

글로벌 라이프로그 미디어 클라우드 개발 및 구축

Global lifelog media cloud development and deployment

□ 송혁, 최인규*, 이영한, 고민수, 오진택**, 유지상* / 전자부품연구원, *광운대학교, **판도라티비

요약

글로벌 라이프로그 미디어 클라우드 서비스를 위하여 네트워크 기술, 클라우드 기술 멀티미디어 App 기술 및 하이라이팅 엔진 기술이 요구된다. 본 논문에서는 미디어 클라우드 서비스를 위한 개발 기술 및 서비스 기술 개발 결과를 보였다. 하이라이팅 엔진은 표정인식기술, 이미지 분류기술, 주목도 지도 생성기술, 모션 분석기술, 동영상 분석 기술, 얼굴 인식 기술 및 오디오 분석기술 등을 포함하고 있다. 표정인식 기술로는 Alexnet을 최적화하여 Alexnet 대비 1.82% 우수한 인식 성능을 보였으며 처리속도면에서 28배 빠른 결과를 보였다. 행동 인식 기술에 있어서는 기존 2D CNN 및 LSTM에 기반한 인식 방법에 비하여 제안하는 3D CNN 기법이 0.8% 향상된 결과를 보였다. (주)판도라티비는 클라우드 기반 라이프로그 동영상 생성 서비스를 개발하여 현재 테스트 서비스를 진행하고 있다.

I. 서론

멀티미디어 콘텐츠를 분석하여 이해하는 이미지 분석, 동영상 분석 및 오디오 분석 기술은 오래전부터 많은 연구자들이 관심을 가져온 영역이다. 오랜시간동안 꾸준한 연구를 통하여 그 성능은 지속적으로 향상되어 왔으나 실용화되기에는 부족한 수준이었다. 최근 딥러닝을 이용한 영상 분석 기술이 하드웨어 기술의 발전에 힘입어 급속히 발전되고 널리 퍼지고 있다. 구글은 고양이 얼굴을 1,000개의 컴퓨터를 이용하여 판별해내기 시작하였으며, P.sermanet et al.은 Kaggle competition에서 개와 고양이를 구분하는 수준으로 성능을 높였다[1-2]. 음성분야에서 딥러닝을 이용한 성능은 매우 우수하며 Deng, L et al.은 음성인식 오류율을 1.5% 이하로 줄였다[3].

* 본 논문은 미래창조과학부 SW컴퓨팅산업원천기술개발사업 (과제번호 R0190-16-1115)을 지원받아 수행한 결과입니다.

딥러닝을 이용한 멀티미디어 콘텐츠 분석 기술을 이용한 응용 분야는 매우 다양하다. 방송, 모바일 응용, 로봇, 자동차, 소셜미디어 등에서 활용 가능하며 본 논문에서는 미디어 클라우드 플랫폼 상에서 멀티미디어 분석 기술을 통한 라이프로그 서비스를 제공하는 시스템을 설명한다. 미디어 라이프로그 서비스는 개인 또는 대용량 클라우드에 저장된 개인의 사진, 동영상 및 오디오 정보를 활용하여 자동으로 데이터의 분류 및 멀티미디어 데이터를 생성하여 제공하는 서비스를 의미한다. (주)판도라티비와 전자부품연구원에서는 글로벌 미디어 클라우드 플랫폼 기반 개인 멀티미디어 콘텐츠를

자동으로 인식/처리하고 테마 기반 플랫폼을 활용한 디지털 라이프로그 형태의 서비스를 개발하였다.

본 논문은 본론에서 라이프로그 플랫폼에 포함된 지능형 하이라이팅 엔진과 이를 처리하기 위한 클라우드 서비스를 설명하며 결론에서 연구 결과를 상용화하기 위한 방안을 제시한다.

II. 본 론

본 연구에서 <그림 1>에서 보는 바와 같이 클라우



<그림 1> 글로벌 라이프로그 미디어 클라우드 서비스 구성도(Global lifelog media cloud service)

드 서비스를 위한 기능을 가지는 플랫폼을 개발하였다. 본 서비스의 기능은 크게 하이라이팅 기능, 편집 기능 그리고 뷰어 기능이다. 하이라이팅 엔진은 멀티미디어 분석 기술을 이용하여 이미지, 동영상 및 오디오의 콘텐츠에 포함된 내용을 이해하여 활용 가능한 데이터를 추출하는 모듈이며 편집 기능은 서버 및 모바일 단말에서 사용자에 자동으로 편집 동영상을 제공 또는 직접 편집 가능하도록 하며 뷰어에서는 사용자와의 관계 및 라이프로그를 보여준다. 하이라이팅 기능은 지능형 분석 하이라이팅 엔진에서 기능을 수행하며 편집 및 뷰어 기능은 단말 및 서버에서 기능을 수행한다.

1. 지능형 분석 하이라이팅 엔진

1) 하이라이팅 모듈

하이라이팅 모듈은 크게 클라이언트와 클라우드 서버로 구성되어 있으며 각 모듈의 기능은 <그림 2>와 같다.

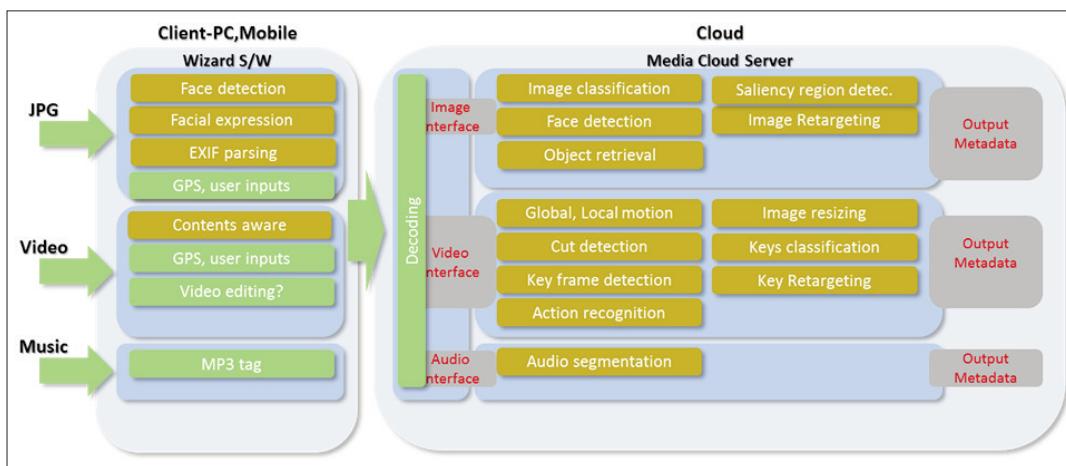
클라이언트는 모바일 디바이스, 또는 PC이며 멀티미디어 데이터의 획득 및 저수준의 영상 분석을 수행한다. 클라이언트에서 수행하는 기능은 이미지 분석, 비디오 분석 및 오디오 분석이며 이미지 분석은 얼굴 검출, 표정 인식 및 메타데이터 분석이다. 비디오 분석은 비디오 파일에 포함된 메타데이터를 추출한다. 오디오 분석은 오디오 파일에 포함된 태그를 추출한다.

클라우드 서버 역시 이미지 분석, 비디오 분석 및 오디오 분석 기능이며 대부분의 기능은 클라우드 서버에서 수행한다. 클라우드 서버에서 수행하는 세부 기술은 영상 분류, 주제도 지도 생성, 객체 검출, 모션 분석, 컷 검출, 키프레임 분석, 행동 인지, 표정 인식, 그리고 오디오 분할이다. 대부분의 기술은 딥러닝 기술을 활용하여 구현하였다.

2) 이미지 분석 모듈

- Image classification

미디어 클라우드 하이라이팅 서버에서 이미지 분



<그림 2> 하이라이팅 모듈 기능도(Functions in highlighting module)

석 모듈의 기능은 입력된 이미지 데이터를 분석하여 이미지 데이터를 분류하고 이미지 내에 포함된 객체를 인식하며 주목도 지도를 생성한다. 결과 데이터는 메타데이터로 저장이 되며 영상 생성 모듈에서 분류된 데이터를 활용하여 동영상을 생성한다. 이미지 분석 모듈은 CNN 기반 학습 모델을 활용하여 정확도를 높였다. Pretrain된 학습 모델은 Imagenet을 이용하여 학습되었으며 목표하는 이미지 카테고리 데이터를 부가 학습하여 성능을 향상시켰다. 얼굴, 해변, 책, 커피, 잔디, 단풍, 눈, 자동차, 개 및 하늘 등 10개의 카테고리로 구분하였다.

– Saliency region detection

라이프로그 동영상을 생성하기 위해 필요한 정보로는 이미지의 분류정보 및 콘텐츠에 포함된 객체 정보뿐 아니라 다양한 테마에 기반한 모션이 포함된 동영상을 생성하기 위해서 주목도 지도가 필요하다.

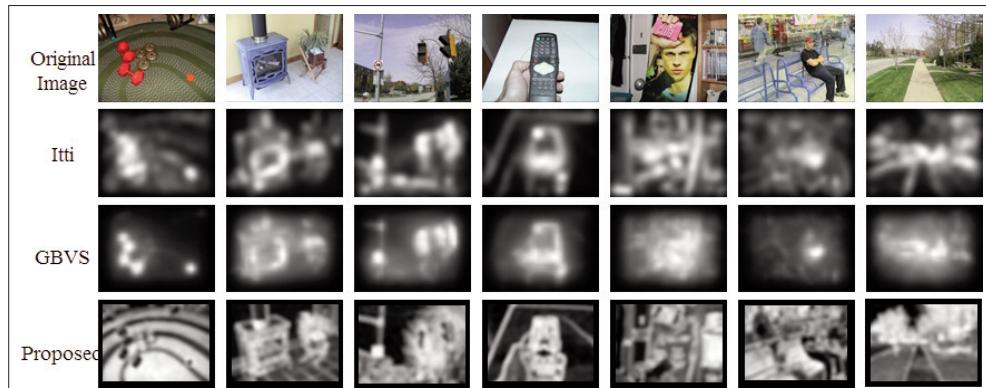
주목도 지도는 Local Image Patch를 이용하여 연산되며 79개의 basis 함수를 이용한다[4]. Basis 함수들은 infomax ICA(Independent Component Analysis)기법을 이용하여 다수의 자연영상에서 계

산하여 얻는다. 입력 영상의 모든 영역에 basis 함수를 곱해 79개의 특징지도를 얻는다. 커널 밀도 추정에 의해 특징지도의 각 계수값들을 이용하여 확률 분포를 얻을 수 있으며, 이를 이용하면 각 화소의 우도값을 계산할 수 있다. 각 특징지도들의 같은 화소 위치의 우도값들을 이용하여 결합 우도를 계산하며, 이는 Shannon의 정리에 의해 self-information 값으로 변환할 수 있다. 이 self-information 값이 각 화소 위치의 주목도 값이 된다. 해상도별 화질별 다양한 영상을 이용하여 〈그림 3〉과 같이 결과 테스트를 진행하였다.

– Image retrieval

영상 내에서 객체 검색 기법은 기존의 Convolutional layer와 Fully connected layer로 이루어진 CNN 구조에서 Fully connected layer를 제거하고 마지막 Convolutional layer의 2D 특징맵만을 이용한다. 식 (1)과 같이 마지막 특징맵들에서 최고값만을 추출하여 K차원의 특징벡터를 만든다[5].

$$f_{\Omega} = [f_{\Omega,1} \cdots f_{\Omega,i} \cdots f_{\Omega,K}]^T, \text{ with } f_{\Omega,i} = \max_{p \in \Omega} X_i(p) \quad (1)$$



〈그림 3〉 주목도 지도 생성 결과(Results of saliency region detection)

여기서 f_o 는 특징벡터를 나타내며, $f_{o,i}$ 는 i 번째 특징맵의 최고값을 나타낸다. 또한 $X_i(p)$ 는 i 번째 특징맵의 p 좌표에서의 값을 나타내며, K 는 마지막 convolution layer의 채널수를 나타낸다. 이러한 특징벡터를 query와 reference 영상에 대해 각각 계산하고 cosine similarity를 이용하여 query에 가장 가까운 reference 영상을 찾게 된다. 식 (2)는 cosine similarity를 계산하는 식을 나타낸다.

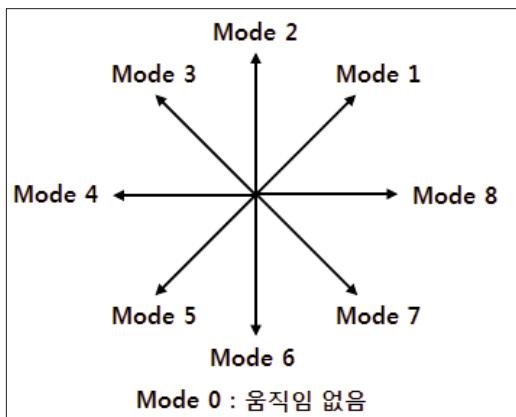
$$\text{similarity} = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} \quad (2)$$

여기서 A, B 는 유사도를 비교할 두 벡터를 나타낸다. 위와 같은 방법은 기존에 영상 분류를 위해 학습된 모델을 추가 학습 없이 사용할 수 있다는 장점이 있다.

3) 동영상 분석 모듈

– Motion analysis

동영상의 모션을 분석하고 이를 이용하여 영상의 카메라의 움직임을 추정한다. 카메라의 움직임은 ‘움직임 없음, 줌인, 줌아웃, 카메라이동, 객체중심 카메라이동’의 5가지로 분류한다. 먼저 동영상 프



〈그림 4〉 움직임에 따른 모드(Modes of motions)

레임간의 optical flow를 계산하여 움직임을 추정한다[6]. 〈그림 4〉와 같이 추정된 모션들 중 일정 크기보다 작은 모션들은 움직임이 없는 경우로 결정하고 나머지 움직임이 있는 경우는 360° 를 8방향으로 등분한 대표방향으로 양자화시킨다. 이렇게 9가지 값으로 변환된 움직임들은 각 모드에 대해 누적시켜 움직임 히스토그램을 얻을 수 있다[7].

움직임 히스토그램을 통해 카메라의 움직임을 추정하는 방법은 다음과 같다.

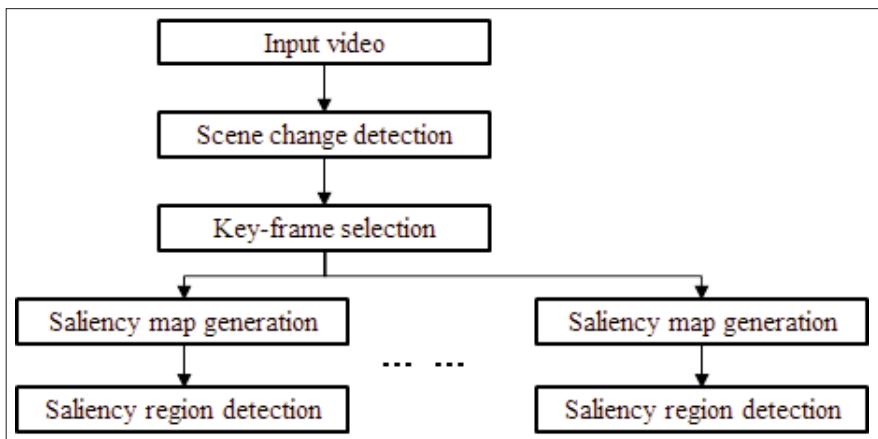
- 1) 움직임 없음 : Mode 0의 히스토그램값이 가장 클 경우
- 2) 줌인, 줌아웃 : 1)의 경우가 아닐 때, 8방향의 히스토그램값 모두가 일정 비율 이상의 값을 가질 경우
 - 2-1) 줌인 : 영상을 세로축으로 2등분하여 위 영역의 mode 1~3과 아래 영역의 mode 5~7의 값들의 합이 위 영역의 mode 5~7과 아래 영역의 mode 1~3값들의 합보다 클 때
 - 2-2) 줌아웃 : 2-1)이 아닌 경우
- 3) 카메라이동, 객체중심 카메라이동 : 1), 2), 모두 해당되지 않는 경우
 - 3-1) 객체중심 카메라이동 : 8방향 히스토그램 값의 합과 mode 0의 값의 비율이 일정 이상일 때
 - 3-2) 카메라이동 : 3-1)이 아닌 경우

– Scene change detection & key frame detection

키프레임 동영상에서 대표적인 프레임을 의미하며 비디오 콘텐츠의 요약 내용을 포함하고 있다. 키프레임을 사용함으로 비디오 데이터의 주 내용을 표현할 수 있다. 〈그림 6〉은 제안하는 알고리즘을 나타낸다[8].



〈그림 5〉 모션 분석 결과(Results of motion analysis)



〈그림 6〉 비디오 분석 알고리즘(Video analysis algorithm)

본 연구에 사용된 알고리즘은 고정 크기의 연속된 프레임에서 비유사도를 계산하기 위하여 Eigen value를 구하였다. SVD기반의 영상 압축 알고리즘을 이용하여 입력 영상을 변환하였다. 본 연산은 식 (3)과 같이 표현된다. A는 이미지 데이터 매트릭스를 의미하며 U는 Left singular 벡터, V는 Right

singular 벡터를 의미한다.

$$A = U \sum V^T = [u_1 \ u_2 \ \dots \ u_n] \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_n \end{bmatrix} \begin{bmatrix} q_1^T \\ q_2^T \\ q_3^T \\ q_4^T \end{bmatrix} \quad (3)$$

$$u_1 \sigma_1 q_1^T + u_2 \sigma_2 q_2^T + \dots + u_n \sigma_n q_n^T$$

4) 표정 인식 모듈

표정 인식은 멀티미디어 라이프로그 데이터에 포함된 감정표현을 추출하는 대표적인 방법이며 콘텐츠 제공을 위한 테마 선택에 활용 가능한 정보이다. 표정인식 모듈은 서버 및 단말에서 동작하며 단말에서의 동작을 위하여 최적화 설계되었다. 표정 데이터 중대를 위하여 얼굴 각도를 여러 각도로 변환하여 학습 데이터 셋의 수를 늘렸다. 기본적인 학습 모델은 Alexnet을 활용하여 설계되었다. Alexnet 모델을 활용하여 표정 인식 모듈을 구동할 경우 소형 단말기기에서의 동작은 상대적으로 긴 동작 시간이 요구된다. 따라서 모델의 최적화가 필수적이다. Alexnet의 구조를 각 레이어별로 최적화하여 <표 1>과 같은 최적화 결과를 도출하였다[9].

표정인식의 평균 인식률은 <표 2>와 같이 96.88%를 보였으며 이를 모바일 단말기인 삼성 갤러시 S6에서 동작하였을 경우 0.8초 이내 동작 속도를 보여 Alexnet을 구동하였을 경우 22초 동작

한 결과에 비하여 속도 개선을 하였다.

5) 행동 인식 모듈

멀티미디어 콘텐츠 내에 다양한 행동을 분석하여 객체간 행동 정보를 활용하여 콘텐츠 분석 및 동영상 생성에 활용한다. 행동인지를 위한 구조는 Donahue, et al.의 Long term recurrent convolutional network 기법을 활용하여 <그림 7>과 같이 3D CNN 구조를 적용하였다[10]. 제안하는 방법은 2D CNN구조를 3D CNN으로 변형하여 활용하였다. 독립적인 3D CNN으로 입력되는 연속적인 프레임의 수를 N이라 하고 각 3D CNN을 통해 나오는 값을 참조하는 단계의 수를 T라고 정의한다. 연속적인 프레임에서 $N \times T$ 개의 프레임을 선택할 때 시작 프레임의 간격을 Stride로 표현한다.

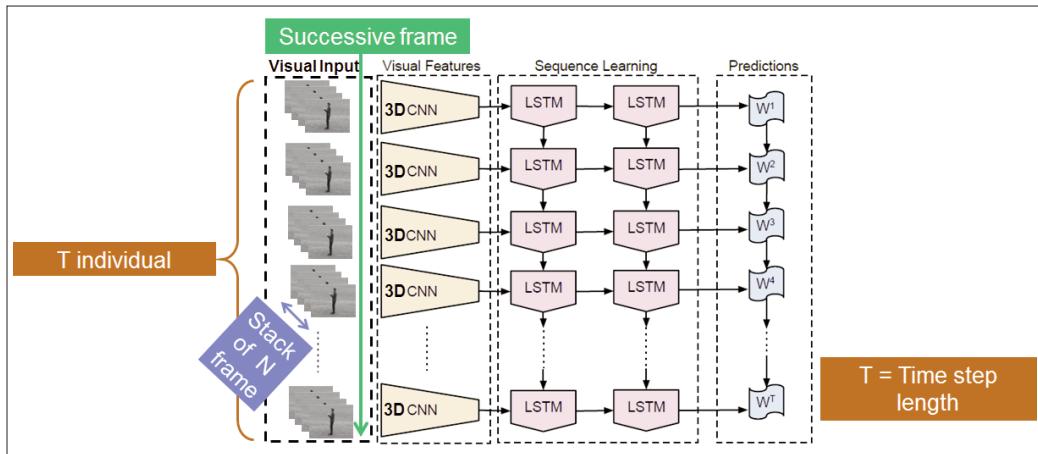
행동 인식률은 <표 3>에서 보는 바와 같이 기존 Donahue의 기법에 비하여 0.8% 향상된 결과를 보였다.

<표 1> 표정인식 결과(Results of facial expression recognition)

	Feature map channel/ Nodes	Data Augmentation	Ratio(%)	Size of Parameter(MB)
Intermediate1	96-256-384-384-256 / 4096-4096-6	O	95.05	217
Intermediate2	36-96-144-96-128 / 1024-1024-6	O	96.88	6.2
Proposed	36-96-144-96-128 / 1024-1024-6	X	88.8	6.2

<표 2> 카테고리별 인식 결과(Confusion matrix of facial expression recognition)

	Neutral	Happy	Sadness	Angry	Surprise	Disgust	Total	Ratio(%)
Neutral	82	3	1	2	0	0	88	93.18
Happy	0	88	0	0	0	0	88	100
Sadness	2	0	44	0	0	0	46	95.65
Angry	0	1	1	51	0	2	55	92.73
Surprise	0	0	0	0	57	0	57	100
Disgust	0	0	0	0	0	50	50	100
							384	96.88



〈그림 7〉 3D convolution과 LSTM을 이용한 행동인식 기법(Action recognition using 3D convolution and LSTM)

〈표 3〉 행동인식 결과(Results of action recognition)

CNN	Training Method	Input size	Data	Ratio (%)
2D (AlexNet)	Fine-tuning	224x224	T : 16 Stride : 8	93.7
2D (AlexNet)	Training	224x224	T : 16 Stride : 8	89.7
2D CNN (Intermediate)	Training	80x120	T : 16 Stride : 8	89.9
3D CNN (Intermediate)	Training	224x224	N : 3 T : 12	86.0
3D CNN (Proposed)	Training	80x120	N : 9 T : 7	96.35

6) 오디오 모듈

오디오 파일은 동영상 생성시에 있어서 장면전환이나 효과의 시작과 끝을 동기화하여 활용한다. 동영상의 비트를 추출하기 위하여 Mel-frequency cepstral coefficient와 Harmonic patch class profile을 사용하였다. 이후에 오디오 데이터를 Smoothing 기법과 Normalizing 기법을 통하여 분할하여 0.53의 F1 정확도를 추출하였다.

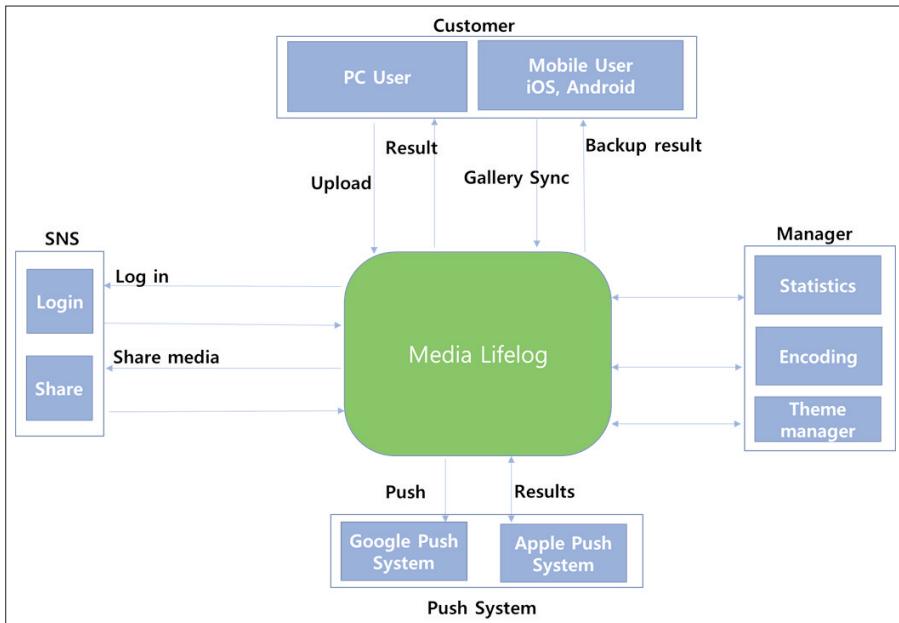
2. 클라우드 기반의 미디어 라이프로그 플랫폼 연구

앞선 영상 분석 기능을 이용한 하이라이팅 모듈

을 활용하여 사용자의 관심사별로 분류, 재가공된 비디오 로그들을 퍼스널 빅데이터와 융합하여 시각화 및 아카이빙 해주는 글로벌 미디어 클라우드 서비스를 위한 주요 기술은 아래와 같다. 크게 미디어 관제 시스템과 NLE 렌더링 사이의 전처리, 스트리밍을 위한 딜리버리 시스템, 콘텐츠 및 메타데이터 분석 처리시스템 그리고 인코딩 센터로 구분된다.

1) 미디어 관제 시스템과 NLE 렌더링 사이의 전처리 프로세스

① NLE에서 영상의 필요 구간으로 직접 Seek 인식 처리로 렌더링 과정의 프로세스 단계가 줄어들어 효율적임



〈그림 8〉 글로벌 라이프로그 미디어 클라우드 콘텍스트 구조(Global lifelog media cloud development and deployment Context Diagram)

② 필요한 구간을 미리 잘라내어 NLE에 전달하여 처리함으로써 NLE에서는 입력 영상의 포맷과 Codec을 고려하지 않아도 됨

2) 모바일 어플리케이션 스트리밍에 적합한

딜리버리 시스템

- ① 모바일에서 미디어 서버로 요청이 가능할 때 사용자는 딜리버리 시스템으로 영상 재생 요청을 할 수 있으며 5초 이내에 영상 시작
- ② Http Live Streaming 방식 : mp4 파일을 chunk 단위로 다운 받아 재생하는 방식
 - Mp4를 hls로 변환 시 서버의 리소스가 증가 될 것으로 예상되지만 변환 작업은 파일로딩 후의 연산에 대한 내용이기 때문에 CPU 사양으로서 커버가 가능함
 - 전체 파일에 대한 다운로드 없이 Chunk 단

위의 요청이기 때문에 서버의 부하도 그만큼 낮아짐

3) 콘텐츠 및 메타데이터 분석 처리 시스템

① 콘텐츠 분석 처리

- 영상 분석 과정에서 생성된 이미지는 카테고리 분류를 위한 썸네일이며 영상 분석 결과를 바탕으로 생성된 이미지는 영상의 정확한 구간별 대표 썸네일로 추출함
- Message queue 방식은 분석 요청이 많아질 때를 대비해서 신뢰성, 안정성, 확장성을 보장

② 메타데이터 분석처리 시스템

- 콘텐츠 분석기를 통해 분석된 결과를 메타데이터 수집기를 통해 분석된 메타데이터 (GPS, 날짜)를 이용하여 날씨정보와 메타데

이터로 사용자가 감수성을 느끼게 하는 텍스트 정보 생성

- 분석을 위해 MongoDB의 Map & Reduce, Aggregation을 이용하여 관리자에게 실시간 데이터를 조회하고 여러가지 프레임워크의 조합 없이 단순한 구조로 관리자가 원하는 데이터를 MongoDB의 Aggregation을 통해 빠르고 유연하게 대처
- 단순한 구조로 인하여 배포의 용이성, 유지 보수성이 향상됨

4) 인코딩 Job 분할 분산 인코딩 센터

- ① Encoding 엔진에서 입력 영상을 Keyframe 단위로 분할하고 분할된 영상은 시스템의 물리적 core 개수 만큼의 Encoding thread에 분산시켜 처리하여 하드웨어 성능을 최대한 활용, Encoding 속도를 최대화
- ② GPU 가속(nvidia cuda, Intel media sdk)을 이용하여 Video stream을 인코딩하고 CPU는 Audio stream을 인코딩하는 하이브리드 방식을 적용하여 빠른 인코딩을 구현

III. 결 론

본 연구를 통하여 전자부품연구원과 (주)판도라티비는 글로벌 라이프로그 미디어 클라우드 서비스를 오픈하였으며 현재 내부 테스트를 진행하고 있다. 전자부품연구원의 하이라이팅 엔진은 이미지, 동영상 및 오디오 콘텐츠를 분석하여 사람이 인식할 수 있는 대부분의 정보를 분석하여 메타데이터로 추출하였으며 각 세부 모듈의 인식률은 기존 연구와 비교하여 우수한 결과를 보였다. 하이라이팅 분석 엔진은 글로벌 라이프로그 미디어 클라우드 서비스 뿐 아니라 보안, 방송, 엔터테인먼트, 차량, 교육 등 다양한 분야에 활용될 수 있을 것으로 보인다.

현재 하이라이팅 엔진의 대부분 모듈이 딥러닝 기반 지능형 알고리즘을 구현되어있으나 일부 모듈에 있어 아직 연구가 미진한 부분이 있어 이에 대한 연구가 진행되어야 하며 상용화를 위해서 각 모듈별 최적화 작업을 진행중이다.

참 고 문 헌

- [1] Le, Quoc V. "Building high-level features using large scale unsupervised learning." 2013 IEEE international conference on acoustics, speech and signal processing. IEEE, 2013.
- [2] Sermanet, Pierre, et al. "Overfeat: Integrated recognition, localization and detection using convolutional networks." arXiv preprint arXiv:1312.6229 2013.
- [3] Deng, Li. "Deep learning: from speech recognition to language and multimodal processing." APSIPA Transactions on Signal and Information Processing. 2016.
- [4] N.D.B. Bruce, J.K. Tsotsos, "Saliency Based on Information Maximization," Advances in Neural Information Processing Systems, 18, pp. 155-162, June 2006.
- [5] Tolias, Giorgos, Ronan Sicre, and Hervé Jégou. "Particular object retrieval with integral max-pooling of CNN activations." arXiv preprint arXiv:1511.05879 (2015).
- [6] Zach, Christopher, Thomas Pock, and Horst Bischof. "A duality based approach for realtime TV-L 1 optical flow." Joint Pattern Recognition Symposium. Springer Berlin Heidelberg, 2007.

- [7] H. Kim, S. Lee and A. C. Bovik, "Saliency Prediction on Stereoscopic Videos," in IEEE Transactions on Image Processing, vol. 23, no. 4, pp. 1476-1490, April 2014.
- [8] Min Soo Ko, Hyok Song, "Video Analysis Algorithm based on Saliency Region Detection from Selected Key-frames", ITC-CSCC 2016.
- [9] In Kyu Choi, Hyok Song, Jisang Yoo, "Convolutional Neural Networks for Facial Expression Recognition", KOSBE, 11. 2016.
- [10] Donahue, Jeffrey, et al. "Long-term recurrent convolutional networks for visual recognition and description." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.

필자 소개



송 혁

- 1999년 2월 : 광운대학교 제어계측공학과 학사
- 2001년 2월 : 광운대학교 전자공학과 석사
- 2013년 2월 : 광운대학교 전자공학과 박사
- 2000년 ~ 현재 : 전자부품연구원 근무
- 주관심분야 : 영상 인식, 딥러닝, 영상 보안



최 인 규

- 2014년 2월 : 광운대학교 전자공학과 학사
- 2016년 2월 : 광운대학교 전자공학과 석사
- 2016년 2월 ~ 현재 : 광운대학교 전자공학과 박사과정
- 주관심분야 : 컴퓨터 비전, 영상 인식, 딥러닝



이 영 한

- 2005년 2월 : 광운대학교 전자공학과 학사
- 2007년 2월 : 광주과학기술원 정보통신공학과 석사
- 2011년 8월 : 광주과학기술원 정보통신공학부 박사
- 2011년 9월 ~ 2014년 12월 : LG전자기술원 선임연구원
- 2015년 1월 ~ 현재 : 전자부품연구원 근무
- 주관심분야 : 음성/오디오 신호처리, 머신러닝

필자소개



고민수

- 2010년 2월 : 광운대학교 전자공학과 학사
- 2012년 2월 : 광운대학교 전자공학과 석사
- 2016년 2월 : 광운대학교 전자공학과 박사
- 2016년 ~ 현재 : 전자부품연구원 근무
- 주관심분야 : 영상신호처리, 머신러닝



오진택

- 2000년 2월 : 부산과학기술대학교 산업디자인학과
- 2011년 8월 : 고려사이버대학교 경영학과 학사
- 2001년 4월 ~ 2005년 3월 : 유비정보시스템 근무
- 2005년 7월 ~ 2008년 3월 : 한일네트워스 근무
- 2011년 6월 ~ 2013년 12월 : 네오플러스 근무
- 2014년 1월 ~ 현재 : 판도라티비 근무
- 주관심분야 : 영상 처리 및 관제시스템, 클라우드 서비스



유지상

- 1985년 : 서울대학교 전자공학과 공학사
- 1987년 : 서울대학교 전자공학과 공학석사
- 1993년 : Purdue Univ., EE, PH. D.
- 1997 ~ 현재 : 광운대학교 교수
- 주관심분야 : 방송기술, 컴퓨터 비전, 영상처리