

얼굴 교체를 위한 영상 블렌딩 방법

장성길, 송중석, *한동훈, 박종일¹

한양대학교 컴퓨터소프트웨어학과, *LG 전자 Convergence 연구소

sgzhang@mr.hanyang.ac.kr, jssong@mr.hanyang.ac.kr, *donghoon5.han@lge.com*

jipark@hanyang.ac.kr¹

The image blending method for face swapping

Xingjie Zhang, Joongseok Song, *Donghoon Han, and Jong-Il Park¹

Hanyang University Computer Software, *LG Electronics Convergence R&D Lab

요 약

최근 들어 얼굴 교체와 같은 영상 합성 기술들이 많은 관심을 받고 있다. 일반적으로 영상을 합성할 때, 영상간 뚜렷한 명암 차이로 인해 부자연스러운 경계가 발생하는데 이를 자연스럽게 제거하는 블렌딩 기술이 필요하다. 본 논문에서는 이러한 문제를 해결하기 위해 적응적 가중치 기반의 영상 블렌딩 방법을 제안한다. 실험 결과, 본 논문에서 제안하는 방법이 얼굴 합성시 발생하는 뚜렷한 경계 부분을 자연스럽게 제거하고, 합성하고자 하는 얼굴의 눈, 코, 입과 같은 주요 부위를 잘 보존하는 것을 확인할 수 있었다.

1. 서론

최근 들어, 영상 내의 얼굴을 다른 사람의 얼굴로 교체하려는 시도가 많은 관심을 받고 있다. 이때, 두 얼굴의 자연스러운 영상 합성이 필요한데 이를 구현하기 위해서는 효과적인 얼굴 블렌딩 기술이 필요하다. 본 논문에서는 영상 내의 정보에 적응적인 가중치를 기반으로 한 블렌딩 기법을 사용하여 두 얼굴을 자연스럽게 합성하는 방법을 제안한다. 다음 그림 1 과 같이, 목표 영상에 참조 영상을 자연스럽게 합성하기 위해서는 참조 영상의 중요 얼굴 요소 부위인 눈, 코, 입의 특징을 보존해주어야 한다. 이를 위해, 우선 참조 영상과 목표 영상으로부터 얼굴 영역을 검출해내고 블렌딩을 적용할 합성 영역을 추출한다. 이후, 참조 영상의 안면 부분을 목표 영상의 안면 부분에 덮어 씌운다. 이때, 본 논문에서 제안하는 적응적 가중치를 이용한 블렌딩 방법을 사용하여 두 얼굴의 자연스러운 합성을 수행한다. 제안하는 방법은 얼굴의 크기에 따라 자동으로 블렌딩 범위를 설정한다. 실험을 통하여 제안하는 방법이 참조 영상의 얼굴을 주요부위(눈, 코, 입)를 보존하면서 목표 영상의 얼굴에 효과적으로 합성시킬 수 있음을 확인하였다. 본 논문의 구성은 다음과 같다. 2 절에서는 전체적인 블렌딩 알고리즘의 흐름과 방법에 대해서 설명하고 3 절에서는 제안하는 방법에 대한 실험 결과를 나타내고 있다. 마지막으로 4 절에서는 본 논문에 대한 결론을 나타낸다.

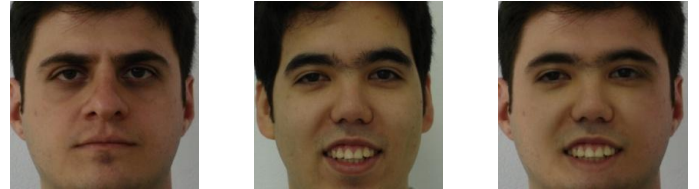


그림 1. (가) 목표 영상, (나) 참조 영상, (다) 본문에서 제안하는 블렌딩이 적용된 결과 영상.

2. 영상 블렌딩 방법

2.1 알파 블렌딩 방법

Thomas Porter²와 Tom Duff 가 제안한 알파 블렌딩 [4] 방법은 원본 영상과 교체하려는 영상의 명암 값을 일정한 비율로 섞어서 융합시키는 방법이다. 그림 2 에서 보는 바와 같이 α 값이 커질수록 결과 영상의 명암 값 중에서 참조 영상의 명암 값이 점유하는 비율이 커지면서 참조 영상의 특징이 목표 영상의 특징에 비해 상대적으로 많이 표현된다. 그림 2-(다) 에서 나타난 것과 같이 α 가 큰 값을 가질 때 블렌딩이 적용되는 합성 영역과 배경 사이에 피부색, 조명 등의 차이로 인해 뚜렷한 경계가 발생하는 것을 볼 수 있다. 그러므로 참조 영상의 안면 특징이 잘 표현되면서 경계를 자연스럽게 제거할 수 있는 효과적인 블렌딩 기술이 필요하다.

¹ 교신 저자

- 본 연구는 LG 전자(주)의 지원으로 수행하였음.



그림 2. (가) 가중치 α 를 0.1로 적용한 결과 영상, (나) 가중치를 0.2로 적용한 결과 영상, (다) 가중치를 0.9로 적용한 결과 영상.

2.2 이미지 전처리

블렌딩을 하기 전에 먼저 블렌딩을 적용할 합성 영역을 추출해야 한다. 우선 Viola-Jones가 제안한 Haar-like feature를 이용한 haar cascade[1]를 사용하여 목표 영상과 참조 영상 내 얼굴의 위치를 찾아내고 영상을 각각 저장한다. 생성된 안면 영상에 대하여 수동으로 합성 영역을 추출하고 목표 영상에 덮어씌운다. 두 얼굴을 자연스럽게 교체하기 위해 눈의 위치를 추출하여 두 눈의 중심점을 기준으로 블렌딩을 진행한다.

2.3 가중치 변화에 따른 블렌딩 방법

본 논문에서는 합성 영역의 각 화소의 명암 값에 가중치를 부여할 때 그림 3-(나)에 나와 있듯이 합성 영역을 핵심 영역과 일반 영역 두 개 층으로 나누어 처리하는 방법을 제안한다. 참조 영상의 특징을 잘 표현하면서 안면 영역이 목표 영상에 융합되게 하기 위하여 눈, 코, 입 등 안면의 주요 부위를 포함한 영역을 핵심 영역으로 지정하고 핵심 영역에 포함되는 참조 영상 명암 값에 가중치의 최대치 $W_{max} = 1$ 를 부여한다. 동시에 핵심 영역을 제외한 부분을 일반 영역으로 지정하고 각 화소의 위치와 경계선 사이의 거리 d 에 비례하여 d 가 크면 참조영상 화소 값의 가중치를 크게, 반대로 d 가 작으면 가중치를 작게 부여한다. 그림 3-(가)에 나와 있듯이 거리 d 는 상, 하, 좌, 우 네 방향에 있는 경계선까지 거리 중에서 제일 작은 값을 의미 한다.

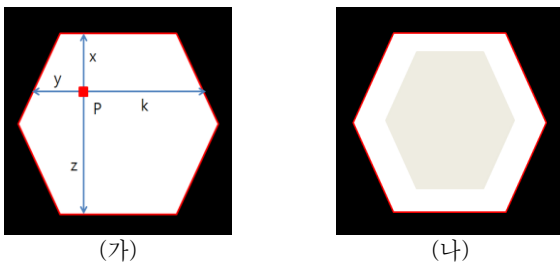


그림 3. (가) x, y, z, k 는 화소 P로부터 상, 하, 좌, 우 경계선까지의 거리이다. (나) 붉은색 번두리 영역은 합성영역, 밝은 영역은 일반 영역, 회색 영역은 핵심 영역이다.

이는 다음 식(2)로 표현할 수 있다. 여기서 d_i 는 핵심 영역과 일반 영역을 결정하는 참고거리이며 임의의 화소에 대응되는 거리 d 가 d_i 보다 클 경우 핵심 영역으로 분류된다. d_i

는 합성 영역의 너비와 높이 중 작은 값에 비례하여 결정된다.

$$R(d) = \begin{cases} S_1(d) \times \frac{d}{d_i} + S_2(d) \times (1 - \frac{d}{d_i}) & , 0 < d \leq d_i \\ S_1(d) & , d > d_i \end{cases} \quad (2)$$

3. 실험 및 구현결과

본 논문에서는 MS Visual C++을 Compiler로 사용하고 openCV 라이브러리들을 사용하여 실험을 진행하였다. 실험에

사용된 데이터는 The FEI face database 이다. d_i 는 합성 영역의 너비와 높이 중 작은 값을 취해 0.2를 곱한 값을 사용하였다. 그림 4 중 위쪽의 영상들은 실험에 사용한 목표 영상들이고 가운데 영상들은 참조 영상들이며 아래쪽 영상들은 본문에서 제안하는 블렌딩 방법을 적용한 결과이다.

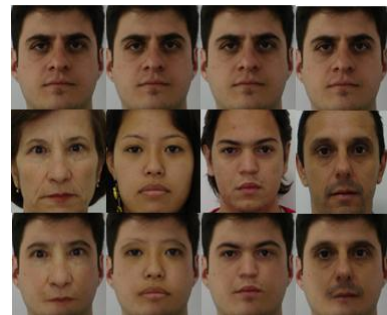


그림 4. 실험 결과.

4. 결론

본 논문에서는 자연스러운 얼굴 합성을 위해 적응적 가중치를 이용한 블렌딩 방법을 제안한다. 실험 결과, 참조 영상의 안면 영역이 주요 특징인 눈, 코, 입 등을 보존하면서 자연스럽게 목표 영상의 안면 영역에 합성되게 하였다. 또한, 두 얼굴을 합성할 때 발생하는 경계가 자연스럽게 제거됨을 확인할 수 있었다. 하지만, 조명에 따라 발생하는 전반사 문제나 서로 다른 피부 톤의 일치와 같은 문제는 아직 해결되지 않았기 때문에 향후 보완할 계획이다.

참고문헌

- [1] Viola, Paul, and Michael J. Jones. "Robust real-time face detection." International journal of computer vision 57.2 (2004): 137-154.
- [2] Milborrow, Stephen, and Fred Nicolls. "Active Shape Models with SIFT Descriptors and MARS." VISAPP 1.2 (2014): 5.
- [3] Milborrow, Stephen, and Fred Nicolls. "Locating facial features with an extended active shape model." Computer Vision- ECCV 2008. Springer Berlin Heidelberg, 2008. 504-513.
- [4] Porter, Thomas, and Tom Duff. "Compositing digital images." ACM Siggraph Computer Graphics. Vol. 18. No. 3. ACM, 1984.