

조건 목적을 사용한 딥러닝 기반의 인지적 영상 복원

박승호 / 서울대학교 정보신호처리연구실

I. 서론

영상 초고해상화(Super-Resolution, SR) 기술은 낮은 해상도의 이미지를 고해상도로 향상시키는 것으로서 비디오 스트리밍, 의료 영상, 위성 영상 등과 같은 다양한 응용 분야에서 중요한 전처리 단계로 사용되기도 하고 객체 감지, 이미지 분할, 인식과 같은 많은 컴퓨터 비전 분야에서의 작업 성능을 크게 향상시킬 수 있다.

II. 인지 지향 초고해상화 (Perception-Oriented Super-Resolution)

본 논문에서 지향하는 내용은 단순히 픽셀 왜곡을 줄이는 것이 아닌 시각적 화질을 증대하는 것이 주요 목표이다. 초기 딥러닝 기반 초고해상화 학습부터 사용된 전형적인 손실 함수는 픽셀 기반의 왜곡 지향(distortion-oriented) 측정이었다. 하지만 이러한 손실 함수로 학습된 모델의 초고해상화 결과는 <그림 1>의 (c)와 같이 왜곡은 적으나 시각적으로 흐리고 디테일이 적으며 실제 같

지 않다는 단점을 가진다. 이를 극복하기 위해 다양한 인지 관점 손실 함수와 이를 이용한 SR 방법이 제안되었다 [1,2]. <그림 1>의 (d)는 이러한 인지 지향 모델의 선명하고 실제 같은 초고해상화 결과를 보여준다.

III. 제안하는 인지 지향 영역 적응적 초고해상화 방법

1. 제안하는 전체 구조

위의 내용에서 소개한 목적을 위하여 학위 논문에서 제안한 내용은 <그림 2>의 딥러닝 네트워크로서[4], 초고해상화 생성 모델, G_θ 와 영역별 최적 목적을 추정하는 예측 모델, C_ψ 로 구성된다.

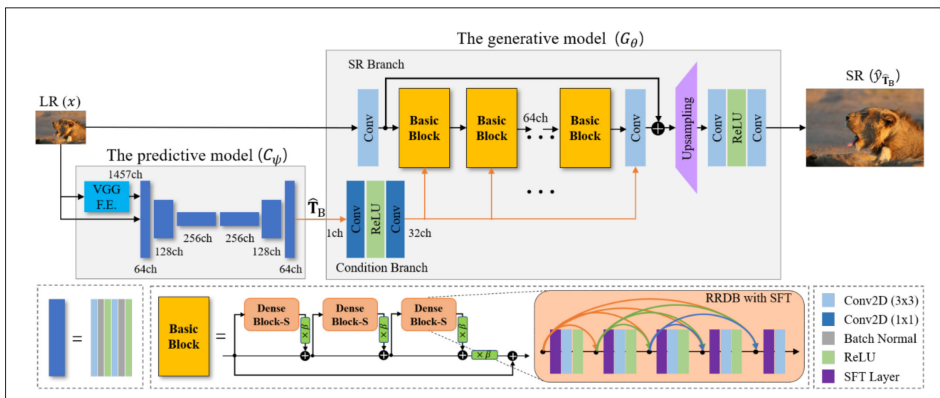
2. 목적 집합을 학습하는 생성 모델

<그림 3>은 생성 모델 학습 측면에서 기존과 제안 모델 간 학습법 차이를 보여준다. 기존 방법에서는 객체 세

졸업논문
소개



<그림 1> 왜곡 지향과 인지 지향 모델들 간의 초고해상화 결과 비교



<그림 2> 제안 방법의 전체 블록도

Conventional Fixed Objective	Proposed Conditional Objective
Training Phase	
$\hat{y} = G_{\theta}(x \psi)$ <i>(ψ: prior information)</i>	$\hat{y}_{T_t} = G_{\theta}(x T_t)$ <i>(t: objective control parameter)</i>
$L = \lambda_{rec} \cdot L_{rec} + \lambda_{adv} \cdot L_{adv} + \sum_{per_l} \lambda_{per_l} \cdot L_{per_l}$ $\theta^*_{\lambda} = \arg \min_{\theta} \mathbb{E}_{x,y \sim \mathcal{P}_{x,y}} [L(\hat{y}, y, \lambda)]$	$L(t) = \lambda_{rec}(t) \cdot L_{rec} + \lambda_{adv}(t) \cdot L_{adv} + \sum_{per_l} \lambda_{per_l}(t) \cdot L_{per_l}$ $\theta^* = \arg \min_{\theta} \mathbb{E}_{x,y \sim \mathcal{P}_{x,y}} \mathbb{E}_{t \sim \mathcal{P}_t} [L(\hat{y}_{T_t}, y, \lambda(t))]$
※ $t \sim \mathcal{P}_t$ ※ $T_t = 1 \times t$	
Testing Phase	
$\hat{y} = G_{\theta}(x \psi)$	$\hat{y} = G_{\theta}(x T)$

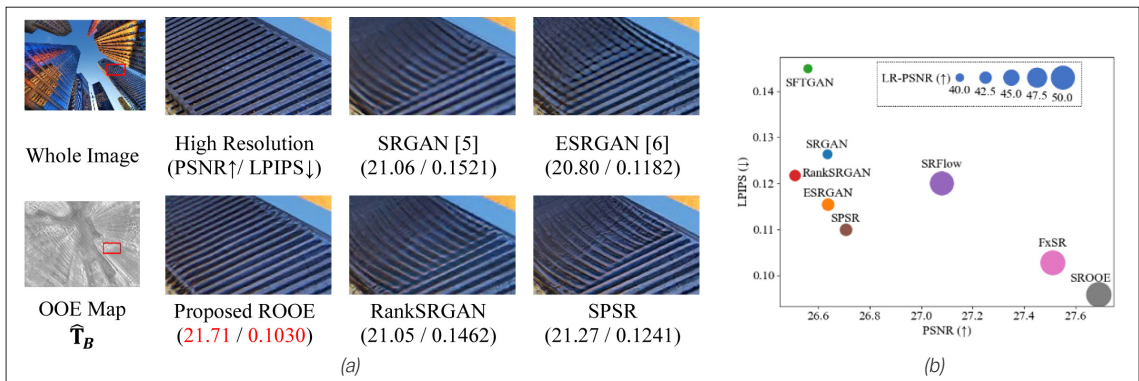
<그림 3> 기존 고정 손실 함수와 제안하는 조건 손실 함수 비교

그랜테이션 등의 정보 맵을 생성 모델에만 입력했지만, 제안 방법에서는 목적 맵을 생성 모델 뿐 아니라 손실 함수에도 동시에 전달한다.

3. 실험 결과

제안하는 최적 목적 기반의 초고해상화 방법은 주로

사용되는 5개 벤치마크 셋에서 왜곡 기반 측정 방법인 PSNR, SSIM과 인지 기반 측정법인 LPIPS, DISTSI의 4가지 정량적 측정 모두에 대해서 최신 인지 지향 초고해상화 방법들에 비해 가장 우수한 결과를 보였다. 또한 LR-PSNR은 45dB를 초과하여 생성된 초고해상화 결과가 입력 저해상도 이미지와 일관성 있음을 보여준다. <그림 4>는 이러한 시각적 비교와 정량적 비교 결과를 보여준다.



<그림 4> 최신 인지 지향 초고해상화 방법들과 제안 방법 간의 (a) 시각적, (b) 정량적 비교[10]

IV. 결론

주어진 이미지에 대한 영역별 최적 목표를 예측 적용하여 초고해상화 결과를 생성하는, 지각적으로 정확한 고해상도 복원을 위한 새로운 Single-image SR(SISR) 프레임

워크를 제시했다. 제안된 방법은 구조적 왜곡 및 부자연스러운 디테일 같은 시각적 부작용을 줄이고 정량 및 정성적 측면에서 최첨단 인식 지향 초고해상화 방법들에 비해 향상된 결과를 달성한다는 것을 보여준다. 제안 방법은 또한 다양한 기존 SISR 네트워크 아키텍처에 적용될 수 있다.

참 고 문 헌

- [1] Simonyan, K., and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 3rd International Conference on Learning Representations (ICLR 2015), Computational and Biological Learning Society, 2015.
- [2] Johnson, Justin, Alexandre Alahi, and Li Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14, Springer International Publishing, 2016.
- [3] Park, Seung Ho, Young Su Moon, and Nam Ik Cho, "Flexible style image super-resolution using conditional objective," IEEE Access 10 (2022): 9774-9792.
- [4] Park, Seung Ho, Young Su Moon, and Nam Ik Cho, "Perception-Oriented Single Image Super-Resolution using Optimal Objective Estimation," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.



박 승 호

- 2005년 ~ 2015년 : 삼성전자 DMC 연구소
- 2016년 ~ 현재 : 삼성전자 VD 사업부
- 2023년 8월 : 서울대학교 전기정보공학과 박사
- 주관심분야 : 영상 처리, 컴퓨터 비전, 딥러닝 기반 실시간 알고리즘 개발