

영상 다운 스케일링 기법 분석

*박재현, **조성인

동국대학교 멀티미디어공학과

*pjh0011@dongguk.edu, **csi2267@dongguk.edu

Survey on Image Down Scaling Techniques

*Jae Hyeon Park, **Sung In Cho

Department of Multimedia Engineering, Dongguk University

요 약

영상 다운스케일링 기법은 입력 영상의 해상도를 목적으로 하는, 입력영상의 해상도보다 낮은 해상도로 조정하는 기법을 일컫는다. 본 논문에서는 기존의 다운스케일링 기법들을 커널 기반의 방법과 프레임 단위의 최적화 기반 기법으로 분류하고 각 방법들의 대표 기법들에 대해서 분석한다. 이후, 최종적으로 그 결과를 도출하여 각 방법들의 특성을 도출한다.

1. 서론

최근 디스플레이 관련 기술의 비약적인 발전에 따라서 디스플레이 패널의 해상도, 종횡비 등이 다양화되고 있다. 이에 따라서, 기존의 이미 생성된 영상들의 해상도와 종횡비를 디스플레이 패널의 다양한 해상도와 종횡비로 조정하는 기법들에 대한 수요가 증대되고 있다. 이에 영상의 종횡비를 조정하기 위한 리타겟팅 기법 [1-5] 과 해상도를 조정하기 위한 스케일링 기법 [6-18] 등이 개발되어 왔다.

리타겟팅의 경우 일반적으로 영상 내 중요도에 따라서 영상의 그리드를 조정하는 방법으로 영상의 종횡비를 조정한다. 영상 내 중요 영역의 추출과 추출된 중요 영역의 종횡비 보존에 집중한다. 따라서, 생성된 새로운 비정규적인 그리드에 따라서 최종 픽셀 값을 보간법을 사용하여 최종 결과를 추출한다. 반면 영상 스케일링의 경우 영상 내 중요 영역을 분리하여 최종 픽셀 값을 결정하기 보다는 주어진 영상의 정보를 효율적으로 전달하기 위한 픽셀 값 생성에 주력한다. 이러한 영상 스케일링은 오랜 기간 연구가 되어온 분야로, 크게 입력 영상의 해상도를 목표 디스플레이 해상도에 맞추어 증대시키는 업스케일링과 입력영상의 해상도를 목표 디스플레이 해상도로 줄여주는

다운스케일링으로 구분될 수 있다. 영상 다운 스케일링에서는 영상의 중요 정보를 최대한 보존하면서 해상도를 낮추는 것에 연구가, 업스케일링에서는 영상의 정보량을 증대시키는 것에 대한 개발이 주로 이루어졌다. 이러한 영상 스케일링 기법 연구는 지난 수년간 이루어져 왔는데, 상대적으로 다운스케일링보다는 업스케일링에 대한 연구가 더 활발하게 이루어졌다. 이는, 저해상도의 영상을 고해상도의 디스플레이에 표현하는 경우가 고해상도를 저해상도 디스플레이에 표현하는 경우보다 기존에는 더욱 빈번하였기 때문이다. 하지만 디스플레이 패널의 해상도 다양화에 맞추어 영상 콘텐츠의 해상도 역시 다양화되고 있다. 따라서 고해상도의 영상을 저해상도의 디스플레이 패널에 맞춰야 하는 상황이 빈번히 발생하고 있어 효과적인 다운 스케일링 기법에 대한 요구가 급격히 증대되고 있다.

임의의 다운스케일링 기법을 성능을 평가하기 위해서는 원본 입력 영상을 업스케일링하고, 업스케일링된 영상에 임의의 다운스케일링 기법을 적용하여 본래 원본 입력 영상과 비교하는 과정이 필요하다. 하지만, 기존에 존재하지 않는 정보를 완벽히 새롭게 생성할 수 있는 업스케일링 기법이 존재하지 않기 때문에 다운스케일링 기법을 평가하기에 적합한 입력을 생성하기 어렵다.

따라서 최근 널리 사용되는 훈련 기반의 스케일링 기법은 다운 스케일링 기법에 적용하기가 용이하지 않았다. 이러한 이유로 업스케일링에 비해 다운스케일링에 대한 기법은 상대적으로 다양성이 부족하였고, 실제 디스플레이 시스템에서는 bicubic 인터폴레이션과 같은 기본적인 커널 기반의 다운스케일링 기법을 다운 스케일링의 기법으로 가장 널리 사용하고 있다. 디스플레이 시스템에서는 연산 복잡도는 임의의 알고리즘을 내재화 할 때, 매우 중요한 고려 요소이기 때문에 이러한 커널 기반의 방법이 크게 선호되고 있다.

2. 기존 영상 다운스케일링 기법

다운 스케일링의 경우 크게 커널 기반의 방법과 프레임 단위의 최적화를 기반으로 하는 방법으로 구분될 수 있다. 각 방법들의 특성과 대표 기법들을 아래와 같이 정리할 수 있다.

2.1 커널 기반의 영상 스케일링 기법

커널 기반의 방법으로는 bilinear, bicubic 인터폴레이션 [11] 과 Lanzos [12] 와 같은 방법이 가장 대표적인 다운 스케일링 방법으로 사용되고 있다. 이들 방법은 다운 스케일링시에 발생할 수 있는 얼리어싱 문제를 해결하기 위하여 저역 통과 필터 기능을 내포하고 있다. 이러한 커널 기반의 방법들은 얼리어싱 문제를 해결하면서, 다운스케일링 이후의 결과에 극심한 왜곡은 효과적으로 방지할 수 있는 장점을 가진다. 하지만, 입력영상의 정보 특성을 고려하지 않고 단순히 저역통과 필터의 기능을 수행하여 입력 영상의 정보가 크게 훼손되는 문제가 있다. 이에 입력 영상의 특성을 고려하여 주요 영역 혹은 주요 정보를 효과적으로 보존하기 위한 다운스케일링 기법들이 제안되기 시작하였다. 기존의 커널 기반의 기법들이 영상의 특성에 관계없이 다운 스케일링을 위한 필터링을 수행하였던 것에서 벗어나 [13] 에서는 아래 수식 1 과 같은 bilateral 필터링에서 사용한 컨셉과 같이 공간 영역과 색 영역의 필터를 조합한 커널을 지역 영역의 특성에 맞추어 조정하여 다운스케일링 이후에 영상의 흐려지는 문제를 해결하고자 하였다.

$$I_p^{out} = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s} (\|p - q\|) G_{\sigma_r} (|I_p - I_q|) I_q,$$

(1)

여기서 p 는 현재 픽셀 값의 위치를, q 는 p 픽셀의 인접 픽셀의 위치를 나타낸다. I 는 픽셀의 출력 값을, W_p 는 정규화를 위한 값을, G_{σ_s} 와 G_{σ_r} 는 거리 차이와 픽셀 값 차이에 대한 커널을 나타낸다. S 는 현재 픽셀 값을 결정하기 위해 참조하는 인접

영역을 나타낸다. 이 방법은 기존 방법보다는 영상의 주요 정보들을 보존하는 성능을 향상시키기는 했으나, 많은 영상 내 정보들이 손실되는 문제가 존재하였다.

2.2 프레임 단위의 최적화 기반의 영상 스케일링 기법

국소 영역의 픽셀 값들을 이용한 커널 기반의 필터링 기법은 입력 영상 전체를 고려할 수 없다는 한계가 존재한다. 이에 입력 영상의 전체적인 정보를 고려하여 다운스케일링을 하기 위한 기법들이 제안되었다. [15] 에서는 기존의 필터 기반의 다운스케일링 기법에서 벗어나 프레임 단위로 결과 영상을 최적화하는 기법을 사용하였다. 입력 영상과 다운스케일링 된 영상의 유사도를 structural similarity index (SSIM)으로 도출하고 이를 최대화는 방법으로 다운 스케일링을 수행한다. 커널 기반의 필터링을 다운스케일링에 사용하지 않은 이들 방법은 영상들의 주요 정보들의 보존력을 향상시켰으나, 영상 노이즈와 같은 원하지 않은 신호가 증폭되는 문제를 야기시켰다. [16] 에서는 두 가지 L0-regularized priors 를 이용하여 원본 영상의 주요 정보를 보존하면서 영상의 해상도를 낮출 수 있는 기법을 제안하였다. 아래 수식은 [16] 에서의 목적함수를 나타낸다.

$$\min_{x,m} \|x - y_l(m)\|_2^2 + \sigma [\mu P(\nabla x) - P(\nabla y)] + \lambda P(m), \quad (2)$$

여기서 x 와 $y_l(m)$ 은 두 다운스케일링된 영상을 의미한다. 구체적으로 x 는 기울기 비율 관련 제약을 반영한 결과를 나타내고 $y(m)$ 은 다운스케일링을 위한 행렬인 m 을 이용한 다운스케일링 결과이다. $P(\cdot)$ 는 L0-Regularized 연산자를 나타내고, μ 는 다운스케일링된 영상과 원본영상에서의 엣지 개수 차이를 계산할 때 픽셀 수 차이를 고려하기 위한 요소이다. σ 와 λ 는 엣지 요소 개수 차이와 m 매트릭스의 sparsity 를 비유으로 간주할 때 이를 밸런싱하기 위한 파라미터이다. 정리하면, 영상 내 엣지의 개수를 보존할 수 있는 L0-Regularized Gradient-Ratio Prior 와 L0-Regularized Reconstruction Prior 를 이용하여 원본 영상의 엣지를 효과적으로 보존하면서 원본 영상의 픽셀값을 유지하면서 영상의 해상도를 줄이게 된다. 이 방법은 다운 스케일링된 영상에서의 선명도를 획기적으로 향상시키면서 기존의 프레임 전체 최적화를 이용하여 스케일링을 수행하는 기법의 성능을 한 단계 올려 놓았다. 하지만 저자가 밝혔듯이 fine edge 보존에 실패하는 경우가 발생할 수 있고, 엣지들의 과도하게 향상시켜 원본의 정보를 왜곡하는 상황이 발생할 수 있다. 특히, 다운 샘플링에 의한 해상도 변화가 크기 않을 때에 이런 왜곡은 더욱 두드러지게 된다.

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2020R1C1C1009662).

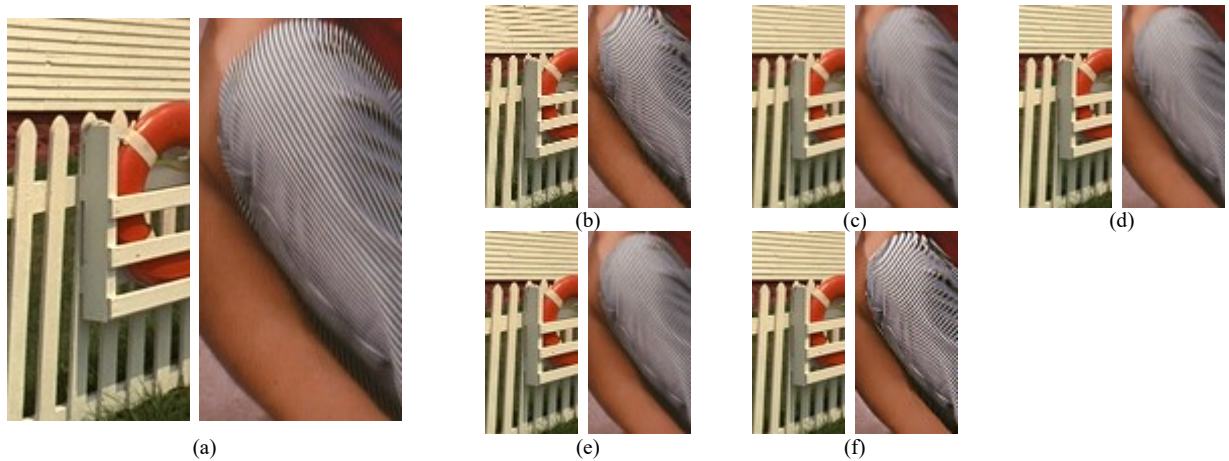


그림 1. 다운스케일링 결과 예제 (다운스케일링 팩터: 0.5) (a) cropped original frames, (b) subsampling, (c) bicubic, (d) Lanczos filter (e) perceptually based downscaler (PS), (f) L0-regularized downscaler (LOS)

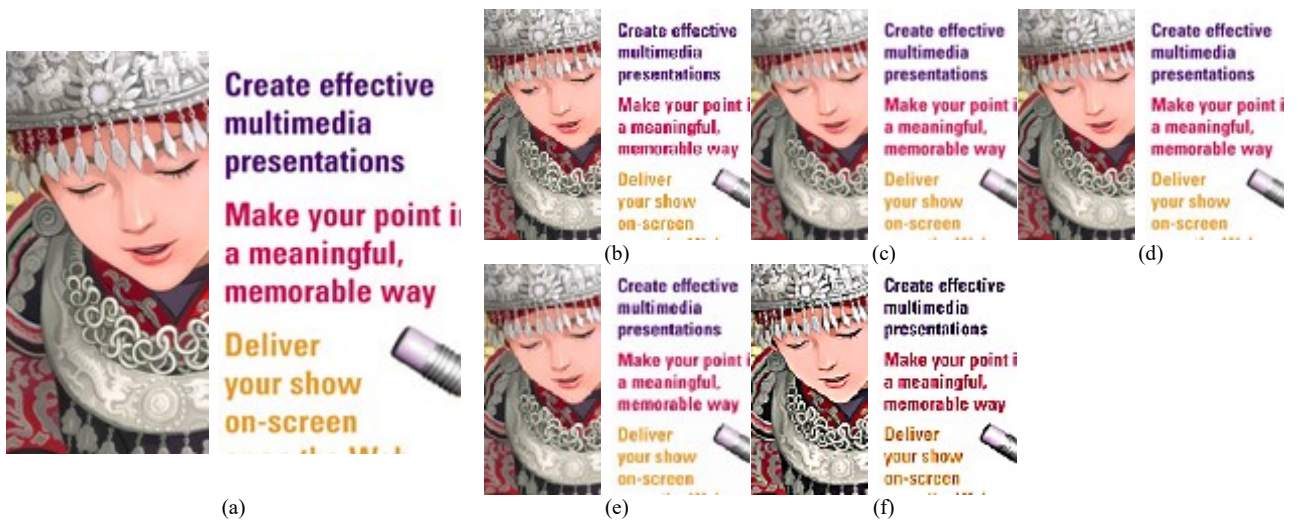


그림 2. 다운스케일링 결과 예제 (다운스케일링 팩터: 0.57) (a) cropped original frames, (b) subsampling, (c) bicubic, (d) Lanczos filter (e) PS, (f) LOS

표 1. 다운스케일링 결과 평가

방법	얼리어싱 왜곡정도	엣지 선명도	자연스러움
subsampling	상	중상	중
bicubic	하	하	상
Lanczos	하	중하	상
PS	하	중하	상
LOS	상	상	하

3. 실험 결과

본 논문에서는 기존 방법들의 다운스케일링 결과를 영상으로 확인하고 분석하고자 한다. 비교 방법은 다운스케일링에 가장 기본이 되는 방법인 subsampling 기법과 커널 기반의 다운스케일링 기법인 bicubic, Lanczos filter 를 사용하였다. 또한 프레임 단위의 최적화 기법인 perceptually based downscaler (PS)와 L0-regularized downscaler (LOS)를 비교를 위한 방법으로 사용하였다. 그림 1 과 2 는 각각 다운스케일링

팩터 0.5 와 0.57 에 대한 결과 영상을 보여주고 있다. 먼저 그림 1 에 존재하는 반복 패턴에서의 결과를 살펴보면 subsampling 기법은 대부분의 반복 패턴에서 극심한 얼리어싱 왜곡을 발생시키는 것을 볼 수 있다. LOS 기법 역시 상대적으로 얼리어싱 왜곡이 일부 패턴에서 극심하게 발생할 수 있다. 나머지 방법들은 얼리어싱 왜곡은 상대적으로 적은 것을 확인할 수 있다. 나머지 일반 엣지 영역에서의 원본 정보의 보존 정도를 살펴보면, subsampling 기법이 나머지 방법들 대비 더 나은 선명도를 보이는 것을 볼 수 있다. LOS 방법은 엣지 성분을 가장 잘 부각시키는 것을 볼 수 있다. 다만 이러한 부각이 왜곡으로 느껴질 여지가 존재한다. 전체적인 평가를 진행하면 표 1 과 같이 나타낼 수 있다. 눈에 띄는 엣지 보존을 요구하는 경우에는 LOS 기법을, 얼리어싱 왜곡 없이 가장 안정적인 결과를 도출하고자 할 경우에는 기존 커널 기반 방법을 사용하는 것이 적절할 것으로 예상된다.

4. 결론

본 논문에서는 기존의 다운스케일링 기법들을 크게 두 가지 접근 방법으로 분류해보고 각 방법들의 특성에 대해서 분석하여 보았다. 실험 결과에서 볼 수 있듯이, 커널 기반의 방법들이 엣지 보존에는 실패하면서도 얼리어싱 문제에 대해서는 상대적으로 강건한 결과를 내 놓는 것에 반하여, 프레임 단위의 최적화 기법은 엣지 보존의 성능을 우수하였으나, 얼리어싱 문제와 영상의 자연스러움을 유지하는 것에는 한계를 도출하는 것을 확인할 수 있었다.

5. 참고문헌

- [1] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Trans. Graphics*, vol. 26, no. 3, pp. 10-18, 2007.
- [2] B. Li, L.-Y. Duan, J. Wang, R. Ji, C.-W. Lin, and W. Gao, "Spatiotemporal grid flow for video retargeting," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1615-1628, Apr. 2014.
- [3] S. J. Lee, S. I. Cho, and S.-J. Kang, "Object detection-based video retargeting with spatial-temporal consistency," *IEEE Trans. Circuits Syst. Video Technol.*, [Online]. Available: <https://ieeexplore.ieee.org/document/9043574>
- [4] S. I. Cho and S.-J. Kang, "Temporal incoherence-free video retargeting using foreground aware-extrapolation," *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 4848-4861, Dec. 2020.
- [5] S. I. Cho and S.-J. Kang, "Extrapolation-based video retargeting with backward warping using image-to-warping vector generation network," *IEEE Signal Proc. Lett.*, vol. 27, no. 1, pp. 446-450, Dec. 2020.
- [6] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1-8.
- [7] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861-2873, Nov. 2010.
- [8] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2014, pp. 184-199.
- [9] J. S. Choi and M. Kim, "Super-interpolation with edge-orientation-based mapping kernels for low complex 2x upscaling," *IEEE Trans. Image Process.*, 2016, vol. 25, no. 1, pp. 469-483.
- [10] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2016, Vol. 38, No. 2, pp. 295-307.
- [11] D. P. Mitchell and A. N. Netravali, "Reconstruction filters in computer graphics," in *Proc. ACM SIGGRAPH*, 1988, pp. 221-228.
- [12] C. E. Duchon, "Lanczos filtering in one and two dimensions," *J. Appl. Meteor.*, vol. 18, no. 8, pp. 1016-1022, Aug. 1979.
- [13] J. Kopf, A. Shamir, and P. Peers, "Content-adaptive image downscaling," *ACM Trans. Graph.*, vol. 32, no. 6, 2013, Art. no. 173.
- [14] N. Weber, M. Waechter, S. C. Amend, S. Guthe, and M. Goesele, "Rapid, detail-preserving image downscaling," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 1-6, Nov. 2016.
- [15] C. Öztireli and M. Gross, "Perceptually based downscaling of images," *ACM Trans. Graph.*, vol. 34, no. 4, 2015, Art. no. 77.
- [16] J. Liu, S. He, and R. W. H. Lau, "L0-regularized image downscaling," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1076-1085, Mar. 2018.
- [17] H. Kim, M. Choi, B. Lim, and K. M. Lee, "Task-aware image downscaling," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 399-414.
- [18] Y. Li, D. Liu, H. Li, L. Li, Z. Li, and F. Wu, "Learning a convolutional neural network for image compact-resolution," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1092-1107, 2019.