



웨어러블 라이프로그 기반 상황 인식과 응용 시나리오

Wearable Lifelog-based Context Awareness and its Application Scenario

이충연, 이상우, 곽동현, 장병탁

한국정보과학회 추계학술발표회 2014

2014. 12. 19

서울대학교 컴퓨터공학부 바이오지능연구실

연구 목적 및 필요성

■ 라이프로그(Lifelog)

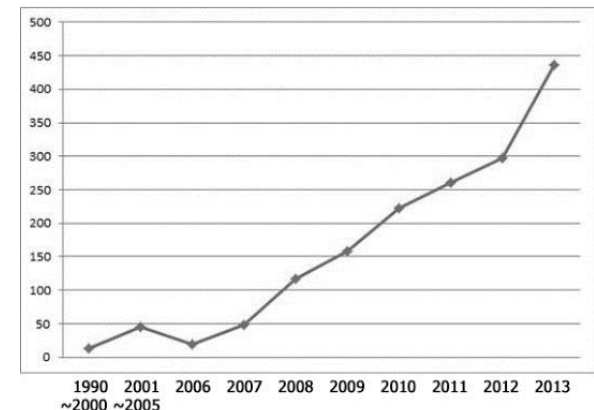
- 개인의 일상생활에 대한 디지털 기록
- 일인칭 기준 시점(Egocentric)에서 기록되는 영상, 음성, 신체 반응과 움직임, GPS를 통한 위치정보, 디바이스 사용 로그, 스케줄 내역, 이메일, 메신저 사용, 방문한 사이트 기록 등 → 데이터 범위가 매우 넓고 다양한 특성을 가짐
- 유비쿼터스(Ubiquitous) 컴퓨팅 환경에서 각종 지능형 계측 장비를 통해 수집
- 활용 가능성
 - 사물인터넷 환경에서 능동적인 맞춤형 서비스 제공을 위한 개인의 특성 명세 데이터로 활용 가능
 - 특정한 상황에서 개인의 의도 및 취향을 파악하는데 매우 중요한 정보

■ 라이프로그 연구 트렌드

- 2000년대 후반 이후 라이프로그 관련 연구가 급격하게 증가
- 스마트폰의 보급으로 라이프로그 수집이 용이해지게 됨

연도	1990 ~2K	2001 ~5	2006	2007	2008	2009	2010	2011	2012	2013
논문 편수	14	45	19	49	117	159	222	261	298	437

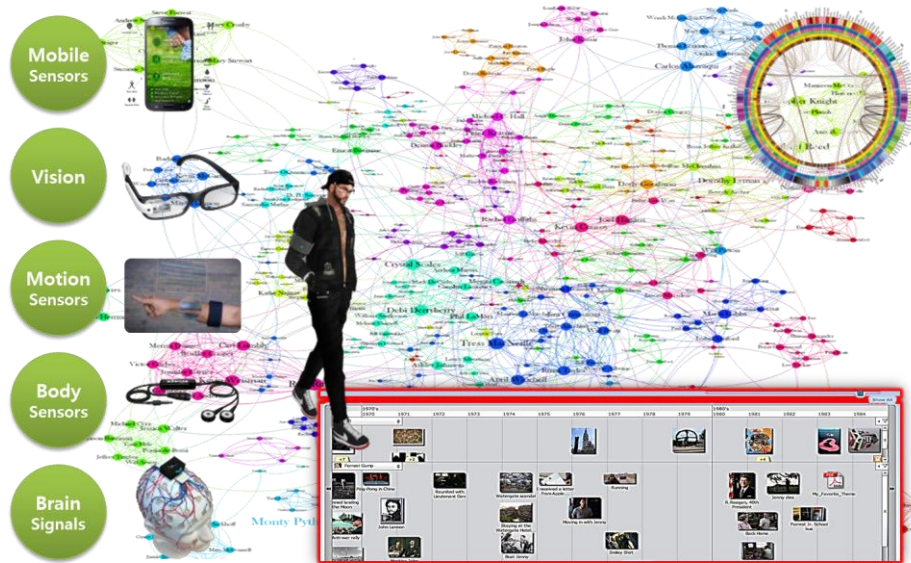
출처: Google Scholar (<http://scholar.google.com/>)



Wearable Devices and IOT

■ More Natural and Genuine Life-logging

- Wearable Device: 일상에서 몸에 걸치고 다닐 수 있는 형태로 제작된 정보 기기 → **Human**
 - Google Glass, Galaxy Gear, Fitbit, Misfit Shine 등
- Internet of Things: 각종 사물에 컴퓨터 칩과 통신 기능을 내장하여 인터넷에 연결 → **Environment**
- 실제로 사용자가 접하는 다양한 정보를 보다 자연스럽게 기록하기 위해 사용 가능
 - 스마트폰을 통해 기록 가능한 정보는 실제 사용자의 일상 정보를 충분히 기록하기에는 아직 부족
 - 특히 영상과 음성 정보는 보통 주머니에 들어가 있는 스마트폰으로 기록하는 것이 자연스럽지 않음



Wearable Lifelog 획득 방법

■ 영상 데이터 전처리 및 특징 추출

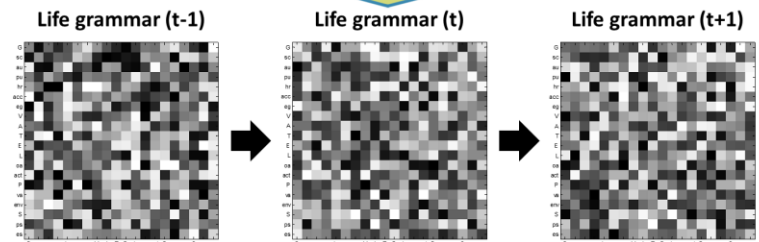
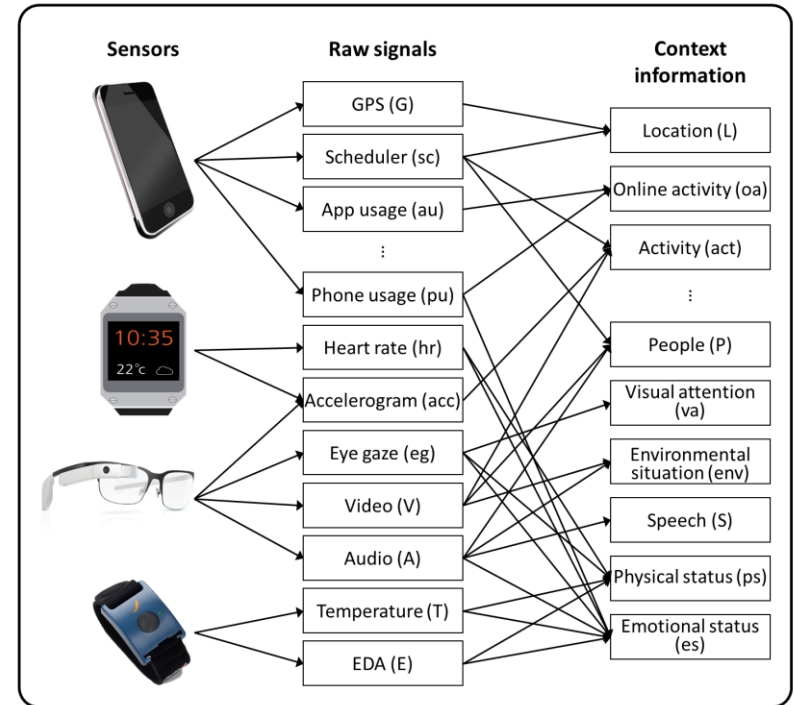
- Wearable Device들로부터 센서 정보와 이를 가공하여 얻을 수 있는 상황 정보를 추출

■ 1차, 센서 데이터

- 스마트안경: 일인칭시점 동영상, 시선추적정보 등
- 스마트워치: 심박수, 가속도, 체온, 피부전도도 등
- 스마트폰: 위치정보(GPS), 스케줄러, 통신기록 등

■ 2차, Context 정보

- 주변 환경 정보: Object, Scene, 밝기, 색상 등
- 사용자 정보: 활동, 대화, 주의 집중 정보 등



연구 내용 (1/3)

수집 데이터 통계

- 사용자가 구글 글래스를 착용하고 일상생활을 하면서 1인칭 영상과 소리를 수집

수업	여가	휴식	연구	도서관	이동
6,232	1,064	1,576	8,430	983	13,629

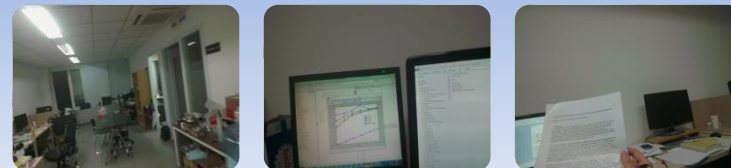


각 일상행동별 영상의 예

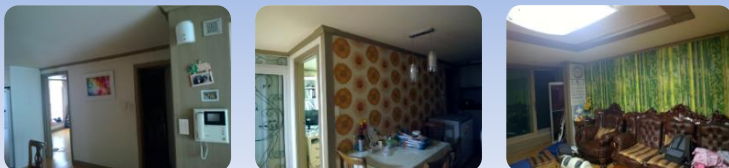
Take a class



In the lab



Rest at home



Moving



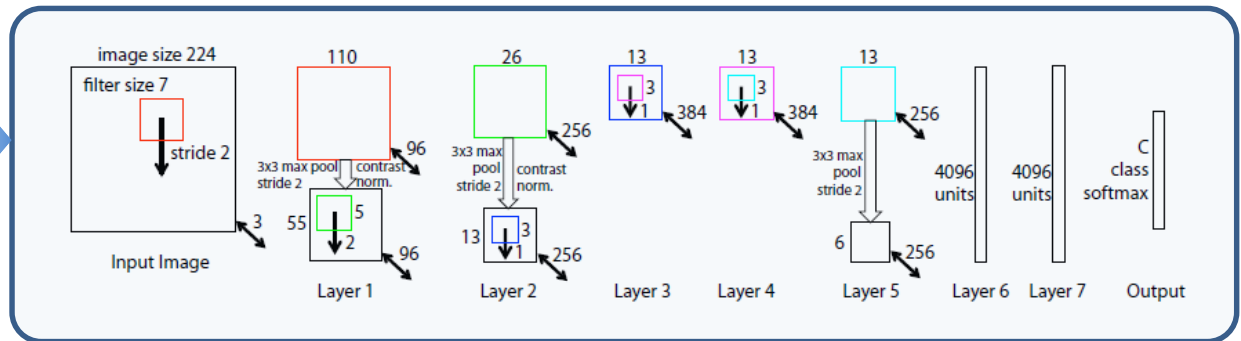
연구 내용 (2/3)

■ 딥러닝 객체인식 기반 확률적 영상 인자 추출

- 딥러닝 기법을 적용된 Convolutional neural network (CNN)을 이용하여 영상의 객체 및 특성을 표현하는 태그를 추출하여 확률적인 영상인자로 사용.
- Convolutional deep neural network은 영상에 정방향 윈도우로 스캔하면서 영상 표현에 필요한 국부적 요소들을 학습하는 기법으로 최근 영상처리에 널리 이용되고 있으며 영상의 특징 표현에 강력한 힘을 발휘하고 있음.

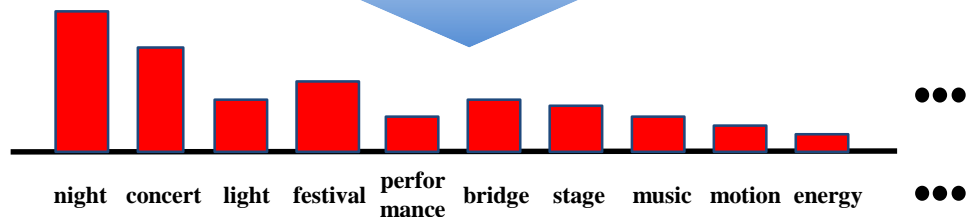


(입력) 일상 영상



사용된 8-계층 Convolutional deep neural network

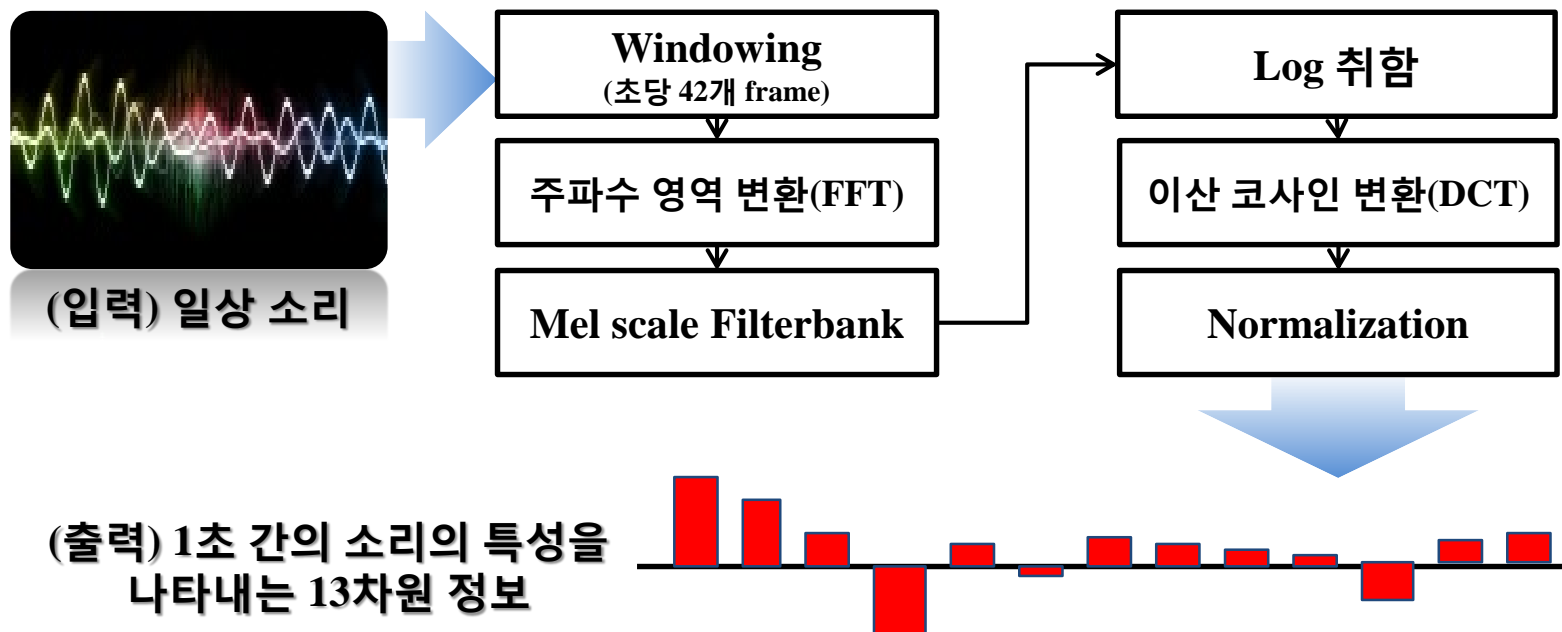
(출력) 영상을 표현하는
태그 형식의 Descriptor와 각 확률



연구 내용 (3/3)

MFCC 기반 소리 인자 추출

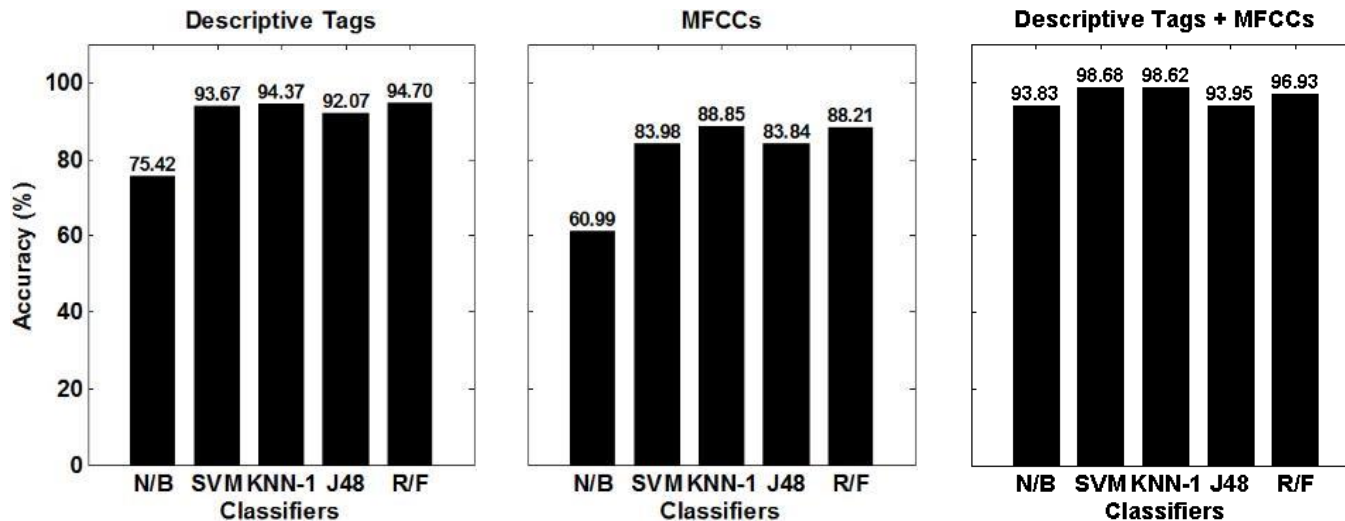
- MFCC 계수 추출 (13개 계수)
- 2048개 샘플로 구성된 프레임을 1024개 간격으로 이동시키면서 추출 (50% 중첩, 총 42개)
- Hamming windowing (불연속 지점 오류 최소화), FFT (주파수 영역 변환)
- Mel-scale filterbank (파워스펙트럼 추출), DCT (중첩된 필터뱅크 출력데이터 간의 상관관계 제거)



실험 결과

■ 실험 결과 및 논의

- 영상 데이터의 설명 태그와 음향 데이터의 MFCC 계수들을 특징값으로 사용하여 사용자의 활동 내역을 분류
- 분류기: Naïve Bayes (N/B), Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Decision Tree (J48, Random Forest)
- Naïve Bayes를 제외한 나머지 분류기에서 모두 높은 정확도를 보임
- 특히 영상 데이터 설명 태그는 90%가 넘는 정확도를 보여 높은 상황 설명력을 가지고 있음이 검증됨.
- 두 특징값을 모두 사용하여 분류한 결과 분류 정확도가 향상되는 결과를 나타냄.



논의 및 결론

■ 향후 연구 방안 및 활용

- 동적인 환경에서 실시간 멀티모달 센서 데이터에 기반을 둔 Lifelong 학습 및 추론 모델을 연구하기에 적합
- 산업적 응용 관점에서 웨어러블 디바이스가 스마트폰과 차별화되는 킬러앱으로 포지셔닝하는데 기여

문제	연구 주제 및 응용 시나리오
상황인식	<ul style="list-style-type: none">• 상황의 변화 및 예외적(abnormal) 상황을 감지• 학교에 가는 버스 탑승, 친구를 만나 대화 시작, 수업 시작 등의 상황을 실시간으로 인식하고, 사용자의 주의가 필요한 예외적 상황을 감지하여 알림
일상요약	<ul style="list-style-type: none">• Temporal clustering/segmentation• Micro/Macro 스케일의 상황 및 일상 요약, 기억 증강
상황예측	<ul style="list-style-type: none">• 재생성 학습 기반 상황 시뮬레이션 및 예측 추론 기법• 이동 경로 및 활동, 만날 가능성이 있는 사람, 나누게 될 대화 주제 등을 예측

■ 결론

- 웨어러블 라이프로그는 특정 시간대에서 개인이 얻게 되는 대다수의 정보를 포함
- 웨어러블 라이프로그는 인간의 기억과 학습에 사용되는 실제 데이터를 대표하는 것으로 볼 수 있음.
- 사람과 유사하게 학습 가능한 지능형 에이전트를 개발하는데 웨어러블 라이프로그가 효과적일 것으로 기대됨.
- 일인칭 시점의 영상과 음향 데이터를 이용한 상황 인식 실험을 통해 이러한 가능성을 확인