Keynote: Affordable and Practical Home Context Recognition with "Image as a Document" Approach

Masahide Nakamura
Graduate School of System Informatics
Kobe University
Kobe, Hyogo Japan
masa-n@cs.kobe-u.ac.jp

ABSTRACT

Technologies for home context recognition have been studied for many years in the field of ubiquitous computing. The traditional ubiquitous computing employs ambient sensors (e.g., temperature, humidity, presence), wearable sensors (e.g., accelerometer, heart rate), and indoor positioning systems. In more recent years, the emerging deep learning allows the system to recognize multimedia data. Since image, voice, and text data usually contain richer information than the conventional sensor data, it is promising to use such multimedia data for recognizing the home contexts.

Unfortunately, however, these existing technologies are yet far from practical use in general households, since they require expensive resources at home. It is difficult for ordinary users to operate and maintain proprietary systems at home on a daily basis. One may try to recognize home contexts via image recognition based on deep learning. However, constructing a custom recognition model dedicated for a single house requires a huge amount of labeled datasets and computing resources. Thus, there is still a big gap between research and real life.

To achieve more affordable and practical home context recognition, we present a novel technique that integrates image-based cognitive API and light-weight machine learning. The cognitive API receives an image from an external application, recognizes specific information within the image, and returns the information as a set of text words called tags. Our key idea is to use these tags as features of the image, and apply light-weight machine learning techniques to infer the target context. Since every image can be considered as a document, the expensive deep learning is no more needed. We call this idea the image-as-a-document approach.

To demonstrate the practical feasibility, we have conducted an experiment that recognizes contexts within our laboratory. The experimental results showed that the overall accuracy of the recognition model was 0.929. We also present further approach to improve the accuracy, by exploiting multiple cognitive APIs.

SPEAKER BIOGRAPHY

Masahide Nakamura received the B.E., M.E., and Ph.D. degrees in Information and Computer Sciences from Osaka University, Japan, in 1994, 1996, 1999, respectively. From 1999 to 2000, he has been a post-doctoral fellow in SITE at University of Ottawa, Canada. He joined Cybermedia Center at Osaka University from 2000 to 2002. From 2002 to 2007, he worked for the Graduate School of Information Science at Nara Institute of Science and Technology, Japan. He is currently an associate professor in the Graduate School of System Informatics at Kobe University. In 2015, he worked for Universite Grenoble Alpes as a visiting professor. In 2018, he starts working for RIKEN, Center for Advanced Intelligence Project (AIP), as a visiting researcher. His research interests include the service computing, cloud computing, gerontechnology, smart home, smart city, and life log. He is a member of the IEEE, ACM, IEICE and IPSJ.