**1. 상황 인지의 정의에서 목표 범위 좁히기**

상황 인식은 (로봇의) 환경 중 상황 정보 혹은 객체의 상태를 인식하는 것으로 좁혀질 수 있다.

상황은 객체의 상태를 규정할 수 있는 어떤 정보도 포함한다. 객체는 사용자와 어플리케이션, 그리고 그 사이의 상호 작용과 관련되어 고려해야 할 사람, 장소, 사물을 가리킨다.

- Dey and Abowd(1999), Towards a Better Understanding of Context and Context-Awareness,

→ 상황 = 객체의 상태에 대한 정보

특정 상황에 관련된 객체가 지정되면 상황 정보를 추출 가능

- 류영달(2008), 상황인식 컴퓨팅의 현황과 전망

**2. 상황 인식 컴퓨팅 (Context-Aware Computing)과 지능형 로봇 (Intelligent Robot)**

상황 인식 컴퓨팅은 유비퀴터스 컴퓨팅에서 넘어온 개념으로 보임.

**2.1. 상황 인식 컴퓨팅**

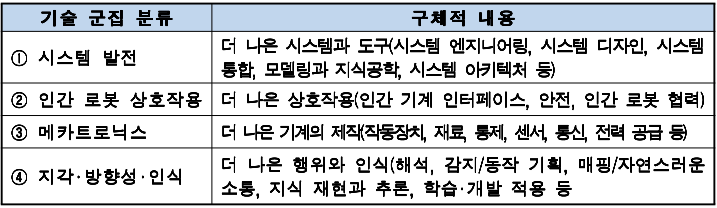
사용자의 업무와 관련 있는 적절한 정보, 지식, 통화, 또는 콘텐츠 서비스를 사용자에게 제공하는 공정에서 특정 사용자의 특정 상황'을 감지하고 인식한 다음 그의 요구대로 정보와 호출을 처리하고 나서 즉시 그의 단말기를 통해 표현해 주는 컴퓨팅

**2.2. 지능형 로봇**

외부환경을 인식(Perception), 스스로 상황을 판단(Cognition), 자율적으로 동작(Manipulation)하는 로봇

**3. 지능형 로봇**

**3.1. 로봇 기술의 목적별 분류**

출처: [주요국의 로봇 기술 정책 (2015?)](http://mdon.co.kr/news/download.html?no=3694&atno=7783)

**3.2. 4대 중점 돌파 기술**

지능형 로봇 = 기존 로봇 + 상황 판단 + 자율동작

- 상황판단 기능 = 환경인식 + 위치인식

- 자율동작 기능 = 조작제어 + 자율이동

**4. 환경 인식 > 위치 인식**

위치 인식 기능 보다는 환경인식 기능에 집중 (?)

“제품형 로봇에서는 카메라가 아닌 레이저 공간센서와 같은 측정기술을 삼차원으로 적용하는 형태가 될 것

예: 바닥을 스캔하여 이물질을 판단 후, 간단한 3차원 모델로 물체의 종류를 판단”

**4.1. 환경 인식/물체 인식 기술**

\* (미리 학습한 지식정보를 바탕으로) 물체의 영상에서 종류/크기/방향 위치 등 3D 공간정보를 실시간으로 파악

\* 로봇분야에서는 이동 중 인식기능 등에 특화, 예: 집안에서 특정물건을 구분 → 주인에게 배달 (심부름 로봇)

\* 인간과 같이 사물을 판별하는 기술은 앞으로 10년 이후에도 완벽히 재현하기 어려운 고난이도의 기술

→ 제품형 로봇에서는 카메라가 아닌 레이저 공간센서와 같은 측정기술을 삼차원으로 적용하는 형태가 될 것

예: 바닥을 스캔하여 이물질을 판단 후, 간단한 3차원 모델로 물체의 종류를 판단

**4.2. 위치인식 기술**

\* 위치인식률 50%정도로 앞으로 많은 연구가 필요한 분야

\* 가정용 로봇(청소/가사 등) 상용화를 위해 가장 먼저 풀어야 기술.

\* 센서기반, 마크기반, 스테레오 비전 기반 (현재 상용화 기술은 마크기반 정도)

- 예: 삼성 하우젠 (청소로봇) - 천정의 윤곽으로 위치를 인식

“임의의 공간에서 위치를 판별하는 기능 또한 현재로선 인공지능이라는 지식핸들링기술의 비밀이 벗겨지지 않는한 당장 실현될 것으로 보이지 않는다.”

“ 제품형 로봇에서는 RF기반 위치센서 등이 실내 GPS역할을 하며 위치를 판별할 수 있을 것으로 보인다.”

**5. 능동적 vs 수동적 상황인식 (Active vs. Passive Context Awareness)**

새로운 상황이나 상황 의 변경이 발생할 경우,

\* 자동으로 이와 관련된 동작을 실행하도록 구성 → 능동적 상황인식

\* 사용자에게 이에 대한 정보를 제공하거나 차후 검색이 가능하도록 정보를 유지 → 수동적 상황인식

능동성의 문제 = 인지의 기술적 어려움 + 낮은 정확도 + 사용자 선호도가 낮음

=> 수동성을 우선 + 특수 상황에서 능동성을 조금 넣는 시나리오(?)

=> 스타워즈의 Droids를 참고. C-3PO보다 R2-D2에 가까운 시나리오를 짜야함(?).

**6. 상황 정보의 4개 주요 성격**

객체의 정보를 4개 주요 성격으로 세분화

1. Identity: 객체에 식별자를 부여

2. Location: 2차원에서의 위치 정보 및 방향과 고도, 객체간 위치 관계, 내용물을 포함

3. Status/Activity: 객체가 가지고 있는 고유의 특성 중 센싱이 가능한 부분

4. Time

스마트하다고 느끼려면 과거 상황을 기억할 필요가 있음. 같은 내용이 반복되면 바보같이 느껴짐.

예: 안녕하세요, 회장님 처음 뵙겠습니다. (30초 후) 안녕하세요, 회장님 처음 뵙겠습니다. (계속 반복)

→ DB에 일정 분량의 과거 이력을 저장할 필요

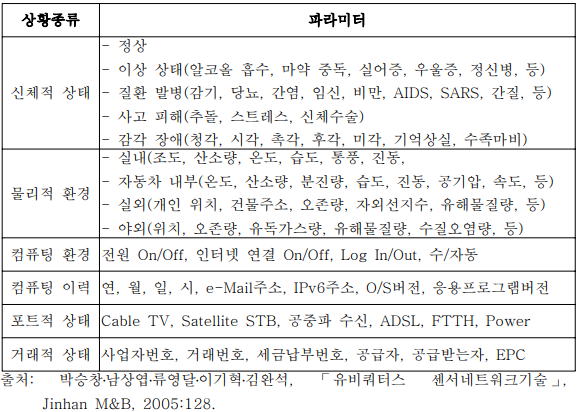
**7. 상황 정보의 파라미터**

\* 특정 개인의 상황정보를 정확하게 추출하기 위해서 해상도와 공차범위 내에서 감지.

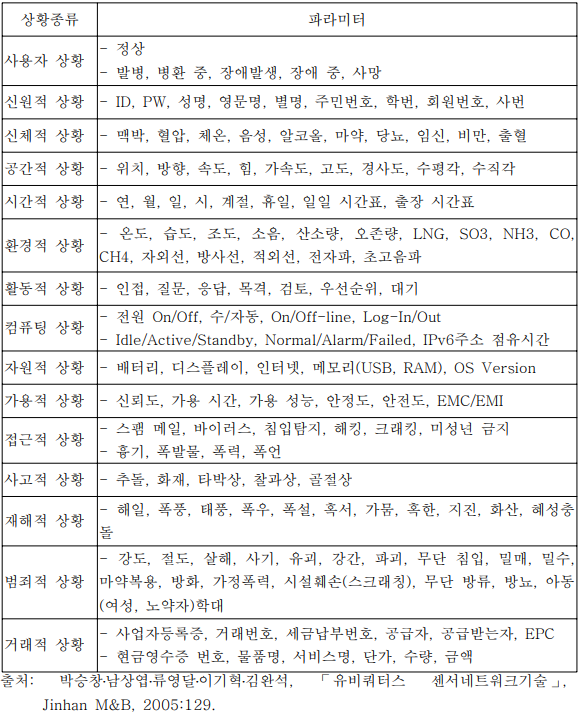
\* 로봇의 하드웨어로 처리 가능한 파라미터를 고려해서 구현 가능한 상황에 대한 시나리오를 좁힌다.

\* 상황 정보를 위한 서비스 파라미터를 참고.

**7.1. 참고용: 유비티즌의 상황인식 파라미터**



**7.2. 참고용: 가정 공간의 상황인식 파라미터**



**8. 자율주행 로봇의 홈 환경 상황 인지 시나리오**

인지-행위 프로세스(Perception-and-Action Process)를 목적에 따라 다르게 구성

**8.1. 목표 시나리오**

1. 출퇴근 여부, 어떻게 퇴근해서 왔다

2. 부엌에서 설겆이 중

**8.2. 대체 시나리오**

\* 연기가 자욱한 상황

1. 상황 인지: 온도/유해물질/카메라 → 새로운 상황 발생 → 카메라|화재/연기 탐지

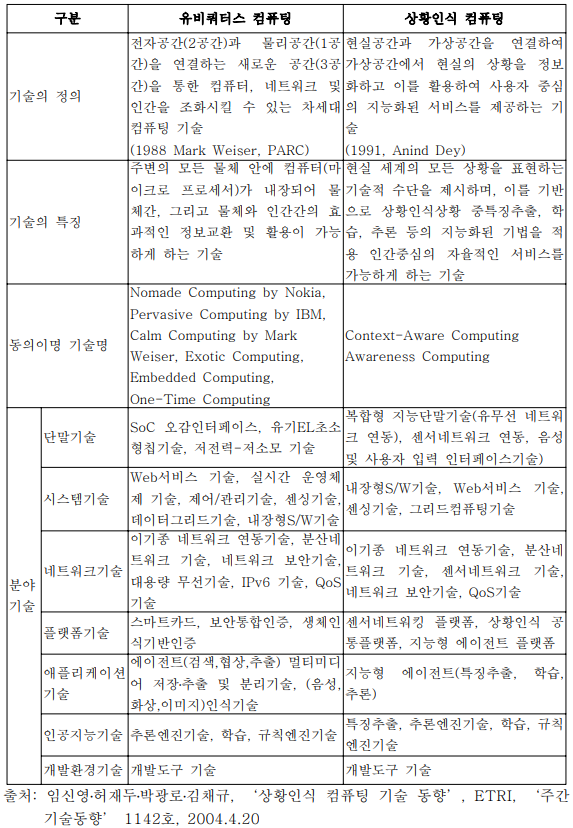
2. 행위 프로세스

- 능동적: [발화] 연기가 자욱하네요. 비상구로 안내합니다. 따라오세요.→[자율 주행]→[방해물 회피]→[도착]

- 수동적: [발화] 연기가 자욱하네요. 비상구로 안내해드릴까요? → [대기]

If yes→[비상등 불][사이렌 소리][발화] 비상구 안내합니다. 저를 따라오세요.→[자율 주행]→[도착]

**부록. 유비쿼터스 컴퓨팅 vs 상황인식 컴퓨팅**



**9. 구현 기술**

\* Image Captioning

\* Human Action Recognition and Prediction

**9.2.1. Image Captioning 관련 Dataset**

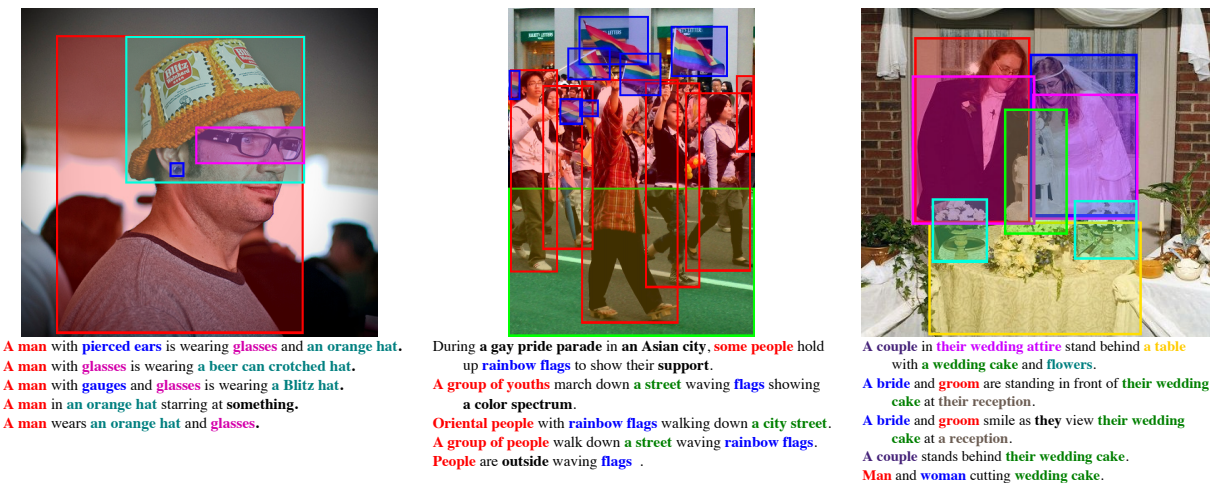
\* COCO Dataset

- [MS COCO Dataset](http://cocodataset.org/)의 [Captions-2015 dataset](http://cocodataset.org/" \l "captions-2015) (Detection | Keypoints | Stuff | Panoptic | Captions)

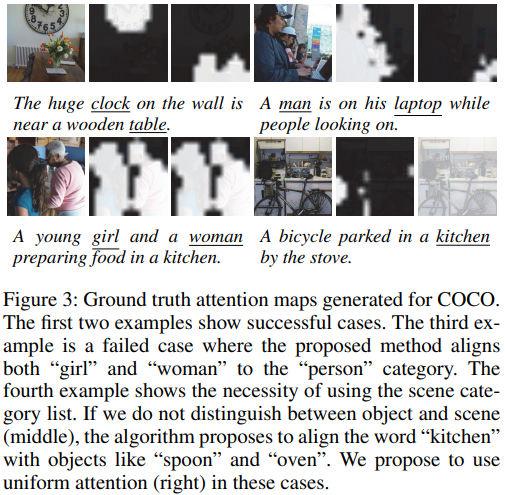


\* Flickr30k

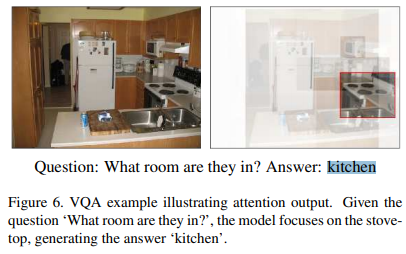
- [Flickr30k Entities: Collecting Region-to-Phrase Correspondences for Richer Image-to-Sentence Models](https://arxiv.org/pdf/1505.04870), Plummer, 2016.

- [Flickr30K Entities Dataset, Github](https://github.com/BryanPlummer/flickr30k_entities)

**9.2.2. 관련 논문 발췌**

[Attention Correctness in Neural Image Captioning](https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14246/14297), Liu et.al., AAAI, 2017

[Reference Based LSTM for Image Captioning](http://openaccess.thecvf.com/content_cvpr_2018/papers/Anderson_Bottom-Up_and_Top-Down_CVPR_2018_paper.pdf), Anderson et.al., CVPR, 2018



[Image Captioning with both Object and Scene Information](http://www.jdl.ac.cn/doc/2011/201711421115678568_p1107-li.pdf), Li et.al., 2016

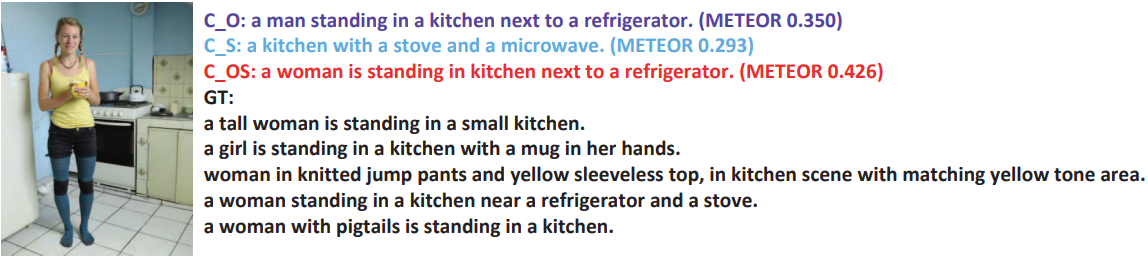
Figure 3: Image descriptions generated with different kinds of information on MS COCO (a) and CPC (b). The METEOR score of each generated sentence is evaluated over the ground truth sentences.



Figure 4: Visualization of generated captions and image attention maps on the COCO dataset. Different colors show a correspondence between attended regions and underlined words. First 2 rows are success cases, last rows are failure examples. Best viewed in color.

**9.3. Human Action Recognition and Prediction**

[미래 로봇형 컴퓨터(FRC) 시장 동향 및 전망](http://www.itfind.or.kr/admin/getFile.htm?identifier=02-001-110628-000003), ITFIND, 한억수 등, 2011.

\* 독일 karlsruhe 공대 ISAS 연구실에서는 가정환경(부엌)에서 사람의 이동 경로 및 냉장고 문을 여는 등의 행동에 따라 로봇이 종합적으로 상황을 판단하고 사람의 의도를 인식하는 확률 모델을 연구하고 있다.

\* 이러한 사용자 행동 패턴에 기반한 사용자 모델 생성을 통해 로봇이 최적화된 서비스 시나리오를 제공하도록 하는 것이 목적이다.

[KIT, ISAS (Intelligent Sensor-Actuator-Systems)](https://isas.iar.kit.edu/)

\* [The KIT Robo-Kitchen Data set for the Evaluation of View-based Activity Recognition Systems](https://isas.iar.kit.edu/pdf/Humanoids11_Rybok.pdf), Rybok et.al., Humanoids 2011, 2011.

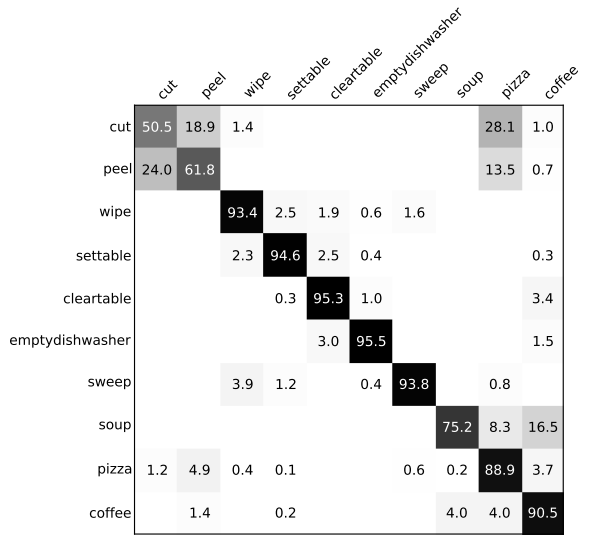
\* A Mobile Service Robot for the Hospital and Home Environment, Schmidt et.al., 1997.

THE KIT ROBO-KITCHEN DATA SET

\* Video data (10 s~4 min)

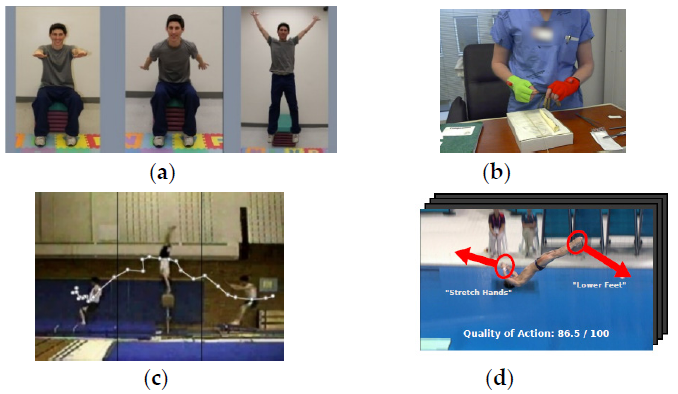
\* 17 different subjects of different age, gender, cultural background, and household skills

\* Activities



[A Survey of Vision-Based Human Action Evaluation Methods](https://arxiv.org/pdf/1806.11230), Lei et.al., 2019.

\* Healthcare and rehabilitation, skill training, sports activity scoring,



[Human Action Recognition and Prediction:A Survey](https://arxiv.org/pdf/1806.11230.pdf), Kong and Fu, 2018.

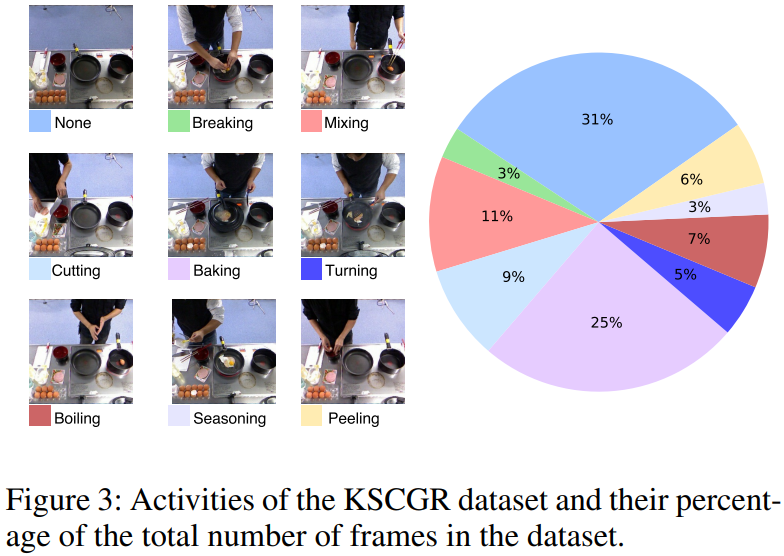


[A Deep Neural Architecture for Kitchen Activity Recognition](https://pdfs.semanticscholar.org/cbdd/4f611f5840b70fcbb5d36e1ea2cc237e59ff.pdf?_ga=2.136625904.1796248851.1575528478-179055882.1570780735), 2017,

\* KSCGR (Kitchen Scene Context based Gesture Recognition) Dataset, Challenge in ICPR 2012

- 5 menus (ham and eggs, omelet, scrambled egg, boiled egg, kinshi-tamago) done by 7 subjects

- 8 cooking gestures: breaking, mixing, baking, turning, cutting, boiling, seasoning, peeling, none,



Human Activity Recognition in Smart Home With Deep Learning Approach, 2019.

\* healthcare and security of smart homes

\* DMLSmartActions Datasets

A Survey of Deep Learning-Based Human Activity Recognition in Radar