

기계 소비 기반의 비디오 부호화를 위한 시공간적 재표본화 기술 연구

안은빈 / 연세대학교 Visual Communication Lab.

머신 비전 기술의 발전으로 디지털 비디오의 소비 주체가 인간에서 기계로 전환되고 있다. 기존 인간 중심의 비디오 부호화 기술은 기계 소비를 위해 생성된 비디오 데이터를 압축하는 것에 한계가 있다. 본 연구는 기계 중심 비디오 소비의 특성을 분석하고, 그 특성을 반영하여 기계를 위한 새로운 부호화 기술을 제안한다.

먼저, 인간의 소비를 고려하여 정의된 기존의 비디오 데이터의 중복성 정의와는 차별화된 기계 소비 기반의 비디오 데이터 중복성인 ‘객체 분포 기반 중복성’과 ‘객체 움직임 기반 중복성’을 정의한다. 각각의 중복성에 대하여 공간적 및 시간적 재표본화 기법을 적용하여 기계 소비 관점에서 불필요한 중복된 비디오 데이터를 효율적으로 제거하는 방법을 제시한다. 또한 기계의 임무 성능 유지 대비 얼마나 효율적으로 비디오 데이터를 절감했는지를 평가하기 위하여 DRAER (Data Reduction to Accuracy Error Ratio) 지표를 새롭게 제안한다.

특히, 객체 움직임 기반 중복성을 효과적으로 제거하기 위하여 객체 추적 정보를 활용한 적응형 시간적 재표본화 기법을 개발하였다. 본 기법은 각 프레임으로부터 수집된 객체 추적 정보를 기반으로 프레임 간 유사도를 분

석하여 대표 프레임을 제외한 나머지 프레임을 적응적으로 제거한다. 프레임 유사도를 기반으로 재구성된 비디오를 구조적으로 표현하기 위하여 새로운 픽처 구조인 GoS (Group of Similar Pictures) 및 SGoS (Set of GoS)를 설계하였다. GoS는 객체 추적 정보를 기반으로 유사하다고 판단된 픽처들을 하나로 묶은 단위를 의미하며, 시간적 재표본화된 하나의 GoS는 한 픽처로 구성된다. SGoS는 독립적으로 복호화될 수 있는 GoS의 묶음을 표현하기 위하여 개념적으로 설계된 구조이다.

제안된 기법은 MPEG VCM (Video Coding for Machines) 시스템에 통합되어 종단간 압축 프레임워크로 구현하고, 기계 임무 정확도 기반의 BD-rate로 성능을 평가하였다. 또한, 시간적으로 적응형 재표본화된 비디오의 임의 접근 호환성 문제를 해결하기 위하여 모조 픽처를 삽입하는 방안을 함께 제시하였다. VCM CTC (Common Test Condition)에 따라 다양한 데이터셋과 코딩 환경에서의 실험이 진행되었다. 제시된 실험을 통해 BD-rate 절감과 기계 인식 정확도 유지 측면에서의 효과가 입증되었다.



안 은 빈

- 2016년 8월 : 연세대학교 정보통신공학 학사
- 2025년 8월 : 연세대학교 전산학 박사
- 주관심분야 : 기계를 위한 비디오 부호화, 몰입형 미디어