

이미지를 사용한 가상의상착용 알고리즘들의 성능 분석⁺

(Performance Evaluation of VTON (Virtual-Try-On) Algorithms using a Pair of Cloth and Human Image)

따이 트안 투안¹⁾, 미나르 마드올 라흐만²⁾, 안 희 준³⁾*
(Thai Thanh Tuan, Matiur Rahman Minar, and Heejune Ahn)

요 약 가상착용기술(VTON: Virtual try-on)은 의상의 온라인 유통을 활성화를 위하여 중요한 기술이다. 그러나 3차원 그래픽스기반 방식은 의상과 인체의 3차원 정보의 확보가 필요하여 범용화에 어려움이 있고, 이러한 제약을 해소하기 위해 개발되는 이미지 기반 방식들의 연구들은 그 기술적 한계가 불명확하다. 구체적으로 VITON (Virtual image try-on) 과 CP-VTON (Content preserving VTON)등은 가능성 위주의 매우 단편적인 결과만을 제시하고 있다. 본 논문은 이미지기반 기술의 상용화의 한계를 파악하기 위해, 세 가지 대표적 방식(SCMM 기반의 비-딥러닝 방식, 딥러닝기반 VITON 과 CP-VTON)에 대하여 인물의 자세 및 체형, 의상의 가려짐 정도, 의상의 특성 등에 따라 분석을 하였다. 객관적인 평가를 위하여 변형단계와 합성단계의 성능을 각각 IoU와 SSIM로 평가하였고, 상대적인 비교 분석을 하였다. 그 결과, CP-VTON이 가장 좋은 성능을 보이지만, 자세와 의상의 복잡도에 따라 성능의 한계가 크게 차이가 남을 보였다. 그 주 원인은 2차 기하변형의 한계와 GAN을 통한 합성 기술의 한계로 파악되었다.

핵심주제어 : 가상착용, 딥러닝, 이미지 기반, 2D 변형, 성능 평가

Abstract VTON (Virtual try-on) is a key technology that can activate the online commerce of fashion items. However, the early 3D graphics-based methods require the 3D information of the clothing or the human body, which is difficult to secure realistically. In order to overcome this problem, Image-based deep-learning algorithms such as VITON (Virtual image try-on) and CP-VTON (Characteristic preserving-virtual try-on) has been published, but only a sampled results on performance is presented. In order to examine the strength and weakness for their commercialization, the performance analysis is needed according to the complexity of the clothes, the object posture and body shape, and the degree of occlusion of the clothes. In this paper, IoU and SSIM were evaluated for the performance of transformation and synthesis stages, together with non-DL SCM based method. As a result, CP-VTON shows the best performance, but its performance varies significantly according to posture and complexity of clothes. The reasons for this were attributed to the limitations of secondary geometric deformation and the limitations of the synthesis technology through GAN.

Keywords: Virtual-Try-On, Deep-learning, Image-based, 2D deformation, Quality evaluation

* Corresponding Author: heejune@seoultech.ac.kr

+ 이 논문은 2018년도 정부(교육부)의 재원으로 한국연구재단의 기초연구사업으로 수행되었습니다 (No.2018R1D1A1B07043879)

Manuscript received October 02, 2019 / revised October 29, 2019 / accepted October 29, 2019

1) 서울과학기술대학교 전기정보공학과, 제1저자

2) 서울과학기술대학교 전기정보공학과, 제2저자

3) 서울과학기술대학교 전기정보공학과, 제3저자, 교신저자

1. 서론

최근 논문 (Ahn, 2018a; Ahn 2018b)에 따르면 온라인 패션 시장 규모는 매년 급속히 증가하여, 미국은 2016년 720억 달러로, 2022년에 1조 230억 달러 규모로 성장할 것으로 보고 있다. 국내의 경우 2017년 추산 약 6000억 원으로 전체 온라인시장의 10%에 달하고, 매년 5~10% 정도의 증가율을 보인다. 기능과 성능의 규격화가 용이한 전자제품 등과 달리 의류는 형태와 색상, 재질이 매우 다양하며, 개인적인 선호도의 차이가 크다. 따라서 온라인 정보만으로는 구매결정이 상당히 어렵다. 따라서 판매품목의 고가화를 위해서는 온라인 가상 착용기술이 절실하다.

기존의 가상착용기술은 3차원 애니메이션 기술 중의 하나로 주로 의상 디자인에 사용되었고, 이를 거의 그대로 사용자를 위한 기술로 가져온 것이다. 그러나 일반인이 의상과 사람 신체의 3차원 모델을 확보하기 어렵기 때문에 2차원 이미지를 사용한 가상착용기술에 관심이 증가하고 있다(Ahn, 2018a; Ahn, 2018b; Han et al., 2018; Wang et al. 2018). Fig. 1에 대표적인 가상 착용 기술에 대하여 예시를 보인다.

이미지 기반 가상 착용기술로 현재까지 발표된 연구는 (Ahn, 2018a; Ahn, 2018b)를 포함하여 딥러닝을 중심으로 한 VITON (Virtual image try-on) (Han et al., 2018), CP-VTON (Content preserving-virtual try-on) (Wang et al. 2018), SwapNet (Raj et al. 2018) 등이 있다. 이들은 이차원 이미지를 입력으로 사용하고 알고리즘도 2차원 영상처리 및 딥러닝 방식을 사용하는 방법들이다. 세부적으로는 사람의 자세 예측, 이미지 분할 기술, 2차원 기하변환, 2차원 이미지 블렌딩 기술이 필요하며, 전통적인 규칙기반 방식과 딥러닝에 의한 학습기법이 같이 사용되고 있다. 한편, 입력으로 이미지를 사용하지만 3차원 방식과 결합하는 방식들도 최근에 발표 되었다 (Zafir et al., 2018; Weng et al., 2019) 이 방식들도 기본적으로는 2차원방식과 유사한 단계를 거치나, 이미지에서 3차원 모델 복원기법을 일부 사용하여 3차원 복원기법을 정보를 활용한다.

본 연구에서는 2차원 VTON방식들을 대상으로

성능 분석을 실시한다. 그 이유는 3차원방식의 연구는 아직 아이디어 차원으로 결과 품질 면에서 이차원 방식에 비하여 떨어지기 때문이며, 현실적으로 재실험을 할 수 있는 환경에 대한 설명이 부족하고, 코드나 데이터 셋이 제공되고 있지 않기 때문이다.

각 알고리즘의 제안 논문에서도 실험 결과를 제시하고 예를 보였으나, 실제로 저자들이 공개한 코드를 바탕으로 재실험을 해 본 결과 논문에 제시된 결과들은 체계적인 기준이 없이 성공적인 경우를 위주로 하였음이 확인되었다. 일반적으로 모든 경우에 상용화하여 적용할 만한 결과를 보이는 알고리즘은 드물다. 하지만 알고리즘 분석을 통하여 해당 알고리즘이 잘 동작하는 영역을 찾는다면 알고리즘을 실용화 할 수 있는 가능성이 생길 수 있다. 이러한 관점에서 본 논

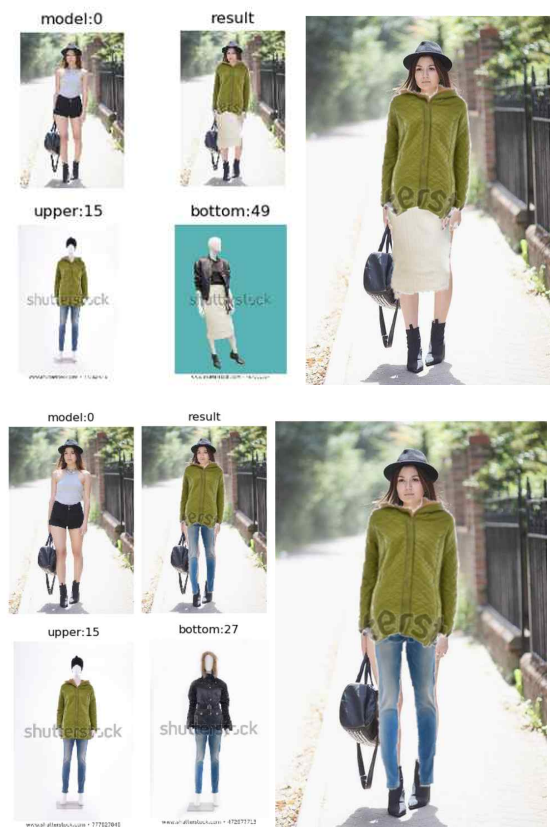


Fig. 1 The Input and Final VTON Image (enlarged) (Left: upper & skirt, Right upper & pants)(Ahn, 2018b)

문에서는 실험조건을 의상의 가려짐 정도, 자세의 복잡성, 의상의 복잡도 등으로 나누어 실험 결과를 분석하여 제시한다. 또한 대상이 되는 알고리즘들이 모두 형태 변형과 합성 단계로 구성되는 점을 고려하여 각 단계별 결과를 제시하고, 가능한 정량화하여 비교할 수 있도록 동일 의상 재착용 실험을 고안하여 IoU와 SSIM 수치를 제공한다.

본 논문의 구성은 다음과 같다. 제 2절에서는 본 논문의 분석 대상인 SCMM, VITON, CP-VTON 알고리즘에 대하여 간략히 비교 설명한다. 제 3절에서는 실험조건과 방법, 평가 기준을 제시한다. 제 4절에서는 이를 바탕으로 실험한 결과를 제시하고 그 성능을 분석한다. 마지막으로 5절에서는 분석결과를 바탕으로 현재 시스템을 활용할 수 있는 방안 에 대하여 설명한다.

2. 2차원 이미지기반 가상착용기술

다음에 설명하는 구체적인 가상착용 알고리즘들은 기술들은 공통적으로 사람의 자세와 의상을 인식할 필요가 있다. 자세 추정과 영상 분할 알고리즘은 매우 다양하고 다른 방식을 사용할 수도 있으나 자세예측에는 OpenPose (Cao et al., 2017) 방식과 사용자 분할에는 LIP (Liang et al., 2018) 방식이 대표적으로 사용된다. 다음 알고리즘을 설명할 때 입력으로 이 알고리즘 또는 이에 대등한 알고리즘으로 사전에 처리되었음을 가정으로 한다. 각 알고리즘은 VITON과 CP-VTON은 github (<https://github.com>)에 저자들이 공개된 코드와 데이터를 그대로 사용하였으며, 기 학습된 파라미터를 사용하였다. SCMM 기반의 경우는 저자의 SCMM Matlab 코드를 바탕으로 직접 구현하였다.

2.1 SCMM 기반 (Ahn, 2018b)

Fig. 2에서 제시한 이 방식은 가상착용을 적용하는데 있어서 비-딥러닝 방식을 사용한다. 알고리즘의 전체적인 구성은 Ahn (2018a, 2018b)에 기초하나, 여기서는 in-shop 의상을 사용해야 하기 때문에 형태 매칭에서 조인트를 이용하는 대신

Shape-context Matching (Belongie et al., 2002)를 이용한 방법을 사용하여 매칭 및 변형을 하게 된다. 기존의 연구에서는 ARAP (As rigid as possible)을 사용하여 변경하였으나, 본 논문에서는 다른 알고리즘과 비교를 위하여 동일한 TPS 알고리즘을 사용하였다.

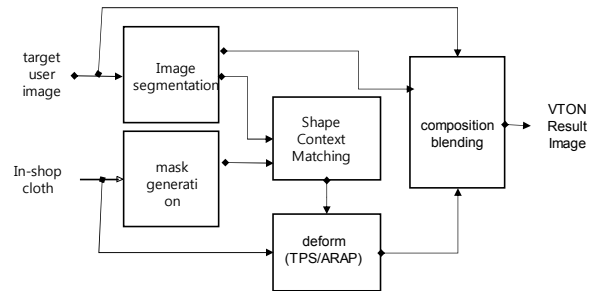


Fig. 2 SCMM based Try-On (Ahn, 2018b)

목적 인물에 맞게 변형된 의상은 인물의 분할된 다른 영역들과 고려하여 블렌딩 합성을 하여 최종결과 이미지를 생성한다.

2.2 VITON (Han et al., 2018)

Fig. 3의 VITON 알고리즘의 첫 단계는 대상인물의 실루엣과 조인트정보를 사용하여 의상과의 상관성을 학습시켜 변형된 목적 의상 마스크와 일차 착용결과를 생성한다. 그리고 이렇게 얻어진 마스크와 SCMM를 사용하여 변형된 의상을 생성한다. 두 번째 단계는 일단계의 결과를 이용하여 알파-블렌딩 맵을 만들어 이를 사용하여 최종 이미지를 생성한다. 이렇게 기하변형과 합성의 두 단계를 거치는 이유는 CNN 이 일반적으로 화소정도의 위치를 변형하는 기능이 어렵기 때문이다.

2.3 CP-VTON (Wang et al., 2018)

Fig. 4의 CP-VTON도 기본 구조는 VITON과 매우 유사하다. 의상의 기하변환을 행하는 GMM 단계와 의상을 합성하는 TON으로 구분된다. GMM 단계에서 VITON에서 처럼 마스크를 만드는 대신 STN (Spatial transfer network)을 사용하여 직접 변형된 결과를 생성한다. TON(Try-on)

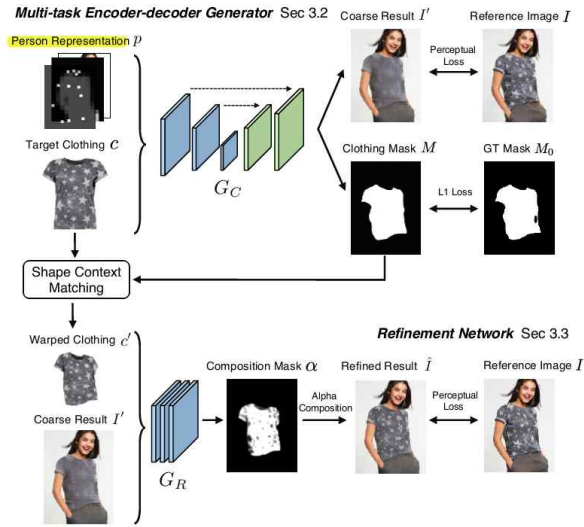


Fig. 3 VTON NN Pipeline (Han et al., 2018)

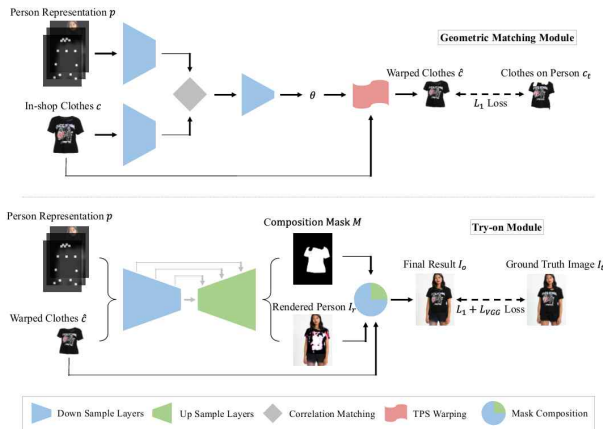


Fig. 4 CP-VTON NN Pipeline (Wang et al., 2018)

단계에서는 1단계 이미지를 사용하지 않고 원본 입력과 변형된 의상을 통하여 마스크를 구하고 이 마스크를 사용하여 최종 이미지를 생성한다. 논문에서 저자들은 CP-VTON이 의상의 텍스처를 VTON에 비하여 잘 유지한다고 주장하였다.

3. 실험 및 평가 방법

각 알고리즘이 제시된 논문들에서도 데이터

셋을 바탕으로 타 방식들과 비교한 실험결과를 제시하고 있다. 특히 CP-VTON 논문에서는 VTON을 포함한 다양한 방식과의 결과를 비교하였다. 그러나, 앞선 논문들에서의 실험결과 제시에는 크게 두 가지 문제점을 가지고 있다. 첫째는 사용 데이터가 상용화 시스템에 필요한 특성을 갖고 있는지 고려하고 있지 못하고, 실제로 데이터 셋이 충분한 대표성을 갖추고 있지 못하며 단순한 조건에 편중된 것을 볼 수 있다. 예를 들어 짧은 팔의 단색의 티셔츠가 주를 이룬다. 두 번째 문제는 결과를 객관적인 수치로 성능을 평가하고 있지 못하다.

본 논문은 이 두 문제를 가능한 해소하려 한다. 데이터 셋의 대표성 문제는 어느 누구도 의상의 특징에 대한 통계치를 정하기는 어려운 상황이므로, 통계적인 결과를 제시하는 대신, 테스트 샘플의 특성을 기준으로 분류하여 대표적인 결과를 제시하겠다. 객관적 성능평가의 문제는 동일한 의상을 사용한 실험을 통하여 기준 영상을 통한 성능을 IoU와 SSIM을 사용하여 제시하겠다. 또한 새로운 의상을 착용하는 실험은 현재로는 정답(Ground-truth)이 없으므로, 대신 동일한 결과를 주관적으로 비교하여 평가하는 것으로 하였다. 동일 모델이 두 개 이상의 의상을 착용한 데이터베이스가 추가로 구축되면 객관적 성능평가가 가능하다.

3.1 실험 데이터 분류

우선 실험 샘플들을 구분하는 기준은 Table 1에서 보는 바와 같이, 의상의 가려짐 정도, 대상 인물의 자세, 의상 자체의 복잡도로 구분하였다. 가려짐 정도는 변형의 대상의 정확성에 영향을 주는 요소이고, 자세는 변형의 정도이며, 의상 복잡도는 의상 자체의 처리 복잡도를 의미한다. 단, 구분의 범위에는 들어 있으나 실제 실험에서 제외된 경우는 괄호로 표시하였다. 제외한 조건의 경우는 테스트 데이터에 포함되어 있지 않거나, 현 기술로는 평가가 의미가 없을 정도로 복잡한 경우로 판단되는 것들이다. 이를 기준으로 다음과 같이 6가지 경우로 구분하였다.

- B: 가려짐이 거의 없고 자세가 정자세 (긴 팔, 짧은 팔)
- OP: S와 같으나 머리카락, 팔에 의하여 일부 가려짐
- OB: 하의에 의해 옷의 상당부분 가려짐이 있는 경우
- OF: 팔 등에 의하여 의상 전면이 가려짐이 있는 경우 (긴 팔 옷, 짧은 팔 옷)
- P: 큰 자세 변형이 있는 경우 (팔의 큰 움직임 또는 뒤틀리거나 옆 자세)
- S: 큰 체형 변화가 있는 경우 (전체 또는 일부가 굵은 체형이거나 임신)

참고로, 사용된 데이터의 의상은 (깃이 없는) T-셔츠 등으로 의상의 형태가 대체로 단순한 형태로 구성되어 있다. 또한 대부분의 의상이 단색이나 단조로운 무늬들로 되어 있는 등 제한적인 상황이라는 점도 염두에 둘 필요가 있다.

3.2 평가 방법과 기준

평가는 대상 인물사진이 입고 있는 동일한 상의를 다시 입히는 동일의상 착용실험과 새로운 의상을 착용하는 새 의상 착용실험으로 구분하였다. 동일의상 착용실험을 추가한 이유는 정답이 있는 경우를 통하여 정량적인 성능지표를 사용하여 결과를 비교해 보기 위함이다. 여기서 명확히 할 점은 동일한 의상을 수행하더라도 3

Table 1 Test Dataset Category

Criteria	Complexity Level
occlusion	external partial (hair, arms)
	external large (covered by pants) self-occlusion (long sleeves, legs-crossed)
pose	Standing front
	standing side
	fatness/pregnant (etc, backside, sitting, laying down)
shape complexity	short sleeves
	long sleeves (etc. multi-pieces, complex shape)

개의 알고리즘 모두 해당의상의 정보를 직접 사용하지 않는다. 물론 이 경우에 기존 의상 영역과 현재 의상의 외형이 유사함으로 인하여 다른 의상을 적용하는 경우보다는 착용결과가 우수할 수 있으나 상대적으로는 동일한 조건이다.

3개의 알고리즘 모두 의상을 형태를 변형하는 GMM (Geometric matching and manipulation) 단계와 변형된 의상을 대상에 입히는 TON 단계로 구성되므로, 이 두 가지 단계의 성능을 각각 확인하도록 한다. GMM 성능은 변형된 마스크와 목적 영역의 IoU (Intersection of union, 수식 (1))를 사용한다.

$$IoU = \frac{\tilde{C} \cap C_{GT}}{\tilde{C} \cup C_{GT}} \quad (1)$$

여기서 \tilde{C} 는 변형된 의상, C_{GT} 는 목표 이미지에서 의상이 위치해야하는 영역을 의미한다. 최종적으로 TON에 의하여 생성된 이미지의 성능은 평균 SSIM (Structural Similarity, 수식 (2))을 사용한다.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (2)$$

4. 실험 결과

실제 실험에 사용한 이미지는 각 종류별로 수개에서 수십 개를 사용하였으나, 본 논문에는 지면관계상 모두 포함하지 못하고, Fig. 5와 Fig. 6에 설명을 위하여 대표적인 이미지들만을 제시하였다.

4.1 동일의상 착용실험

동일 의상 실험의 경우는 IoU, SSIM과 시각적인 결과를 기준으로 성능을 비교하였다.

- B: 짧은 팔 의상의 경우는 세 가지 알고리즘 모두 높은 성능 보였다. 특히 SCMM의 경우 높은 성능을 보였다. 그러나 긴

팔의 경우 팔부분의 변형 결과가 좋지 않았다. 이는 매칭과 변형이 주로 전체 실루엣을 바탕으로 이루어지기 때문으로 보인다. VITON와 CP-VTON은 합성과정에서 일부 이를 보완하여 합성이 됨을 보여주고 있다. 골격정보를 포함하면 이를 보완할 수도 있으나, 현재는 의상의 골격정보를 자동으로 추출해주는 알고리즘이 개발되어 있지 않다.

- OP: 머리카락 등에 의하여 부분적인 가려짐이 발생하면 이 부분이 목적의상 영역에서 제외됨으로 의상의 변형에 일부 영향을 미침을 보였다. VITON과 CP-VITON도 사람 표현에 머리와 피부를 포함하고 있기 때문에 마찬가지로 문제점을 보였다. 따라서 GMM 파트에는 머리카락 등 요소는 제거하고 신체의 형태만을 사용하도록 하면 개선이 가능할 것으로 예측된다.
- OB: 하의에 의한 가려짐이 발생하는 경우, SCMM 알고리즘들은 변형과 가려짐을 구별하지 못한다. 따라서 IoU에 저하를 보이고, 특히 긴 팔이 줄어들어 피부가 들어나는 문제를 보인다. VITON, CP-VITON은 의상자체가 아니라 체형 정보를 사용하므로 이러한 영향은 줄어드는 것을 보인다.
- OF: 팔에 의하여 전면부가 가려지는 경우는 사람 표현을 이용하면 옷의 영역과 팔의 영역을 구분할 수 있지만 딥러닝 합성시에 이 부분이 선명하지 못한 경우가 발생하는 경우가 있다.
- P: 큰 포즈 변화가 있는 경우 세 알고리즘 모두 의상 변형에 큰 오류가 발생한다. 이는 이차원 알고리즘을 사용하는 큰 한계점이라고 판단된다.
- S: 동일의상의 데이터의 경우, 의상 자체가 크게 준비 되어있어서 오류가 크지 않았다.

4.2 새 의상 착용실험

새 의상을 착용하는 경우가 사실상 궁극적으로 응용에 필요한 결과이다. 앞서와 같이 객관적인 평가가 가능하면 좋겠으나, 현재 한 모델

이 두 개 이상의 의상을 착용한 데이터셋이 구축되지 못하여 이를 제시하지 못하였다. 제한적이기는 하지만 시각적인 결과를 바탕으로 각 알고리즘의 상대적 차이점을 비교가 가능하다.

- B: 의상변형 자체는 동일의상의 결과와 유사하게 좋은 결과를 보이나, 합성 시에 알고리즘에 따른 약간의 차이를 보인다. 긴팔에서 짧은 팔로의 교체 시에 SCMM은 팔의 피부색을 복원하지 못하고 VITON과 CP-VTON은 이를 생성해내는 것을 보인다. 특히 CP-VTON은 생성 능력이 우수함을 보인다. 이것이 비딥러닝 방식에 비하여 pix2pix기반 딥러닝을 사용하는 장점이라고 볼 수 있다.
- OP: 동일 의상의 결과와 유사한 정도의, 중간적인 성능을 보였다.
- OB: 동일영상과 같은 특성, 즉, SCMM의 경우 가려짐과 변형을 구별하지 못하는 문제를 보였다.
- OF: 동일의상과 같은 특성을 보였다. 여기서도 특히 SCMM의 경우 짧은 팔에서 긴 팔로 교체 시 합성 알고리즘에 개선이 필요한 것으로 판단되었다.
- P: 동일의상의 경우와 같이 변형에 큰 오류가 있는 것을 보였다.
- S: 체형변화에는 어느 정도 잘 적응하는 것을 보였다. 특히 VITON 과 CP-VTON은 체형 변화에 매우 잘 적응하는 것을 보였다.

4.3 종합적인 결과 분석

앞서 각 조건별 분석이외에 종합적인 특성에 대하여 분석하면, 다음과 같은 큰 특징이 있음을 볼 수 있다. 첫째, GMM에 의한 의상 형태 변형은 현재 착용의상에 따른 영향이 있음을 확인하였다. 그 이유는 SCMM은 현재의상의 영역을 사용하고 있으며, 또한 딥러닝 방식들은 체형자체를 사용하기는 하나 학습과정에서 사용한 정답 마스크에 현재 의상의 영역이 반영된다.

이러한 현상의 OB 경우에 특히 심각하며, OP의 경우에도 영향을 줄 수 있다. 따라서 이

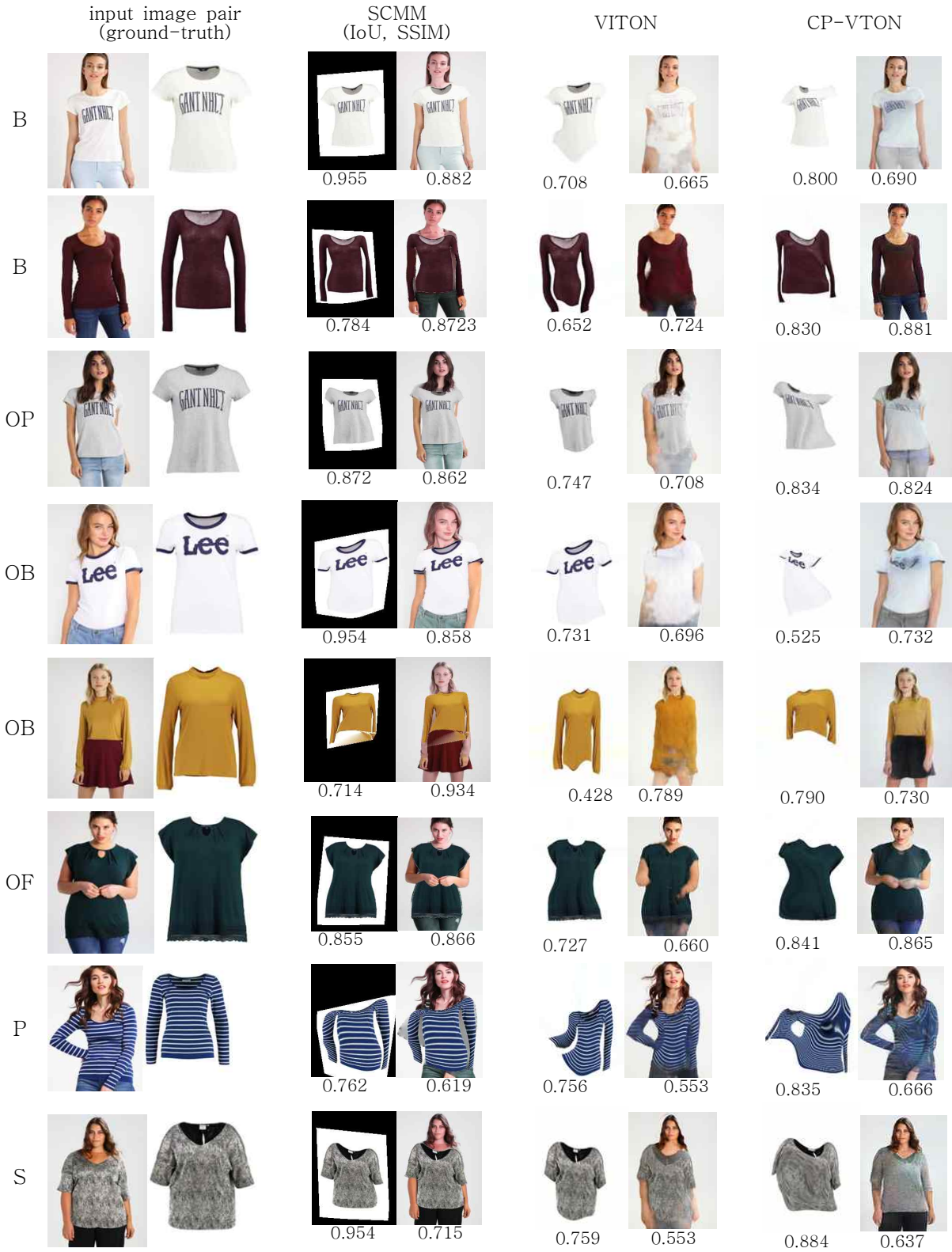


Fig. 5 VTON Results with the Original Cloth

Performance Evaluation of VTON (Virtual-Try-On) Algorithms using a Pair of Cloth and Human Image

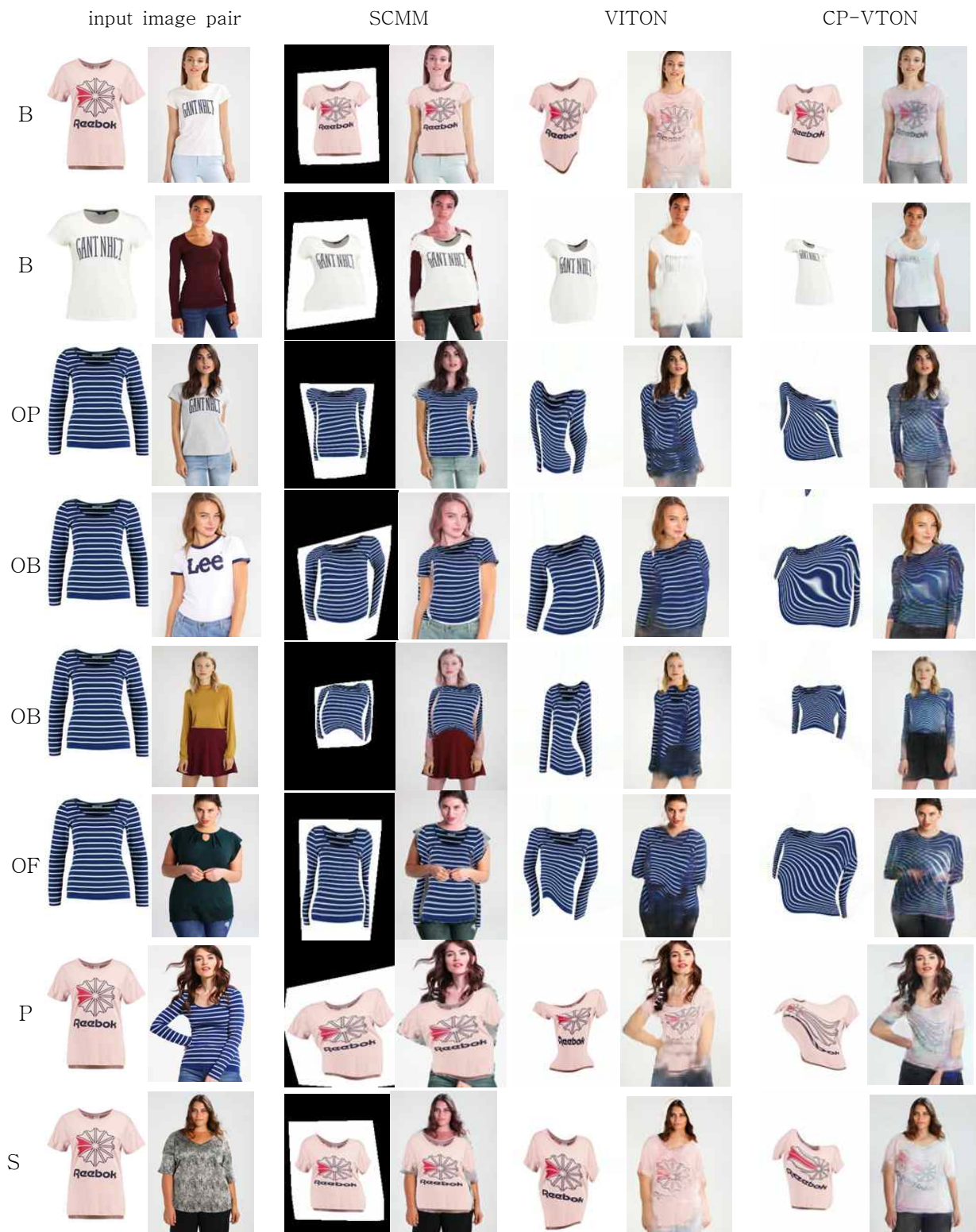


Fig. 6 VTON Results with a New Cloth

부분은 사람표현방식과 GT에 이를 반영하는 것이 필요하다. 둘째, 딥러닝 방식은 합성단계에서 알과 블렌딩을 사용하고 있는데 이 때문에 전체적으로 의상의 영역의 선명도가 SCMM방식에 비하여 떨어지는 것을 볼 수 있다. 시각적으로는 느끼기 어려운 차이가 발생하는 경우도 SSIM으로는 상당한 차이로 검출되는 것을 확인할 수 있다. 따라서, 객관적 성능수치와 주관적인 결과는 차이가 발생할 수 있음을 성능평가지 고려하여야 할 것이다. 셋째, TPS 알고리즘과 같이 2차원 변형알고리즘을 사용해서는 의상자체 가려짐이나 의상이 크게 변형되는 자세는 전혀 처리가 되지 않음을 확인할 수 있다. 이는 2차원 변형알고리즘 자체의 한계라고 볼 수 있다.

5. 결 론

본 논문에서는 최근 발표된 이미지기반 딥러닝 기반 가상착용기술인 VITON 및 CP-VTON과 비-딥러닝 방식의 SCMM 방식을 실용적인 측면에서 분석하기 위하여, 의상의 복잡도, 대상 자세 및 체형, 의상의 가려짐 정도 등에 따라 분석하였다. 기존 논문에 제시하지 못하였던 정량적인 분석을 위하여 현재 목적 인물이 입고 있는 동일한 의상을 다시 입히는 경우의 IoU와 SSIM를 비교하였고, 새로운 옷을 입히는 경우는 각 알고리즘을 결과를 상대비교를 통하여 분석하였다. 그 결과, 동일 의상을 착용하는 실험에서는 SCMM 방식과 CP-VTON이 비슷한 정도로 가장 좋은 성능을 보였고, 새로운 의상을 착용하는 경우는 CP-VTON이 가장 좋은 주관적인 성능을 보였다. 그러나, 기존 논문에서 주장하는 바와 달리, 의상의 가려짐이 거의 없고, 자세가 변형이 거의 없는 경우를 제외하면 실용적인 기준에 미치지 못하는 것으로 평가되었다. 그 주요한 원인은 실제 3차원 의상변형의 2차원 기하변형 근사함에 있어서의 한계와 GAN을 통한 합성 기술의 한계로 파악 되었다.

이 결과를 바탕으로 보면 2D VTON 방식은 적용할 수 있는 이미지의 조건이 상당히 제한적

임을 확인하였다. 따라서, 현재 기술을 실용화하기 위해서는 제한된 조건을 만족하는 입력 이미지만을 허용하게 하거나, 제한된 조건을 만족하도록 사전 처리하는 방법을 고려할 수 있다. 후자의 방법에 관련되어 Weng et al. (2019)에서 입력 이미지의 사람을 SMPL (Skinned multi-person linear) 기반의 3D 모델을 만들고, 이를 이용하여 자세 변환이 가능함을 보였다. 이 기법을 이용하여 목적 사용자를 2D VTON 방식이 잘 동작하는 조건의 기본 포즈로 변형 하고, 기존이 알고리즘으로 가상착용을 실행한 후, 다시 원상태로 복원을 하게 된다면 실용적인 측면에서 큰 발전이 있을 것으로 보인다.

References

- Ahn H., (2018a). Online Virtual Try On using Mannequin Cloth Pictures, *Journal of the Korea Industrial Information Systems Research*, 23(6), 29 - 38.
- Ahn H., (2018b). Image-based Virtual Try-On System, *Journal of Korean Computer Game Society*, 31(3), 37-45.
- Belongie, S., Malik, J., and Puzicha, J. (2002). Shape Matching and Object Recognition using Shape Contexts, *IEEE Transactions on PAMI*, 25(4), 509-522.
- Cao, Z., Simon, T., Wei, S. E., and Sheikh, Y. (2017). Realtime Multi-person 2d Pose Estimation using Part Affinity Fields, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7291-7299.
- Han, X., Wu, Z., Wu, Z., Yu, R., and Davis, L. S. (2018). Viton: An Image-based Virtual Try-on Network, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7543-7552.
- Liang, X., Gong, K., Shen, X., and Lin, L. (2018). Look into Person: Joint Body Parsing & Pose Estimation Network and a

New Benchmark. *IEEE Transactions on PAMI*, 41(4), 871-885.

Raj, A., Sangkloy, P., Chang, H., Lu, J., Ceylan, D., and Hays, J. (2018). Swapnet: Garment Transfer in Single View Images, *Proceedings of the European Conference on Computer Vision*, pp. 666-682.

Wang, B., Zheng, H., Liang, X., Chen, Y., Lin, L., and Yang, M. (2018). Toward Characteristic-preserving Image-based Virtual Try-on Network, *Proceedings of the European Conference on Computer Vision*, pp. 589-604.

Weng, C. Y., Curless, B., and Kemelmacher-Shlizerman, I. (2019). Photo Wake-up: 3d Character Animation from a Single Photo, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5908-5917.

Zanfir, M., Popa, A. I., Zanfir, A., and Sminchisescu, C. (2018). Human Appearance Transfer, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5391-5399.



파이 트안 투안 (Thai Thanh Tuan)

- HUTECH (베트남) 컴퓨터공학과 학사 (2010)
- HCMVNU (베트남) 컴퓨터공학 석사 (2015)
- (현재) 서울과학기술대학교 전기정보공학과 박사과정
- 관심분야 : 컴퓨터비전, 딥러닝, 데이터마이닝



미나르 마드올 라흐만 (Matiur Rahman Minar)

- BUET (방글라데시) 컴퓨터공학과 학사 (2015)
- Automation Solutionz Inc., (캐나다) 2014-2018, 방글라데시 원격근무 프로그래머
- (현재) 서울과학기술대학교 전기정보공학과 석사과정
- 관심분야 : 컴퓨터비전, 딥러닝, 데이터마이닝



안 희 준 (Heejune Ahn)

- 종신회원
- KAIST 전기정보공학과 박사 (2000)
- (주) LG전자 차세대단말연구소 선임연구원(1998-2002)
- (주) Tmax 소프트 책임연구원 (2002-2004)
- (현재) 서울과학기술대학교 전기정보공학과 교수
- 관심분야 : 컴퓨터비전, 컴퓨터 통신, 데이터마이닝