

최적 전송을 이용한 웨이블릿 변환 기반 실사적 스타일 변환

배 경 민*, 박 대 영°

Wavelet Transform-Based Photorealistic Style Transfer Using Optimal Transport

Gyeongmin Bae*, Daeyoung Park°

요 약

스타일 변환(style transfer)은 콘텐츠 이미지(content image)의 공간 정보를 유지하면서도 콘텐츠 이미지의 스타일을 스타일 이미지(style image)의 스타일로 바꾸는 과정이다. 기존 연구로는 스타일 분포(style distribution)의 second-order statistics가 같아지도록 그람 행렬(Gram matrix)을 이용하는 방법과 instance normalization을 이용하는 방법 등이 있다. 본 논문은 이 방법들을 일반화시킨 최적 전송(optimal transport)을 웨이블릿(Wavelet) 변환에 적용해서 사진과 비슷한 이미지를 생성하는 실사적 스타일 변환(photorealistic style transfer)에서 스타일 변환 성능을 개선하는 방법을 제안한다.

Key Words : Photorealistic style transfer, Spatial information, Style distribution, Optimal transport

ABSTRACT

Style transfer is the process of changing the style of a content image to the style of a style image while maintaining its spatial information. Conventional studies include the Gram matrix or the instance normalization so that the second-order statistics are the same. This paper applies a more generalized method based on the optimal transport to the wavelet transform in order to improve the style conversion performance for photorealistic style transfer that creates images similar to photos.

I. 서 론

스타일 변환(style transfer)은 콘텐츠 이미지(content image)가 공간 정보를 유지하면서도 스타일 이미지(style image)의 스타일 분포를 따르도록 변환시키는 과정이다¹⁻⁸⁾. 콘텐츠 이미지의 변화 정도에 따라 크게 예술적 스타일 변환(artistic style transfer)과

실사적 스타일 변환(photorealistic style transfer)으로 분류할 수 있다. 예술적 스타일 변환은 콘텐츠 이미지에 인위적인 현상을 남겨 좀 더 예술적인 결과물을 얻는다. 반면에 실사적 스타일 변환은 인위적인 현상을 줄여서 실제 사진과 같은 이미지를 얻는다.

최적 전송(optimal transport)은 확률 분포 간의 거리를 최소화시키는 최적의 경로를 구하는 매핑

* 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (2020-0-01389, 인공지능융합연구센터지원(인하대학교)). 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2019R1A2C1005512).

• First Author : Department of Electrical and Computer Engineering, Inha University, bgm1425@gmail.com, 학생회원

° Corresponding Author : Department of Information and Communication, Inha University, dpark@inha.ac.kr, 종신회원

논문번호 : KICS202104-079-C-RU, Received April 7, 2021; Revised June 14, 2021; Accepted June 24, 2021

(mapping)인 Monge 맵을 구하는 기술이다⁸⁾. 가우스 측도(Gaussian measure)을 이용한 최적 전송은 닫힌 형태(closed form)의 Monge 맵(map)으로 구할 수 있다⁹⁾. 본 논문에서는 가우스 측도 기반 최적 전송을 웨이블릿 변환에 적용하여 기존 실사적 스타일 변환 모델의 스타일 손실을 최소화하는 방안을 제시한다.

II. 스타일 변환

Convolutional neural network (CNN)을 이용한 스타일 변환의 최초 연구는 VGG 모델을 이용하였다. 콘텐츠 이미지와 스타일 이미지를 VGG에 입력으로 넣어 특징(feature)을 추출하고, 콘텐츠 이미지의 특징을 변환시켜서 얻은 그람 행렬(Gram matrix)이 스타일 이미지의 그람 행렬과 비슷해지도록 콘텐츠 이미지의 특징을 변환시켰다¹⁾. 그람 행렬 사이의 손실 함수를 최소화시키는 것이 2차 다항식 커널을 이용한 maximum mean discrepancy(MMD)를 최소화하는 것과 같다는 것이 증명되면서, 그람 행렬에 대한 손실 함수를 최소화하는 것과 스타일 분포로 변환시키는 것이 동등한 것이 밝혀졌다²⁾. 그러나 그람 행렬을 이용하는 방법은 연산 속도가 매우 느리기 때문에 이를 개선하기 위해 instance normalization의 개념을 이용한 adaptive instance normalization(AdaIN)이 연구되었다³⁾. 인코더/디코더 구조를 이용하여, 인코더 맵의 출력으로 얻는 특징 공간(feature space)에서 콘텐츠 분포의 평균과 분산을 스타일 분포의 평균과 분산과 같도록 변환시킨다. 우선, 입력 값과 똑같은 출력을 내도록 인코더와 디코더를 학습하고, 원하는 콘텐츠 이미지와 스타일 이미지를 인코더로 통과시켜 얻는 특징의 평균과 분산을 구해서

$$AdaIN(x,y) = \sigma(y)\left(\frac{x-\mu(x)}{\sigma(x)}\right) + \mu(y) \quad (1)$$

와 같이 출력 이미지 특징을 구하고 이것을 디코더의 입력으로 넣어 스타일 변환된 이미지(style-transferred image)를 얻는다³⁾.

그러나 위 식 (1)은 평균과 분산만 맞추는 것으로 second-order statistics가 갖지 않기 때문에 공분산까지 맞추는 whitening and coloring transforms(WCT)가 연구되었다⁴⁾. 인코더를 통과한 콘텐츠와 스타일의 특징 $f_s \in R^{C \times H_s \times W_s}$, $f_c \in R^{C \times H_c \times W_c}$ 를 이용하여 그람 행렬의 고유 값 분해 $f_s f_s^H = \Sigma_s \Lambda_s \Sigma_s^H$, $f_c f_c^H = \Sigma_c \Lambda_c \Sigma_c^H$

를 구한다. 그 후

$$\hat{f}_c = \Sigma_c \Lambda_c^{-\frac{1}{2}} \Sigma_c^H f_c \quad (2)$$

$$\hat{f}_{cs} = \Sigma_s \Lambda_s^{\frac{1}{2}} \Sigma_s^H \hat{f}_c \quad (3)$$

로 변환하여 스타일 분포의 공분산을 맞춘다. 한 번에 좋은 성능이 나오지 않아 5개의 층에서 이 과정을 반복하는 다층 구조(multi-level)를 가진다. 다층 구조는 연산속도가 느리고 에러가 증폭되는 단점이 있다⁴⁾. 지금까지 소개한 방법은 예술적인 결과물을 얻을 수 있지만 콘텐츠 이미지의 구조 정보가 손실되어 실사적이지 않다.

실사적 스타일 변환 모델로서 DPST⁵⁾, PhotoWCT⁶⁾ 등이 연구되었다. PhotoWCT 모델은 VGG 모델에서 구조 정보의 손실을 줄이기 위해 maxpooling된 부분을 기억했다가 unpooling할 때 사용하고, 성능을 향상시키기 위해서 여전히 다층 구조를 가진다. 따라서, 연산속도가 여전히 느리고 에러가 증폭되며 모델의 크기가 매우 커진다는 단점을 가진다. 그리고 결과물의 공간 정보의 인위적(artifact) 현상을 처리하기 위해서 콘텐츠 이미지의 채널 간 공간 정보에 대한 정규화 식(regularization term)을 추가하고 평활화(smoothing)하는 후처리 과정을 추가했다. 후처리 과정으로 인해 연산 속도가 느릴 뿐 아니라 필요한 메모리가 크게 증가하였으며, 평활화된 결과를 얻는다는 단점이 존재한다⁶⁾.

다른 실사적 변환 모델인 WCT² (wavelet corrected transfer based on whitening and coloring transform)¹⁾는 기존 연구의 반복적 모델의 단점과 CNN기반 모델에서 많이 사용하는 maxpooling의 단점을 보완하였다⁷⁾. WCT²는 maxpooling 대신 wavelet corrected transform을 사용하고 각 층마다 whitening and coloring transforms (WCT)를 이용하여 특징 변환을 수행한다. 웨이블릿 pooling은 저주파 통과 필터(low-pass filter)와 고주파 통과 필터(high-pass filter)

$$L^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad H^T = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \end{bmatrix} \quad (4)$$

1) 참고로, WCT는 특징 변환 매핑인 whitening and coloring transforms 뿐 아니라 스타일 변환 모델인 wavelet corrected transfer의 약자에 해당한다.

를 외적하여 LL, LH, HL, HH 총 4개의 커널(kernel)을 가진다. 저주파 통과 필터는 부드러운 표면과 질감 정보를 포착하고, 고주파 통과 필터는 수직적, 수평적, 대각선의 모서리 정보를 추출한다. 웨이블릿 unpooling은 pooling의 반대과정으로써 원본 이미지를 복원한다. VGG 모델의 각 convolutional 층 다음에 WCT를 추가하여 하나의 인코더/디코더 구조를 구성했다. WCT는 DPST에 비해 연산속도가 매우 빠르고 PhotoWCT의 절반 정도의 메모리만을 사용했으며, SSIM 지수와 그램 손실 측면에서 기존 연구보다 나은 성능을 보였다⁷⁾. 본 논문은 웨이블릿 변환에 WCT가 아닌 최적 전송(optimal transport)를 적용하는 스타일 변환을 제안한다.

III. Optimal Transport를 이용한 스타일 변환

기존에는 스타일을 변화시키기 위한 특징 변환(feature transform)으로써 AdaIN과 WCT이 사용되었다. 본 논문에서는 최적 전송(optimal transport)하는 Monge 맵을 특징 변환으로 사용한다.

인코딩 맵. 콘텐츠 이미지 I_c 와 스타일 이미지 I_s 를 n 개의 맵 $F_j: R^d \rightarrow R^m, j=1, \dots, n$ 으로 특징을 추출한다. 여기에서 d 는 이미지의 차원을 나타내고 m 은 특징의 차원을 나타낸다. 이를 이용해서 인코딩 맵을

$$E: I \in R^d \rightarrow \nu_I = \frac{1}{n} \sum_{j=1}^n \delta_{F_j(I)} \in P(R^m) \quad (5)$$

와 같이 정의한다. 위 식 (5)에서 $P(R^m)$ 은 R^m 에서 경험적 측도(empirical measure)의 공간을 나타낸다. 그리고 $\nu = E(I)$ 를 경험적 분포(empirical distribution)라고 한다. 예를 들어, F 는 이미지를 $R^{C \times HW}$ 로 매핑해주는 VGG 모델이라 할 수 있으며, $m = C, n = HW$ 이다. 스타일 이미지 특징의 차원은

$$F_c \in R^{C \times H_c \times W_c}, F_s \in R^{C \times H_s \times W_s} \quad (6)$$

와 같이 표현할 수 있다.

디코딩 맵. 인코더 E 가 가역적이어서 $D(E(I)) = I$ 라 가정하자. 많은 경우에 가역적인 VGG 모델을 이용해 인코더/디코더를 구성한다^{3,4,6,7)}.

$\nu_c = E(I_c)$ 를 $\nu_s = E(I_s)$ 로 변환하는 최적의 매핑은

$$\begin{aligned} \operatorname{argmin}_T \int \|x - T(x)\|_2^2 d\nu_c(x) \\ \text{s.t. } T_{\#}(\nu_c) = \nu_s \end{aligned} \quad (7)$$

를 풀어서 얻는 Monge 맵과 같다^[8]. 식의 최적값(optimal value)은 ν_c 와 ν_s 의 2-Wasserstein distance $W_2^2(\nu_c, \nu_s)$ 이다^[8]. CNN 기반 모델의 특징 공간은 실제로는 가우시안이 아니지만, 중심극한정리를 적용하여 가우시안이라 가정한다. ν_c 와 ν_s 가 가우시안이라면 평균과 공분산에 의해서 2-Wasserstein distance가 닫힌 형태(closed-form)로 주어진다. 따라서, 식 (7)의 최적해인 Monge 맵은

$$T_{\nu_c \rightarrow \nu_s}(x) = \mu_s + A_{c \rightarrow s}(x - \mu_c) \quad (8)$$

$$A_{c \rightarrow s} = \Sigma_c^{-\frac{1}{2}} \left(\Sigma_c^{\frac{1}{2}} \Sigma_s \Sigma_c^{\frac{1}{2}} \right)^{\frac{1}{2}} \Sigma_c^{-\frac{1}{2}} \quad (9)$$

와 같으며, μ_c, Σ_c 는 ν_c 의 평균과 공분산을 나타내고 μ_s, Σ_s 는 ν_s 의 평균과 공분산을 나타낸다^[10].

본 논문에서 최종적으로 제안하는 스타일 변환은

$$I_{c \rightarrow s} = D(T_{\nu_c \rightarrow \nu_s}(E(I_c))) \quad (10)$$

와 같다. Wavelet corrected transfer 모델에 식 (7)의 최적 매핑함수를 적용하는 것이므로 WCT-OT(wavelet corrected transfer based on optimal transport) 모델이라 부른다.

WCT-OT의 최적 맵은 다음과 같이 기존 방식의 준 최적인 맵을 유도할 수 있다.

1) 만약 식 (8)에서 공분산의 곱셈에 대해서 교환법칙이 성립하여 $\Sigma_c \Sigma_s = \Sigma_s \Sigma_c$ 를 만족한다면, 최적 맵은

$$T_{\nu_c \rightarrow \nu_s}^{WCT}(x) = \mu_s + \Sigma_s^{\frac{1}{2}} \Sigma_c^{-\frac{1}{2}} (x - \mu_c) \quad (11)$$

와 같이 whitening and coloring transforms (WCT)^[4]로 표현 할 수 있다.

2) 특징이 서로 상관없이 없어서(uncorrelated) 공분산 행렬이 대각 행렬인 경우에는 최적 맵이

$$T_{\nu_c \rightarrow \nu_s}^{AdaIN}(x) = \mu_s + \sqrt{\sigma_s} \odot \frac{(x - \mu_c)}{\sqrt{\sigma_c}} \quad (12)$$

이 되어 AdaIN^[3]의 경우에 해당한다.

위 내용을 정리하면 WCT-OT는 기존 연구의 특징 변환 맵들을 포괄하는 매우 일반화된 방법임을 알 수 있다. 또한 식 (11), 식 (12)은 공분산 행렬이 특정한 경우에만 최적 맵인 식 (8)과 일치하므로 WCT와 AdaIN은 최적이지 아닌 준최적이라 할 수 있다.

본 논문에서 제안한 WCT-OT는 최적 Monge 맵을 이용하여 콘텐츠 이미지 분포를 스타일 이미지 분포로 매핑한다. 기존 연구들이 단순히 평균과 분산 그리고 공분산을 맞추는 것과 비교하면, 제안하는 WCT-OT는 분포 자체를 맞추기 때문에 실험 결과에서 정량적인 평가 수치와 정성적인 이미지의 품질 면에서 향상된 스타일 변환을 달성할 것으로 예상할 수 있다.

IV. 실험 결과

학습된 WCT² 모델에서 특징 변환을 본 논문에서 제안하는 식 (8)의 최적 전송을 사용하여 WCT-OT를 구현하였다. 식 (8)의 $\mu_c, \mu_s \in R^C$ 는 각각 콘텐츠와 스타일 이미지를 인코더에 통과시켜서 얻은 식 (6)으로 주어지는 특징 행렬의 열벡터의 평균으로 구한다. 그리고 $\Sigma_c, \Sigma_s \in R^{C \times C}$ 는 식 (6)의 특징 행렬의 열벡터에 앞서 구한 평균을 빼 것의 그람 행렬의 평균으로 구한다.

그림 1은 WCT²와 WCT-OT를 이용하여 콘텐츠 이미지의 스타일을 변화시켜 얻은 그림을 비교한다. 각 모델은 LL 커널을 통과한 것에 대해서만 특징 변환을 수행한다. 스타일 이미지는 전체적으로 어두우며 노란색 불빛이 건물 외벽에 반사되고 있다. WCT²를 통해

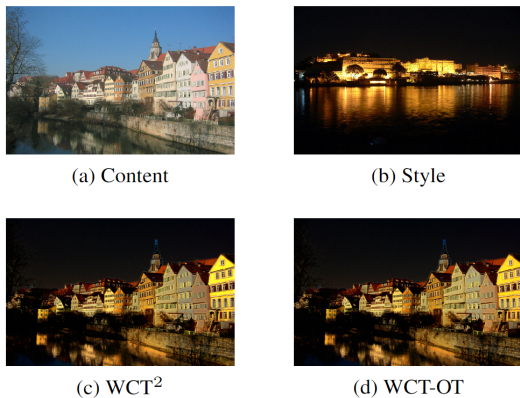


그림 1. 풍경 사진의 스타일 변환 비교
Fig. 1. Style transfer comparison of a landscape image.

스타일 변환된 이미지에서 가장 우측 노란색 건물의 밝기가 다소 차이가 있는 것을 눈으로 확인할 수 있다.

그림 2는 노란 꽃 사진을 음영이 존재하는 빨간색 장미의 스타일 변환된 이미지를 비교한다. WCT²는 스타일 이미지의 그림자 부분을 잘 표현하지 못한다. 반면, 제안 하는 모델인 WCT-OT는 그림자 부분을 잘 표현하고 있다.

그림 3은 그림 2의 꽃 사진의 스타일 변환된 이미지와 스타일 이미지의 히스토그램을 비교한다. 히스토그램은 컬러 이미지를 그레이 스케일로 변환시킨 후 각 밝기(intensity)에 해당하는 픽셀의 빈도수를 나타내며, 픽셀의 밝기에 대한 분포를 파악할 수 있다^[11]. WCT-OT로 스타일 변환된 이미지가 WCT²로 스타일 변환된 이미지보다 히스토그램이 스타일 이미지와 비슷한 분포를 나타내는 것을 확인할 수 있다. 이미지 픽셀의 히스토그램이 균일(uniform)한 분포에 가까울수록 밝은 부분과 어두운 부분의 차이인 콘트라스트(contrast)가 커져 물체의 윤곽을 알아보기 쉽다^[11].

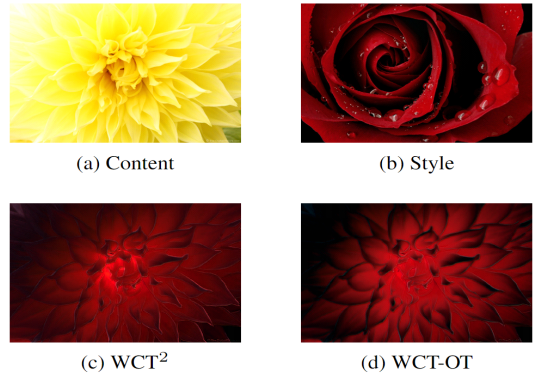


그림 2. 꽃 사진의 스타일 변환 비교
Fig. 2. Style transfer comparison of a flower image.

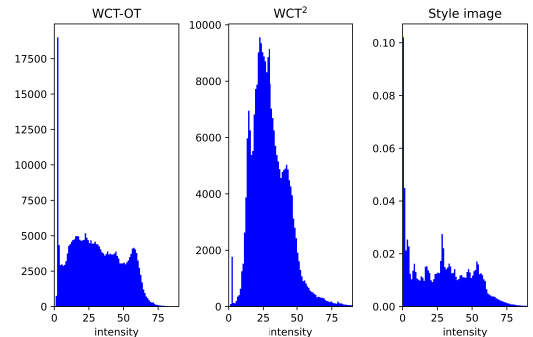


그림 3. 꽃 사진에 대한 각 모델의 출력 이미지와 스타일 이미지의 히스토그램
Fig. 3. Histogram of output image of each model and style image for a flower image.

WCT-OT는 WCT²보다 히스토그램이 더 균일한 분포에 가까우므로 콘트라스트가 더 클 것으로 예상할 수 있다.

그림 4는 위의 결과를 뒷받침하기 위하여 그림 2의 꽃 사진의 스타일 변환된 이미지와 스타일 이미지의 밝기 값의 평균 제곱근 (Root Mean Square (RMS)) 콘트라스트를 비교한다^[12]. 이미지 픽셀의 개수 n 개인 경우 평균 제곱근 콘트라스트는

$$RMS = \left[\frac{1}{n} \sum_{i=1}^n (x_i - \tilde{x})^2 \right]^{\frac{1}{2}} \quad (13)$$

으로 정의된다. 여기서 x_i 는 표준화된 밝기를 의미한다. 즉, $0 \leq x_i \leq 1$ 의 범위를 가진다. 그리고 \tilde{x} 는

$$\tilde{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (14)$$

와 같이 표현할 수 있고 표준화된 밝기의 평균값을 의미한다. WCT-OT로 스타일 변환된 이미지는 WCT²로 스타일 변환된 이미지보다 평균 제곱근 콘트라스트의 값이 크다. 그림 3과 그림 4를 통해서 이미지 히스토그램이 좀 더 균일한 분포를 갖는 WCT-OT가 WCT²보다 콘트라스트 값이 크다는 것을 확인할 수 있다.

그림 5는 각 모델에서 얻은 스타일 변환된 이미지를 콘텐츠 이미지 및 스타일 이미지와의 Frechet

Root Mean Square (RMS) Contrast		
WCT-OT	WCT ²	Style image
0.62	0.44	0.84

그림 4. 꽃 사진에 대한 각 모델의 출력 이미지와 스타일 이미지의 평균 제곱근 콘트라스트.

Fig. 4. Root mean square contrast of each model and style image for a flower image.

Frechet Inception Distance (FID)		
Model	Content image	Style image
WCT-OT	104.7	179.3
WCT-OT(multi-level)	124.0	181.8
WCT ²	107.3	179.5
WCT ² (multi-level)	124.6	185.7

그림 5. 스타일 변환된 이미지 집합과 콘텐츠 이미지 집합의 FID 및 스타일 변환된 이미지 집합과 스타일 이미지 집합의 FID.

Fig. 5. FID between the style-transferred images and the content images and FID between the style transferred images and the style images.

Inception Distance(FID)^[13]를 비교한다. 다층 구조는 동일한 모델을 연속으로 3번 통과시키는 것을 의미한다^[7]. FID는 2개의 이미지의 픽셀 또는 2개의 자료 집합의 분포 거리를 나타내며, 각 이미지를 인셉션 모델 (Inception model)^[14]에 입력하여 얻는 특징 벡터의 평균과 분산을 이용하여 계산한다. 서론에서 언급한 것과 같이 스타일 변환의 목적은 스타일 변환된 이미지가 콘텐츠 이미지의 구조를 유지하면서 스타일 이미지의 분포와 같아지도록 하는 것이다. 스타일 변환된 이미지가 콘텐츠 이미지와의 FID가 작다면 콘텐츠를 스타일 이미지를 스타일 변환시키는 과정에서의 왜곡이 적은 것을 의미하며, 스타일 이미지와의 FID가 작다면 스타일 이미지와 비슷한 스타일을 가지는 것을 의미한다. 따라서 스타일 변환된 이미지가 콘텐츠 이미지 및 스타일 이미지와의 FID가 모두 작다면 스타일 변환이 잘 되었다고 할 수 있다. 그림 5에서 WCT-OT의 스타일 변환된 이미지가 WCT²의 스타일 변환된 이미지보다 콘텐츠 이미지와의 FID 및 스타일 이미지와의 FID가 모두 작은 것을 확인할 수 있다. 이러한 현상은 다층 구조의 WCT-OT 및 다층 구조 WCT²를 비교하여도 동일하게 관찰할 수 있다.

그림 6은 스타일 변환된 이미지를 structural similarity(SSIM) 지수와 그람 손실 측면에서 스타일 변환 성능을 비교한다. SSIM 지수는 이미지의 유사성을 나타내며, 콘텐츠 이미지와 스타일 변환된 이미지의 에지 응답(edge response)^[15]에 대해서 비교한다. 0이면 두 이미지가 완전히 다른 것을 의미하고 1이면 완전히 동일한 것을 의미하므로 수치가 높을수록 두 이미지의 구조 정보가 유사하다는 것을 나타낸다. 한

SSIM Index and Gram Loss

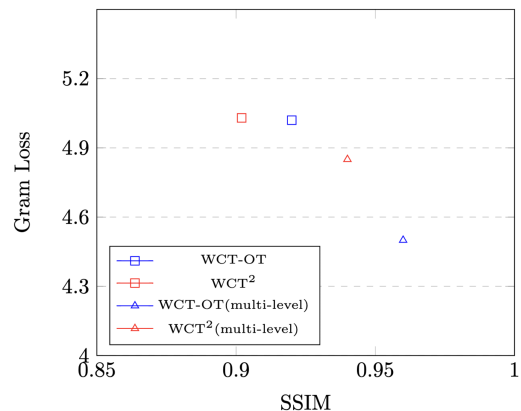


그림 6. SSIM 지수와 그람 손실.
Fig. 6. SSIM index and Gram loss.

편, 그람 손실은 두 이미지 간의 스타일의 차이를 나타내며, VGG 모델에 스타일 이미지와 스타일 변환된 이미지를 통과시켜 얻은 특징 벡터들의 그람 행렬의 mean squared error(MSE)로 계산하며 낮을수록 좋은 것을 나타낸다. WCT-OT는 WCT²보다 SSIM 지수를 높이면서도 그람 손실을 낮추었다. 다층 구조의 경우에도 WCT-OT가 WCT²보다 SSIM 지수와 그람 손실 측면 모두에서 더 나은 스타일 변환 성능을 보였다.

V. 결 론

본 논문은 인코더/디코더 구조의 wavelet corrected transfer (WCT)에 최적 전송(optimal transport)을 특징 변환으로 사용하는 WCT-OT를 제안하였다. 제안하는 최적 전송으로 얻은 맵을 이용하면 기존의 준최적 맵을 유도할 수 있다. WCT-OT는 WCT²보다 스타일 변환된 이미지의 히스토그램이 스타일 이미지와 비슷한 분포를 가질 뿐 아니라, 히스토그램이 더 균일한 분포에 가까워서 좀 더 높은 콘트라스트를 갖는다. 또한, WCT-OT는 이미지 구조 정보의 왜곡을 나타내는 SSIM 지수와 이미지 간의 스타일의 차이를 나타내는 그람 손실 측면에서 모두 WCT²보다 더 우수한 스타일 변환 성능을 나타내었다. 뿐만 아니라, WCT-OT는 WCT²보다 콘텐츠 이미지 및 스타일 이미지와의 FID도 낮은 것을 확인할 수 있었다. 따라서, 제안하는 웨이블릿 기반의 스타일 변환에 최적 전송을 적용한 WCT-OT는 FID, 그람 손실, SSIM 측면에서 WCT²보다 나은 성능을 나타내었다.

References

[1] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *Proc. CVPR*, pp. 2414-2423, 2016.

[2] Y. Li, N. Wang, J. Liu, and X. Hou, "Demystifying neural style transfer," in *Proc. IJCAI*, pp. 2230-2236, 2017.

[3] X. Huang and S. Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization," in *Proc. ICCV*, pp. 1501-1510, 2017.

[4] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu, and M. -H. Yang, "Universal style transfer via feature transforms," in *Proc. NIPS*, pp.

386-396, 2017.

[5] F. Luan, S. Paris, E. Shechtman, and K. Bala, "Deep photo style transfer," in *Proc. CVPR*, pp. 4990-4998, 2017.

[6] Y. Li, M.-Y. Liu, X. Li, M.-H. Yang, and J. Kautz, "A closed-form solution to photorealistic image stylization," in *Proc. ECCV*, pp. 453-468, 2018.

[7] J. Yoo, Y. Uh, S. Chun, B. Kang, and J. Ha, "Photorealistic style transfer via wavelet transforms," in *Proc. ICCV*, pp. 9036-9045, 2019.

[8] Y. Mroueh, "Wasserstein style transfer," in *Proc. AISTATS*, pp. 842-852, 2020.

[9] G. Bae and D. Park, "Wavelet transform-based style transfer using optimal transport," in *Proc. KICS Winter Conf.*, pp. 85-86, Feb. 2021.

[10] J. A. Cuesta-Albertos, C. Matran-Bea, and A. Tuero-Diaz, "On lower bounds for the l2-Wasserstein metric in a Hilbert space," *J. Theoretical Probability*, vol. 9, no. 2, pp. 263-283, 1996.

[11] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd Ed., Prentice-Hall, Inc., 2006.

[12] E. Peli, "Contrast in complex images," *J. Opt. Soc. Am. A*, vol. 7, pp. 2032-2040, 1990.

[13] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local Nash equilibrium," in *Proc. NIPS*, pp. 6629-6640, 2017.

[14] C. Szegedy, et al., "Going deeper with convolutions," in *Proc. CVPR*, pp. 1-9, 2015.

[15] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proc. ICCV*, pp. 1395-1403, 2015.

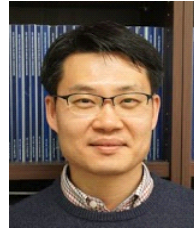
배 경 민 (Gyeongmin Bae)



2020년 8월 : 인하대학교 정보
통신공학과 졸업
2020년 8월~현재 : 인하대학교
전기컴퓨터공학과 석사과정
<관심분야> 머신러닝, 제어공
학, 통신공학

[ORCID:0000-0002-7499-1869]

박 대 영 (Daeyoung Park)



2004년 : 서울대학교 전기컴퓨
터공학부 박사
2004년~2007년 : 삼성전자 책임
연구원
2007년~2008년 : University of
Southern California, 방문
연구원

2008년~현재 : 인하대학교 교수
<관심분야> 이동통신, 다중안테나 시스템
[ORCID:0000-0001-8573-3526]