

딥러닝 기반 저화질 미디어의 고품질 변환시스템에 적용하기 위한 영상특성기반 콘텐츠 분석기 구현

라상중, 조숙희, 김나연*
한국전자통신연구원, *과학기술연합대학원대학교
sjna@etri.re.kr

Implementation of video characteristic-based content analyzer to apply to high-quality conversion system of low-quality media based on deep learning

Sang-Jung Ra, Sukhee Cho, Na-Yeon Kim*
Electronics and Telecommunications Research Institute, *University of Science & Technology

요 약

본 논문은 저화질 영상을 고품질로 변환하는데 필요한 요소기술을 영상의 특성에 맞게 적용하기 위하여 원영상의 특성을 분석하는 콘텐츠 분석기에 대하여 설명하고, 구현된 영상특성 기반 콘텐츠 분석기의 동작 결과를 확인한다. 시간해상도 및 공간해상도 증가, 색영역 및 명암비 확장 등의 화질 향상을 위한 요소기술을 영상 특성에 맞게 적용하기 위하여, 원영상에 대해 장면단위로 시공간 복잡도를 분석한 후 복잡도별로 구분한 장면에 대해 요소기술을 선택적으로 융합 적용하여 화질 개선을 위한 성능을 향상시킬 수 있도록 한다.

I. 서 론

한류 콘텐츠의 세계시장 확산과 함께 과거에 제작된 구작 콘텐츠에 대한 국내외 수요가 급격하게 증가하고 있는 추세인 반면, 대형화된 디스플레이에 비해 열악한 원본 화질은 시청자의 만족도를 반감시키는 요인이다. 또한, 원본 영상의 리마스터링을 통한 개선된 변환 영상이 원본 영상의 특성을 충분하게 반영하지 못하여 원작의 느낌을 충실하게 전달하지 못하는 단점이 있다.

본 논문에서는 이러한 문제점을 극복하기 위하여 개발중인 저화질 영상을 고품질로 변환하기 위한 시스템을 구성하는 모듈 중 영상특성기반 콘텐츠 분석기에 대한 구조를 정의하고 동작 결과를 분석한다.

II. 고품질 변환을 위한 시스템 구조 설계

저화질 영상을 고품질로 변환하기 위한 시스템은 그림 1 과 같이 원영상의 특성을 분석하는 영상특성 분석 모듈과 분석된 특성에 따라 변환요소기술을 적용하여 화질을 개선하는 고품질화 모듈을 포함하여 구성된다.

고품질 영상 변환을 위한 시스템 내의 영상특성 분석 모듈은 장르분석 유닛과 시공간복잡도 분석 유닛을 포함한다. 우선 장르분석 유닛은 원본 영상의 콘텐츠 장르를 분석하여 구분하는 유닛으로 영화/드라마/스포츠/뉴스 등의 카테고리에 따라 장르를 분류한다. 다음의 시공간복잡도 분석 유닛은 입력된 원본

영상을 장면(scene) 단위로 분할하고, 분할 장면에 따라 시간 및 공간해상도에 대한 복잡도를 계산하여 설정된 임계값(threshold)을 기준으로 구분한다.

시스템을 구성하는 다른 모듈인 고품질화 모듈은 고해상화(Super Resolution) 유닛, HFR(High-Frame Rate) 유닛, HDR(High-Dynamic Range) 유닛, WCG(Wide-Color Gamut) 유닛 및 화질개선 유닛으로 구성된다. 먼저 고해상화 유닛은 SD(480p)의 경우 최대 4K(2160p) 해상도로 변환하며, HD(1080p)의 경우 최대 8K(4320p) 해상도로 변환하여 해상도를 확대하는 기능을 하고, HFR 유닛은 60i(인터레이스 방식)를 60p(프로그레시브 방식)로 변환하거나 60p 를 최대 240p 로 변환하여 시간당 처리하는 프레임의 수를 늘려 더욱 자연스러운 화면 처리를 제공하는 기능을 한다. HDR 유닛은 SDR(Standard Dynamic Range) 영상을 HDR(High Dynamic Range) 영상이나 HLG(High Log Gamma) 영상으로 변환하여 더욱 선명한 화면을 제공하는 기능을 하고, WCG 유닛은 BT.601 을 최대 BT.2020 으로 색영역을 확장하여 최대한 자연색에 가까운 화면으로 변환하는 기능을 한다. 마지막으로 화질개선 유닛은 압축왜곡 개선 기능을 적용한 영상의 부호화 과정을 통해 발생하는 왜곡을 개선하는 기능을 수행하며, 디블러 기능을 통해 영상의 색 번짐을 줄이고 디노이즈 기능을 통해 영상의 잡음 성분을 제거한다.

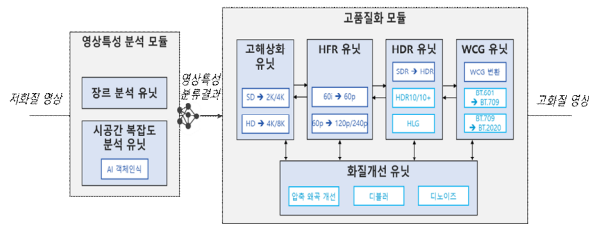


그림 1. 고품질 변환을 위한 시스템 구조

III. 영상특성기반 콘텐츠 분석기 설계 및 동작 분석

그림 2 는 고품질 영상 변환을 위한 영상특성기반 콘텐츠 분석기의 구성도를 개념적으로 나타낸 것이다. 입력된 동영상 파일에서 오디오, 비디오를 추출하고, 추출된 비디오에 대해 복호화, 장면(scene)단위의 샷 분할, 시공간 복잡도 분석, 동영상 콘텐츠 A/V 동기화 및 부호화 등을 통해 최종 출력 파일을 생성한다. 이러한 영상특성기반 콘텐츠 분석기는 입력된 비디오에 대하여 동작을 수행하며, 오디오는 입력 오디오에 대한 별도의 동작을 수행하지 않는다.

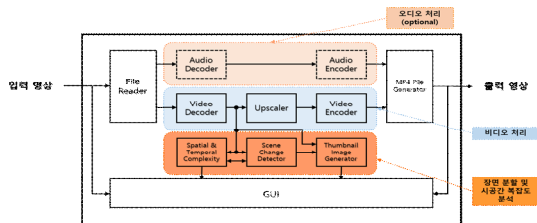


그림 2. 영상특성기반 콘텐츠 분석기 구조

영상특성기반 콘텐츠 분석기는 입력되는 저화질 영상파일을 처리하기 위한 File Reader, File Reader 모듈에 의해 추출된 저화질 비디오 ES 를 복호화 하기 위한 Video Decoder, 저화질 영상에 대한 업샘플링을 처리하기 위한 Upscaler, 장면 전환 검출을 위한 Scene change detector, 영상의 시공간 복잡도를 파악하기 위한 Spatial & Temporal Complexity, 썸네일 이미지 생성을 위한 Thumbnail Image Generator, 업샘플링된 결과 영상을 부호화하기 위한 Video Encoder, 고품질 영상 출력파일을 만들기 위한 MP4 File Generator 와 이를 동작 제어하기 위한 GUI 로 이루어져 있다.

그림 3 은 영상특성기반 콘텐츠 분석기의 동작 화면을 나타낸다. 분석기에 영상을 입력한 후 동작시키면 입력 영상에 4 배 업스케일링하여 변환 영상을 생성하고, 입력 영상에 대한 시간 및 공간복잡도 등을 분석한 특성 정보가 별도의 창을 통해 표시된다. 동작 화면 하단의 타임라인에서는 입력 및 변환 영상을 장면 단위로 분할한 영상이 표시된다.



그림 3. 영상특성기반 콘텐츠 분석기 동작 화면

입력 영상 및 변환 영상과 관련하여 현재는 단순 4 배 업스케일링만을 적용하여 동작하지만, 향후 변환요소기술이 적용된 변환 영상을 출력할 수 있도록 기능을 확장한다.

그림 4 는 영상특성기반 분석기 동작 후 별도 저장된 분석결과를 나타낸다. 장면단위로 분할된 각 영상에 대해 시작프레임, 종료프레임, 시간 및 공간복잡도 등을 별도 파일로 저장하여 분석할 수 있다.

그림 4. 영상특성 분석 후 저장 파일

별도 저장 파일에 분류된 시간 및 공간복잡도 분류 값(상/중/하 혹은 상/하로 구분)에 따라 장면을 구분하여 변환 요소기술을 선택적으로 융합적용 할 수 있도록 한다.

IV. 결론

본 논문에서는 딥러닝 기반 저화질 미디어의 고품질 변환 시스템을 구성하는 모듈 중 영상특성기반 콘텐츠 분석기의 구조를 정의하고, 입력 영상에 대한 특성분석 결과를 분석하는 것을 목적으로 한다. 향후 이러한 연구결과를 활용하여 상용화 수준의 통합 시스템을 구현하고 저해상도의 동영상에 대한 화질 개선을 수행하여 저화질 영상의 고품질 복원 및 녹화된 감시카메라 영상 개선을 통한 범죄 예방 등 다양한 분야에서 중요한 역할을 하기를 기대한다.

ACKNOWLEDGMENT

“이 논문은 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2021-0-00087, SD/HD 급 저화질 미디어의 고품질 변환 기술 개발)”

참 고 문 헌

- [1] Woonsung Park, Munchurl Kim, “Deep Predictive Video Compression using Mode-Selective Uni- and Bi-directional Predictions based on Multi-frame Hypothesis”, IEEE Access., vol. 9, pp. 72-85, Dec. 2020
- [2] Sehwan Ki, Jeonghyeok Do, Munchurl Kim, “Learning-based JND-directed HDR Video Preprocessing for Perceptually Lossless Compression with HEVC”, IEEE Access., vol. 8, pp. 228605-228618, Dec. 2020
- [3] ISO/IEC 23090-8:2020, Information technology — Coded representation of immersive media — Part 8: Network based media processing, 2020