

얼굴 잡티 제거를 통한 화장 전이 성능 개선

권소예, 이재구*
국민대학교

*jaekoo@kookmin.ac.kr

Improving Makeup Transfer with Removal of Blemishes on Face

Soye Kwon, Jaekoo Lee*
College of Computer Science, Kookmin University

요약

화장 전이 과업(makeup transfer)은 참조 사진(reference image)의 화장 특징을 원본 사진(source image)으로 전이하는 과업이다. 화장을 직접 체험하지 않고 원본 사진과 참조 사진만으로 화장을 가상으로 체험할 수 있는 일은 코스메틱(cosmetic) 시장에서 큰 가치를 가진다. 이러한 과업을 수행하기 위해 기존 화장 전이(makeup transfer)모델은 피부 화장의 특징을 전이하는 과정에서 참조 사진의 피부 특징이 잘 전이되지 않고 원본 사진의 잡티가 그대로 남아있는 문제점이 존재한다. 따라서 원본 사진의 잡티를 제거하는 모듈을 추가함으로써 원본 사진의 잡티가 그대로 남아있는 문제점을 해결하고 참조 사진의 피부 화장을 더 잘 따라가도록 하여 화장 전이 성능을 개선했다.

I. 서론

사용자의 가상 체험(virtual try on)서비스에 대한 수요가 커짐에 따라 기업의 수요 또한 커지고 있다. 이러한 가상체험의 발전에 따라, 화장 가상 체험의 수요 또한 증가하고 있다.[1] 이러한 수요에 맞춰 인공지능 기반의 화장 전이가 많이 제안되어 왔는데, 본 논문은 CycleGAN[2]기반 화장 전이의 SOTA(state-of-art)인 EleGANt[3]의 성능 개선 방안을 제안한다.

여기서 화장 전이 과업(makeup transfer)이란 원본 사진(source image)의 얼굴에 참조 사진(reference image)의 화장 특징을 전이하는 것을 목표로 하는데, 기존 연구[3]는 [그림 1]과 같이 원본 사진에 피부 잡티가 있는 경우 목표 사진의 피부 화장을 잘 따라가지 않고 결과 사진에 피부 잡티가 그대로 남아있는 문제점이 존재한다. 피부 화장은 일반적으로 잡티나 트러블을 가려주는 기능이 있기 때문에 원본 사진의 잡티가 그대로 남아있는 것은 서비스적으로도 치명적인 문제가 될 수 있다.

따라서 본 논문은 잡티 제거 모듈을 추가함으로써 결과 사진에 잡티가 남아있는 문제점을 해결하고 결과 사진이 참조 사진의 피부 화장을 더 잘 따라가도록 하였다. 잡티 제거 모듈은 적응형 이진화를 통해 잡티를 감지하고 감지된 잡티를 주위 피부색으로 바꿈으로써 잡티를 제거한다.

제시한 방법의 성능 평가를 위해 SSIM 이라는 평가 지표를 이용하여 정량적으로 평가한 결과 2.1% 향상된 결과를 보였으며 기존 방법과 제시된 방법의 시각적 비교를 통해 정성적 평가를 수행하여 결과의 신뢰도를 높였다.



그림 1 기존 화장 전이 문제점

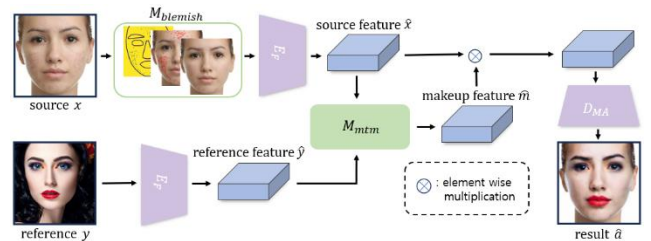


그림 2 본 논문에서 제안한 모델 구조

II. 본론

본 논문에서 제안하는 모델의 구조는 기존 모델[3]에 잡티 제거 모듈 $M_{blemish}$ 을 추가한 [그림 2]와 같다. 원본 사진 x 와 참조 사진 y 가 입력으로 주어지면 출력으로 원본 사진의 얼굴에 목표 사진의 화장 특징이 전이된 얼굴 \hat{a} 가 생성된다.

기존 모델[2]은 원본 사진 특징(source feature) 맵 \hat{x} 이 잡티의 정보를 가진 채 화장 특징이 전이되기 때문에 [그림 1]의 결과 사진처럼 잡티가 그대로 존재하는 문제점이 발생한다. 따라서 원본 사진 특징 맵 \hat{x} 이 이 잡티의 정보를 가지고 있지 않도록 사전 전처리 단계에서 잡티를 감지하고 제거할 필요가 있다.

따라서 두 사진이 입력으로 주어지면 원본 사진은 잡티 제거 모듈 $M_{blemish}$ 을 거쳐 잡티가 제거된다. 우선 잡티 제거 모듈 $M_{blemish}$ 은 잡티를 감지하기 위해 눈, 코, 입을 제외한 피부에 적응형 이진화를 수행한다. 적응형 이진화는 사진의 부분마다 조명 조건이 다를 수 있는 점을 반영하기 위해 사진을 작은 영역으로 나누어 각각에 적합한 임계 값을 기준으로 이진화를 수행한다. 적응형 이진화를 수행한 결과는 [그림 3]의 왼쪽 그림과 같다.



그림 3 적응형 이진화를 수행하고 잡티를 감지

잡티 제거 전 잡티 감지 결과 잡티 제거 결과

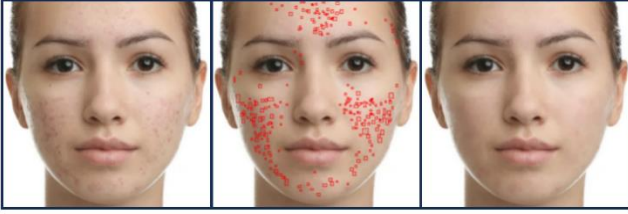


그림 4 잡티를 감지하여 제거한 결과

[그림 3]의 왼쪽 그림을 보면 잡티는 점의 형태를 띄기 때문에 이진화 된 결과에서 점의 형태만 감지하면 [그림 3]의 오른쪽 그림과 같다. 잡티로 감지된 위치를 원본 사진에 표시하면 [그림 4]의 잡티 감지 결과와 같다. 이후 감지된 영역을 주위 피부색으로 변환하고 스무딩 필터(smoothing filter)를 적용하면 [그림 4]의 잡티 제거 결과와 같다.

그 후 인코더(encoder) E_F 를 통과해 원본 사진 특징 맵 \hat{x} 이 생성되고 참조 사진은 잡티 제거 모듈을 거치지 않고 바로 인코더 E_F 를 통과해 참조 사진 특징(reference feature) 맵 \hat{y} 이 생성된다.

$$\hat{x} = E_F(M_{blemish}(x)) \quad (1)$$

$$\hat{y} = E_F(y) \quad (2)$$

다음 두 특징 맵 \hat{x} , \hat{y} 은 화장 전이 모듈 M_{mtm} 을 거친다. 화장 전이 모듈 M_{mtm} 은 두 특징 맵 \hat{x} , \hat{y} 을 어텐션(attention)연산하여 어텐션 맵을 생성하고 이 어텐션 맵과 참조 사진 특징 맵 \hat{y} 를 요소별 곱셈(element-wise multiplication)하여 화장 특징 맵 \hat{m} 을 생성한다.

$$\hat{m} = M_{mtm}(\hat{x}, \hat{y}) \quad (3)$$

마지막으로 원본 사진 특징 맵 \hat{x} 과 화장 특징 맵 \hat{m} 을 요소별 곱셈한 후 디코더(decoder) D_{MA} 를 통과하여 결과 사진 \hat{a} 를 생성한다.

$$\hat{a} = D_{MA}(\hat{x} \odot \hat{m}) \quad (4)$$

III. 실험

제안 방법의 성능을 평가하기 위해 기존 방법과 성능을 비교하는 실험을 진행한다. 실험에 사용한 데이터 집합은 MT(Makeup Transfer)[4] 데이터 집합으로 1115 개의 화장하지 않은 사진과 2719 개의 화장한 사진을 포함한다. 변인을 통제하기 위해, 우리의 제안 방법과 기존 방법은 동일한 실험 환경에서 학습된다.

정량적 성능을 평가하기 위해 Structural Similarity Index Map(SSIM) 평가지표를 사용한다. SSIM 평가지표는 사진의 밝기 차이와 고주파 정보인 세부사항의 유사성을 판단하므로 밝기 차이에 해당하는 잡티와 세부사항에 해당하는 화장의 텍스처를 비교하는 데 유용하다. SSIM 평가 지표를 통해 결과 사진과 참조 사진의 유사성을 비교하며 이를 통해 화장 전이된 결과 사진이 잡티가 잘 제거되고 참조 사진의 화장 특징을 잘 따랐는지를 평가한다. [표 1]을 보면, SSIM 가 기존보다 2.1% 향상된 것을 통해 우리가 제안한 방법(Ours)이 기존의 방법(Baseline)보다 좋은 성능을 내는 것을 알 수 있다.

사용자 선호 조사를 통해 제안 모델의 화장 전이 성능의 우수성을 뒷받침하고자 18 명을 대상으로 기존 모델과 제안 모델의 결과 사진 중 어느 사진이 피부 화장이 잘 전이되었는지 조사하였다. [표 1]을 보면, 조사한 결과 88.14%의 사용자가 제안 모델의 결과 사진이 참조 사진의 피부 화장을 더 잘 따랐다고 응답하였다.

표 1. 모델 간의 정량적 성능 비교 지표

	Baseline	Ours
SSIM(↑)	0.339	0.346
user study(↑)	11.86%	88.14%

원본 사진 참조 사진 기존 모델 제안 모델

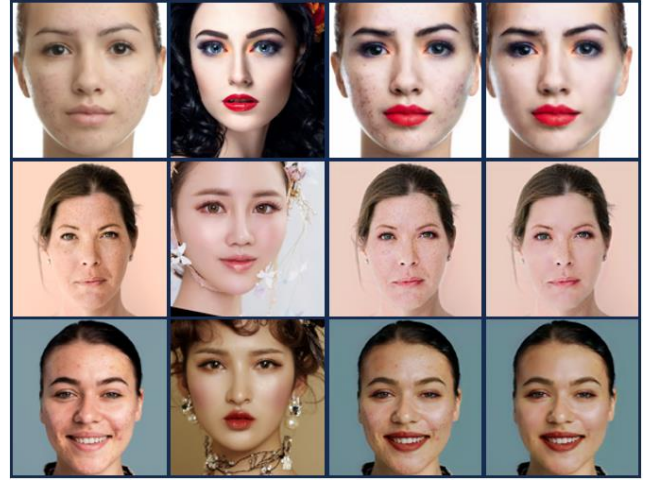


그림 5 모델 결과 비교

[그림 5]는 기존 방식(3 열)과 우리가 제안한 방식(4 열)의 정성적 성능의 결과를 보여준다. 3 열은 원본 사진의 잡티가 그대로 남아있는 반면에 4 열은 잡티가 잘 제거된 것을 볼 수 있다. 이로써 우리의 제안 방식이 잡티를 잘 제거함으로써 참조 사진의 피부를 더 잘 따른다고 할 수 있다.

IV. 결론

본 논문에서는 화장 전이 모델인 EleGANt[3]로 화장전이를 진행한 결과 원본 사진의 잡티가 그대로 남아있어 참조 사진의 피부 화장을 잘 따라가지 못하는 문제점을 해결하기 위해 기존 모델의 구조에 잡티 제거 모듈을 추가하는 방법을 제안하였다. 그리고 정량적, 정성적으로 기존 방식과 비교 실험을 진행하여 제안 방법이 문제점을 해결하고 성능 향상에 기여한다는 것을 보여주었다.

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.RS-2023-00212484, 복잡한 실제 주행환경에서 설명 가능한 움직임 예측).

본 연구는 2022 년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음"(2022-0-00964)

참고 문헌

- [1] Dezaki, Fatima Taheri, et al. "Automated material properties extraction for enhanced beauty product discovery and makeup virtual try-on." (2023).
- [2] Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." *Proceedings of the IEEE international conference on computer vision*. 2017.
- [3] Yang, Chenyu, et al. "Elegant: Exquisite and locally editable gan for makeup transfer." *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 2022.
- [4] T., Qian, R., Dong, C., Liu, S., Yan, Q., Zhu, W., Lin, L.: Beautygan: Instancelevel facial makeup transfer with deep generative adversarial network. In: Proceedings of the 26th ACM International Conference on Multimedia. pp. 645-653 (2018)