2. Related Work

2-1. Key Frame Extraction

동영상에서 하이라이트 이미지를 추출해내기 위한 방법을 제시한다. Key Frame Extraction 방법으로 6가지 중 해당 논문은 Clustering을 사용하는 방법을 제시한다. Clustering을 사용한 Key Frame Extraction은 구현이 비교적 쉽고, 속도가 빨라 실시간으로 적용할 수 있다는 이점이 있다.

Key Frame Extraction하는 과정은 다음과 같다. 우선 Global Level Histogram을 사용하여 프레임 간의 색상 유사도를 측정하고, GLCM(Grey-Level Co-occurrence Matrix)을 사용하여 프레임 간의 텍스처(질감) 유사도를 측정한다.

최종적으로 두 유사도 측정값을 가중치를 사용하여 최종 유사도 측정값으로 통합한다. 이 최종 유사도 측정값을 임계 값과 비교하여 Cluster의 밀집도를 제어한다. 최종 유사도 측정값이 임계 값보다 낮다면 이는 기존 Cluster에 추가되기에는 유사도 값이 낮다는 것을 의미하므로, 새로운 Cluster를 생성한다.

Clustering이 끝나면, Cluster들 중 해당 Cluster가 갖은 Frame 수가 전체 Frame 수의 10%가 넘으면 해당 Cluster가 비중 있는 Cluster라고 판단하여 해당 Cluster의 대표 Frame을 Key Frame으로 선택한다. 해당 Cluster의 대표 Frame은 Cluster 내의 Frame의 HSV와 GLCM의 중심에 가장 가까운 Frame을 Key Frame으로 선택한다.

2-2. InstaGAN

기존 GAN 모델은 이미지 내의 다수 인스턴스를 변화시킬 때, 혹은 이미지 내의 인스턴스의 모양을 변화시킬 때 낮은 성능을 보이는 단점을 가지고 있다. InstaGAN에서는 이러한 단점을 해결하기 위해 모델을 학습할 때 사용하는 학습 데이터 셋에 변환하고자 하는 인스턴스의 Segmentation Mask를 추가하는 방식을 통해 인스턴스의 모양을 변화시키는 모델의 성능을 높였다.

2-3. CartoonGAN

실제 현실 이미지를 만화 스타일의 이미지로 변환하는 것을 목표로 하는 모델로, GAN 모델을 기초로 한다.

만화 스타일의 이미지는 실제 현실 이미지와 다르게 만화 스타일이 갖는 고유한 특성이 존재하며, 현재 널리 사용되는 GAN 모델의 손실함수로는 이를 처리하기 힘들다. CartoonGAN에선 이를 위해, 기존 손실함수를 기반으로 윤곽선에 대한 손실도를 포함하는 손실함수를 정의하여 해당 모델의 성능을 높였다.

추가적으로, CartoonGAN에서는 기존의 GAN 모델이 비선형이기 때문에, 무작위로 초기화할 경우 지역 최소점에 갇혀 원하는 방향으로 학습되지 않을 가능성이 커, 이를 방지하기 위해 semantic content loss를 사용하여 사전에 생성자를 학습하였고, 이것이 결과에 유의미한 차이가 있음을 보였다.