

Defocused 프레임과 영상 품질 평가 결과의 상관성 연구

이성배, 김규헌*

경희대학교

rhee@khu.ac.kr, *kyuheonkim@khu.ac.kr

A Study on the Correlation between Defocused Frame and Image Quality Assessment Results

Rhee Seong Bae, Kim Kyu Heon*

Kyung Hee University

요약

현재 딥러닝 기반 영상 처리는 객체 검출, 영역분할, 깊이 정보 추정, 해상도 향상 등 다양한 분야에서 높은 정확도를 나타내고 있으며, 최근에는 동영상을 위한 딥러닝 알고리즘이 제안되고 있다. 초기 동영상에 대한 딥러닝 알고리즘은 단순히 동영상의 모든 프레임에 대하여 이미지의 과정을 반복하는 방식으로 수행했지만, 최근 Optical Flow, Frame Interpolation, Video Outpainting 등과 같은 분야에서는 프레임마다 독립적으로 알고리즘을 적용하는 것이 아닌, 프레임 간의 temporal 정보를 활용하는 방식으로 수행되고 있다. 이처럼 연속된 프레임의 temporal 정보를 활용하는 것은 동영상에 대한 딥러닝 알고리즘의 효율성과 성능을 높일 수 있었지만, 동영상의 모든 프레임이 같은 영상 품질을 갖지 않을 수 있기에 왜곡된 temporal 정보가 생성될 수 있고, 이에 따라서 머신 태스크(Machine Task)의 성능이 저하될 수도 있다. 일례로, 한 카메라를 통해 촬영된 동영상에서는 촬영 중간에 카메라의 초점이 맞지 않아 흐린(Blur) 프레임이 생성될 수 있고, 해당 프레임은 왜곡된 temporal 정보를 생성할 수 있다. 이와 같은 왜곡된 temporal 정보는 Deblurring 알고리즘을 적용함으로써 줄일 수 있지만, 어떤 프레임에서 흐린 프레임이 생성되었는지에 대한 탐색이 필요하다. 이에 본 논문에서는 영상 품질 평가 방법을 통하여 동영상 프레임 내에서 생성된 흐린 프레임을 검출하는 것이 가능한지에 대한 연구를 진행한다.

I. 서론

딥러닝 기반 이미지 처리는 객체 검출, 영역분할, 깊이 정보 추정, 해상도 향상 등 다양한 분야에서 높은 정확도를 나타내고 있으며, 최근에는 동영상을 위한 딥러닝 알고리즘이 제안되고 있다. 초기 동영상에 대한 딥러닝 알고리즘은 이미지의 알고리즘을 프레임마다 반복하여 수행하는 방식으로 진행되었지만, 최근에는 프레임 간의 temporal 정보를 활용한 동영상 처리 알고리즘이 제안되고 있다[1, 2].

그러나 동영상의 경우, 촬영된 시점마다 촬영 환경이 다를 수 있기에, 프레임마다의 영상 품질이 달라질 수 있고, 프레임 간의 영상 품질이 다른 것은 프레임 간의 왜곡된 temporal 정보를 생성할 수 있다[1]. 일례로, 한 카메라에서 촬영된 영상을 통해서 slow motion 영상을 생성할 때, 동영상 프레임 내에 흐린 프레임이 존재하는 경우 frame interpolation[2]의 성능은 크게 저하될 수 있다. 이처럼 왜곡된 temporal 정보를 생성할 수 있는 흐린 영상은 deblurring[3] 알고리즘을 통해 영상 품질을 향상시키는 경우, 왜곡이 줄어든 temporal 정보를 생성할 수 있기에 동영상 프레임 내 흐린 영상을 검출하는 방법이 필요하다.

II. 본론

본 논문에서는 프레임 간의 temporal 정보 생성에 있어 오류를 발생시킬 수 있는 흐린 프레임을 검출하기 위하여 Non-reference IQA(Image Quality Assessment)를 활용하는 것이 가능한지를 실험으로 검증하고자 한다. 본 논문에서 실험하는 Non-reference IQA는 가장 대표적인 BRISQUE[4]와 최근에 제안된 CONTRIQUE[5], MANIQA[6], ARNIQA[7] 알고리즘을 통해 실험을 진행한다.

Non-reference IQA는 단일 영상에 대하여 평가를 진행하며, 각각의 방법에 따라서 평가지표가 클수록 좋은 영상으로 평가하거나 작을수록 좋은 영상으로 판단할 수 있다. 평가지표 자체는 영상 품질을 비교하는 용도로는 활용될 수 있지만, 영상 품질 자체의 척도로는 활용되기 제한될 수 있다.

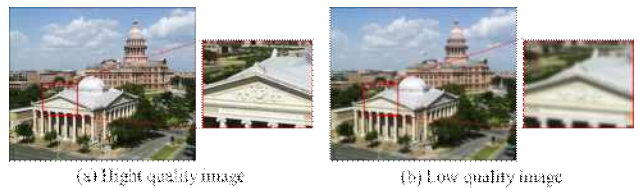


그림 1. 고품질 영상 및 저품질 영상

일례로, 그림 1의 (a)와 (b)에 나타나는 바와 같이, 고품질 영상과 저품질 영상에 대한 Non-reference IQA 평가지표는 표 1과 같이 나타난다. BRISQUE와 CONTRIQUE는 평가지표가 낮을수록 좋은 품질의 영상이라고 판단하는 것과 달리, MANIQA와 ARNIQA는 평가지표가 높을수록 좋은 품질의 영상이라고 판단한다.

표 1. Non-reference IQA 평가 결과

Method	그림 1 (a)	그림 1 (b)
BRISQUE[4] (↓)	16.8884	52.5584
CONTRIQUE[5] (↓)	16.7713	56.6484
MANIQA[6] (↑)	0.5081	0.1424
ARNIQA[7] (↑)	0.7899	0.2635

이와 같은 Non-reference IQA는 각 실험에서 흐린 영상에 대한 알고리즘의 평가지수와 평가자들의 평가지수 간에 상관도를 검증하였다. 그러나 해당 평가에서는 흐린 영상에 대하여 점수 규정이 없이, 평가자가 임의로 판단한 영상 품질 점수와와의 상관성을 실험하였다. 본 논문에서는 흐린 영상에 대하여 temporal 정보를 왜곡시키는 정도에 따라 점수를 차등적으로 부여하고, 해당 점수와 Non-reference IQA와의 상관성을 검증함으로써, 동영상 프레임 내 흐린 프레임을 검출할 수 있는지를 실험한다.

본 논문에서 흐린 영상에 대한 점수는 프레임 간의 temporal 정보 생성에 왜곡을 줄 수 있는 정도에 비례하여 할당하였다. 구체적으로, 카메라의 움직임으로 인해 모든 영상 영역에 움직임 왜곡이 발생한 프레임은 0점, 모든 배경 영역에 왜곡이 생겼지만, 객체는 보존된 경우는 1점, 배경 영역은 보존되었지만 객체 단위 또는 지역 단위로 왜곡이 발생한 프레임인 경우는 2점, 정상적인 프레임이면 3점을 부여했다.

실험에서는 DAVIS dataset[8]에 있는 11개의 시퀀스로, 총 876장의 동영상 프레임에 대하여 상기와 같은 흐린 점수 규정을 적용하여 정답 점수 데이터를 구성하였다. BRISQUE, CONTRIQUE, MANIQA, ARNIQA를 통해 평가한 결과와 흐린 영상의 정답 점수의 상관도는 표 2에 나타나는 바와 같다.

표 2. Non-reference IQA와 흐린 영상 상관계수

Method	PLCC(↑)	SRCC(↑)	KRCC(↑)	RMSE(↓)
BRSIQUE	0.4423	0.4508	0.3541	0.3776
CONTRIQUE	-0.6961	-0.6331	-0.5118	0.6174
MANIQA	0.6908	0.6512	0.5231	0.2866
ARNIQA	0.5506	0.5196	0.4181	0.3143

상기 표 2에서 나타나는 바와 같이, 흐린 영상의 정답 점수와 가장 유사하게 평가한 기술은 MANIQA로 나타난 것을 확인할 수 있다. 전체적인 상관도가 일반적인 Non-reference IQA 실험에서의 상관도보다 낮게 나타나는 것을 확인할 수 있는데, 이는 흐린 영상의 정답 점수는 특정한 조건에 따라서 양자화된 점수인 것과 다르게, Non-reference IQA는 0과 1 사이의 값으로 정의된 것으로 인한 영향이 있으며, 그리고 흐린 영상의 정답 점수는 흐린 정도에 대하여 차등적인 점수를 할당하지 않은 것이 원인으로 판단된다.

한편, 표 2의 CONTRIQUE는 음의 상관도를 나타내는 것을 확인할 수 있다. 이는 해당 평가에서 저품질로 판단한 영상이 실험에서 정의한 정상 프레임을 선택하는 것을 나타낸다. 해당 평가 결과를 반대로 진행하는 경우, 높은 정확도로 흐린 프레임을 선택할 수 있는 것을 확인할 수 있다.

III. 결론

본 논문에서는 동영상을 위한 딥러닝 알고리즘에서 프레임 간의 temporal 정보를 생성하는데, 오류를 발생시킬 수 있는 흐린 프레임을 검출하기 위하여 Non-reference IQA가 활용될 수 있는지에 대한 가능성을 실험하였다. 프레임의 흐린 양상에 따라 temporal 정보를 생성하는데 왜곡시키는 정도가 다르기에, 본 논문에서는 왜곡을 발생시키는 정도에 따라 흐린 영상에 차등적인 점수를 부여하였다. 해당 흐린 점수를 정답 점수로 하여, 각각의 Non-reference IQA로 평가를 진행하여 상관도를 측정하였다.

실험에서는 MANIQA가 가장 정확하게 흐린 영상을 검출할 수 있는 것으로 나타났으며, 정상 프레임에서 평가한 값을 기준으로 일정 비율 이상 차이 나는 프레임을 선택할 시, 높은 정확도로 흐린 프레임을 검출할 수

있는 것으로 나타났다. 이에 HDR 영상을 구성함에 있어, 동영상 내 흐린 프레임 검출에 Non-reference IQA가 활용될 수 있음을 실험을 통해 검증하였다.

ACKNOWLEDGMENT

This work was supported by Korea Research Institute for defense Technology planning and advancement(KRIT) grant funded by the Korea government (DAPA(Defense Acquisition Program Administration)) (21-106-A00-007, Space-Layer Intelligent Communication Network Laboratory, 2024) and supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2024-2021-0-02046) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation

참 고 문 헌

- [1] Jin, Meiguang, Zhe Hu, and Paolo Favaro. "Learning to extract flawless slow motion from blurry videos." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.
- [2] Jiang, Huaizu, et al. "Super slo-mo: High quality estimation of multiple intermediate frames for video interpolation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
- [3] Zamir, Syed Waqas, et al. "Restormer: Efficient transformer for high-resolution image restoration." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.
- [4] Mittal, Anish, Anush Krishna Moorthy, and Alan Conrad Bovik. "No-reference image quality assessment in the spatial domain." IEEE Transactions on image processing 21.12 (2012): 4695-4708.
- [5] Madhusudana, Pavan C., et al. "Image quality assessment using contrastive learning." IEEE Transactions on Image Processing 31 (2022): 4149-4161.
- [6] Yang, Sidi, et al. "Maniqa: Multi-dimension attention network for no-reference image quality assessment." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.
- [7] Agnolucci, Lorenzo, et al. "ARNIQA: Learning Distortion Manifold for Image Quality Assessment." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2024.
- [8] F. Perazzi, J. Pont-Tuset, B. McWilliams, L. Van Gool, M. Gross, and A. Sorkine-Hornung. A benchmark dataset and evaluation methodology for video object segmentation. In Computer Vision and Pattern Recognition, 2016