

깊이 정보를 활용한 사람의 피부영역 검출

장 석 우*, 박 영 재**, 김 계 영**

Human Skin Region Detection Utilizing Depth Information

Seok-Woo Jang*, Young-Jae Park**, Gye-Young Kim**

요 약

본 논문에서는 입력되는 3차원의 정지 또는 동적인 입체영상으로부터 색상과 깊이 특징을 결합하여 인간의 피부 색상 영역을 강건하게 추출하는 새로운 방법을 제안한다. 제안된 방법에서는 먼저 스테레오 정합 기법을 이용하여 입력된 좌우 영상으로부터 카메라와 물체 사이의 거리를 나타내는 깊이 특징을 강건하게 추출한다. 그런 다음, 유사한 깊이 특징을 가지는 화소들을 레이블링하고, 레이블링한 영역 중에서 인간의 피부색상 분포를 가지는 영역들을 실제적인 피부색상 영역이라고 판단한다. 실험에서는 2차원 위주의 기존의 피부영역 추출 방법과 제안된 3차원의 특징을 활용한 방법의 성능을 정확도 측면에서 비교 및 평가하였다. 그 결과 제안된 방법은 색상 특징과 깊이 특징을 효과적으로 결합함으로써 기존의 배경 영역에서 부정확하게 검출되는 피부색상 영역의 오류를 상당수 제거하는 효과를 가지며, 이로 인해 전반적으로 보다 정확하게 피부영역을 추출하였다.

▶ Keyword : 피부색상 영역, 3차원 영상, 깊이 정보, 색상 특징

Abstract

In this paper, we suggest a new method of detecting human skin-color regions from three-dimensional static or dynamic stereoscopic images by effectively integrating depth and color features. The suggested method first extracts depth information that represents the distance between a camera and an object from input left and right stereoscopic images through a stereo matching technique. It then performs labeling for pixels with similar depth features and determines the labeled regions having human skin color as actual skin color regions. Our experimental results show that the suggested skin region extraction method outperforms existing skin detection methods in terms of skin-color region extraction accuracy.

▶ Keyword : Skin-Color Region, Three-Dimensional Image, Depth Information, Color Feature

• 제1저자 : 장석우 • 교신저자 : 김계영

• 투고일 : 2012. 02. 15, 심사일 : 2012. 03. 14, 게재확정일 : 2012. 03. 27.

* 안양대학교 디지털미디어학과(Dept. of Digital Media, Anyang University)

** 송실대학교 컴퓨터학부(School of Computing, Soongsil University)

1. 서론

최근 들어, 저가의 디지털 카메라, 그리고 카메라가 내장된 스마트폰이나 테블릿 컴퓨터 등의 빠른 보급으로 인해 정지영상이나 동영상을 촬영하기가 매우 용이해졌으며, 이로 인해 영상 자료들의 양도 기하급수적으로 늘어나고 있다. 더불어, 이런 자료들을 효과적으로 관리하기 위한 다양한 영상 분석 기술에 대한 관심도 꾸준히 증가하고 있다.

이런 영상 분석 분야에서 수행되는 중요한 연구주제들 중의 하나는 주어진 영상에서 특정한 색상을 가지고 있는 영역을 찾는 것이다. 이 주제들 중에서 특히 인간의 피부색상 (skin color)을 나타내는 영역을 정확하게 추출하는 기법은 입력되는 정지영상 또는 동영상으로부터 사람을 검출하는데 필요한 의미 있는 단서를 제공하므로 매우 중요하다. 그리고 이런 피부색상의 추출은 여러 응용분야, 예를 들어 이동 물체의 감지 및 추적, 손 영역 검출을 이용한 제스처 인식, 얼굴 인식, 내용기반의 영상 검색, 유해 콘텐츠 탐지 및 필터링 등의 분야에서 매우 유용하게 사용된다[1,2].

관련 문헌에서 2차원 위주의 피부색상 영역을 추출하는 기존의 여러 방법들을 확인할 수 있다. Lee는 YCbCr 공간에서 특별한 조명 효과에 기인한 색상 치우침을 견딜 수 있는 피부색상 모델들을 이용해 피부 색상을 분할했다[3]. 그런 다음, 여러 가지 특징들을 이용하여 분할된 피부 영역의 진위 여부를 파악했다. Cho는 HSV 칼라공간을 이용하여 어떤 정해진 임계값을 기준으로 영상 전체의 색상 및 밝기에 따라 임계값을 적용적으로 이동시켜 얼굴 영역을 찾는 방법을 제안하였다 [4]. 이 방법은 한 인종(황인종)만의 피부색상 영역 검출을 시도하였으므로, 다른 인종(백인 또는 흑인)의 피부색상이 존재할 경우에는 문제가 발생하였다. Hsu는 EyeMap과 MouthMap을 이용하여 얼굴을 검출하는 새로운 방법을 제안하였다[5]. 이 방법에서 얼굴의 구성요소는 피부색상 모델로부터 유도된 특징맵을 사용하여 검출하였으며, 검출된 얼굴 구성요소들의 기하학적인 특징을 기반으로 서로의 관계가 정의되었다. 그러나 이 방법은 간단한 기하학적인 관계만을 사용하였기 때문에 유연성에 여러 가지 제한이 있었다. Fang은 얼굴 검출을 위한 새로운 칼라 히스토그램 기반의 방법을 제안하였다[6]. 이 방법에서는 얼굴의 서로 다른 영역에 대한 칼라 히스토그램을 연결하여 이 영역들 사이의 공간적인 관계를 설정하는 벡터를 형성하였고, 이 벡터를 이용하여 효과적으로 얼굴을 검출하는 알고리즘을 개발하였다. 위에서 언급한 방법들 이외의 다른 방법들도 계속해서 문헌에 소개되고 있다[7].

이런 기존의 알고리즘들은 피부 색상 영역 추출 알고리즘에 여전히 많은 문제를 내포하고 있다. 즉, 인간의 피부 색상은 개인 간의 차이나 인종(race) 사이의 차이 등으로 인해 촬영된 영상에 포함된 피부 색상은 기본적으로 동일하지 않다. 뿐만 아니라, 색조화장, 분장, 사용된 카메라, 조명의 변화 등의 여러 가지 다른 환경조건으로 인해서 입력영상 내에 존재하는 피부 색상은 조금씩 다르다. 특히, 색상, 텍스처, 모양, 그리고 기하학적인 관계 등의 2차원적인 특징만을 사용하여 피부색상을 추출하는 것은 제한점이 존재한다.

따라서 본 논문에서는 2차원 위주의 피부색상 추출 방법에서 가장 많이 사용하는 색상 특징과 3차원적인 거리(depth) 정보를 결합하여 보다 정확하게 피부색상 영역을 검출하는 방법을 소개한다. 따라서 본 논문은 기본적으로 3차원의 입체영상을 입력 받는다. 일반적으로, 3차원의 입체 디스플레이는 다시점 스테레오스코픽(stereoscopic) 비전 기술을 적용하여 2차원 영상에 깊이 정보를 부가하고, 이 깊이 정보로 인하여 시청자가 마치 영상이 제작되고 있는 현장에 있는 것 같은 생동감 및 현실감을 느낄 수 있게 하는 차세대 신기술이다.

그림 1은 본 논문에서 제안한 피부색상 영역 추출 방법의 전체적인 개요도를 블록 다이어그램의 형태로 보여준다. 제안된 방법은 먼저 입력된 칼라 영상을 그레이 영상으로 변환한 후, 스테레오 정합 기법을 이용하여 좌우 영상으로부터 카메라와 물체 사이의 거리를 나타내는 3차원의 깊이 특징을 추출한다. 그런 다음, 유사한 깊이 특징을 가지는 화소들을 레이블링하고, 레이블링한 영역 중에서 2차원의 피부색상 분포를 가지는 영역들만을 실제적인 피부색상 영역이라고 판단한다.

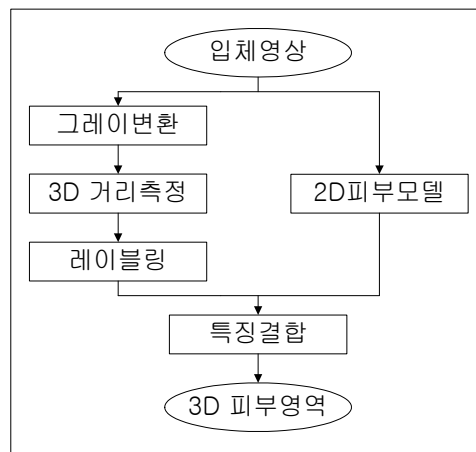


그림 1. 전체 개요도
Fig. 1. Overall Flow

1장에서는 본 연구를 수행하게 된 동기 및 배경, 그리고 전체적인 개요를 기술하였다. 2장에서는 스테레오 정합 알고리즘을 통해 3차원 입체 영상으로부터 깊이정보를 추출하는 방법에 대해 기술한다. 3장에서는 피부색상 분포 모델을 이용하여 2차원적인 피부색상 영역을 추출하는 방법을 설명한다. 4장에서는 색상 기반의 2차원적인 피부영역과 깊이정보를 결합하여 최종적인 3차원의 피부영역을 검출하는 방법을 설명한다. 그리고 5장에서는 본 논문에서 제안된 방법의 성능을 평가하기 위해 수행한 실험결과를 보이며, 마지막으로 6장에서는 결론 및 향후 연구방향을 제시한다.

II. 깊이정보 획득

3차원의 깊이정보 추출을 위한 스테레오(stereo) 정합은 컴퓨터 비전 분야의 고전적인 문제 중의 하나이며, 로봇 네비게이션, 3차원 모델링, 영상 기반 렌더링 등의 많은 응용 분야에서 유용하게 사용된다. 스테레오 정합의 목적은 동일한 장면에 대한 두 개 이상의 영상이 주어질 때 참조(reference) 영상에 대한 변이(disparity) 맵을 계산하는 것인데, 여기서 변이는 대응하는 두 개 화소 사이의 위치의 차이를 의미한다. 그리고 변이 맵을 계산하기 위해서는 각 화소에 대한 대응성(correspondence) 문제를 해결해야 한다[8].

일반적으로, 양안 스테레오인 경우 두 개의 입력영상은 보정된 카메라로 촬영되었고 교정(rectification)되었다고 가정하면 에피폴라 라인들은 수평이 된다. 그러나 이런 제약사항에도 불구하고 스테레오 정합의 ill-posed 특성 때문에 변이를 정확하게 결정하는 것은 여전히 어려운 문제로 남아 있다. 특히, 중첩되거나 텍스처가 부족한 영역에서는 더욱 어렵다[9].

보통, 스테레오 정합 알고리즘은 크게 지역적인 방법과 전역적인 방법으로 분류될 수 있다. 지역적인 방법은 대응점 검색의 구분력을 증가시키기 위해서 정합을 위한 일정한 크기의 윈도우를 사용한다. 이 방법에서 대응점은 지역적인 윈도우 안에서 다양한 정합 척도를 사용하여 명암값을 비교함으로써 추출될 수 있다. 이 방법은 매우 빠르나 지역적으로 중첩되거나 텍스처가 부족한 영역, 그리고 변이가 불연속적인 경계 부분에서는 정확도가 저하된다는 단점이 있다.

전역적인 방법은 스테레오 정합의 ill-posed 문제를 해결하기 위해서 완만성(smoothness) 제한조건을 가진 에너지 최소화 알고리즘을 사용한다. 따라서 이 방법은 텍스처가 부족한 영역에서도 대응점을 찾을 수 있다. 그러나 완만성 제한조건으로 인해 변이정보의 불연속성을 제거시킨다. 따라서 불연속성을 보존하는 완만성 제한조건을 사용하는 스테레오 정

합 알고리즘이 소개되었다. 전역적인 방법은 다양한 최소화 기법에 의해 대응점을 추출하는데, 최근에는 그래프 컷(graph cuts)과 신뢰 전파(belief propagation) 기반의 알고리즘들이 우수한 성능 때문에 많이 사용되고 있다. 그러나 상당수의 전역적인 방법들은 중첩 문제를 명확히 고려하지 않고 있다.

본 논문에서는 그래프 컷(graph cuts) 기반의 스테레오 매칭 기법을 사용하여 거리를 측정한다. 이 방법의 경우 색상 정보는 고려하지 않고 그레이 영상을 사용하므로 먼저 식 (1)을 이용하여 입력된 스테레오 칼라 영상을 그레이 영상으로 변환한다. 식 (1)에서 r, g, b 는 입력영상의 r, g, b 칼라값을 의미하고, $gray$ 는 흑백으로 변환시킨 화소의 명암값을 나타낸다.

$$gray = R \times 0.299 + G \times 0.587 + B \times 0.114 \quad (1)$$

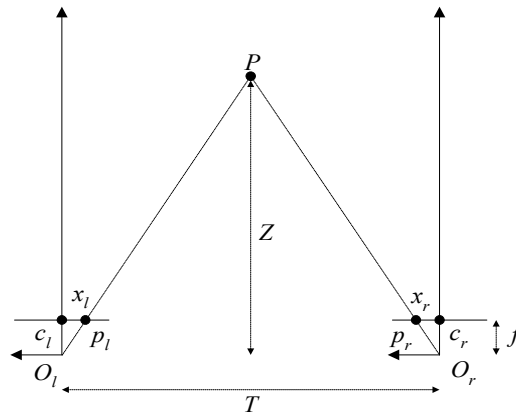
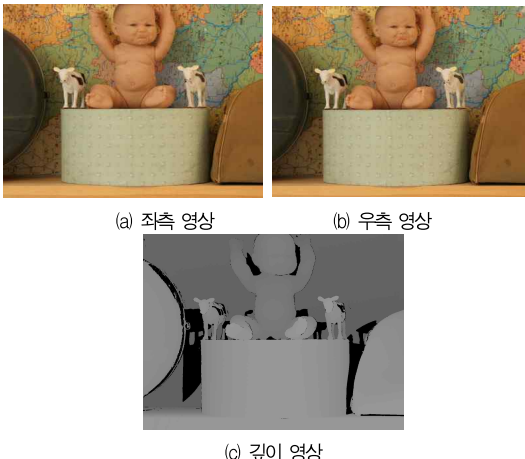


그림 2 스테레오 정합
Fig. 2. Stereo Matching

입력 영상을 그레이 영상으로 변환한 후에는 좌우 영상 간의 정합을 통해 깊이정보를 추출한다. 좌우 영상으로부터 깊이 정보를 추출하는 기본적인 개념은 그림 2와 같다. 그림 2와 같이 P는 실제계의 한 점, x_l 과 x_r 은 P가 좌우 영상에 맺힌 x 좌표, f 는 카메라의 초점거리, T 는 카메라의 baseline, 그리고 Z 는 추출하고자 하는 깊이 값이라고 하자. 그러면 두 개의 삼각형 (P, P, P)과 (O_l, P, O_r)은 닮은 형태이므로 식 (2)가 성립하고, 식 (2)를 Z 에 대해 전개하면 식 (3)과 같이 정리가 되어 깊이 정보인 Z 를 추출할 수 있게 된다.

$$\frac{T + x_l - x_r}{Z - f} = \frac{T}{Z} \quad (2)$$

$$Z = f \times \frac{T}{d} \quad (3)$$



(c) 깊이 영상

그림 3. 깊이정보 추출
Fig. 3. Extracting Depth Information

본 논문에서는 깊이 정보를 추출한 후에 깊이 값이 유사한 화소들을 레이블링하여 영역 단위로 분할한다. 그림 3은 스테레오 영상으로부터 깊이를 추출한 예를 보여준다. 그림 3의

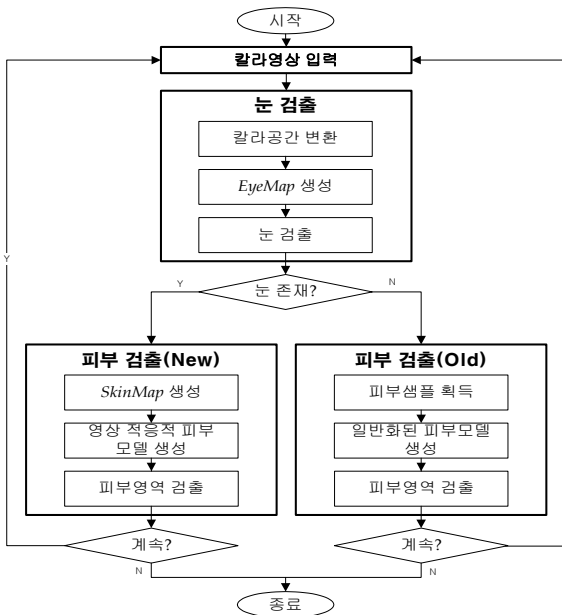


그림 4. 적응적인 피부검출의 블록다이어그램
Fig. 4. Block Diagram of Adaptive Skin Detection

(a)와 (b)는 스테레오 영상의 왼쪽과 오른쪽 영상을 나타내고,

(c)는 좌우 입력영상으로부터 스테레오 정합 알고리즘을 적용하여 깊이 정보를 추출한 결과를 보여준다.

III. 색상 기반의 피부영역 분할

본 논문에서는 색상을 이용하여 영상으로부터 피부색상 영역을 획득한 후 이전 단계에서 추출한 깊이정보와 결합하여 최종적인 피부색상 영역을 추출한다.

일반적으로, 기존의 대부분의 피부색상 추출 방법은 사전 학습을 통해 미리 정의된 피부색상 분포 모델을 사용하지만 여전히 본질적인 문제를 가지고 있다. 즉, 개개인의 고유한 피부색상의 차이 및 인종 간의 피부색상의 차이로 인해 사람마다 피부색상 자체가 동일하지 않다. 또한, 특수 분장이나 색조화장, 촬영 시 사용하는 광학기기, 조명 효과 등으로 인해 획득된 영상의 피부색상 영역이 동일하지 않다. 따라서 사전에 정의된 피부모델을 이용하는 기존의 알고리즘은 위에서 언급한 여러 가지의 변화를 극복하기가 어렵다. 이런 문제를 해결하기 위한 최적의 솔루션은 서로 다른 영상이 입력될 때마다 사람별로 고유한 피부색상 샘플을 신뢰성 있게 선택하여 입력영상에 적응적인 피부모델을 생성한 후, 이 피부모델을 이용해서 입력영상의 피부색상 영역을 추출해야 한다.

이를 위해, 본 연구에서는 입력영상으로부터 얼굴의 주요 구성요소인 눈을 찾고, 찾았진 눈 영역 주변의 피부색상 샘플을 신뢰성 있게 추출하여 입력영상에 최적으로 적합한 피부모델을 생성한 후, 이 모델을 이용하여 입력영상으로부터 피부영역을 추출하는데, 그 단계는 다음과 같다.

- 입력된 영상의 칼라공간을 식 (4)와 같이 피부색상 추출에 가장 적합하다고 알려진 YCbCr 공간으로 변환한다.

$$Y = 0.299R + 0.587G + 0.114B \quad (4)$$

$$C_b = R - Y$$

$$C_r = B - Y$$

- 색상과 명암 기반의 맵을 혼합한 EyeMap[5]를 이용하여 양쪽 눈 영역을 검출한다.
- 추출된 눈 영역에 해당하는 최소 포함 사각형(MER)의 5배 확장된 영역에 위치한 샘플을 선정하고, 선정된 샘플로부터 식 (5)와 같은 피부 맵(skin map)을 생성한다. 식 (5)에서 \hat{C}_r 과 \hat{C}_b 는 각각 일반적인 피부 색상의 Cr과 Cb 값을 나타낸다. 그리고 피부 맵은 0에서 255사

이의 값의 범위를 가지는데, 선택된 샘플이 평균적인 피부 색상에 근접할수록 255와 가까운 값을 가진다.

$$SkinMap = 255 - \frac{|\widehat{C}_r - C_r| + |\widehat{C}_b - C_b|}{2} \quad (4)$$

- 피부 맵을 이진화한 후 피부색상이라고 판단된 피부 맵만을 이용하여 피부색상 모델을 생성한다.
- 위의 과정을 통해 생성된 피부색상 모델을 이용해서 전체 영상으로부터 피부색상 영역을 강건하게 추출한다.

그림 4는 현재까지 기술한 입력영상에 적응적인 피부모델을 이용하여 피부색상 영역을 강건하게 추출하는 방법의 전체적인 개요도를 보여준다. 그림 4에서 확인할 수 있듯이 입력 영상에서 사람의 눈이 존재하지 않거나 사람이 눈을 감아서 눈을 검출할 수 없는 경우에는 기존의 방법에서처럼 사전에 정의된 피부 모델을 사용하여 피부색상 영역을 추출하므로 제안된 방법의 실행 안정성에는 전혀 문제가 없다.

IV. 3차원 피부영역 추출

본 논문에서는 2장에서 스테레오 정합을 통해 추출한 거리 영상과 3장에서 검출한 색상 기반의 피부영역을 결합하여 입력영상에서 3차원적인 피부색상 영역을 검출한다. 이를 위해, 먼저 거리 영상으로부터 동일한 레벨을 가지는 거리 영역을 레이블링하는데, 일반적으로 동일한 객체에서는 여러 가지 레벨의 거리가 측정된다. 그런 다음, 식 (5)와 같이 레이블링된 거리 영상과 입력 영상에 적응적인 피부색상 모델을 사용하여 추출한 이진화된 피부색상 영상을 AND 연산하여 2차원 피부영역에 대응하는 거리영상만을 선택한다.

$$I_{depth}(x,y)' = I_{depth}(x,y) \text{ AND } I_{bi_skin}(x,y) \quad (5)$$

다시 말해, 위의 수식 (5)는 아래와 같이 기술할 수 있는데, $I_{depth}(x,y)$ 는 깊이 영상을 의미하고, $I_{bi_skin}(x,y)$ 는 색상 기반의 이진화된 피부영역을 의미하며, $I_{depth}'(x,y)$ 는 식 (5)의 AND 연산을 통해 획득한 2차원 피부영역에 대응하는 거리영상을 의미한다.

IF $I_{bi_skin}(x,y) == 0$ THEN

$$I_{depth}(x,y)' = I_{depth}(x,y)$$

ELSE

$$I_{depth}(x,y)' = 0$$

그러면 $I_{depth}'(x,y)$ 와 색상 기반으로 추출한 피부영상인 $I_{skincolor}(x,y)$ 로부터 깊이 특징에 대한 레이블별로 피부색상 영역의 후보군을 추출할 수 있다. 그런 다음, 최종적인 3차원 피부영역을 선택하기 위해서 거리 레이블별 피부영역 후보군의 질감 복잡도(texture smoothness)를 평가한다. 일반적으로, 피부색상 영역은 질감이 거칠지 않고 완만하다는 특성을 가지고 있으므로 3차원 피부영역 후보군의 질감 복잡도를 측정하여 질감이 거친 영역들은 제거하고 완만한 영역들만을 실제적인 피부영역으로 판단한다. 본 논문에서는 질감 복잡도 측정을 위해서 해당하는 영역 내에 소벨(Sobel) 에지 연산자를 적용하여 에지의 정도를 추출한다. 즉, 해당 영역 내에 에지가 많이 존재할수록 질감이 거칠다는 것을 나타내며, 에지가 적게 존재할수록 질감이 완만하다는 의미이다. 보통, 소벨 에지 연산자는 에지를 검출하는 미분 연산자로 x축과 y축으로 각각 한 번씩 미분을 수행하는데, 소벨 연산자에 해당하는 회전(convolution) 마스크는 다음의 그림 5와 같다.

-1	-2	-1
0	0	0
1	2	1

(a) 수평 마스크

-1	0	1
-2	0	2
-1	0	1

(b) 수직 마스크

그림 5. 소벨 연산자의 마스크
Fig. 5. Mask of Sobel Operator

깊이 영상과 이진화된 피부영상과의 AND 연산을 통해 추

출한 i 번째 후보 피부영역 R_i 에 대한 질감 복잡도 $T(R_i)$ 는 식 (6)과 같이 추출할 수 있다. 식 (6)에서 $I_{gray}(x,y)$ 는 x 와 y 위치에서의 명암값을 의미하고, $w_h(x,y)$ 와 $w_v(x,y)$ 는 소벨 에지 연산자의 수평과 수직 마스크를 의미한다. 그리고 $E(x,y)$ 는 x 와 y 위치에서의 에지 정도를 나타내고, $N(R_i)$ 는 영역 R_i 에 속한 화소의 개수를 의미한다.

$$T(R_i) = \frac{1}{N(R_i)} \sum_{x \in R_i} \sum_{y \in R_i} E(x,y) \quad (6)$$

$$E(x,y) = \sum_{j=-1}^1 \sum_{i=-1}^1 w_h(i,j) \times I_{gray}(x+i,y+j) + \sum_{j=-1}^1 \sum_{i=-1}^1 w_v(i,j) \times I_{gray}(x+i,y+j)$$

제안된 방법에서는 식 (6)을 이용하여 각 후보 피부영역에 대한 질감 복잡도를 계산한 후 적절한 임계치 범위 내에 포함된 후보영역만을 실제 피부영역이라고 판단한다. 본 논문에서는 다양한 입체 영상을 이용한 실험을 통해서 질감 복잡도에 대한 임계치를 결정하였다.

V. 실험결과

본 논문의 실험을 위하여 사용한 컴퓨터는 인텔 코어 i7-2600의 3.40GHz CPU와 8GB의 메모리를 사용하였고, 운영체제는 마이크로소프트사의 Windows 7을 사용하였다. 그리고 응용 소프트웨어의 구현을 위한 컴파일러로는 마이크로소프트사의 Visual C++ 2008을 이용하여 제안된 피부색상 영역 추출 알고리즘을 구현하였다. 또한, 실험에 사용할 영상 데이터베이스의 구축을 위해서 특정한 제약조건이 주어지지 않은 일반적인 환경에서 촬영한 다양한 3차원의 정지 및 동영상상을 수집하여 활용하였다.

그림 6은 제안된 방법을 이용하여 입력된 영상으로부터 피부영역을 검출한 예를 보여준다. 그림 6 (a)는 입력된 스테레오 영상 중에서 좌측 영상을 보여주고, 그림 6 (b)는 입력영상으로부터 스테레오 정합을 통해 측정된 거리영상을 보여준다. 그리고 그림 6 (c)는 2차원 피부색상 모델을 이용하여 피부영역을 추출한 이진 결과영상상을 보여주며, 그림 6 (d)는 거리영상과 색상 기반의 피부영역을 결합하여 최종적으로 검출한 3차원적인 피부영역을 나타낸다.

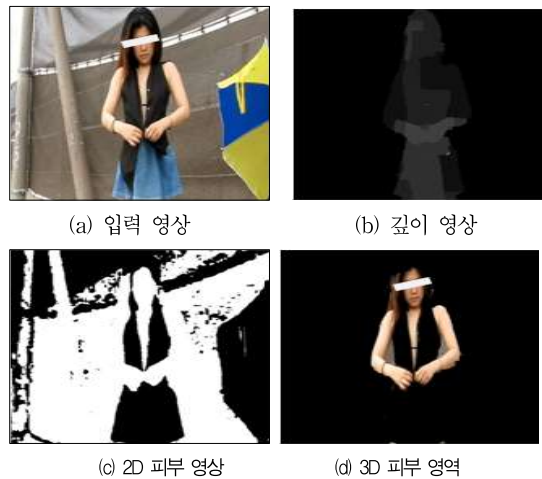


그림 6. 피부영역 추출
Fig. 6. Extracting Skin Regions

그림 6의 (c)에서 확인할 수 있듯이 2차원적인 피부색상 모델을 이용하여 추출한 피부색상 영역은 많은 오 검출을 포함하고 있다. 즉, 사람의 피부영역이 아닌 배경 영역에서 피부색상과 유사한 영역이 존재할 경우 피부영역으로 오 검출한다. 그러나 2차원적인 피부색상 특징에 깊이 특징을 결합하여 검출한 3차원적인 피부색상 영역은 보다 정확하게 피부영역만을 추출함을 확인할 수 있다.

제안된 방법에서는 입력영상에서 텍스처가 매우 빈약하거나 없는 영역에서는 3차원의 깊이 정보를 추출하기가 어려우므로 피부색상 추출에 어려움이 존재할 수 있다.

제안된 피부영역 검출 알고리즘의 성능을 정량적으로 평가하기 위해서 식 (7)과 같은 오차의 제곱 평균 제곱근(RMSE) 척도를 정의하였다. RMSE는 영상 화질의 일반적인 면에서도 구체적인 관점을 다루는 척도로 실제적인 값과 측정값의 차이를 측정하기 위해 종종 사용되는데, 정확성 측정에 있어 좋은 척도라고 알려져 있다[10].

$$RMSE = \sqrt{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N |OB(i,j) - RB(i,j)|^2} \quad (7)$$

식 (7)에서 M과 N은 영상의 가로와 세로의 길이를 나타내고, i 와 j 는 영상의 위치를 나타내는 행과 열의 인덱스이다. 그리고 $OB(i,j)$ 는 그라운드 트루스(ground truth) 이진영상을 의미하며, $RB(i,j)$ 는 피부영역을 검출한 이진결과영상을 나타낸다.

성능을 보다 정확히 평가하기 위해서 원래의 테스트 영상

을 피부 영역과 비피부 영역으로 수작업으로 변환한 후 결과 영상과 비교 분석하였다. 그리고 피부영역 추출에는 두 가지 종류의 에러, 즉 false positive와 false negative가 존재하므로 다음 식 (8)과 같은 척도 overall을 정확도 평가에 사용한다[11].

$$overall = recall \times \left(2 - \frac{1}{precision} \right) \quad (8)$$

$$precision = \frac{|True\ Positive|}{|True\ Positive| + |False\ Positive|}$$

$$recall = \frac{|True\ Positive|}{|False\ Negative| + |True\ Positive|}$$

표 1과 그림 7은 기존의 2차원 피부영역 검출 방법과 본 논문에서 제안된 3차원의 피부영역 검출 방법을 비교 평가한 성능의 결과를 표와 그림으로 각각 보여준다.

표 1. 성능비교표
Table 1. Table of Performance Comparison

성능	기존 방법	제안 방법
Precision	0.986000	0.947333
Recall	0.702947	0.922128
Overall	0.692966	0.870863

그림 7에서 확인할 수 있듯이 제안된 방법은 2차원 피부영역 검출 결과를 기반으로 최종적인 피부영역을 검출하기 때문에 precision은 제안된 방법이 조금 낮은 결과를 얻었다. 그러나 배경영역에 피부색상과 유사한 색상이 분포되어 있는 경우 2차원 피부영역 검출은 오 검출이 많이 발생하므로 다소 낮은 recall이 나온다. 이에 비해, 제안된 방법은 거리영상을 기준으로 영역을 분리하여 피부영역을 검출하기 때문에 오 검출율을 현저히 낮출 수 있으므로 precision과 recall 모두 좋은 결과를 얻을 수 있었다.

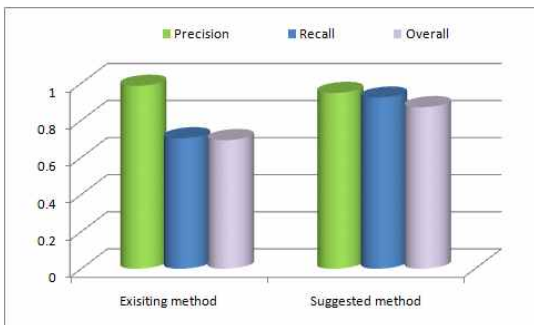


그림 7. 성능 비교 차트
Fig. 7. Performance Comparison Chart

VI. 결론 및 향후연구

본 논문에서는 깊이 특징과 색상 특징을 결합하여 입력되는 3차원의 동영상에 존재하는 인간의 피부색상 영역들을 강건하게 추출하는 방법을 제안하였다. 제안된 방법에서는 먼저 스테레오 정합 기법을 이용하여 입력된 좌우 영상으로부터 카메라와 물체 사이의 거리인 3차원의 깊이 정보를 추출하였다. 그런 다음, 유사한 깊이 특징을 가지는 영역들을 군집화하고, 군집화된 영역 중에서 2차원의 피부색상 분포를 가지는 영역들을 실제적인 피부색상 영역이라고 판단하였다.

향후에는 보다 다양한 종류의 실험영상을 가지고 제안된 방법의 성능을 비교 분석할 예정이며, 스테레오 정합을 통해 3차원의 깊이 정보를 추출하는 알고리즘에 대한 정확성을 최대한 향상시킬 예정이다. 또한, 칼라 이외의 특징들을 좀 더 추가하여 피부색상 영역 추출의 정확도를 향상시킬 계획이다.

참고문헌

- [1] A. Drosou, D. Ioannidis, K. Moustakas, and D. Tzovaras, "Spatiotemporal Analysis of Human Activities for Biometric Authentication," *Computer Vision and Image Understanding*, Vol. 116, No. 3, pp. 411-421, 2012.
- [2] S.-W. Jang, Y.-J. Park, G.-Y. Kim, and S.-Y. Lee, "Skin Region Extraction Combining 3D Depth and Color Features," In *Proc. of the Winter Conference on the Korea Society of Computer and Information*, Vol. 20, No. 1, pp. 201-204, 2012.
- [3] J.-S. Lee, Y.-M. Kuo, P.-C. Chung, and E.-L. Chen, "Naked Image Detection based on Adaptive and Extensible Skin Color Model," *Pattern Recognition*, Vol. 40, No. 8, pp. 2261-2270, 2007.
- [4] K.-M. Cho, J.-H. Jang, and K.-S. Hong, "Adaptive Skin-Color Filter," *Pattern Recognition*, Vol. 34, No. 5, pp. 1067-1073, 2001.

[5] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face Detection in Color Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, pp. 696-706, 2002.

[6] J. Fang and G. Qiu, "A color histogram-based approach to human face detection," In Proc. of the International Conference on Visual Information Engineering, pp. 133- 136, 2003.

[7] K. M. Lee, "Component-based Face Detection and Verification," Pattern Recognition Letters, Vol. 29, No. 3, pp. 200-214, 2008.

[8] N. Baha and S. Larabi, "Accurate Real-Time Neural Disparity MAP Estimation with FPGA," Pattern Recognition, Vol. 45, No. 3, pp. 1195 - 1204, 2012.

[9] Y.-M. Paik, H.-J. Choi, Y.-H. Seo, and D.-W. Kim, "A Study on the Outlier Improvement Method Using Cost Function," In Proc. of the Fall Conf. of the Korean Society of Broadcasting Engineers, pp. 269-272, 2009.

[10] G.-J. Liu, X.-L. Tang, H.-D. Cheng, J.-H. Huang, and J.-F. Liu, "A Novel Approach for Tracking High Speed Skaters in Sports Using a Panning Camera," Pattern Recognition, Vol. 42, No. 11, pp. 2922-2935, 2009.

[11] H.-H. Do, S. Melnik, and E. Rahm, "Comparison of Schema Matching Evaluations," Lecture Notes in Computer Science, Vol. 2593, pp.221-237, 2003.



박 영 재
 2008년 2월 : 숭실대학교 컴퓨터학과
 (공학석사)
 2008년 3월 - 현재 : 숭실대학교 컴퓨터학과
 박사과정
 관심분야 : 컴퓨터 비전, 영상처리, 패턴인식
 등
 Email : plylj1@ssu.ac.kr



김 계 영
 1996년 2월 : 숭실대학교대학원 컴퓨터학과
 (공학박사)
 2001년 3월 - 현재 : 숭실대학교 컴퓨터학과
 교수
 관심분야 : 컴퓨터 비전, 형태인식, 생체인식,
 증강현실, 신호처리 등
 Email : gykim11@ssu.ac.kr

저 자 소 개



장 석 우
 2000년 8월 : 숭실대학교대학원 컴퓨터학과
 (공학박사)
 2009년 3월 - 현재 : 안양대학교 디지털미
 디어학과 교수
 관심분야 : 로봇비전, 증강현실, HCI, 게임,
 비디오 색인 및 검색, 이라닝 등
 Email : swjang@anyang.ac.kr