상 중에서 가장 좌표 값이 가까운 손 영상 N개를 찾고(그림 2-(i)), N개 중에서

그림 2. 구현된 인터페이스의 흐름도. (a) 입력 영상, (b) 손 영역 검출, (c) 손을 포함하는 최소 영역, (d) 손 영상, (e) 배경을 제거한 손 영상, (f) 거리 변환된 손 영상, (g) 최댓값을 포함하는 일정 영역, (h) 분리되고 정규화된 일정 영역, (i) 가장 비슷한 N-개의 손 영상, (j) 히스토그램 계산, (k) 히스토그램, (l) 손의 정면 영상, (m) 손 조각, (n) 손 형상 분해, (o) 손가락 검출, (p) 손가락 포즈 검출 결과, (q) 가상 객체 증강.

Fig. 2. The overall flow of the implemented interface. (a) captured scene, (b) hand region detection, (c) minimal region including hand, (d) hand image, (e) hand image excluding background, (f) distance transformed hand image, (g) pre-defined region including max pixel value, (h) segmented and normalized pre-defined region, (i) N most similar hand images, (j) histogram calculation, (k) histogram, (l) hand's frontal view, (m) hand segments, (n) hand shape decomposition, (o) finger detection, (p) palm pose estimation, (q) virtual object rendering.

i번째 손바닥 포즈에 해당하는 손 영상이 포함된 수를 k_i 라고 정의한다. 두 번째 데이터인 그래디언트 히스토그램을 구하기 위해서, 우선 거리 변환된 손 영상 안에서 가장 큰 화소 값을 가지는 좌표를 구하고, 그 좌표를 중심으로 일정 크기의 영역을 분리 한 뒤, 정규화(normalization)한다(그림 2-(g)와 (h)). 그리고 분리한 영역의 각 화소의 각도(θ)와 크기(m)을 구한다(그림 2-(j)). 그리고 이를 각도에 따른 히스토그램으로 변환 한다. 히스토그램의 x축은 각도를 나타내고, y축은 각 각도를 가지는 화소의 크기 합이다(그림 2-(k)). 그리고 데이터베이스의 모든 손 영상에도 같은 방법을 적용하고, 각 손바닥 포즈별 평균 히스토그램을 구한다. 입력 손 영상의 히스토그램과 각 손바닥 포즈(i)의 평균 히스토그램 간의 차이를 HistoDi라고 정의 한다. 여기서 각 손바닥 포즈(i)에 대한 입력 손 영상의 유사도(s_i)는 다음과 같은 공식으로 나타낼 수 있다.

$$s_i = \frac{k_i}{N} \times e^{-HistoDi \times f} \quad (1)$$

위의 식에서 f 는 사용자 정의 상수이며, 본 논문에서는 10으로 설정되었다. 마지막으로 가장 높은 유사도를 가지는 손바닥 포즈(i)를 입력 손 영상의 손바닥 포즈라고 정의 한다(그림 2-(p)).

3. 손가락 동작 인식

손바닥 포즈 변화에 강인하게 손가락 동작을 인식하기 위해서, 손가락 동작 인식 이전에 계산된 손바닥 포즈 정보를 기반으로 손 영상을 3차원 회전시킴으로써 손의 정면

영상을 획득한다(그림 2-(l)). 다양한 손바닥 포즈별 손 정면 영상 획득의 결과는 그림 3과 같다.

획득한 손의 정면 영상을 이진화한 다음, 손목부터 손가락 끝 방향으로 진행하면서 진행 방향과 수직인 방향으로 화소의 끊어진 횟수를 센다. 바꾸어 말하면, 화소 값이 1에서 0으로 변한 다음 다시 1이 변하는 횟수를 센다. 이때 손의 영역을 구분하기 위해서 만약 끊어진 횟수가 변할 경우에는 손의 정면 영상을 위와 아래 부분으로 분리시킨다. 이런 과정을 거치면 그림 2-(m)과 같이 손의 정면 영상은 여러 가지 조각으로 나뉘는데, 이때 손가락을 하나의 조각으로 합치기 위해서 조각끼리의 상관관계를 통해서 다시 조각을 연결해준다. 상관관계를 구하기 위해서 우선 임의의 두 조각을 연결하고, 연결한 두 조각의 폭이 갑자기 변하는 부분이 있는지 없는지를 찾은 다음, 갑자기 변하는 부분이 있다면 연결한 조각을 다시 원래 상태로 되돌리고 없다면 연결한 채로 둔다. 본 논문에서는 100×100 해상도의 손 영상에 대해서 폭이 15 화소 이상 차이 날 경우에 원래 상태로 되돌렸다. 이와 같은 과정을 거친 결과는 그림 2-(n)과 같으며, 그림에서 보이는 것과 같이 손가락과 손바닥이 분리가 되는데, 이때 손가락은 가늘고 긴 조각을 찾음으로써 검출이 된다(그림 2-(o)).

구현된 인터페이스는 손가락 위치뿐만 아니라 그 길이도 검출할 수 있다. 그림 4는 손가락을 완전히 편 손 모양과 살짝 구부린 손 모양, 두 종류의 손 모양 인식 과정 및 결과를 보여준다. 그림 4-(d)와 (h)를 비교해보면 알 수 있듯이, 손가락을 완전히 편 것을 살짝 구부린 것보다 더 길게 검출함으로써, 손 모양을 섬세하게 구분할 수 있는 것을 알 수 있다.

그림 3. 손의 정면 영상 획득

Fig. 3. Hand's frontal view acquisition

4. 각 단계의 병렬 처리

구현된 인터페이스의 두 번째 단계인 손바닥 포즈 계산은 데이터베이스와의 반복적인 비교를 기반으로 하기 때문에 시간이 오래 걸린다. 따라서 그림 5와 같이 손바닥 계산을 다른 단계들과 병렬적으로 연산되도록 하였다. 하나의 스레드(thread)가 입력 영상으로부터 손 영상을 검출하고, 다른 스레드에게 손 영상을 넘기면 다른 스레드는 손바닥 포즈를 계산한다. 그리고 이때 손 영상을 검출한 스레드는

손 정면 영상 획득 및 손동작을 인식한다.

IV. 실험 및 검증

인터페이스를 실제로 사용하기 위해선 인터페이스의 안정성이 무엇보다 중요하며, 구현된 인터페이스와 같이 인식 기반 인터페이스의 경우 이는 인식률에 비례한다. 따라서 본 논문에서는 구현된 인터페이스의 인식률 실험을 수

그림 4. 손 형상 분해. (a, e) 손의 정면 영상, (b, f) 손의 조각, (c, g) 상관관계에 의해서 연결된 조각, (d, h) 손가락 검출

Fig. 4. Hand shape decomposition. (a, e) hand's frontal view image, (b, f) segments of hand, (c, g) merged segments based on the connectivity, (d, h) finger detection

그림 5. 구현된 인터페이스의 병렬 처리

Fig. 5. Parallel processing of the proposed interface

그림 6. (a) 검출된 손의 예, (b) 부적절하게 검출된 손의 예
Fig. 6. (a) An example of detected hand, (b) an example of inappropriate detected hand

행하였다. 구현된 인터페이스에서는 그림 6-(b)와 같이 엄지손가락의 뿌리가 완전히 접히는 경우 손을 검출하지 못하는 경우가 있다. 따라서 본 논문의 인식을 실험에서는 기본적으로 엄지손가락의 뿌리가 완전히 접히지 않으며, 또한 손가락이 서로 붙어 있지 않다고 가정한다. 인식을 실험은 다양한 입력 영상에서 손을 찾고, 손바닥 포즈를 구하고, 손 모양 인식까지를 뜻하며, 구현된 인터페이스는 그림 7과 같은 16 종류의 손바닥 포즈를 포함하는 800 개의 영상에 대한 인식을 실험에서 평균적으로 83%의 인식을 보였다. 이 실험에서 손바닥 포즈는 피치 방향으로 $-60^{\circ} \sim 60^{\circ}$, 요 방향으로 $-45^{\circ} \sim 60^{\circ}$, 롤 방향으로 $-180^{\circ} \sim 180^{\circ}$ 의 범위를 가지며, 그림 8은 구현된 인터페이스에서 사용된 방법과 일

반적인 PCA 기반 손동작 인식 방법의 인식률을 손바닥 포즈 별로 비교한 것이다. 이때, PCA 기반 손동작 인식 방법은 구현된 인터페이스와 동일한 데이터베이스를 이용해서 PCA 공간을 구축한다. 그리고 입력 영상으로부터 손 영상을 검출하는 것 까지는 구현된 인터페이스와 동일하고, 검출한 손 영상을 구축된 PCA 공간상에 투영 시켜서 좌표값을 얻은 뒤, 데이터베이스안의 손 영상의 데이터와의 비교를 통해서 가장 가까운 좌표값을 가진 손 영상을 찾아서 인식한다. 모든 포즈에 대해서 구현된 인터페이스에 사용된 방법이 월등히 높은 인식률을 가졌다. 다만, 요 방향으로 크게 회전하면 인식률이 급격히 저하되는 것을 확인할 수 있는데, 이는 손가락이 서로 겹쳐서 구분이 안 되기

그림 7. 실험에 사용된 손 영상
Fig. 7. Hand images used in experiment

때문이다. 구현된 인터페이스는 SAMSUNG Galaxy S2(1.2 GHz dual core processor, 1GB RAM, 그리고 Android 2.3 OS)상에서 최대 5.4fps에서 최소 4.4fps의 속도를 보였으며, 평균적으로 5fps의 속도로 동작했다.

그림 8. 피치와 요 방향 회전에 대한 인식률
Fig. 8. Recognition rates for pitch and yaw rotations

본 논문에서는 구현된 인터페이스의 효용성을 검증하기 위해서, 다음과 같은 다양한 손 기반 어플리케이션을 구현하였다.

손바닥 포즈 변화를 이용한 어플리케이션: 구현된 인터페이스는 기본적으로 손바닥 포즈를 계산하기 때문에, 손바닥 포즈 변화를 이용한 어플리케이션에 적용 가능하다. 그림 9는 손바닥 포즈 변화를 이용한 가상 객체와의 상호작용의 예를 보여준다. 사용자가 자신의 손바닥을 기울일 때, 손바닥 포즈의 변화를 감지하고 포즈가 일정 수준에 달하면 가상 객체는 손바닥에서 미끄러진다.

그림 9. 손을 기울일 때 손에서 미끄러지는 안드로이드 캐릭터와의 상호작용 ((a)부터 (d) 순서로 진행)
Fig. 9. Slide interaction with the android character when the hand is tilting (the procedure is from (a) to (d))

그림 10. 손을 흔들 때 점프하는 안드로이드 캐릭터와의 상호작용 ((a)부터 (d) 순서로 진행)
Fig. 10. Jump interaction with the android character when the hand is shaking (the procedure is from (a) to (d))

손 위치 변화를 이용한 어플리케이션: 구현된 인터페이스는 손의 위치 변화에 따라서 가상 객체가 반응하는 어플리케이션에도 적용 가능하다. 그림 10은 손의 위치 변화를

이용한 가상 객체와의 상호작용의 예를 보여준다. 사용자가 자신의 손을 흔들면, 손목의 위치 정보가 변화하게 되고 어플리케이션은 이를 감지한다. 만약 손목의 위치가 갑자기 위쪽으로 움직이면, 가상 객체는 이에 맞춰서 점프를 한다.

손가락 동작을 이용한 어플리케이션: 구현된 인터페이스는 손바닥의 포즈와 손의 위치뿐만 아니라 손가락 동작도 인식할 수 있기 때문에, 그림 11과 같이 손가락을 이용한 상호작용도 가능하다. 구현된 인터페이스는 손바닥 포즈와 함께 손가락의 위치와 길이를 검출하기 때문에, 몇 개의 손가락이 구부러졌으며 구부러진 손가락이 어디에 위치하고 어디를 가리키는가를 알 수 있다. 그러므로 그림 11과 같은 하나의 손가락을 구부려서 메뉴의 항목을 선택하는 상호작용을 제공할 수 있다. 또한, 손가락끼리의 상대적 위치를 이용하면 사용자가 어떤 손가락을 구부렸는지까지 인식할 수 있다. 이를 이용하면 보다 다양하고 세밀한 상호작용을 제공할 수 있을 것이다.

그림 11. 손가락을 이용한 메뉴 선택. (a) 메뉴 선택 화면, (b) 손의 정면 영상과 손가락 동작 검출, (c) 검출된 손 영상
Fig. 11. Menu selection using fingers. (a) Menu selection, (b) hand's frontal view and finger gesture recognition, (c) detected hand

V. 결 론

본 논문에서는 휴대폰 환경에서 동작하는 손동작 기반 증강현실 인터페이스를 구현하였다. 구현된 인터페이스는 다양한 손바닥 포즈와 손가락 동작을 빠르고 정확하게 인식하며, 사용자의 손바닥 위에 증강된 가상 객체와의 다양하고 자연스러운 상호작용을 제공할 수 있다. 현재 실행 속도 면에서의 개선을 위해서, 손바닥 포즈 계산 단계에서 유사도 공식을 간략화 하는 등 여러 가지 방안을 모색하고 있다. 또한, 촉감을 제공하기 위한 진동 센서와의 결합을 통해 보다 실감나는 상호작용이 가능하도록 발전시킬 계획 이다.

참 고 문 헌

- [1] M. W. Krueger, *Artificial Reality*, 2nd ed., Addison-Wesley: Redwood City, CA, 1991.
- [2] T. Miyaki and J. Rekimoto, "GraspZoom: zooming and scrolling control model for single-handed mobile interaction," In proceedings of MobileHCI'09, 2009.
- [3] A. T. Campbell, T. Choudhury, S. Hu, H. Lu, M. K. Mukerjee, M. Rabbi, and R. D. S. Raizada, "NeuroPhone: Brain-mobile phone interface using a wireless EEG headset," In proceedings of MobiHeld'10 2010.
- [4] A. Haro, K. Mori, T. Capin, and S. Wilkinson, "Mobile camera-based user interaction," In proceedings of ICCV, Workshop on HCI'05, pages 79-89, 2005.
- [5] T. Nagamatsu, Michiya Yamamoto, and Hiroshi Sato, "MobiGaze: Development of a Gaze Interface for Handheld Mobile Devices," In proceedings of CHI'11, 2011.
- [6] J. An and K. Hong, "Finger gesture-based mobile user interface using a rear-facing camera," In proceedings of ICCE'11, pages 303-304, 2011.
- [7] O. Gallo, S. M. Arteaga, and J. E. Davis, "A camera-based pointing interface for mobile devices," In proceedings of ICIP'08, 2008.
- [8] T. Lee and T. Höllerer, "Handy AR: Markerless Inspection of Augmented Reality Objects Using Fingertip Tracking," In Proceedings of ISWC'07, 2007.
- [9] B.-K. Seo, J. Choi, J. Han, H. Park, and J.-I. Park, "One-handed interaction with augmented virtual objects on mobile devices," In proceedings of VRCAI'08, 2008.
- [10] J. Choi, H. Park, and J.-I. Park, "Hand shape recognition using distance transform and shape decomposition," In proceedings of ICIP'11, 2011.
- [11] D. G. Lowe, "Object recognition from local scale-invariant features," IJCV, volume 60, number 2, pages 91-110, 2004.

저 자 소 개

최 준 영

- 2007년 : 한양대학교 전자전기컴퓨터공학부 학사
- 2009년 : 한양대학교 전자컴퓨터통신공학과 석사
- 2009년 ~ 현재 : 한양대학교 전자컴퓨터통신공학과 박사 과정
- 주관심분야 : 증강현실, 컴퓨터그래픽스/비전, 3차원 영상처리, 인간컴퓨터 상호작용

박 한 훈

- 2000년 : 한양대학교 전자통신전파공학과 학사
- 2002년 : 한양대학교 전자통신전파공학과 석사
- 2007년 : 한양대학교 공과대학 전자전기컴퓨터공학과 박사
- 2007년 ~ 2008년 : 한양대학교 BK21 박사 후 연구원
- 2008년 ~ 현재 : NHK STRL 박사 후 연구원
- 주관심분야 : 가상현실, 컴퓨터그래픽스/비전, 3차원 영상처리, 인간컴퓨터 상호작용

박 정 식

- 2010년 : 한양대학교 전자전기컴퓨터공학부 학사
- 2010년 ~ 현재 : 한양대학교 전자컴퓨터통신과 석사 과정
- 주관심분야 : 3차원 영상처리, 증강현실, 컴퓨터 비전, GPU 컴퓨팅

박 종 일

- 1987년 : 서울대학교 전자공학과 학사
- 1989년 : 서울대학교 전자공학과 석사
- 1995년 : 서울대학교 전자공학과 박사
- 1992년 ~ 1994년 : 일본 NHK 방송기술연구소 객원연구원
- 1995년 ~ 1996년 : 한국방송개발원 선임연구원
- 1996년 ~ 1999년 : 일본 ATR 지능영상통신연구소 연구원
- 1999년 ~ 현재 : 한양대학교 공과대학 컴퓨터공학부 교수
- 주관심분야 : 가상현실, 컴퓨터그래픽스/비전, 3차원 영상처리, 인간컴퓨터 상호작용