Abdulmotaleb El Saddik
M. Shamim Hossain
Burak Kantarci *Editors*

Connected Health in Smart Cities

Springer

Connected Health in Smart Cities

Abdulmotaleb El Saddik • M. Shamim Hossain •
Burak Kantarci
Editors

Connected Health in Smart Cities

Springer

Editors
Abdulmotaleb El Saddik
School of Electrical Engineering and Computer Science
University of Ottawa
Ottawa, ON
Canada

M. Shamim Hossain
Computer and Information Sciences
King Saud University
Riyadh, Saudi Arabia

Burak Kantarci
School of Electrical Engineering and Computer Science
University of Ottawa
Ottawa, ON
Canada

ISBN 978-3-030-27843-4 ISBN 978-3-030-27844-1 (eBook)
https://doi.org/10.1007/978-3-030-27844-1

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The book *Connected Health in Smart Cities* seeks to provide an opportunity for researchers, academics, and practitioners to explore the relationship between connected health techniques, theoretical foundations, essential services, and recent advances of solutions to problems, which may arise in a variety of problem domains of connected health in a smart city context. This book can serve as a repository of significant reference material.

This book aims to report the theoretical foundations, fundamental applications, and the latest advances in various aspects of connected services in health, more specifically the state-of-the-art approaches, methodologies, and systems in the design, development, deployment, and innovative use of multisensory systems, platforms, tools, and technologies for health management for the success of smart cities ecosystem.

The title of each of the book chapters is self-explanatory and a good hint to what is being covered. The overview of each chapter is as follows:

Chapter "Image Recognition-Based Tool for Food Recording and Analysis: FoodLog"—Maintaining food consumption and habits and analyzing food records is indispensable for the well-being of the citizen in a smart city context. To this end, FoodLog, a smart phone-based image recognition tool, is used for food recording and analysis from digital food image through image recognition or searching. FoodLog's application can be used for the management of food-related data of the athletes or sports activities. This chapter also has better insights related to improved health for healthy diet selection to control various diseases.

Chapter "A Gesture Based Interface for Remote Surgery"—At present, specially equipped vehicles or air-lifting to nearest hospitals/clinics is not affordable for the citizens in emergency cases or inadequate for areas with a large population that is remote from emergency surgical services. These vehicles can only serve a few patients or citizens every day. In this situation, there is a need for remote surgical services by skilled surgeons. Considering the above facts, this chapter discusses the application of gesture-based interactive user interfaces in performing remote endovascular surgery. The conducted experiments in the chapter demonstrate the feasibility of the approach and also the accuracy of the robotic controller at the base of the catheter, before it enters an artery.

Chapter "Deep Learning in Smart Health: Methodologies, Applications, Challenges"—Today, deep learning is one of the emerging theoretical foundations of connected health that can support healthcare professionals to find out the hidden opportunities in healthcare data and its pattern to assist doctors in order to have better analysis for improved health care for the citizens of smart cities. Keeping the above benefits in mind, this chapter presents very good insights of how deep learning techniques can be used for smart health data analysis, processing, and prediction. It also discusses about the emerging applications of deep learning techniques in smart health from cancer diagnosis to health status predictions.

Chapter "Emotional States Detection Approaches Based on Physiological Signals for Healthcare Applications: A Review"—Emotional health is one important consideration for improving citizens' quality of life and well-being in the smart cities. With these issues in mind, this chapter discusses existing emotional state approaches using machine and/or deep learning techniques, the most commonly used physiological signals in these approaches, and existing physiological databases for emotion recognition and highlights the challenges and future research directions in this field. It also discusses about how to incorporate accurate emotional state detection wearable applications (e.g., patient monitoring, stress detection, fitness monitoring, wellness monitoring, and assisted living for elderly people) within the smart cities so that it can aid to alleviate mental disorders, stress problems, or mental health.

Chapter "Toward Uniform Smart Healthcare Ecosystems: A Survey on Prospects, Security, and Privacy Considerations"—Security and privacy consideration is of paramount importance in the connected healthcare applications for the citizens' safety and well-being in smart cities. To this end, this chapter explores the latest trends in connected healthcare applications along with enabling technologies (e.g., sensing, communication, and data processing) and solutions (e.g., low-power short-range communication, machine learning, and deep learning) that might be driving forces in future smart health care. It reports the latest cyber-attacks and threats, which could be major vulnerabilities and weaknesses of the future smart healthcare ecosystem. It concludes with the proposed solutions and their associated advantages and disadvantages of each solution and analyzes their contribution to the overall security as an integral part of the connected healthcare system.

Chapter "Biofeedback in Healthcare: State of the Art and Meta Review"—This chapter begins by discussing the scope of utilizing biofeedback technology in smart healthcare systems. It presents a brief history of biofeedback technology and highlights the sensory technology in biofeedback systems by presenting the different types of sensors and their features. Recent research of biofeedback-based healthcare systems will be explored by presenting a range of applications in different fields. A set of challenges/issues that affect the deployment of biofeedback in healthcare systems will be discussed.

Chapter "Health 4.0: Digital Twin for Health and Well-Being"—With the advances in wearable computing, smart living, and communication technologies, personalized healthcare technology has entered a new era of healthcare industry to provide personalized proactive and preventive care in real time without being in close proximity. Digital Twins is an emerging technology to revolutionize healthcare and clinical processes. A digital twin virtualizes a hospital to have more personalized care. This chapter gives an overview of the existing literature and aims to provide an overview of existing literature on digital twins for personal health and well-being—key terminologies, key technologies, key applications, and the key gaps.

Chapter "Incorporating Artificial Intelligence into Medical Cyber Physical Systems: A Survey"—The emerging Medical Cyber-Physical Systems (MCPS) can revolutionize our connected healthcare system with high-quality, efficient, and continuous medical care for citizens of smart cities by providing remote patient healthcare monitoring, accelerate the development of new drugs or treatments, and improve the quality of life for patients who are suffering from different medical conditions, among other various applications. This chapter starts with the general description of the MCPS components and then discusses (1) how multisensory sensor devices and body sensor networks can assist in healthcare data acquisition, aggregation, and preprocessing and (2) how machine intelligence algorithms process the medical data from the previous steps to facilitate monitoring through connected healthcare systems and make self-directed decisions without much involvement of healthcare staff in a secure way to preserve the privacy of the citizens of smart cities.

Chapter "Health Promotion Technology and the Aging Population"—One of the important aspects for the success of connected health is the use of emerging healthcare technologies, which are of paramount importance in connected health services to the aging population in cities to improve the quality of care. To this end, this chapter provides an overview of assisted technologies and a survey of how the technology can be used to affect the elderly population to integrate healthier habits into their lives. The variety of accessible technologies allows individuals to use them in conjunction for their desired outcomes.

Chapter "Technologies for Motion Measurements in Connected Health Scenario"—The proactive and efficient care is one of the utmost requirements for connected health or technology-enabled care (TEC) in smart cities. For such care, smart sensing technology-based wearable solutions are essential for human motion tracking, rehabilitation, and remote healthcare monitoring. In such a context, this chapter presents an unobtrusive sensing solution (e.g., the Internet of Things (IoT)-enabled sensing) based on key enabling technologies with the aim of providing human motion measurement accompanied by motion measurement-related research and open issues. Finally, it demonstrates how the human motion measurements in motion tracking can contribute to the remote health monitoring system based on IoT and publish/subscribe communication paradigm.

Chapter "Healthcare Systems: An Overview of the Most Important Aspects of Current and Future m-Health Applications"—With the increasing number of aging population and the widespread use of mobile devices and communication technologies, citizens in smart cities would like to access the connected health service from anywhere at any time. In this respect, the mobile health care (m-healthcare) can provide affordable care for people in a convenient, accessible, and cost-effective manner. This chapter reports an overview of a generic m-Health application along with its main functionalities and components. The use of a standardized method for the treatment of a massive amount of patient data is necessary to integrate all the collected information resulting from the development of m-Health devices, services, and applications. To this end, this chapter discusses about the requirements of a standardization in healthcare, which is supported by related international and European healthcare projects.

Chapter "Deep Learning for EEG Motor Imagery-Based Cognitive Healthcare"—Owing to the massive amounts of complex healthcare data being produced in environments, such as smart cities, deep learning and cognitive capability are necessary to the idea of connected health. Deep learning-based cognitive systems can help various stakeholders, such as medical experts, healthcare professionals, and patients to develop insights into medical data that can help improve health care and provide a better quality of life to smart city residents. Hence, this chapter leverages deep learning techniques for understanding MI EEG data. The improvement in classification accuracy for motor imagery can help impart cognitive intelligence to machines and enable smart city residents to control the environment through sensors attached to their heads. This chapter proposes novel techniques for cross-subject accuracy and achieves outstanding improvement that can usher in new concepts about these complex brain signals.

The target audience of this book includes researchers, research students, and health practitioners in digital health. The book is also of interest to researchers and industrial practitioners in healthcare industry and smart city. We would like to express our great appreciation to all the contributors, including the authors, reviewers, and Springer staff, for their kind support and considerable efforts in bringing this book to reality.

We hope that the chapters from this book will serve as a repository of significant reference material and contribute to the roadmap of emerging use of services, techniques, and technologies for connected healthcare in smart cities.

Ottawa, ON, Canada — Abdulmotaleb El Saddik
Riyadh, Saudi Arabia — M. Shamim Hossain
Ottawa, ON, Canada — Burak Kantarci

Contents

Image Recognition-Based Tool for Food Recording and Analysis: FoodLog

Kiyoharu Aizawa

Abstract While maintaining a food record is an essential means of health management, there has long been a reliance on conventional methods, such as entering text into record sheets, in the health medicine field. Food recording is a time-consuming activity; hence, there is a need for innovation using information technology. We have developed the smartphone application "FoodLog," as a new framework for food recording. This application uses digital pictures and is supported by image recognition and searches. It is available for general release. In this paper, we present an overview of this framework, the data statistics obtained using FoodLog, and the future prospects of this application.

Keywords Food recognition · Image processing · Text search · Visual search

1 Food Recording Tool Using Analysis: FoodLog

We have developed and constructed a system, known as FoodLog, as a technical platform to record and utilize daily consumption data using multimedia information [1]. Although there are many tools to record details of the food consumed daily, the vast majority of these involve the input and output of text. Although there is much effort required for input, the records generated cannot be intuitively understood with just a single glance. The greatest distinguishing characteristic of FoodLog from existing tools is that it supports image recognition and searches through image-based recordings. Additionally, an important advantage of images is that the records can be perused and grasped with just a single glance. Initially, this was developed as a web-based system; now, this system has also been developed as a smartphone application [2, 3]. The functionality has gradually been expanded. The smartphone version released in 2013 supported image searches; since June 2016, an

K. Aizawa (✉)
Department of Information and Communication Engineering, The Faculty of Engineering, University of Tokyo, Tokyo, Japan
e-mail: aizawa@hal.t.u-tokyo.ac.jp

A. El Saddik et al. (eds.), *Connected Health in Smart Cities*,
https://doi.org/10.1007/978-3-030-27844-1_1

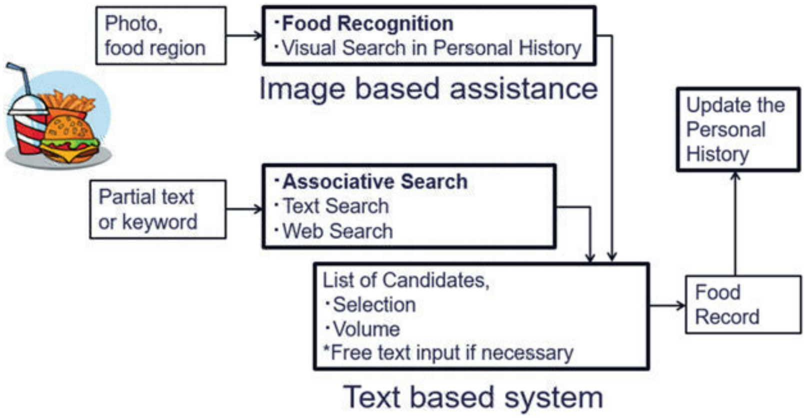

Fig. 1 Outline of FoodLog supported by food recognition and search

(a) (b) (c)

Fig. 2 Screenshots of FoodLog app. (**a**) Diary view, (**b**) food records, and (**c**) energy view

update supports image recognition as a method of record entry (image recognition is currently limited to the iPhone version).

An overview of the current functionality is shown in Fig. 1. A representative screenshot of the application's functionality is shown in Fig. 2. This framework supports record input based on image recognition and searches; notably, keywords are input with text only when image support is insufficient, significantly reducing the effort required to input data. The typical input procedures performed are described as follows:

1. The user photographs the food.

2. On inputting the record, the user starts up the FoodLog app. The FoodLog application automatically distinguishes between various images and presents the food photos from the album. Indeed, it is also possible to manually select the images from the album.
3. The users specify the food area in which they would like to record the presented food photo.
4. Automatic recognition of the food occurs, and the top 20 food items are presented in terms of probability. Additionally, an image search is conducted simultaneously on the individual history, and the top 20 consumed food items are presented in descending order of similarity.
5. If there is a desired item in the presented list, it is selected and the quantity is specified. This step completes the recording.
6. If a food item close to the desired item is included in the presented list (for example, a hamburger instead of a cheeseburger), an associative search is performed; then, the presented list is updated, and if the desired item appears, that item is selected and the quantity is specified. These steps complete the recording.
7. In case the desired item has still not appeared in the list, you can enter a keyword and update the candidate list; also, at the location of the desired item, this entered information is selected and the quantity specified. These steps complete the recording.
8. In case the target food item cannot be found with image or text search, you can describe the food name in free text as a newly appearing item; once this information is entered, it will be available for future searches.

It is not necessary to perform steps 1 and 2 at the same time. It is my personal practice to first take the picture and enter the record later. The photo is an essential mnemonic while inputting the record at a later time.

In case of the app screenshot shown in Fig. 2, (a), (b), and (c) show a list of the food records in calendar format, the record content of the individual food pictures, and a format in which just the calories are superimposed on the images. Furthermore, Fig. 3(a) and (b) shows the food area specified on the screen when entering the food records and a display of the recognition results, respectively. In the published version, this includes the detailed nutritional values of approximately 2000 foods that are typically found in Japan; of these details, only the calories are displayed. Datasets, with detailed nutritional values can, where necessary, be switched to those with high variation. At present, there are approximately 400 food items that have a high number of record registrations that can be detected. The number of food items detected can be easily increased through data quantity, without affecting the performance of the recognizer device. Using associative searches, this app searches a total of about 2000 typical foods based on similarity of food name, nutritional value, or recipe [4].

FoodLog is used by Diabetics [5]; notably, this app is under investigation at Tokyo University as a self-diagnostics tool for diabetes. It is also being used in the application Gluco Note, which Tokyo University is also evaluating [6].

(a) (b)

Fig. 3 Screenshots of FoodLog app. (**a**) Food area specified on the screen and (**b**) the list of candidates by image recognition

2 Trends Visible from FoodLog Data: (1) Frequent Foods

Since the implementation of the FoodLog application in 2013, more than 200,000 users have uploaded at least one photo, and 6 million food items have been recorded so far. The variation in food names registered by the users themselves has greatly exceeded expectations, evidenced by the fact the number of unique food names has exceeded 250,000 items.

Owing to the fact that there is a wide variety of food names registered by all users, it is necessary to summarize these appropriately to produce meaningful statistics. For example, we want to consider "yogurt" and "plain yogurt" as the same food item. To achieve this, we had to perform normalization processing on the recorded names using compression representation [7]. Specifically, (1) the respective foods are broken down into vocabulary words, (2) food names are selected from similar vocabulary words, a vocabulary graph is generated, and (3) representative food names are set based on the collections of shortest path vocabulary words.

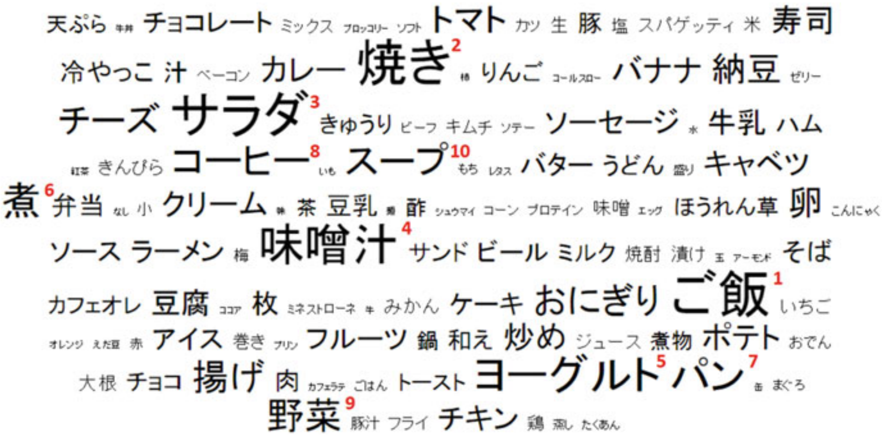

Fig. 4 Top 100 frequent food records of all users of FoodLog

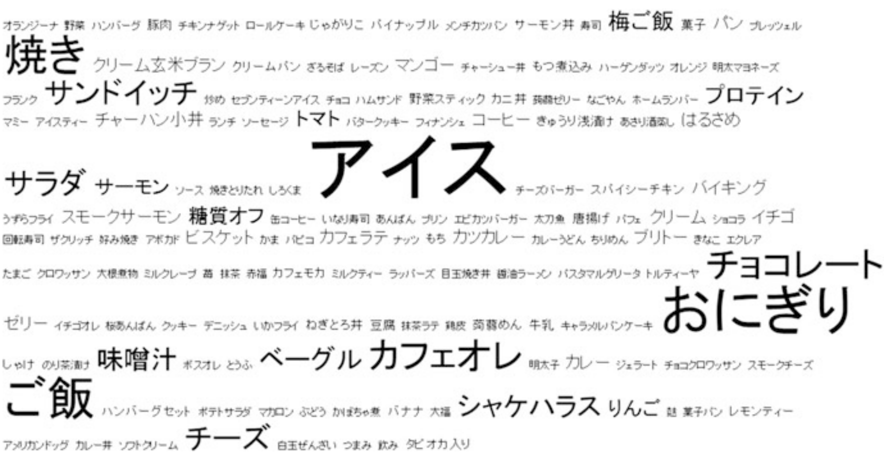

Fig. 5 Top 100 frequent food records of a particular user for its 6 weeks

We analyzed data trends using the first year of data (∼1 million items) generated by the FoodLog application. The representative food names that appear frequently among the users are shown in Fig. 4. The frequently occurring items for all users are food names that are very familiar to all users. At the top of the list, names such as “baked” and “boiled” appear; however, because baked items and boiled items are overcompressed in the automatic processing for the normalization process, names of the cooking methods also appear.

In contrast to this, the same processing performed in relation to 6 weeks of data for a particular user is shown in Fig. 5, and this displays the order of frequency of records for just a single user. It is highly evident that this is very different to the trend for all users, which is very interesting. These food trends for individual users can be intuitively grasped.

3 Trends Visible from FoodLog Data: (2) Food Frequency Temporal Changes

The amount of food records greatly depends on the degree to which the user likes food. Food itself is also greatly influenced by seasonal factors and specific events. Additionally, consumption of newly emerging food items may increase as a result of mass media advertising.

We attempted to investigate changes in food frequency over time using the records recorded by FoodLog. Using the record frequency of food in FoodLog as an indicator, we used the ratio of the number of people that recorded a specific food item among all people who recorded consumption for that day. For reference purposes, we also investigated changes over time using search words on Google (Google Trend). Although Google Trend can intuitively grasp social concerns, the granularity of food is coarse; hence, while looking at the top 500 items in terms of frequency in FoodLog, only $\sim$300 of these could be surveyed. Among these 300, there will always be words included that are not necessarily limited to food; thus, there is no certainty that they are expressing a particular interest in food. However, from $\sim$300 items, frequency changes in 42 of the food names demonstrated a strong correlation ($\geq$0.7) between FoodLog and Google Trend. From these, we selected items of particular interest and demonstrated fluctuations over 3 years, as shown in Fig. 6. Each of these was normalized to a maximum value of 1 within the period.

For all four examples in Fig. 6, trends of interest in terms of frequency trends in FoodLog records or Google search terms are extremely similar. Figure 6(a) and (b) are examples in which the frequency trends change considerably depending on the season or events. (a) Yudofu (boiled bean curd) appears frequently during winter, while (b) fried chicken peaks appear only during Christmas. In contrast to this, natto (fermented bean curd) in Fig. 6(c) is a food eaten daily, and its consumption, for both indices, gently increases over the period of approximately 3 years. (d) Chicken salad only became well known at the end of 2013, and from that point on exhibited a rapidly increasing trend for both indices. The fact that is interesting about (c) and (d) is that they are both health-oriented food items. In this manner, compared to Google, FoodLog, despite having a significantly small user base, is a tool that can be used

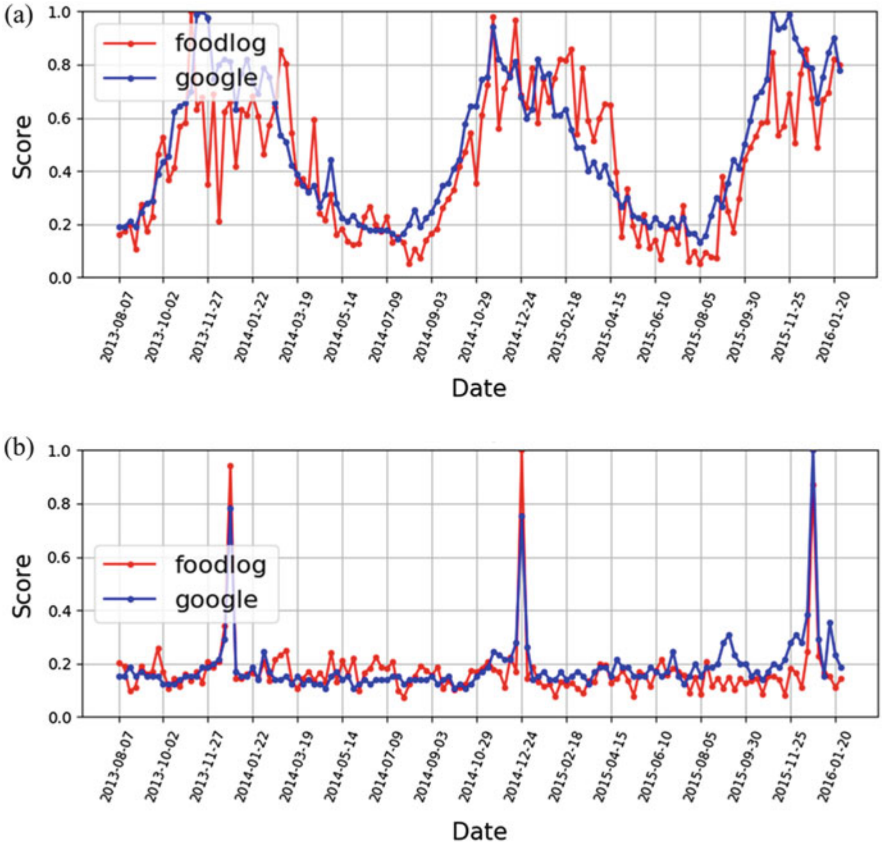

Fig. 6 Temporal changes of frequency of food records. (**a**) Yudofu (湯豆腐): Highly seasonal changes. (**b**) Fried chicken: Highly event-dependent changes. (**c**) Natto (納豆): Growing trend of a dairy food. (**d**) Salad Chicken: Growing trend of newly appearing food. It started widely selling in November 2013 at Seven Eleven Stores

to look at food-related fluctuations in detailed categories. Among these, items that relate to the interests of all users match with similar information from Google to a significantly higher degree.

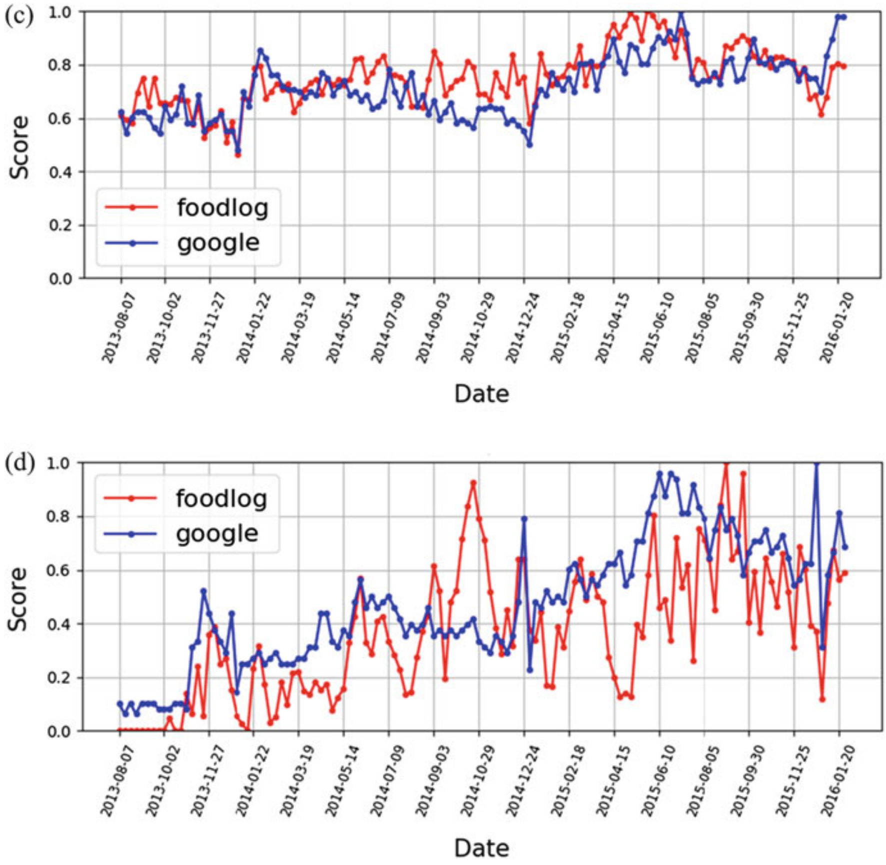

Fig. 6 (continued)

4 Conclusion

In this chapter, we presented an overview of our research and development of FoodLog as a tool to record information concerning food consumption in which the use of images is maximized. Additionally, from an analysis of the recorded data, we introduced examples of visualized food trends, both overall and individual, and changes over time for the frequency of food consumption over long periods.

As a tool, FoodLog is still in the developmental stage and we would like it to be increasingly convenient. For example, it would be useful if it could create a record by just inserting the picture, without the need to specify the food area. In actual fact, great efforts are being made in research and development for this purpose, and many innovations are planned for the interface in the next version of FoodLog.

The trends seen from the data in this paper, as demonstrated in the two examples, show its importance not only as a tool for recording but also as a platform for analyzing food records. Food surveys, which traditionally have taken months from data collection to publication, can be produced in real time through the use of FoodLog. Additionally, based on data from a large number of users, we are addressing the task of estimating records over a long period based on only few days' data for a particular user [8]. In addition, for FoodLog's application in the fields of self-management for health and dietary purposes, its functionality is also being expanded for the management of food-related data and information from other activities of sports athletes.

References

1. K. Aizawa, M. Ogawa, FoodLog: Multimedia tool for healthcare applications. IEEE Multi Media **22**(2), 4–9 (2015)
2. FoodLog. http://app.foodlog.jp/
3. K. Aizawa, K. Maeda, M. Ogawa, Y. Sato, M. Kasamatsu, K. Waki, H. Takimoto, Comparative study of the routine daily usability of FoodLog: A smartphone-based food recording tool assisted by image retrieval. J. Diabetes Sci. Technol. **8**, 203–208 (2014)
4. S. Amano, S. Horiguchi, K. Aizawa, K. Maeda, M. Kubota, M. Ogawa, Food search based on user feedback to assist image-based food recording systems, in *ACM Multimedia Workshop on Multimedia Assisted Dietary Management (MADiMa)*, (2016), pp. 71–75
5. K. Waki, K. Aizawa, S. Kato, H. Fujita, H. Lee, H. Kobayashi, M. Ogawa, K. Mouri, T. Kadowaki, K. Ohe, DialBetics with a multimedia food recording tool, FoodLog: Smartphone-based self-management for type 2 diabetes. J. Diabetes Sci. Technol. **9**(3), 534–540 (2015)
6. GlucoNote. http://uhi.umin.jp/gluconote/ (2016)
7. S. Amano, K. Aizawa, M. Ogawa, Food category representatives: Extracting categories from meal names in food recordings and recipe data, in *IEEE International Conference on Multimedia Big Data, At Beijing, China*, (2015), pp. 48–55
8. A. Tsubakida, S. Amano, K. Aizawa, M. Ogawa, Prediction of individual eating habits using short-term food recording, in *International Joint Conference on Artificial Intelligence (IJCAI) Workshop CEA* (2017), pp. 45–48

A Gesture-Based Interface for Remote Surgery

Irene Cheng, Richard Moreau, Nathaniel Rossol, Arnaud Leleve, Patrick Lermusiux, Antoine Millon, and Anup Basu

Abstract There has been a great deal of research activity in computer- and robot-assisted surgeries in recent years. Some of the advances have included robotic hip surgery, image-guided endoscopic surgery, and the use of intra-operative MRI to assist in neurosurgery. However, most of the work in the literature assumes that all of the expert surgeons are physically present close to the location of a surgery. A new direction that is now worth investigating is assisting in performing surgeries remotely. As a first step in this direction, this chapter presents a system that can detect movement of hands and fingers, and thereby detect gestures, which can be used to control a catheter remotely. Our development is aimed at performing remote endovascular surgery by controlling the movement of a catheter through blood vessels. Our hand movement detection is facilitated by sensors, like LEAP, which can track the position of fingertips and the palm. In order to make the system robust to occlusions, we have improved the implementation by optimally integrating the input from two different sensors. Following this step, we identify high-level gestures, like push and turn, to enable remote catheter movements. To simulate a realistic environment we have fabricated a flexible endovascular mold, and also a phantom of the abdominal region with the endovascular mold integrated inside. A mechanical device that can remotely control a catheter based on movement primitives extracted from gestures has been built. Experimental results are shown demonstrating the accuracy of the system.

I. Cheng · N. Rossol · A. Basu (✉)
Department of Computing Science, University of Alberta, Edmonton, AB, Canada
e-mail: locheng@ualberta.ca; nrossol@ualberta.ca; basu@ualberta.ca

R. Moreau · A. Leleve
Ampere Lab, INSA, Lyon, France
e-mail: rmoreau@insa-lyon.fr; leleve@insa-lyon.fr

P. Lermusiux · A. Millon
CHU Hospital, Lyon, France
e-mail: patrick.lermusiaux@chu-lyon.fr

A. El Saddik et al. (eds.), *Connected Health in Smart Cities*,
https://doi.org/10.1007/978-3-030-27844-1_2

Keywords Gesture recognition · Remote control of catheter · Endovascular surgery

1 Introduction

Despite the increasing popularity of computer- and robot-assisted surgeries, there is still a need for remotely performing surgical tasks. At present, it is difficult to perform surgeries in remote locations. Very often patients in emergency situations need to be driven long distances in specially equipped vehicles or airlifted to nearest hospitals. These emergency transportations are not only very expensive but are also very demanding on fragile patients resulting in a negative effect on their chance of survival. Also, these special transportations are far from guaranteed. For example, the STARS air ambulance in Alberta, Canada, needs to hold several fundraising lotteries every year to support their operating costs. Furthermore, the air ambulance can only serve a few clients every day. While an ad hoc solution like this is somewhat functional for a sparsely populated region with low demand, it is completely inadequate for areas with large population that are remote from emergency surgical services. According to a report from the Fraser Institute entitled "The Effect of Wait Times on Mortality in Canada," over a 16 -year period surgical wait times have led to over 44,000 deaths in an advanced country like Canada with a relatively small population. This fact can give us some insight into the very positive effect remote surgeries can have at a global level, if skilled surgeons could collaborate from around the world.

With the above discussion in mind, in this chapter we discuss the application of gesture-based interactive user interfaces in surgery. This new technology, which will provide touch-free environments for interaction with 3D medical data, will improve both the ease of use and efficiency of a number of medical devices and procedures. Despite the large increase in popularity of touch-based devices, such as the iPad and other tablet computers, in recent years, interaction with these devices remains limited to physically touching the screen. There are many scenarios, particularly in healthcare, where the ability to control a device through touch-free mechanisms offers a significant advantage. During medical scans and procedures, for example, a touch-free user interface will help (1) decrease the risk of spreading infection by reducing the need for the device operator to touch potentially contaminated surfaces; (2) eliminate the need for an extra assistant to help view and interact with image data; (3) improve ergonomic conditions for a medical technician by removing the need to concurrently use one hand for a medical probe and the other to operate a computer keyboard; and (4) give surgeons the ability to leave a surgical room and still be able to interact with a surgical procedure without touching any contaminated surface.

A specific application area is Endovascular surgery, which is a type of Minimally Invasive Surgery (MIS), designed to access target regions of the human body through blood vessels [1, 2]. While endovascular surgery is being increasingly deployed to replace classical surgery for improving recovery time and patient safety, there are still challenges to sustain a safe environment for patients as well as doctors. Ongoing research aims at reducing surgery time, and minimizing exposure to radiation. Currently, medical personnel require heavy cumbersome protective clothing which, nevertheless, can cover only part of the body. The ability to perform a surgery while keeping some distance from sources of radiation will improve the safety of surgeons and medical staff. In addition, the ability to interact with medical devices through gestures will increase patient safety by reducing the potential for contamination.

Several studies have examined active endoscopes and the effective control of catheters [3–5]; while the impact of stiffness of surgical manipulators was discussed in [6]. However, how to control a catheter using only gestures has not been addressed before.

Endovascular surgery has become a very important part of the therapeutic arsenal for the treatment of vascular diseases, such as abdominal aortic aneurysms [7]. These techniques are part of mini-invasive surgery and related to a decreased immediate postoperative morbidity and mortality. Yet, all these benefits are balanced by an additional cost and several challenges still remain.

Endovascular surgery exposes the patient and medical staff to a significant amount of radiation during the procedure [8]. Repetitive exposures to radiation increase the risk of cancer [9] and other diseases, such as cataract and skin injuries [10]. Thus, the ability to keep the surgeon and the staff distant from radiations without any loss of efficiency during the procedure is highly desirable in future endovascular surgeries.

Avoiding contact between the surgeon and the patient also reduces the risk of per-operative contamination. Stent-graft [11] infections are rare but associated with high morbidity and mortality rates [12]. Aortic stent-graft infections require, in most of the cases, the surgical explantation of the stent-graft [13].

From another perspective, gesture-based interactive user interfaces for surgery could be used for training of junior surgeons. The role of simulators in medical training is becoming more and more important [14–16]. Using such simulators, junior surgeons would be able to reproduce preoperative gestures during a surgical simulation, improve their skills before performing surgery on real patients, and have their learning evaluated through objective metrics.

2 Materials and Methods Used in Our System

There are various steps involved in our overall system to access the feasibility of using gestures to conduct or assist remotely in a surgery. These include:

- Extracting blood vessels from CT and identifying arteries
- Planning a path to follow during endovascular surgery
- Detecting hand movements and interpreting gestures; and
- Robotically controlling a catheter based on gestures performed remotely.

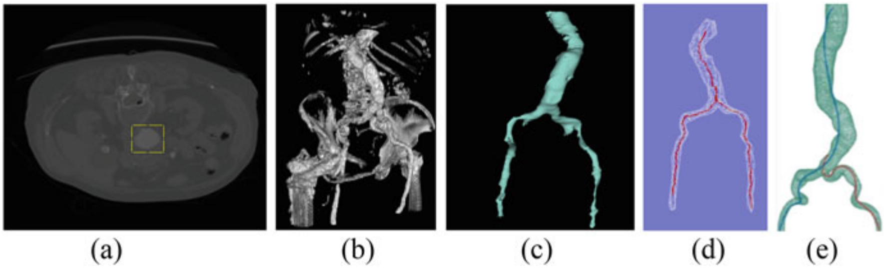

Fig. 1 Various steps in endovascular surgical planning: (**a**) A low contrast CT slice in axial view with the aorta inside the rectangle; (**b**) artery neighborhood detected using the OSIRIX software; (**c**) 3D segmentation using initialization of livewires followed by using the Turtleseg software; (**d**) medial axis generated inside the artery 3D model; and (**e**) artery phantom: with optimal path chosen by a surgeon for catheter navigation

The novelty in this work is in the last two areas mentioned above. However, we will briefly outline some of our work in the first area as well. We extracted the blood vessel and the medial axis (the curve in the middle that is equidistant from all sides) using several segmentation techniques. We used a new scale-space skeletonization algorithm for robust 3D medial axis extraction. Our algorithm is adaptive and can adjust the scale so that both narrow and wide regions of a blood vessel can be processed accurately. Figure 1 describes some of the methods related to preprocessing the medical image data that have been developed by us over the past few years. These include enhancing segmentation algorithms to extract blood vessels and arteries; skeletonization to find the path close to the center of the blood vessels; and designing and fabricating a phantom representing the actual arteries in a CT scan to allow surgical training. Details on some of these procedures are available in our earlier publication in [17, 18].

Endovascular surgery has become a very important part of the therapeutic arsenal for the treatment of vascular diseases, such as abdominal aortic aneurysms [7]. These techniques are part of mini-invasive surgery and related to a decreased immediate postoperative morbidity and mortality. Yet, all these benefits are balanced by an additional cost and several challenges still remain.

Endovascular surgery exposes the patient and medical staff to a significant amount of radiation during the procedure [8]. Repetitive exposures to radiation increase the risk of cancer [9] and other diseases, such as cataract and skin injuries [10]. Thus, the ability to keep the surgeon and the staff distant from radiations without any loss of efficiency during the procedure is highly desirable in future endovascular surgeries.

Avoiding contact between the surgeon and the patient also reduces the risk of per-operative contamination. Stent-graft [11] infections are rare but associated with high morbidity and mortality rates [12]. Aortic stent-graft infections require, in most of the cases, the surgical explantation of the stent-graft [13].

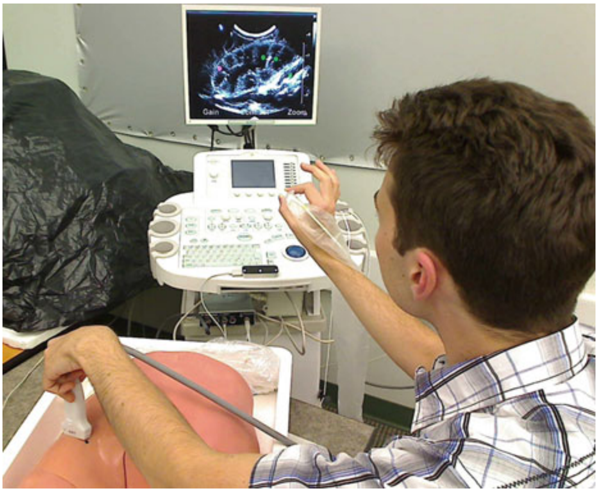

Fig. 2 A hand gesture tracked interface for controlling an ultrasound workstation

From another perspective, gesture-based interactive user interfaces for surgery could be used for training of junior surgeons. The role of simulators in medical training is becoming more and more important [14–16]. Using such simulators, junior surgeons would be able to reproduce preoperative gestures during a surgical simulation, improve their skills before performing surgery on real patients, and have their learning evaluated through objective metrics.

Figure 2 shows our work on detecting gestures being used for controlling the user interface of an ultrasound workstation. In the system we developed, the hand gestures that can be used to adjust the gain, brightness, zoom, and other parameters on the display shown in the figure without the fingers touching the screen.

Our hand gesture detection system uses the LEAP motion sensor. However, there are several challenges in situations with high occlusion. Many skeletal hand pose estimation techniques follow particle-filter approaches, which can get trapped in a local minimum due to inadequate visible data to resolve ambiguity. Incorrect pose may be identified as long as the hand pose remains static. Traditional approaches like the Kalman Filter [5] may not be adequate because the distribution of noise is similar for correct vs. incorrect detection scenarios [20]. Also, for fast movements of the fingers, tracking angular velocities of the fingers to predict future positions may not work.

We address the issues raised by occlusion using multiple sensors that are strategically placed with different viewing angles. We process the skeleton information provided by the Leap motion sensor instead of depth maps. The skeleton information consists of a collection of points with lines, representing a simplified version of the human skeleton of the hand. This type of data is easier to process than 3D depth information that may be available from other sensors, because of the limited amount of data that needs to be considered. Alternative strategies were considered to

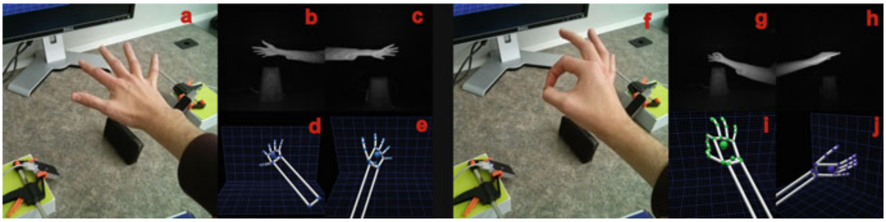

Fig. 3 Summary of our recent work in [19]. Using multiple sensors, we can reduce the effect of occlusion. For example, for the pair of images on the left, both of the sensors can detect the open hand pose. However, for the pair of images on the right, one sensor detects the pinch pose while the other incorrectly detects an open hand. Intelligently combining the results from the two sensors, it is possible to correctly detect the pinch gesture on the right

Fig. 4 Various hand poses used to compare performance against ground truth

determine the best way of combining (or fusing) data from multiple sensors. Figure 3 shows some results of our approach implemented in real time.

To determine the accuracy of our approach, we used a flexible hand phantom as shown in Fig. 4. Various alternative approaches were compared, including (1) Single Sensor, (2) Averaging of Multiple Sensors, (3) Sensor Confidence, (4) Weighted Fusion, and (5) Our Intelligent Fusion Approach. These results are shown in Fig. 5. It can be seen that our results are closest to the best possible if a person already knows which sensor to use.

Following the recording of gestures, we need to control the movement of a catheter potentially at a location that is distant from where the gestures were captured and analyzed. There are at least two scenarios where this approach is

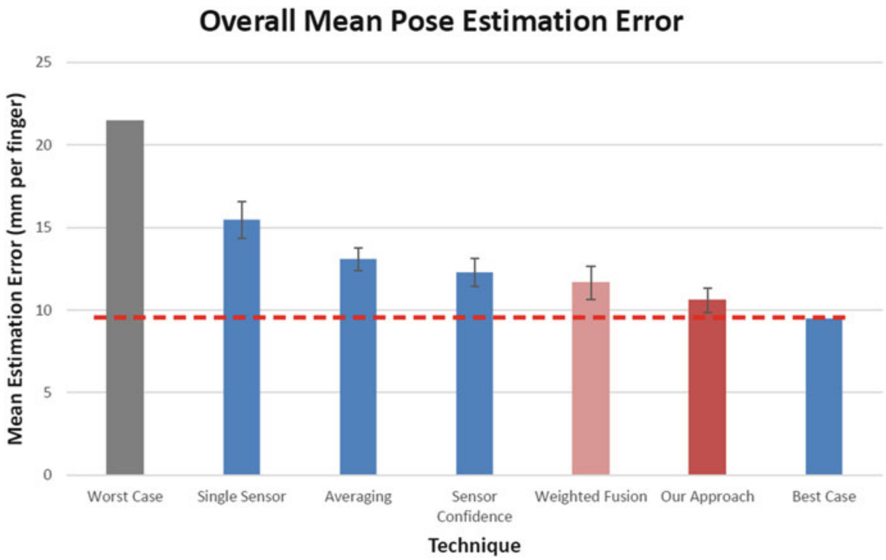

Fig. 5 Comparison of various approaches for hand pose estimation

very valuable. First, creating some distance between a surgeon and a patient in an operating room would reduce, if not eliminate, surgeons being exposed to radiation from intra-operative X-rays during surgery. At present, surgeons need to wear protective clothing with lead barriers to reduce radiation exposure. This type of clothing is heavy and uncomfortable, and still does not eliminate exposure to the face, arms, and legs. The second utility of our approach lies in surgeons being available to perform an emergency surgery even from a remote location. This will support greater access to patients who are unable to travel to a limited number of specialized urban centers that have surgical facility available.

We designed and built an electromechanical system for precise computerized control of a catheter. Our device consists of a motorized XZ Table that can move in two directions, coupled with two conveyor belts. The velocities of the two conveyors can be controlled through a computer or if necessary by a remote joystick. The guidewire is placed in between the two conveyor belts. The XZ Table is connected to one of the conveyors, while the other conveyor is fixed. By moving the two conveyors, we can move the guidewire backward or forward, while rotations of the guidewire can be realized through radial movements of the XZ Table. Figure 6 shows the composition of the electromechanical system built by us for controlling a catheter.

In order to determine the precision of our system we built an artery phantom. The phantom was designed by segmenting the arteries of an actual patient from the CT scans captured before a surgery. Figure 7 demonstrates our system. On the left, the mechanical system and the joystick are shown. On the right, the artery phantom is shown connected to the mechanical system. Inside the artery phantom, the catheter being controlled electronically can be seen.

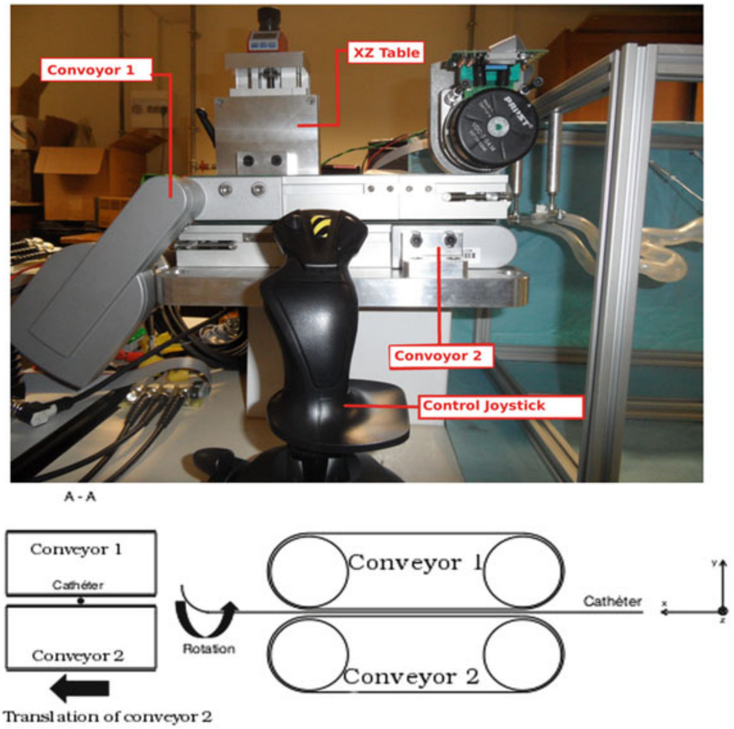

Fig. 6 Our system for controlling catheter movements

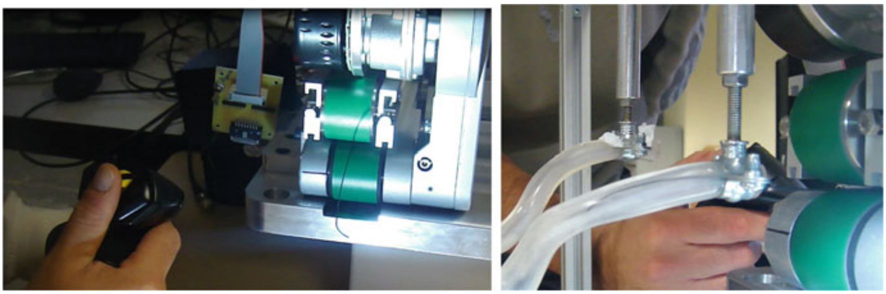

Fig. 7 Moving a catheter electronically into an artery phantom

3 Results in a Simulated Environment

We have already conducted experiments to verify our proposed methodology. Hand gestures were recorded and processed at the University of Alberta, Canada, using the Leap motion sensor. From these recordings, the primitives of the hand gestures were automatically extracted. Finally, these primitives were transmitted to Lyon,