
손과 얼굴의 피부색을 이용한 웹캠 게임 개발

Development of Web-cam Game using Hand and Face Skin Color

오치민, Chimin Oh*, Dhi Aurrahman*, Md. Zahidul Islam*, 김형관, Hyunggwon Kim* 이철우, Chilwoo Lee**

sapeyes@image.chonnam.ac.kr, leecw@chonnam.ac.kr

요약 ~ 소니 아이토이는 Playstation 2에서 개발되었으며 웹캠을 이용하여 사람을 찾고 TV를 통해 게임속의 객체를 직접 터치하여 진행하는 게임이다. 기존의 비디오 게임 방식과 다르게 사용자가 게임 속에 실제적인 주인공으로 등장한다. 거울에 비친 모습처럼 자신의 모습이 게임 안에 나타나게 된다. 비록 이미 상용화되었고 잘 알려진 게임이지만 이러한 게임인터페이스는 사용자에게 흥미를 유도할 수 있는 장점이 있기 때문에 새롭게 개발 응용될 가능성이 많이 있다. 본 논문에서는 이러한 게임을 개발할 수 있는 환경을 개발하기 위해 피부색을 이용해 사람을 찾는 영상 처리 과정, 게임 개발을 위한 그래픽, 게임인터페이스 설계과정을 연구하였다. 설계된 개발환경은 간단한 풍선 터뜨리는 게임을 통해 테스트되었다. 이 개발환경은 영상처리 기능 개선, 제스처인식 추가 등 앞으로 많은 발전 가능성이 있으며 곧 오픈되어 많은 개발자들이 시험적으로 사용할 수 있을 것이다.

Abstract ~ The sony Eytoy is developed on Playstation 2 using webcam for detecting human. A user see his appearance in television and become real gamer in the game. It is very different interface compared with ordinary video game which uses joystick. Although Eyetoy already was made for commercial products but the interface method still is interesting and can be added with many techniques like gesture recognition. In this paper, we have developed game interface with image processing for human hand and face detection and with game graphic module. And we realize one example game for busting balloons and demonstrated the game interface abilities. We will open this project for other developers and will be developed very much.

핵심어: *Eyetoy Playstation, Skin Detection, Segmentation, Web-Cam, Game Software Development Kit*

1. 서론

사람이 직접 게임 속에 들어가 적과 싸우면 얼마나 재미 있을까? 하는 아이디어에서부터 시작된 웹캠기반 게임은 소니 Playstation 2에서 처음 실용화 되었다[1]. 아이토이로 불리는 이 게임은 게임 콘솔에 연결된 웹캠의 영상에서 사람의 얼굴, 손, 몸동작을 인식해 게임의 인터페이스로 사용한다. 기존 조이스틱 방식의 비디오 게임과 다르게 사람이 직접 게임 속에서 몸을 움직여 진행한다. 이미 상용화 된 게임이지만 앞으로 제스처 인식이나 다수의 사용자 환경과 같은 여러 형태로 발전할 가능성이 많이 있다. 본 논문에서는 이를 인터페이스를 설계하여 게임 개발환경을 개발하였다.

본 논문에서는 이러한 웹캠기반 게임을 개발하기 위한 방법을 제시한다. 먼저 게임을 개발할 수 있는 개발환경을 연

구하였다. 이 개발환경은 사람의 움직임을 감지하기 위해 피부색을 추출하여 사람의 영역정보를 얻는 영상인식 모듈이 있다. 그리고 게임 그래픽 모듈, 게임 객체 개발 인터페이스가 있다. 이 모든 모듈은 게임의 영상을 순차적으로 공유하기 위해 파이프라인 버퍼의 개념을 사용하고 있다. 이 게임 개발환경을 통해 소니의 아이토이 게임과 같은 형태의 여러 웹캠 개발이 가능하며 본 논문에서는 단순한 형태의 풍선 터뜨리기 게임을 통해 그 기능을 테스트 해보았다.

본문에서는 먼저 사람의 손과 얼굴을 세그멘테이션 하기 위한 영상처리의 단계를 2장에서 설명한다. 3장에서는 게임 개발을 위한 게임개발 인터페이스 설계를 설명한다. 4장에서는 실험적으로 개발된 풍선 터뜨리기 게임을 소개하며 마지막 장에서는 결론과 함께 향후 계획을 소개한다.

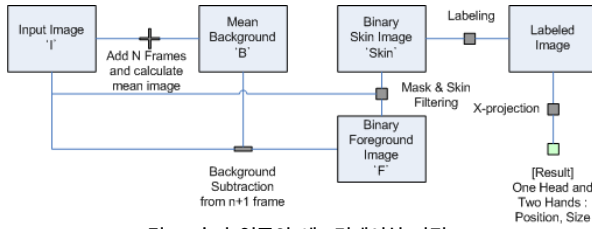


그림 1. 손과 얼굴의 세그멘테이션 과정

2. 손과 얼굴의 세그멘테이션

손과 얼굴의 세그멘테이션은 게임속의 사람의 직접적인 인터페이스 도구로 사용하기 위하여 영상내의 얼굴, 손의 위치와 크기정보를 피부색을 이용하여 얻는 과정이다. 게임 내에서는 손과 얼굴의 위치, 영역정보를 게임을 하는데 가장 중요한 인터페이스 요소로 사용한다. 이러한 사람의 손과 얼굴의 영역은 다음과 같이 영상을 이용해 추출하였다. 먼저 배경을 제거한 후 손과 얼굴의 피부색 영역을 검출하였다. 피부색 영상은 단지 손과 얼굴의 이진영상만 나타내기 때문에 영역과 위치정보를 얻기 위한 레이블링 과정을 거쳐 마무리된다. 본 2장에서는 손과 얼굴의 세그멘테이션 과정을 소개하기 위하여 평균배경영상 및 전경획득, 피부색 분리, 레이블링 및 위치 히스토그램에 의한 손 얼굴 위치와 크기 결정과정을 차례로 소개하였다.

본 논문에서는 평균배경영상을 이용하여 배경에서 검출될 수 있는 피부색 영역을 제거하기 하였다. 배경영상은 입력영상에서 비 움직임 영역인 배경을 제거하기 위하여 사용된다. 움직이는 사람영역만을 배경영상과 입력영상과의 차를 통해 구할 수 있으며 비 움직임 영역에서 나오는 피부색 검출 잡음을 줄일 수 있는 장점이 있기 때문이다.

평균 배경영상은 다음과 같은 순서로 구축된 후 새로운 입력영상과의 차를 통해 전경을 얻는데 사용된다. 먼저 10개의 입력 영상 I는 평균 배경영상 B를 구하기 위해 사용된다. B영상은 10개의 사람이 없는 빈 영상을 모두 더하여 10으로 나눈 평균으로 구한 영상이다. 10프레임 이후 입력영상 I는 전경 F를 구하는데 사용된다. 전경 F는 배경영상B와 입력영상I의 차이값이 30이 넘을 경우 255, 그렇지 않을 경우 0의 값을 갖는 이진영상이다. 최종적으로 이진 전경영상을 통하여 필요한 사람의 영역영상 C를 입력영상 I와 전경영상 I를 mask 연산하여 얻을 수 있다.

↓

$$\begin{aligned}
 B &= \sum_{i=0}^{10} I_i \\
 F &= \text{Threshold}(B - I_{i>10})_{30} \\
 C &= I_{i>10} \wedge F
 \end{aligned}
 \tag{1}$$

$$\begin{aligned}
 C_b &= -0.169R(C) - 0.331G(C) + 0.5B(C) \\
 C_r &= 0.5R(C) - 0.419G(C) - 0.081B(C) \\
 \text{Skin} &= \text{Threshold}(C_b, C_r)_{0 \leq \text{value}(C_b) \leq -50, 10 \leq \text{value}(C_r) \leq 50}
 \end{aligned}
 \tag{2}$$

손과 얼굴의 피부색 영상은 세그멘테이션된 사람영상 C을 YCbCr 컬러모델로 변환된 후 피부색을 나타내는 문턱값을 통해 얻은 이진영상이다. 휘도 성분인 Y는 이용하지 않으며 색채를 나타내는 Cb영상과 Cr영상을 사용한다. 따라서 식2를 통해 Cb, Cr영상을 계산하였다. Skin영상은 색채영상 Cb, Cr의 픽셀값이 각각 -50~0, 10~50 사이에 있을 경우 255의 값을 가지며 아닐 경우 0값을 갖는 이진영상이다[2].

레이블링은 사람의 얼굴과 손으로 추정되는 Skin 영상에서 얼굴과 손의 위치와 영역정보를 구하기 위해 사용된다. 먼저 레이블링을 하기 위해 x축 투영 위치히스토그램을 구한 후 x축에서 일정한 폭을 갖는 객체의 y축 위치와 폭을 구하는 과정을 거쳐 마무리 된다.

피부색 영상 Skin의 x축 투영은 x축 상의 피부색 근접의 위치와 폭의 크기를 얻는데 사용된다. 먼저 영상을 식 3을 통해 X축으로 투영하여 히스토그램 x_hist를 얻는다. 히스토그램 빈은 320×240 영상의 width와 같은 길이이며 y축 상의 피부색 픽셀 개수를 갖는 특징이 있다.

$$x_hist = \int_0^{319} Skin dx
 \tag{3}$$

투영된 히스토그램을 통해 피부색 객체의 위치와 영역은 다음과 같은 순서로 구한다. 먼저 객체의 존재는 히스토그램 x_hist 배열에서 0~319 순서로 탐색한다. 피부픽셀 개수를 나타내는 빈의 값이 15이상인 빈이 연속적으로 10개가 있을 경우 하나의 객체가 존재하는 것으로 정의하였으며 존재 결정 이후 빈의 값이 15미만이 발견된 후 객체의 x축 길이가 결정된다. x축상의 객체의 존재와 위치, 길이가 결정된 후 각 해당 객체의 y축상에서 갖는 가장 작은 픽셀 위치와 가장 높은 픽셀위치를 찾아 객체의 2차원 영상에서의 위치를 최종적으로 결정한다.

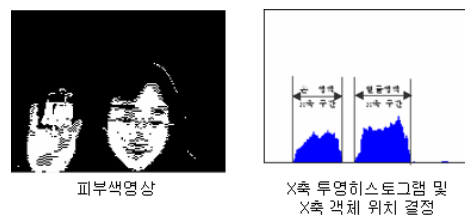


그림 2. 피부색 영상 및 x축 객체 위치 결정

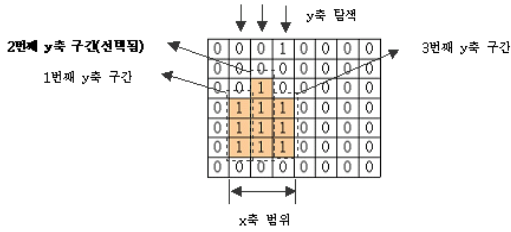


그림 2. y축 객체 위치 결정과정

최종적으로 손과 얼굴은 배경제거와 피부색 검출 그리고 레이블링을 통해 위치와 영역정보를 나타내었다. 그림 3은 그 결과를 나타낸다. 일반적으로 알려진 4또는 8방향 레이블링보다 결과는 좋지 않지만 전경을 구하는 알고리즘이 효과적이지 않아 객체수가 증가하는 오류를 잡기 위해서는 세그먼테이션 객체의 군집을 나타내는 히스토그램이 적절하였다. 현재 세그먼테이션의 개선은 현재 배경모델에 통계모델, 가우시안모델 등 더 고난이도의 이론을 도입하고 y축 투영 계산을 추가함으로써 이뤄질 수 있다.

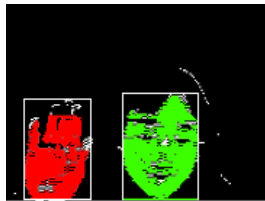


그림 3. 세그먼테이션 된 손과 얼굴 결과

3. 웹캠 게임개발 인터페이스

게임 개발을 위한 인터페이스를 설계하였다. 앞의 세그먼테이션은 영상처리의 과정이고 입력영상을 통해 결국은 손의 2차원 영상에서의 위치와 크기정보를 얻게 된다. 이러한 입력과 결과의 구분을 통해 모듈의 재사용성을 높이기 위한 게임개발 인터페이스를 설계하였다. 따라서 웹캠을 이용한 다양한 게임 개발이 가능하다. 3장에서는 인터페이스 설계시의 사용한 그래픽과 행동과 의미를 가진 게임객체 정의, 사용자의 손, 얼굴을 휴먼객체로 정의, 그리고 게임 알고리즘 동작과 그래픽 드로잉 모듈의 파이프라인을 통한 입력영상의 공유방법을 설명한다.

게임 내에서 휴먼객체는 최대 3개가 존재한다. 즉 두 손과 얼굴이 휴먼객체이다. 휴먼객체는 손과 얼굴을 세그먼테이션 하는 영상처리 모듈의 결과로 생성되며 객체의 영역정보와 크기정보를 알 수 있다. 게임객체는 그래픽 객체를 상속한 것으로 그래픽 드로잉 기술에 필요한 속성을 가지고 있으며 사용자 휴먼객체와 충돌시 발생하는 이벤트에 대한 대응을 정의할 수 있는 인터페이스를 제공하고 있다.

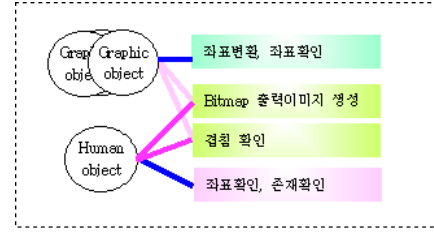


그림 4. 그래픽 객체와 휴먼 객체의 기능

게임객체는 다음과 같은 행위를 정의할 수 있다. 타임아웃, 충돌 이벤트에 대한 행위를 객체마다 다르게 정의할 수 있다. 이 이벤트 중 충돌에 대한 이벤트 구현은 휴먼객체가 게임객체와 겹치면 시스템에 알려줌으로써 이벤트가 발생하며 이때 정의된 이벤트에 대한 행위를 실행하게 된다. 게임객체와 휴먼객체의 영역 비교를 통한 충돌 확인을 나타낸다.

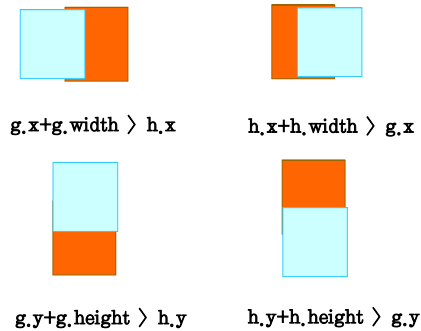


그림 2. 휴먼객체와 그래픽객체 충돌확인

휴먼객체를 얻을 수 있는 영상처리 모듈 그리고 그래픽객체를 상속한 게임객체를 그릴 수 있는 그래픽 모듈은 공통점이 있다. 공통점은 두 모듈 모두 입력영상을 사용하는 것이다. 따라서 입력영상의 순차적인 접근을 위해 모듈 사이에 배열 형태의 파이프라인을 연결함으로써 해결하였다.

웹캠 게임개발 인터페이스는 기존의 영상처리 모듈과 이벤트와 객체지향 프로그래밍 개발론에 근거하여 설계, 개발하였다. 현재 초기 개발단계 수준에 머물고 있지만 이벤트 방식에 의한 객체의 역할 분배, 파이프라인에 의한 입력영상 공유 등 웹캠 개발 공유라이브러리로 개발하기 위한 기초적인 설계 토대는 충분하다고 생각한다. 이를 토대로 간단한 게임을 만들어 보았다.

4. 풍선 터뜨리기 게임

특정 물과 점수를 갖는 3가지 종류의 풍선과 게임의 지시를 위한 이미지의 게임 객체를 정의하였다. 이를 게임 그래픽 모듈에 연결하여 풍선 터뜨리기 게임을 제작하였다.



그림 2. 풍선 터뜨리기 게임

게임의 순서는 다음과 같다. 사용자가 NEW 버튼을 선택한다. 첫 스테이지를 시작하며 풍선을 터뜨리면 종류에 따라 점수를 획득 또는 잃게 된다. 총 4단계의 스테이지가 있으며 스테이지당 500점의 점수를 30초 안에 획득한 후 스테이지가 끝날 때까지 유지하면 다음 스테이지로 넘어갈 수 있다.

게임 제작은 시나리오, 스케치에 의한 프로토타입 작성, 게임 풍선 및 관련 디자인, 게임 객체 룰 지정, 구현 순서로 이루어졌다. 게임 객체들의 룰을 지정하고 게임 객체의 그래픽 디자인을 작성한 후 실제 구현에서 그래픽 연결과 룰을 구현하면 제작을 쉽게 마칠 수 있었다.

5. 결론

단순하지만 제작된 영상처리 모듈은 사람의 손과 얼굴을 찾는데 충분하였다. 주요 문제점은 얼굴과 손의 영역이 완벽

히 검출되지 않는 문제점과 레이블링에서 x축 투영만 고려한 결과 손, 얼굴이 하나의 객체로 인식되는 오류가 있었다. 이는 풍선 터뜨리기와 같은 단순한 터치 게임을 하는데 불편함은 없었지만 차후 보완해야할 점으로 보인다.

이러한 게임 개발환경은 단 한번의 구현으로 끝나지 않게 하기 위해 라이브러리화 하였다. 이는 곧 오픈하여 발전시켜갈 생각이다. 본 풍선 터뜨리기 구현을 통해 게임 제작과정이 그래픽 정의, 게임 룰 정의 및 이를 연결하기 위한 구현과정이 단순하였으므로 새로운 웹캠 게임을 만드는데 도움이 될 것이라고 생각한다.

앞으로 다양한 웹캠 게임 개발을 함으로써 현재 구현된 개발환경의 단점을 고쳐나갈 계획이며 더 사용하기 쉽게 고쳐나갈 생각이다. 향후 개발된 라이브러리는 웹상으로 공개하여 관련 흥미를 가진 개발자들과 공유할 계획이다.

<http://sapeyes.nameob.net>

Acknowledgement

본 연구는 문화관광부 및 한국문화콘텐츠진흥원의 전남대학교 문화콘텐츠기술연구소(CT) 육성사업의 연구비 지원에 의해 수행되었습니다.

참고문헌

- [1] J. Diniz-Sanches, EyeToy:Groove, 2004
- [2] 김경호, 이철우, "손가락 끝 점을 이용한 손 형상 인식", 전남대학교, 2006