조명 및 법선벡터 지도 추정을 이용한 사실적인 가상 화장 시스템

양명현° 신현준✉

아주대학교 라이프미디어융합전공

{shiningy, joony}@ajou.ac.kr

Virtual Make-up System Using Light and Normal Map Approximation

Myung Hyun Yang° Hyun Joon Shin✉

Life Media Department, Ajou University

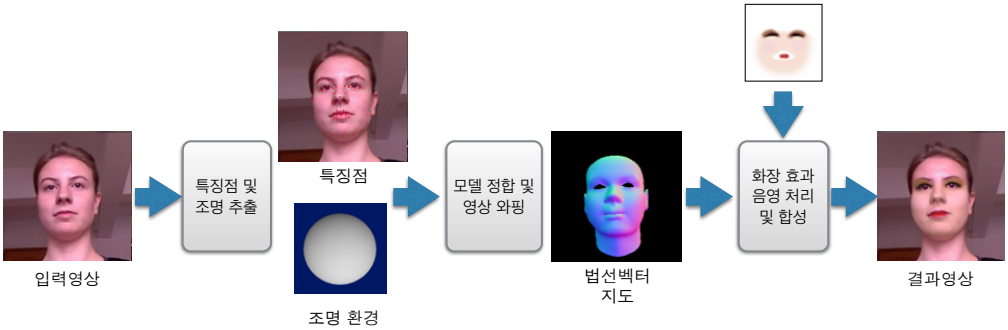

Figure 1: 가상 화장 시스템의 개요

요 약

본 논문은 입력 영상에 자연스러운 가상 화장 효과를 빠르게 적용하기 위한 방법을 다룬다. 특히 얼굴의 굴곡과 조명에 따라 나타나는 자연스러운 음영 효과를 가상 화장에 적용함으로써 기존의 방법에 비해 자연스럽고 다양한 화장 효과를 얻기 위한 방법을 제안한다. 가상 화장 효과에 음영을 적용하기 위해서는 입력 영상의 조명 환경과 입력 영상의 얼굴 면의 각 화소에 대한 법선 벡터가 필요하다. 기존에 영상에서 조명 환경 및 법선 벡터를 생성하는 과정이 비교적 많은 계산량을 필요로 하기 때문에 본 논문에서는 얼굴의 특징점을 찾는 단계적 자세 회귀 방법에 조명 환경을 근사하는 과정을 통합하고, 3차원 표준 모델을 변환, 와핑하여 법선 벡터를 계산하는 과정을 통하여 빠르게 화장 효과를 합성하기 위한 방법을 제안한다. 실험을 통하여 본 논문에서 제안하는 방법으로 근사한 조명 환경 및 법선 벡터는 시각적으로 자연스러운 가상 화장 효과를 표현할 수 있는 것을 보였다.

Abstract

In this paper, we introduce a method to synthesize realistic make-up effects on input images efficiently. In particular, we focus on shading on the make-up effects due to the lighting and face curvature. By doing this, we can synthesize a wider range of effects realistically than the previous methods. To do this, the information about lighting information together with the normal vectors on all pixels over the face region in the input image. Since the previous methods that compute lighting information and normal vectors require relatively heavy computation cost, we introduce an approach to approximate lighting information using cascade pose regression process and normal vectors by transforming, rendering, and warping a standard 3D face model. The proposed method consumes much less computation time than the previous methods. In our experiment, we show the proposed approximation technique can produce naturally looking virtual make-up effects.

키워드: 가상 화장, 조명 추정, 법선벡터 지도

Keywords: Virtual makeup, light approximation, normal map

1. 서론

본 논문에서는 실제 얼굴 영상에 가상적으로 화장(makeup) 효과를 적용한 자연스러운 영상을 빠르게 생성하기 위한 방법을 제안한다. 최근 휴대용 장치(mobile device)에 전면 카메라가 장착되어 촬영자의 영상을 실시간으로 처리하여 다양한 효과를 주기 위한 응용들이 다수 소개되고 있다. 전면 카메라의 방향이 촬영자의 얼굴이라는 점에 착안하여 입력 영상의 얼굴에 효과를 주는 응용, 특히 얼굴면에 잡티 등을 제거하기 위한 응용이나, 가상적으로 화장을 적용하기 위한 응용들이 소개되고 있다. 이중 가상 화장 응용은 화장법을 교육하는 용도, 화장품 제조사에서 자사의 제품의 효과를 미리 확인하기 위한 용도 등에 효과적으로 활용될 수 있기 때문에 다양한 형태의 응용에 적용되고 있다.

대부분의 가상 화장 응용에서는 먼저 입력 영상으로부터 얼굴의 주요 특징점의 위치를 추출하고, 미리 생성된 화장이 적용된 영상을 적절한 위치로 변환/변형 하여 적용시키는 과정을 통하여 이루어진다. 하지만, 이러한 방법은 얼굴의 굴곡에 따라서 생기는 얼굴의 각 부분의 음영(shading)을 적절하게 반영하지 못하고 있다는 한계가 있다. 즉, 입력 영상에 화장 영상을 합성하는 과정에서 화장 영상에 얼굴 곡면에 의한 음영이 적용되지 않기 때문에 화장 영상은 평면적으로 나타난다. 따라서 화장이 얼굴에 밀착되어있는 것이 아니라 들떠 보이는 문제점을 가진다. 일반적인 상용 응용에서는 화장 영상의 투명도를 높여 입력 영상 자체의 음영이 드러나도록 하는 방법이나, 음영 차가 심하게 나타나지 않도록 적은 부위에만 화장을 적용하는 방법 등을 통하여 이와 같은 문제를 해결하고 있다. 하지만, 이런 방법은 블러쉬(blush) 등의 색조 화장에는 적합하지만 파우더 등의 넓은 영역을 커버하면서 투명도가 낮아야 효과가 분명한 화장에는 적용하기 어렵다. 따라서 본 연구에서는 화장 영상을 합성하는 과정에서 화장 영상 자체에 음영 효과를 적용하고 합성함으로써 다양한 화장 기법을 자연스럽게 적용할 수 있는 방법을 제안한다.

화장에 음영을 적용하여 합성하기 위해서는 입력 영상의 각 화소별로 음영을 계산하여야 한다. 본 논문에서는 음영을 계산하기 위해 조명(lighting) 기법을 활용한다. 조명 기법을 통하여 음영을 계산하는 데는 광원의 위치, 방향 등을 포함하는 조명 환경과 더불어 화장 영상이 적용될 각 화소별로 얼굴 면의 법선 벡터가 필요하다. 따라서 본 연구에서는 입력 영상을 분석하여 입력 영상의 조명 환경과 얼굴 부분의 각 픽셀의 법선 벡터를 계산하고 이들을 기반으로 화장 영상에 조명 효과를 적용하고 입력 영상에 적용하므로써 자연스러운 화장 효과를 얻고자 한다.

얼굴 영상으로부터 조명과 법선 벡터를 계산하는 방법이 Kemelmacher-Shlizerman과 Basri에 의해 제안되었다[1]. 이 방법은 얼굴 영상에 새로운 조명을 비추었을 때의 영상을 계산하기 위한 방법으로 역시 입력 영상에서 조명과 얼굴 부분의 각 화소별 법선 벡터를 계산하는 과정이 필요하다. 따라서 Kemelmacher-Shlizerman과 Basri는 입력 영상의 얼굴 부분의 각 화소 정보에 대한 최소 자승 법을 이용하여 조명을 근사하고, 수치적 최적화 기법을 이용하여 법선 벡터들을 계산하는 방법을 제안하였다. 하지만 휴대용 장비 등의 제한적인 속도를 가지는 환경에서는 기존의 방법의 계산 복잡도와 계산량이 매우 크다는 문제점이 있다. 따라서 본 논문에서는 정확한 조명과 법선 벡터를 요구하는 재조명 문제와는 다르게 본 시스템에서는 비교적 정확도가 낮은 조명과 법선 벡터로도 충분한 효과를 얻을 수 있다는 점에 착안하여 이들을 빠르게 계산하기 위한 방법을 제안한다.

본 논문의 기여점은 다음과 같다.

- 가상 화장에 입력된 영상 내의 얼굴의 굴곡과 조명 환경을 고려하여 자연스러운 화장 적용 결과를 얻기 위한 방법을 제안한다.
- 빠르게 얼굴 영상으로 부터 촬영 당시의 얼굴 조명을 근사하는 방법을 제안한다.
- 얼굴 영상에 대응되는 법선 벡터 지도를 빠르게 근사하는 방법을 제안한다.

본 논문은 다음과 같이 구성되어 있다. 먼저 2절에서는 본 논문에서 제안하는 가상 화장 시스템의 개괄과 각 구성 요소의 기술적 세부 사항을 소개한다. 본 논문에서 제안하는 시스템 및 요소 기술에 대한 실험 결과를 3절에서 보이고, 마지막으로 4절에서는 제안한 시스템의 장단점과 향후 연구에 관하여 논한다.

2. 가상화장 시스템

본 논문에서 제안하는 가상 화장 시스템의 가장 중요한 특징은 입력된 얼굴에 화장을 입히는 과정에서 얼굴의 음영을 고려하여 화장을 추가한다는 것이다. 즉, 2차원 반투명 텍스처로 저장되어 있는 화장을 얼굴 위치에 합성하는 일반적인 방법과 다르게, 화장 텍스처를 중첩하는 과정에서 얼굴면과 조명 상태에 따라서 텍스처에 세부적인 음영을 적용하여 보다 자연스러운 화장을 적용한다. 따라서 본 논문에서 제안하는 방법은 화장 텍스처를 위치를 조정하고 변형하기 위해 얼굴의 중요 특징점을 분석하는 과정, 입력 영상으로 부터 조명 상태를 얻는 과정, 얼굴면의 굴곡을 계산하는 과정, 마지막으로 얼굴 굴곡과 조명 상태에 따라서 얼굴의 텍스처에 음영을 적용하여 입력 영상에 합성하는 과정으로 이루어진다.

본 논문에서 조명을 획득하고 얼굴의 굴곡을 얻는 방법은 Kemelmacher-Shlizerman과 Basri가 제안한 얼굴 재조명(relighting)에 관한 연구에서 상당 부분 착안하였다. Kemelmacher-Shlizerman과 Basri가 제안한 방법은 표준 3차원 얼굴 모델을 사용한다. 이 방법에서는 먼저 표준 3차원 얼굴 모델을 특징점을 기반으로 얼굴 영상에 정합한다. 정합된 3차원 얼굴 모델의

법선 벡터와 입력 영상의 음영을 분석하여 조명을 근사한다. 이때 조명은 구면 조화 함수(spherical harmonic function)의 계수(coefficient)[2]로 근사하고, 실제 계수는 얼굴 영역에 해당되는 모든 화소에 대한 최소 자승 해(least squares solution)를 계산하는 과정을 얻는다. 조명이 근사되면 영상의 음영 정보와 3차원 모델의 법선 벡터들, 근사된 조명을 통하여 영상의 화소별 법선 벡터를 계산한다. 화소별 법선 벡터를 계산하는 과정은 수치적 최적화로 해결한다.

이와 같은 방법은 후처리를 통한 얼굴 재조명 등의 응용에는 적합하지만, 가상 화장 등 빠른 응답이 필요한 응용에 적용하기에는 비교적 많은 계산 시간을 요구한다. 따라서 본 논문에서는 영상으로부터 조명 환경을 얻는 과정, 입력 영상의 화소별 법선 벡터를 계산하는 과정을 변형하여 계산 과정을 단순화하기 위한 방법을 제안한다. 특히, 가상 화장 등의 응용에서는 높은 정확도가 요구되지 않는다는 가정에서 각 과정을 단순화한다.

본 논문에서 제안하는 가상 화장 시스템의 흐름은 그림 1과 같다. 먼저 얼굴 영상이 입력 되면 단계적 자세 회귀법(cascade pose regression)을 이용하여 얼굴의 특징점을 분석한다. 이때 얼굴 영상을 분석하는 과정에서 동시에 조명 환경을 계산함으로써 빠르게 조명 환경을 계산한다. 계산된 특징점의 위치를 기반으로 3차원 표준 얼굴 모델을 정합한다. 이때 얼굴 형태의 개인차와 표정 등에 기인하여 얼굴의 각 부분이 정확하게 일치하지 않는다. 따라서 영상 와핑(warping) 기술을 적용하여 2차원 영상 수준에서 3차원 모델을 변형하여 입력 영상과 일치시킨다. 가상 화장에 필요한 정보는 얼굴 내의 각 화소별 법선 벡터이므로 3차원 영상의 색상을 렌더링하는 대신 법선 벡터를 렌더링하여 법선 지도를 생성하고 이를 입력 영상에 맞게 와핑한다. 이런 과정으로 얻은 법선 지도와 조명 정보를 이용하여 화장 텍스처에 음영을 적용한 후 영상에 합성한다. 이와 같은 과정을 통하여 상당한 가속 효과를 얻을 수 있고, 기존의 방법에 비해 보다 자연스러운 화장 효과를 얻을 수 있다.

2.1 얼굴 특징점 및 조명 추정

본 논문에서 제안하는 시스템의 첫 단계는 입력된 얼굴 영상으로부터 얼굴의 특징점과 조명을 분석한다. 얼굴 영상으로부터 특징점을 얻기 위한 다양한 방법들이 존재한다. 본 시스템에서는 단계적 자세 회귀법[3]에 기반한 방법을 사용한다. 단계적 자세 회귀법은 놀라운 속도와 정확도에 기인하여 활발하게 연구되고 사용되고 있다. 단계적 자세 회귀법에 기반한 얼굴 특징점 분석 방법은 다음과 같이 요약할 수 있다[4, 5].

단계적 자세 회귀법은 기계학습의 방법 중 하나인 단계적 학습(boosted learning) 방법에 기반한다[6]. 단계적 학습에서는 초기 추정치로부터 시작하여 단계적으로 추정치를 보정하여 최종적인 결과를 얻는다. 단계적 자세 회귀법에서 각 단계에서 사용되는 보정자를 펀(fern)이라고 한다. 각 펀은 입력 영상에 대해 간단한 색인(index)을 계산하고 학습 영상 중 같은 색인을 가지는 영상들을 찾는다. 선택된 학습 영상들에서 특징점 변화 방향으로 추적하고 있는 특징점을 움직이는 과정을 여러개의 펀에 대해 계속 진행하여 최종 특징점을 얻는다. 이때 각각의 펀에 고유한 색인 방법을 부여하기 위해 각 펀은 얼굴 영상에 대하여 무작위로 선택한 몇 개의 화소 위치 순서쌍을 가진다. 입력 영상(혹은 학습 과정에서 학습 영상)이 입력되면 각 순서쌍에 저장된 위치의 화소 값의 밝기의 차에 따라 이진 부호(binary code)를 한 자리를 부여한다. 각 순서쌍에 대한 이진 부호를 연결하면 각 펀에 고유한 색인을 얻을 수 있다. 입력 영상으로부터 각 펀 별로 영상의 색인을 얻는 과정이 매우 빠르고, 색인이 얻어진 후에는 특징점에 변화량을 더하는 과정이 효율적으로 처리될 수 있기 때문에 단계적 자세 회귀법은 매우 낮은 시간 복잡도를 가진다. 특히, 펀의 개수를 조절하여 결과의 정확도와 계산 시간 사이의 균형(trade-off)를 조정할 수 있다.

본 논문에서는 펀의 색인을 계산하는 과정이 영상의 얼굴 영역을 전반적으로 분석하는 과정이라는 점, 색인과 최종 결과를 얻는 값에 대응관계가 없다는 점에 착안한다. 즉, 조명 정보와 같이 영상에서 얻을 수 있는 정보를 학습 과정에 추가하고, 단계적 자세 회귀법을 적용하면 부가적인 색인 과정 없이 정보를 얻을 수 있다. 따라서 본 연구에서는 학습 과정에서 영상의 조명 정보를 추가하고, 각 펀은 특징점과 더불어 조명 정보를 보정하도록 하여 입력 영상의 조명 정보를 계산한다. 이때 조명 환경을 간단한 숫자로 표현하기 위해 본 논문에서는 Kemelmacher-Shlizerman과 Basri의 방법과 동일하게 구면 조화 함수를 사용한다. 가상 화장의 특성상 매우 정교한 조명 환경이 필요한 것은 아니기 때문에 조명 환경은 4개의 구면 조화 함수의 계수로 표현한다. 학습 과정에 필요한 각 영상에 대한 조명 정보는 Kemelmacher-Shlizerman과 Basri가 제안한 방법으로 계산한다.

2.2 법선 벡터 맵 생성

화장 영상에 음영을 적용하기 위해 조명 환경 뿐 아니라 얼굴 영역의 각 화소에 대한 얼굴 곡면의 법선 벡터가 필요하다. 본 논문에서 제안하는 시스템에서는 표준 3D 모델을 입력 영상에 맞게 강체 변환(rigid transformation)하고 그 결과를 렌더링 한 결과를 2차원 와핑하여 얼굴 입력 영상과 일치시키는 방법으로 법선 벡터 지도를 계산한다. 물론, 정확한 법선 벡터를 계산하기 위해서는 조명 환경과 입력 영상의 각 화소별로 음영의 관계를 분석하여 법선 벡터를 계산하여야 한다. 하지만 각 화소별로 연속성 등을 고려하여 법선 벡터를 계산하는 과정은 수치적 최적화 방법을 사용하여야 하기 때문에 매우 많은 계산량을 요구한다.

따라서 본 논문의 다음과 같은 점에 착안하여 법선 벡터 지도를 빠르게 생성한다. 먼저 대부분의 사람의 얼굴 형태는 유사한 특징

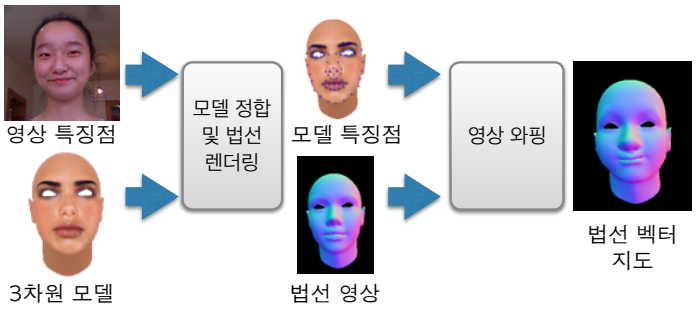

Figure 2: 법선 벡터 지도 생성 방법의 단계

을 가진다. 예를 들면 거의 모든 사람들의 코는 위로 솟아 있으며, 얼굴은 매우 거시적인 관점에서 달걀의 형태와 유사하다. 또한, 가상 화장에서는 얼굴의 굴곡을 근사하는 법선 벡터는 얼굴의 전반적인 특징을 표현하는 정도의 정확성으로도 상당한 효과를 거둘 수 있다. 즉, 코가 돌출되어 있기 때문에 생기는 법선 벡터의 변화를 가상 화장에 적용하여 음영을 계산하는 것으로도 상당한 효과를 얻을 수 있고, 코의 세부적인 높이에 따라 음영이 변화하는 변화 폭은 그리 크지 않다. 따라서 본 논문에서는 표준 모델을 적절히 변형하여 입력 영상에 정합하고 이를 통해 법선 벡터를 계산한다. 본 논문에서 법선 벡터를 계산하는 과정은 그림 2에 나타나있다.

먼저 입력 영상에서 얻은 특징점이 주어지면 표준 3차원 모델의 3차원 위치와 방향을 조절하여 영상에 정합한다. 이와 같은 영상 정합 과정은 일반적으로 알려져 있는 3차원 모델과 대응되는 영상내의 2차원 점들이 주어졌을 때 카메라의 위치와 방향을 찾는 카메라 정합(camera calibration)과정과 동일하다. 즉, 3차원상의 점 $\mathbf{p}_i$와 투영 행렬(projection matrix) $\mathbf{A}$, 영상 내의 점 $\mathbf{x}_i$가 주어지면 다음식을 최소화 하는 회전 행렬 $\mathbf{R}$과 변위 $\mathbf{t}$를 찾는 문제로 나타난다[7, 8].

$$E = \sum_i ||\mathbf{x}_i - \mathbf{A}(\mathbf{R}\mathbf{p}_i + \mathbf{t})||^2 \quad (1)$$

본 논문에서는 이와 같은 문제에서 사용되는 전형적인 방법인 Levenberg-Marquardt 방법을 사용한다. 여기서 $\mathbf{R}$은 역시 일반적인 방법에 따라 Rodrigues 식을 사용하여 지역적으로 선형화하여 계산한다. Levenberg-Marguardt 방법의 특징과 특징점의 개수가 수 십개(실제 실험에서는 74개)로 제한되어 있기 때문에 이와 같은 과정은 개인용 컴퓨터 기준으로 약 2-3 밀리초 안에 계산된다.

이런 방법으로 정렬된 3차원 얼굴 모델을 2차원으로 투영한 결과는 대략적인 얼굴의 방향, 크기, 위치 등은 일치하지만, 사람 얼굴의 다양성과 표정 등에 기인하여 입력 영상과 정확하게 일치하지 않는다. 특히, 눈, 코, 입 등 굴곡이 심한 부분의 위치가 입력 영상과 일치 하지 않기 때문에 이를 직접 법선 벡터 계산에 사용하면 부자연스러운 음영을 얻을 수 있다. 따라서 본 논문에서는 영상 와핑 방법을 통하여 문제를 해결한다.

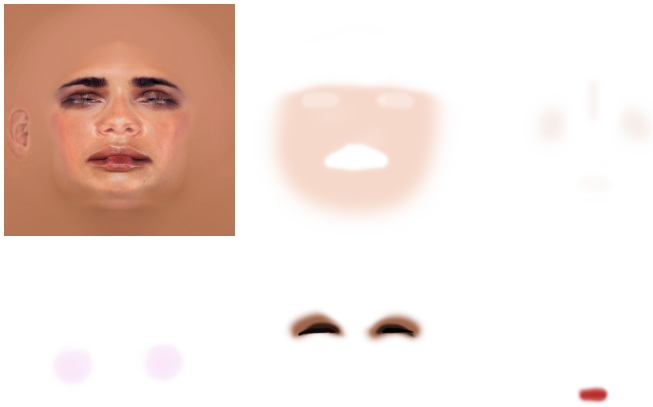

Figure 3: 3차원 표준 모델의 기본 텍스처와 여러가지 화장 효과 텍스처

법선 벡터 지도를 계산하기 위해 먼저 특징점을 기반으로 강체 변환을 통하여 정합된 3차원 모델을 렌더링(rendering) 한다. 이때 실제 원하는 정보가 색상(color)값이 아니라 법선 벡터이므로 색상 버퍼에 법선 벡터를 대신 렌더링한다. 본 논문의 구현에서는 RGB 채널에 각각 법선 벡터의 $x-, y-, z-$축의 값을 저장하였다. 이 영상을 3차원 특징점을 투영한 점들을 입력 영상에서 얻은 특징점에 정합되도록 영상을 와핑한다. 영상 와핑에는 분산 데이터 보간법(scattered data interpolation)을 적용한다. 본 연구에서는 방사 기저 함수(radial basis function)을 이용한 보간법을 사용하였다. 특히 박판 함수(thin-plate function)을 적용하였고, 전체적인 경향은 1차함수를 이용하여 보간하였다.

2.3 가상 화장 적용

가상 화장을 적용하기 위해 먼저 각각의 화장 효과를 영상으로 정의한다. 화장 효과는 얼굴 면에만 적용되어야 하며 입력 영상의 얼굴의 방향에 따라 변화하여야 한다. 따라서 가상 화장은 표준 모델에 텍스처의 형태로 정의된다. 3차원 표준 모델이 주어지면 얼굴 모델의 각 정점 별로 텍스처 좌표를 지정한다. 이후 3차원 얼굴의 특정 부분에 화장을 추가하기 위해 텍스처 영상의 해당되는 부분에 색상을 입힌다. 이때 적절한 투명도를 지정하여 화장의 두께 및 투명도를 조절할 수 있다. 그림 3은 화장 영상의 예를 보이고 있다.

이와 같이 텍스처가 준비되고 입력 영상이 주어지면, 화장을 입히는 과정은 3차원 모델을 렌더링 하는 과정과 동일하게 진행될 수 있다. 먼저 3차원 모델을 입력 영상에 정합한 후 3차원 모델을 렌더링하는 과정에서 법선 벡터를 렌더링 하는 과정과 별도로 화장 효과 텍스처들을 렌더링 한다. 화장 효과를 렌더링 한 영상에도 동일하게 와핑을 적용하고, 미리 와핑 된 법선 벡터, 조명 효과 등을 고려하여 화장 영상의 각 화소에 음영을 추가하여 입력 영상과

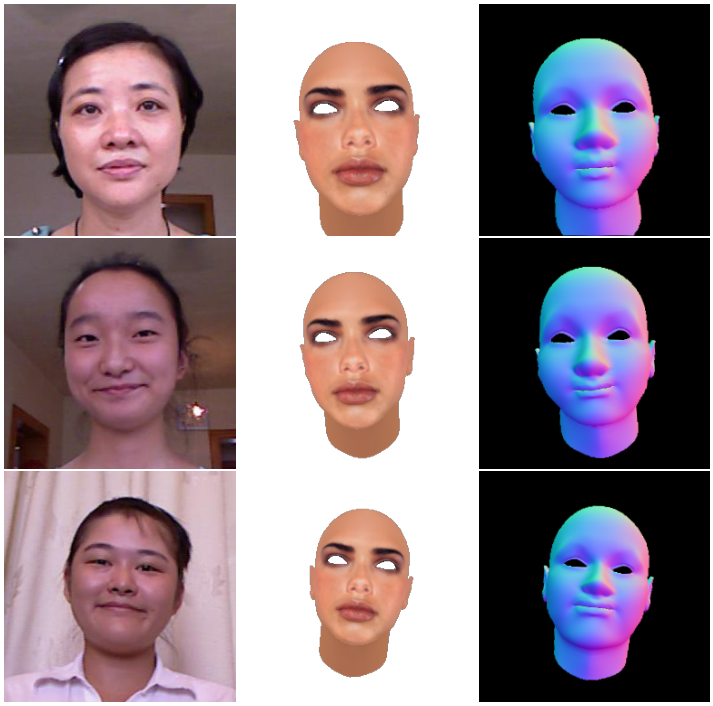

Figure 4: 영상에 정합된 3차원 모델과 와핑된 법선 벡터 지도

혼합하면 화장이 추가된 영상을 얻을 수 있다.

3. 실험 결과

본 시스템의 세부적인 구현 방법은 다음과 같다. 먼저 특징점 추적을 위해서는 Cao 등이 제안한 방법[5]을 사용하였다. 이 방법은 계층적 자세 회귀법을 사용한다. 위에서 언급한 바와 같이 수작업에 의해 지정된 특징점 좌표와 더불어 Kemelmacher-Shlizerman과 Basri가 제안한 방법으로 얻어진 구면 조화 함수의 계수를 추가하여 학습 과정을 진행하였다. 학습에는 Face warehouse 데이터 집합[9]을 사용하였다. 이 데이터 집합에는 5904개의 얼굴 영상에 각각 74개의 특징점이 지정되어 있다. 조명 환경을 나타내는 구면 조화 함수 계수는 4개를 사용하였다.

그림 4은 입력 영상에 표준 3차원 모델을 강체 변환으로 정합한 결과, 특징점을 기준으로 렌더링 된 영상을 와핑 한 결과 및 최종적으로 얻은 법선 벡터 지도를 보이고 있다. 그림에서 나타나듯이 기본적으로 정합 과정에서는 입력 영상 내의 얼굴 방향과 크기, 위치 등을 적절히 근사하고 있으며, 와핑을 통하여 얼굴의 중요한 형태를 복원한 것을 확인 할 수 있다. 여기서 본 실험에 사용한 학습 데이터에는 눈썹 위쪽에 특징점이 없기 때문에 이마 부분의 형태는 정확하지 않다. 특징점이 눈썹 아래쪽 부분에서 법선 벡터 지도는 입력 영상 내의 얼굴의 주요 부분의 굴곡을 그럴듯 하게 표현하고 있는 것을 볼 수 있다. 실제 구현에서는 영상에 와핑을 적용하지 않고, 와핑을 위한 지도를 2차원 영상으로 생성하였다. 지도의 각 화소는 입력 영상의 화소에 대응되며, 각 화소 값은 렌

Figure 5: 가상 화장 적용 결과: (각 행의 왼쪽부터) 원본 영상, 음영이 적용되지 않은 효과, 기존 방법으로 얻은 조명을 적용한 결과, 제안하는 방법을 통해 얻은 결과

더링 된 영상 내에서 법선 벡터 지도, 렌더링 된 화장 효과 영상의 참조 좌표값을 저장한다. 따라서 화소의 참조는 프로그램 가능한 그래픽 파이프라인에서 매우 효율적으로 진행할 수 있다. 또한, 와핑 지도를 입력 영상보다 상당히 작게 생성하고 그래픽 하드웨어에서 빠르게 확대하여 적용하여 보다 효율적으로 와핑 지도를 생성하도록 할 수 있다.

그림 5은 화장을 적용한 영상을 보여주고 있다. 각각의 열은 원본 영상, 음영 효과를 적용하지 않은 결과, Kemelmacher-Shlizerman과 Basri가 제안한 방법을 통하여 얻은 조명을 적용한 결과, 본 논문에서 제안한 방법을 통한 결과를 보이고 있다. 먼저 음영 효과를 적용하지 않은 영상은 얼굴의 굴곡이 변화함에 따라서 화장 효과의 명도가 변화하지 않기 때문에 부자연스러운 결과가 나타나는 경우가 있는 것을 확인할 수 있다. 특히 네번째 행의 립스틱 효과와 같이 투명도가 높은 경우에는 부자연스러움이 덜 하지만, 얼굴 코 주변이나 얼굴 외곽선 등 음영이 심하게 나타나는 부분에 걸쳐 나타나는 화장의 경우 입체감이 줄어드는 현상을 볼 수 있

다. 여기에 음영 효과가 추가되면 보다 자연스러운 결과를 얻을 수 있고, 빠른 방법으로 근사한 조명으로도 충분한 효과를 얻을 수 있음을 확인할 수 있다. 또한 1, 2행에서 볼 수 있듯이 조명의 밝기에 따라서 화장의 밝기를 자동으로 조정함으로써 자연스럽게 조명 환경을 반영하고 있는 것을 확인할 수 있다.

4. 결론 및 향후 연구

본 논문에서는 입력 얼굴 영상에 가상 화장 효과를 추가하는 방법을 제안하였다. 특히 가상 화장에 음영 효과를 추가하여 화장 효과를 자연스럽게 하는 방법을 제안하였다. 이와 같은 효과를 빠르게 생성하기 위해 기존의 조명을 예측하기 위한 방법, 입력 영상의 얼굴 부분의 굴곡에 따라 법선 벡터를 계산하는 방법 등을 근사하여 빠르게 처리하기 위한 방법을 제안하였다. 그 결과 다양하고 자연스러운 가상 화장 효과를 얻을 수 있었다.

본 논문에서 제안하는 방법의 제한점은 먼저 영상에서 얻은 조명이 연속된 영상에 대해 부드럽게 변화하지 않는 경우가 나타난다는 문제점이 있다. 따라서 동영상이나 실시간 입력 영상 등 연속적인 영상에 대해서 화장 효과가 부드럽게 변화하지 않는 경우가 있다. 이는 단계적 자세 회귀법이 기본적으로 가지고 있는 단점으로, 조명이 영상 안에서 급격하게 변화하는 경우가 많지 않기 때문에 적절한 시간 필터링(temporal filtering)을 적용하여 해결할 수 있을 것으로 보인다.

본 시스템에서 가장 많은 시간을 요구하는 부분은 와핑 지도를 만드는 과정이다. 이는 와핑 함수를 생성한 후에 각 화소 별로 변화량을 계산하여야 하기 때문에다. 따라서 이와 같은 과정은 3차원 표준 모델의 정점을 변화시키는 과정을 통하여 와핑을 적용함으로써 가속할 수 있을 것으로 보인다. 단, 2차원 위치의 변화량에 해당하는 3차원 정점의 변화량으로 변환하면서 얼굴 모델의 자연스러움을 유지하는 방법이 필요하다.

Reference

[1] I. Kemelmacher-Shlizerman and R. Basri, "3d face reconstruction from a single image using a single reference face shape," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 2, pp. 394–405, 2011.

[2] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment maps," in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '01. New York, NY, USA: ACM, 2001, pp. 497–500. [Online]. Available: http://doi.acm.org/10.1145/383259.383317

[3] P. Dollár, P. Welinder, and P. Perona, "Cascaded pose regression," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 1078–1085.

[4] M. Valstar, B. Martinez, X. Binefa, and M. Pantic, "Facial point detection using boosted regression and graph models," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2729–2736.

[5] X. Cao, Y. Wei, F. Wen, and J. Sun, "Face alignment by explicit shape regression," *International Journal of Computer Vision*, vol. 107, no. 2, pp. 177–190, 2014.

[6] N. Duffy and D. Helmbold, "Boosting methods for regression," *Machine Learning*, vol. 47, no. 2-3, pp. 153–200, 2002.

[7] D. F. Dementhon and L. S. Davis, "Model-based object pose in 25 lines of code," *International journal of computer vision*, vol. 15, no. 1-2, pp. 123–141, 1995.

[8] H. Araújo, R. L. Carceroni, and C. M. Brown, "A fully projective formulation to improve the accuracy of lowe's pose-estimation algorithm," *Computer Vision and Image Understanding*, vol. 70, no. 2, pp. 227–238, 1998.

[9] C. Cao, Y. Weng, S. Zhou, Y. Tong, and K. Zhou, "Facewarehouse: a 3d facial expression database for visual computing," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 20, no. 3, pp. 413–425, 2014.

[10] K.-C. Lee and B. Moghaddam, "A practical face relighting method for directional lighting normalization," in *Analysis and Modelling of Faces and Gestures*. Springer, 2005, pp. 155–169.