

중요도 지도에 기반한 관심 영역 비디오 압축

김한상, 김창수
고려대학교

{hansangkim, cskim}@mcl.korea.ac.kr

ROI Video Compression Based on Spatiotemporal Saliency Map

Hansang Kim, Chang-Su Kim
Korea University

요 약

본 논문에서는 중요도 지도에 기반한 관심 영역 동영상 압축 방법에 대해 고찰한다. 동영상 압축은 손실 프로세스이기 때문에 관심 영역에서의 정보 손실 최소화가 필요하며, 이를 위해 중요도 감지 과정에서 추출되는 중요도 지도의 신뢰도가 중요하다. 따라서 다양한 다른 기법의 중요도 지도 적용 결과를 비교함으로써 중요도 지도 추출 알고리즘의 요건에 대해 추론하고, 추출된 중요도 지도를 이용하여 적절하게 동영상을 부호화하는 방법에 대해 제안한다. 마지막으로 실험결과를 통해 보완되어야 할 부분을 제시한다.

1. 서론

중요도 감지는 정지영상 및 동영상에서 시각적으로 드러나는 부분을 자동으로 추출해주는 기술로서, 이와 관련된 수많은 연구가 활발히 진행되어 왔으며 물체 인식, 이미지 검색 등 다양한 용도로 활용되고 있다.

특히 동영상에서의 중요도 감지는 정지영상에서의 중요도 감지에 비해 시간적 변이에 따른 특징을 추가적으로 고려해야 하기 때문에 시간적 특징 모델링이 매우 중요하며 공간적 특징과의 결합을 어떻게 하느냐에 따라 그 성능이 좌우된다.

본 논문에서는 동영상에서 중요도 지도를 획득하여 시각적으로 눈에 띄는 영역과 그렇지 않은 영역을 이분화하여 화질을 조절하여 동영상 압축에 적용하고, 압축된 동영상의 화질과 원본 영상을 비교한다. 이를 통해 중요도 지도에 기반한 관심 영역 동영상 압축의 효과를 확인하고자 한다. 또한, 다양한 중요도 지도 획득 알고리즘 중 F-measure, the area under receiver operating characteristic curve (AUC) 등의 성능 측정 기준에서 높은 성능을 발휘한 비교 알고리즘을 적용시킴으로써 관심 영역 동영상 압축 기법 적용에 적절한 중요도 지도의 요건과 역할에 대해 고찰한다.

본 논문의 구성은 다음과 같다. 2 절에서는 두 가지 관심 영역 비디오 압축 제안 기법에 대해 설명하고, 3 절에서는 네 가지 중요도 추출 알고리즘을 통해 추출된 중요도 지도를 이용해 각 제안 기법을 적용한 실험결과를 확인하고, 마지막으로 본 논문에 대한 결론을 맺는다.

2. 제안하는 기법

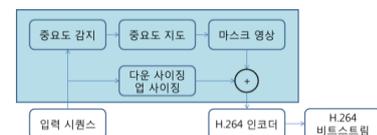
본 논문에서 제안하는 방식은 입력 시퀀스로부터 중요도를

감지하고 중요도 지도를 추출하는 일반적인 중요도 지도 획득 프로세스와 H.264 AVC 인코더를 이용해 입력 시퀀스로부터 비트스트림을 획득하는 부호화 프로세스를 다음 두 방식을 통해 유기적으로 결합하는 것이다.

첫 번째 방식은 전처리 방법으로 중요도 지도로부터 '마스킹 영상'을 만들고 그 이외 영역은 저해상도로 크기를 줄인 후 다시 원래 크기로 되돌림으로써 배경을 의도적으로 화질을 떨어뜨린 '저 화질 영상'을 만든다. 그런 후 두 영상을 결합함으로써 '변경된 입력 시퀀스'를 인코더의 입력으로 넣어준다.

두 번째 방식은 중간처리 방법으로 중요도 지도를 이용하여 양자화 지도를 생성하고, 생성된 '양자화 지도' 정보를 이용해 인코더의 양자화 계수를 조정함으로써 관심 영역의 화질 열화를 최소화하는 방식이다.

전자가 입력 시퀀스의 변경을 통해 간접적으로 관심 영역의 화질을 제어한다면 후자는 인코더의 입력 시퀀스는 변경하지 않고 양자화 지도를 통해 시각적으로 중요한 영역과 그렇지 않은 영역에 다른 양자화 파라미터를 적용함으로써 직접적으로 화질을 제어하는 방식이다.

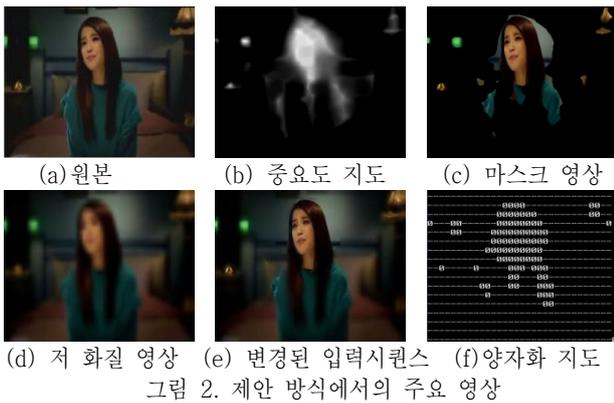


(a) 전처리 기반 방식



(b) 중간처리 기반 방식

그림 1. 두 가지 제안 기법 처리 방식



(a) 원본 (b) 중요도 지도 (c) 마스크 영상



(d) 저 화질 영상 (e) 변경된 입력시퀀스 (f) 양자화 지도

그림 2. 제안 방식에서의 주요 영상

본 논문에서 사용한 중요도 지도는 Random walk with restart (RWR)에 기반하여 공간적 특성뿐 아니라 시간적 특성을 동시에 고려한 알고리즘 [1]의 출력 결과를 활용하였으며, 최근 발표된 논문 중 좋은 성능을 보여주는 세 가지 알고리즘 GMR [2], RBC [3], LSD [4]을 비교 알고리즘으로 선정하였다. 동영상에 대상으로 제안된 알고리즘인 LSD 와 달리 GMR 과 RBC 는 정지영상을 대상으로 제안되었기 때문에 연속적 이미지 간의 의존성이 없어 동영상에 적용 시에는 프레임 별 출력 결과를 누적하여 결과를 추출하였다. 양자화 지도(Q)는 16x16 매크로블록 단위로 계산하며 해당 매크로블록 위치에서의 픽셀 별 중요도 합이 4096 이상이면 관심 영역으로 판단하여 1 로 표시하였다. 또한, H.264(AVC) encoder 는 JM(18.6)을 사용하였으며, 매크로블록 레벨로 양자화 파라미터를 할당하며, 관심 영역과 비관심 영역간 양자화 파라미터를 양자화 지도에 따라 다음과 같이 할당하였다.

$$QP_{mod}(i, j) = \begin{cases} QP_{org}(i, j) - \Delta, & \text{if } Q(i, j) = 1, \\ QP_{org}(i, j) + \Delta, & \text{others,} \end{cases}$$

여기서 QP_{org} 는 JM encoder 가 자동으로 할당하려는 양자화 파라미터이며, QP_{mod} 는 관심 영역 여부에 따라 변경되는 양자화 파라미터이다. i, j 는 매크로블록의 위치좌표이다.

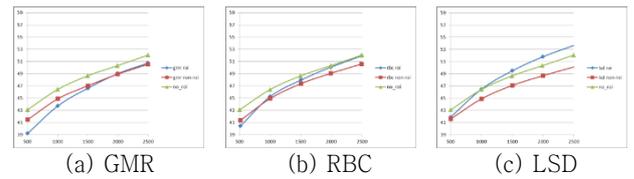
3. 실험 결과 및 결론

전처리 기반 방식의 경우 중요도 지도가 제대로 추출되지 않게 되면 인코더에 입력되기 이전에 화질 열화를 유발한다. 그림 3(a) LSD 결과에서 얼굴이 흐릿한 것과 (b) RBC 결과에서 눈 주위가 흐릿한 것은 모두 중요도 감지의 실패가 그 요인이다. 또한 각 프레임 별 중요도 지도의 시간적 일관성이 보장되지 않을 경우 영상 일부에서 고화질과 저화질이 교차하게 되어 시각적으로 부자연스럽게 된다.



(a) (b)

그림 3. 전처리 기반 방식 변경된 입력 시퀀스(VGA 해상도) 결과 (좌상:RWR, 우상:GMR, 좌하:LSD, 우하:RBC)



(a) GMR (b) RBC (c) LSD

그림 4. 중간처리 기반 방식에 따른 rate-distortion 그래프 ($\Delta=2$, 적:관심영역, 청:비관심영역, 녹색: 관심영역 기법 미 적용)

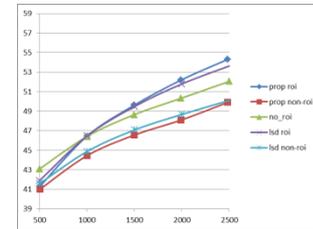


그림 5. LSD 대 RWR 간 rate-distortion 그래프 비교 ($\Delta=2$)

그림 4 와 5 는 중간처리 기반 방식 적용 시 RWR 과 세가지 비교 알고리즘에 대한 rate-distortion 결과를 보여준다. 정지영상 대상 알고리즘인 GMR(그림 4(a))과 RBC(그림 4(b))는 관심 영역, 비관심 영역 모두 제안 기법을 적용하지 않은 결과보다 좋지 않았으며, 특히 RBC 의 경우 1Mbps 이상 비트율이 할당된 경우 비관심 영역의 화질이 오히려 좋아지는 역전현상이 발생하였다. 반면 동영상 대상 알고리즘인 LSD 의 경우 1Mbps 이상에서 관심 영역의 화질이 관심 영역 기법을 미 적용한 경우보다 높음을 확인할 수 있다. RWR 방식의 경우 LSD 와 비슷한 패턴을 보이지만 LSD 에 비해 관심 영역은 더 좋은 화질로 비관심 영역은 더 낮은 화질로 성공적으로 비트량을 할당했음을 보여준다. 하지만 LSD 와 같이 1Mbps 이하에서의 효율은 떨어짐을 확인했다.

이상의 실험결과를 통해 관심 영역 동영상 압축을 위해 사용되는 중요도 지도는 중요도 감지의 신뢰도 뿐 아니라 프레임 간 시간적 일관성 및 변이를 감안한 알고리즘이어야 함을 확인했다. 하지만, 효율적인 압축을 위해서는 단순한 양자화 파라미터의 변경만으로는 부족하다. 1Mbps 이하의 저 화질 영상에서도 적응적으로 적용하기 위해서는 rate-distortion 최적화를 위한 적절한 모델링이 필요함을 확인할 수 있었으며, 특히 이미지 내 관심 영역의 비중 및 비트율과 양자화 파라미터의 관계에 대한 고려도 필요하다.

4. 참고문헌

[1] J.-S. Kim, H. Kim, J.-Y. Sim, C.-S. Kim, and S.-U. Lee, "Video saliency detection based on random walk with restart," in Proc. IEEE ICIP, Sep. 2013, pp. 2465-2469.

[2] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in Proc. IEEE CVPR, 2013, pp.3166-3173.

[3] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Haung, and S.-M. Hu, "Global contrast based salient region detection," in Proc. IEEE CVPR, Jun. 2011, pp. 409-416.

[4] Y. Xue, X. Guo, and X. Cao, "Motion saliency detection using low-rank and sparse decomposition," in Proc. IEEE ICASSP, Mar. 2012, pp.1485-1488.