2017 IEEE Symposium on Service-Oriented System Engineering

IoT Metadata Creation System for Mobile Images and its Applications

Hsin-Yi Tseng, Cheng-Ting Lee,
Cho-An Lin
Dept. of Computer Science
National Tsing Hua University
Hsinchu City, Taiwan
Email: oopsmebo@gmail.com, sunkist0704@hotmail.com, anna81209@gmail.com

Pai H. Chou
Dept. of EECS
Univ. of California, Irvine
Irvine, CA USA
Email: phchou@uci.edu

***Abstract*—This paper proposes the capturing of IoT data and their embedding as metadata in digital snapshots, including photos, audios, and videos. Specifically, we focus on IoT devices based on Bluetooth Low Energy (BLE) Technology, which is used for not only data communication but also proximity sensing and is ubiquitous in wearables, smartphones, personal tags, indoor-navigation beacons, and home-automation devices. The IoT metadata can cover different levels, from the radio signal strength and the media-access controller (MAC) address to device name, indoor location, temperature, humidity, air quality, and advertising messages. These IoT data can readily augment metadata such as timestamp, GPS location, and camera settings already captured by today's cameras and saved in digital photo files. With such IoT metadata, the user will have much richer ways to understand the subject and environment of the scene being captured by the photo or media, such as the subject's fitness condition, the advertised events, and the tagged personal items at the scene. These metadata will enable new ways of searching, querying, and organizing the photos and the associated objects. Moreover, a novel mechanism we propose is the ability to pair with a remote device whose identity and routing information is captured as IoT metadata, so that an authorized user can then pair with it at a later time remotely. We contend that IoT metadata capturing can bring significant benefits to the users.**

Keywords—metadata; Bluetooth Low Energy (BLE); digital snapshots; proximity sensing

1. Introduction

Advances in miniature and low-power electronics with wireless communication capabilities have given rise to the Internet of Things (IoT). It is now possible to embed electronics inside many every-day objects that transmit the state of the physical world as digital data to the cyber world over wireless interfaces. This paper proposes capturing and making use of such free but useful data in the airwaves at the time of taking photo snapshots, embedding them in the digital photos as metadata, and using them afterwards in creative ways.

Digital photography is one of the most powerful ways of capturing the state of the physical world. As a picture is worth a thousand words, there can be a wealth of information in the visual data, including the environment, the subjects in it, the lighting conditions, and symbols. Digital photography goes a step further by being able to embed metadata, including the time and date stamp, the camera settings, and even location stamp, thanks to the location services available on most smartphones. Post-processing makes it possible to embed even more metadata, including the recognized faces, user comments, user scores, event grouping, and additional data to help organize and search. However, such metadata represents only the very beginning of a much greater set of possible data that can be embedded into digital photography and the types of services possible. Thanks to the rise of the Internet of Things (IoT), cameras are no longer limited to embedding data from those sensors that are on the smartphones.

Moreover, much data about everyday objects in the surrounding area can be available for embedding into digital photos, even if they are not in direct view of the camera. These include not only electronic devices such as Bluetooth headphones, remote controls, toys, remote-controllable switches, and weight scales, but also electronically tagged objects. For instance, Bluetooth Low Energy (BLE) tags are being attached to wallets, keychains, pets collars, and other personal items so that the owner can be alerted when they are about to be left behind, or the user can use a smartphone to find the object wirelessly while within RF range. This can be very useful in terms of extending the search capability in less esoteric ways, such as "find the location (of the most recent photo) where my wallet could be detected." The object detection is not limited to one's own personal items but can also include bicycle locks, other cars in proximity, and other people's belongings. This means such digital photos can also be searched for not only lost items but more generally by the content of those IoT messages, including identification and description of a museum piece, shelf numbers and coupons as broadcast by beacons in a store.

In addition, as there are data packets, IoT protocols also include control packets that enable the connection

DOI 10.1109/SOSE.2017.18

and control of the IoT devices. Examples include devices in smart-home applications such as smart lightbulbs, thermostats, timers, and electronic door locks in the area. Most of these devices support remote access, possibly via some Internet gateway, but that does not necessarily mean they are already set up to work with the user's devices at the time of the snapshot. With the broadcast packets captured in such photos, it is possible for the user to detect the proximity of devices and possibly device reference to then enable remote binding and control as if they were right next to the printer. Of course, the proper user interface is needed to make such process natural and easy for the user. We demonstrate the feasibility of this idea and present measurement results to show the effectiveness of our proposed work.

2. Related Work

This section reviews the works related to embedding metadata in multimedia files including photo, video, and audio. Second, we focus on annotation and related technology for photos. Third, we review the technologies on remote control.

2.1. Embedded Metadata in Multimedia

Works on embedding metadata into multimedia and their applications can be divided into three categories based on intended usage: metadata creation, management, and enabling service creation.

Metadata created from time, location or media signal have been created for purposes such as browsing and searching [1]–[3], but metadata can be much more. As the number and scale of IoT grow, even more data will be available and accessible for applications not yet imagined. We propose collecting these data from more general IoT sources and creating and storing them as metadata to enable new real-world applications.

Metadata have been used for management of media data. Metadata such as time and location enable automated assembly of photographs [4]. Our work generalizes the types of metadata that can be embedded and searched, including ambient temperature, people in the surrounding, and objects that are within the RF range.

Metadata can also enable service creation for control applications such as smart homes. One approach is to use a configuration tool that combines semantic metadata and provide users with a visual interface for device control [6]. Control may be from a remote site via a local gateway, but device discovery or setup must be done onsite. Our approach captures the advertising packets of discoverable devices to enable after-the-fact offsite device identification, discovery, and configuration via the gateway.

2.2. Photo Annotation

Photo annotation was driven by the Internet and smartphones with built-in cameras, which have driven photo sharing online. Photo annotation can be done after the time of capture or at the time of capture [8]. Flickr [9] allows annotation of photos in the form of tags or textual tags after users uploading theirs photos to the Flickr, and the tags in Flickr are mostly labeled by the photo's owner mainly for facilitating search of the specific photos. Users can also annotate digital photos at the time of capture [10], where the client-server system assists the user by providing guesses about the content of the photos. Our approach can be viewed as a generalization of these annotated data to general IoT data at the time of capture.

3. Problem Statements

There are two problem statements with our proposed system. The first is the capturing of IoT data and embedding them into media such as photos. The second is a new use of a class of such IoT metadata for the purpose of binding and control of IoT device remotely.

3.1. Capturing of IoT Data

The first problem is to capture data from as many IoT devices in the surrounding environment as possible at the time of taking a photo on the smartphone. In our system, IoT metadata refers to data from surrounding IoT devices that are stored as metadata in photo files at the time the photos are taken. The photo file itself may already contain metadata including other properties determined by the camera, such as location, altitude, and timestamp that smartphones already capture. The new IoT data describes the state of the environment as an enabler for other applications. In the context of this paper, IoT metadata is defined relative to the photo scene as captured by these IoT and non-image sensors, including climate data such as temperature and humidity, proximity tag data that provides identifying information, physiological data such as heart-rate monitors, indoor beacons and the messages they broadcast (such as artwork information in a museum or advertisement in a store).

A main assumption is that unlike cameras with built-in clocks, GPS, and sensors that can be read immediately for existing types of metadata in photos, we assume that the IoT devices have a wireless data interface that is compatible with the smartphone. For the purpose of this work, we further assume Bluetooth Low Energy (BLE) as the wireless interface for its wide use in existing IoT devices. Its use of profiles enables dynamic discovery and querying of typed data in standardized ways. Depending on the IoT device, the capturing process may be passive or active.

Passive capturing means the IoT device simply broadcasts data in its entirety, so that the capturer simply has to listen to the broadcast and record it. An example of passive capturing is with iBeacon, which broadcasts its indoor location and other advertising data to smartphones without pairing. The primary issue with passive capturing has to do with the broadcasting interval, since the capturer needs to listen for long enough around the time of taking the photo in order to capture the broadcast data.

Active capturing, on the other hand, means the capturer needs to initiate a query, possibly after a pairing process, before it can get the data from the IoT device. Active capturing may be a superset of passive because it first has to listen to broadcast to discover the IoT device, determine what profile it supports, and pair with it, before it can query the specific characteristic values. Many wearable devices that are paired with smartphones work this way. One example is the TI SensorTag, which contains an accelerometer, gyroscope, magnetometer, altimeter, temperature sensor, and a light sensor, but it must be paired.

Another issue is raw data vs. processed data. For example, the TI SensorTag can capture the raw values of acceleration and angular velocity, but processing is required to convert the raw data into trajectory. The SensorTag is not powerful enough to perform the processing, the smartphone will capture raw data. A separate post-processing annotator will turn the raw data into trajectory and annotate the photo afterwards.

3.2. Discovery, Binding, and Control of Remote IoT via Photos

The second problem statement is to capture the reference to discoverable IoT devices to enable the viewer of the photo with the embedded device reference as IoT metadata to discover them as if the viewer were at the scene. One example is the BLE tags, which are often attached to items that are easily lost such as wallets, remote control units, and pets. They do not broadcast “data” since they contain no sensors, but they broadcast their own ID so they can be discovered. This would enable someone to find their own lost items or other people’s lost items from their photo collection using the same device discovery process, assuming such advertising packets are captured properly.

Furthermore, if the IoT device still operates properly and can get online via a gateway, the extended problem statement is to enable pairing and control of that IoT device remotely. An example would be for someone to discover a remote-controllable light switch from a photo, pair with it through gateways using the reference embedded in the photo, and turn on or off the light accordingly as if the user were operating their smartphone at the scene.

4. Technical Approach

4.1. System Architecture

Fig. 1 shows a system architecture for capturing IoT metadata and for their remote discovery and control. It consists of three subsystems: a set of IoT edge devices, an optional IoT gateway accessible to the IoT edge device, and a smartmobile (i.e., smartphone or tablet) with a built-in camera and the compatible RF interface. Without loss of generality, we assume BLE.

The IoT edge device is assumed to broadcast periodically either for the purpose of disseminating data or to be discovered. In BLE, many devices may be expected to be paired with the owner’s smartphone, but this poses no difficulty for BLE 4.1 or later, which support the concept of scatternet. That is, a BLE device can join multiple networks. While paired with one master, the node can still advertise and join another network or form its own network.

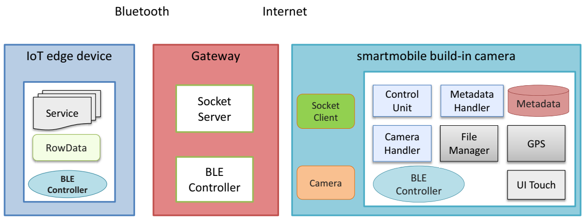


Figure 1. System Architecture

The gateway is a bridge between the edge devices and the Internet. It is optional in the sense that IoT edge devices are often paired with a smartmobile as its gateway to the Internet, but here we define the gateway as a separate base station-like system that serve as the uplink for these edge devices in case the smartmobile that they pair with is absent. In this case, the gateway has a WiFi or Ethernet interface on the Internet side and a BLE controller on the wireless device-area network side.

The smartmobile serves as the camera unit, capturer of IoT data via its built-in BLE interface, metadata storage, and the main controller of all these capturing activities. The smartmobile also contains definitions for the different device profiles so that it can perform passive or active capturing. The same definition is accessible to different apps so that it can be used to schedule the IoT data capturing tasks but also by other apps when making use of the extra metadata. For instance, a photo browser would not need to be hardwired with all possible types of search criteria; instead, it can offer additional search criteria depending on the available metadata definitions, such as temperature range, altitude, etc.

The smartmobile can also serve as the user interface to the second problem statement, namely device discovery through photos and remote connection. To do this, we provide an API layer for devices to access a virtual BLE network interface. Our API would redirect the access via TCP/IP to the associated gateway that then uses its BLE interface to communicate with the target IoT device. This is how the remote device discovery, binding, and control can be realized.

4.2. Device Discovery

We assume the user takes photos with an app that we provide, which performs device discovery to maintain the list of IoT edge devices at the scene of the photo to be taken. We have a software component called the BLE Manager to perform scanning as long as the app is running and possibly while in the background. Since BLE consumes relatively low energy, keeping it running is not expected to have a significant impact on the battery life. Depending on the available definition, the BLE manager will need to decide to perform passive or active capturing for each discovered IoT device.

4.3. IoT Data Capturing upon Snapshot

When the user taps the button to take a snapshot, IoT data captured by the BLE Manager up till the moment will be saved as IoT metadata associated with the photo. An implementation may save the IoT metadata in the app's local document file instead of embedding in the photos in the local camera roll for a number of reasons, including faster retrieval and OS-imposed restrictions. The primary issue is to determine the collection window size.

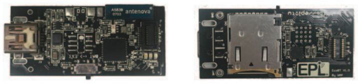
The *collection window size* is the time interval during which the BLE manager captures data from an IoT device for the purpose of synthesizing the corresponding IoT metadata. The collection window size depends on the IoT device type and purpose. For example, an IMU (inertial measurement unit) usually consists of an accelerometer and a gyroscope. Although one can obtain raw data samples using a collection window size equal to the sampling period, an IMU is usually used for a higher level purpose, such as motion tracking or gesture-based input. In the case of trajectory tracking, the collection window size therefore will need to be at least as large as the duration of the trajectory to track. Our system provides the expandable window size for these devices. We assume that the smartphone can discover the IoT device's type as part of its BLE attribute and determine the duration of collecting broadcasting data from it until the user takes the photo snapshot.

4.4. Binding Objects with IoT Metadata

The UITouch component is responsible for creating a relationship between objects in the photo and the IoT metadata. Through detecting the position of the user's finger on the UIView, as many points on the path surrounding the object as possible will be recorded for their X and Y coordinates. After the processing, four points will be selected to represent the location of the object on the picture. The four points representing a quadrilateral will be stored as IoT metadata. Therefore, users can interact with the object through the touchscreen if the touch point is in the range of the quadrilateral. After loading the captured photo, the original photo will be covered by a sublayer, and then the user moves a finger to draw the edges of the object. The detection of touch point is performed by UIKit framework.

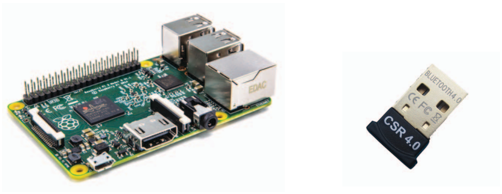
5. Evaluation

This section presents evaluation results from a number of experiments on our prototype IoT metadata creation system. We constructed our prototype using iOS devices as the smartmobiles, a gateway, based on Raspberry Pi 2 (Fig. 3(a)) with a CSR4.0 USB dongle adapter (Fig. 3(b)), and our own BLE nodes (Fig. 2) as IoT edge devices. We evaluate the performance in terms of the time latency and discuss the scalability issue.



(a) Top (b) Bottom

Figure 2. Photos of EcoBT Super



(a) Raspberry Pi 2 (b) BLE Modules

Figure 3. Components of gateway

5.1. Test Cases

Our test cases are two representative examples for the two problem statements (Section 3). They are for searching photos by IoT metadata and remote control.

5.1.1. Searching Photos. In this case, the user takes the photo by using smartmobile, and the smartmobile is going to capture data from IoT edge devices, like iBeacon and wearable devices. After the user saves the photo through BLE Manager in the smartmobile, the app stores the IoT metadata and embeds them into the photo.

One day, if the user tries to find something from the scene in memory. To address this, the user can enter the related information about the photo such as the restaurant's name or the environmental data that is like temperature or humidity of the scene. Then, the app starts searching the IoT metadata associated with the photos and renders the search results.

5.1.2. Remote Control. The second test cases involves first creating IoT metadata in the form of device reference followed by pairing and data transfer, as shown in Fig. 4.

In Fig. 5.1, there are two nodes the user wants to pair and control. The user takes a photo of the two nodes and selects one of the nodes on the smartphone screen (photo in Fig. 5.4), and the BLE Manager extracts the IoT metadata associated with the node from the photo's metadata and shown on screen. If the selected node provides services that can be remotely operated and is shown in Fig. 5.5, the smartphone will try to communicate by using socket communication with the gateway whose reference is included in the IoT metadata. The gateway queries what service the

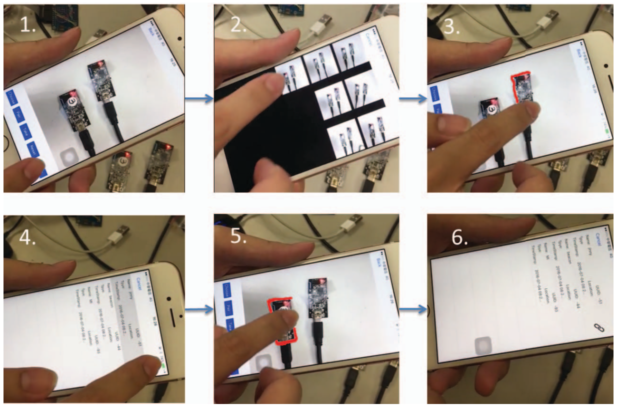


Figure 4. The process of creating IoT Metadata

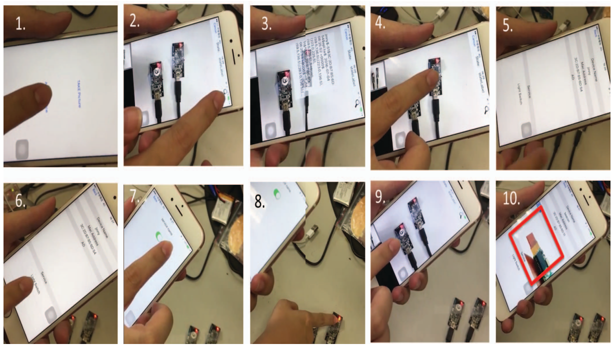


Figure 5. The process of accessing and controlling IoT edge devices

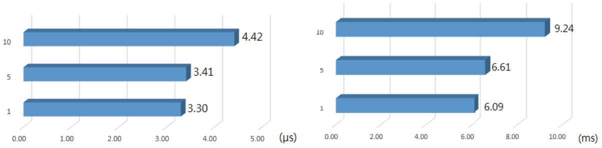
node provides through BLE. The user can choose a service to access and send the service name to the gateway. Then the gateway will query the node about the control parameter according to the service ID and send the query results to the smartphone. Therefore, the user can modify these control parameters of the service provided by node remotely. As the result is shown on Fig. 5.6-8, the user has ability to control the case of the lighting switch through its associated user interface on smart terminal.

5.2. Latency

We measure the latency from the time the smartmobile app sends an instruction to the time when the gateway returns the requested results. The result is shown in Table 1.

Table 1. LATENCY OF INSTRUCTIONS FOR REMOTE CONTROL

Instruction	Latency (sec)
Connect	1.15
UpdateCharacteristic	0.62



(a) (b)

Figure 6. Execution time of collecting data: (a) without accessing MAC address (b) with accessing MAC address of devices

Table 2. STABILITY OF CONNECTION

Experiment \Number of tested devices	1	5	10
1. connect successfully	100%	100%	100%
2. get MAC address successfully	100%	97.2%	86%
3. disconnect before getting MAC	0%	2.8%	14%

We also measure the time of capturing with or without accessing the MAC addressing over different numbers of nodes, i.e., 1, 5, and 10 nodes. In Fig. 6(a), we can see that the capturing times are very short, in units of microseconds, without accessing the MAC address of the nodes.

Fig. 6(b) shows the times with accessing the MAC address by connecting nodes. We find that the capturing time with accessing MAC address grows more 2000 times than capturing time without accessing the MAC address. Note that even though the time of capturing increases significantly percentage-wise, the actual time is still under 1 second.

5.3. Scalability

Scalability is the ability of a computer application or product (hardware or software) to continue to function well as the number of units increase. In our experiments, we test the stability of accessing the MAC address successfully, and the results are shown in Table 2. The probability of *connecting* successfully is 100% even as the number of devices grows. However, some of the devices lose connection before their MAC addresses can be accessed, and the rate of failure grows as the number of devices increases. From the results, we know that all failed cases can be attributed to the disconnection. The solution to this problem is to insert the delay time before building a new connection with the next device. This solution reduces the chance of disconnection due to collision, although the capturing time will increase relatively.

We also measure the number of devices in different real-world environments. The result in Fig. 7 shows that, with few exceptions, many devices exist in daily life and can all be scanned, and some of them even are connectible. Also, the types of devices depend on the environment. For example, the devices founded in the lab mostly are the evaluation board while devices founded in the gym are all smart bracelet such as FitBit.

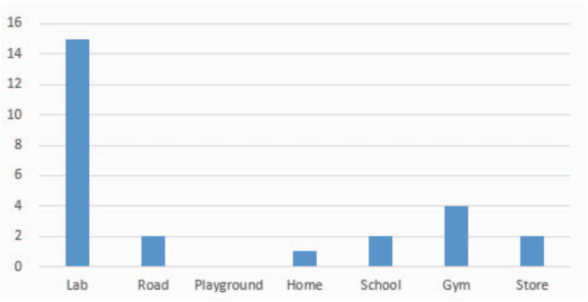


Figure 7. The number of devices in different environments

6. Conclusions

We propose an IoT metadata creation system that uses a novel, natural and efficient way to capture IoT data and the capturing process is transparent to the end user. Also, the feasibility of new applications enabled by the IoT metadata are demonstrated. Unlike other approaches that need the user to generate or choose those data by themselves, our system keeps more filtered and available data in the photo files and can support more dynamic interactions than just adding the tags or static information. With such IoT metadata, the user will have a new way to understand the subject and environment of the scene being captured by the photo or media, such as the subject's fitness condition, the advertised events, and the tagged personal items at the scene. We believe that the use of our system will bring new opportunities for IoT by making them more practical and accessible to more people.

Acknowledgments

This project is sponsored by Ministry of Science and Technology (Taiwan) grant numbers 105-2218-E-007-005- and 105-2218-E-007-011-.

References

[1] R. Sarvas, E. Herrarte, A. Wilhelm, and M. Davis. Metadata creation system for mobile images. In Proceedings of the 2nd international conference on Mobile systems, applications, and services, pages 36–48. ACM, 2004.

[2] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In Proceedings of the SIGCHI conference on Human factors in computing systems, pages 401–408. ACM, 2003.

[3] B. L. Davis, G. B. Rhoads, and W. Y. Conwell. Associating metadata with media signals, and searching for media signals using metadata, Feb. 21 2012. US Patent 8,121,342.

[4] M. Naaman, S. Harada, Q. Wang, H. Garcia-Molina, and A. Paepcke. Context data in geo-referenced digital photo collections. In Proceedings of the 12th annual ACM international conference on Multimedia, pages 196–203. ACM, 2004.

[5] W. Klas and A. Sheth. Multimedia Data Management: Using Metadata to Integrate and Apply Digital Media. March 1998.

[6] S. Mayer, N. Inhelder, R. Verborgh, R. Van de Walle, and F. Mattern. Configuration of smart environments made simple: Combining visual modeling with semantic metadata and reasoning. In Internet of Things (IOT), 2014 International Conference on the, pages 61–66. IEEE, 2014.

[7] M. Lanza, N. Octavian, and M. Rutchik. Embedded metadata in a media presentation, Mar. 27 2007. US Patent App. 11/691,936.

[8] A. Kuchinsky, C. Pering, M. L. Creech, D. Freeze, B. Serra, and J. Gwizdka. FotoFile: a consumer multimedia organization and retrieval system. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pages 496–503. ACM, 1999.

[9] Flickr. http://www.flickr.com.

[10] A. Wilhelm, Y. Takhteyev, R. Sarvas, N. Van House, and M. Davis. Photo annotation on a camera phone. In CHI'04 extended abstracts on Human factors in computing systems, pages 1403–1406. ACM, 2004.

[11] B. Shneiderman and H. Kang. Direct annotation: A drag-and-drop strategy for labeling photos. In Information Visualization, 2000. Proceedings. IEEE International Conference on, pages 88–95. IEEE, 2000.

[12] H. Kang and B. Shneiderman. Exploring personal media: A spatial interface supporting user-defined semantic regions. Journal of Visual Languages & Computing, 17(3):254-283, 2006.

[13] N. Sriskanthan, F. Tan, and A. Karande. Bluetooth based home automation system. Microprocessors and Microsystems, 26(6):281–289, 2002.

[14] G. Song, F. Ding, W. Zhang, and A. Song. A wireless power outlet system for smart homes. IEEE Transactions on Consumer Electronics, 54(4):1688–1691, 2008.

[15] R. Piyare and M. Tazil. Bluetooth based home automation system using cell phone. In Consumer Electronics (ISCE), 2011 IEEE 15th International Symposium on, pages 192–195. IEEE, 2011.

[16] K. Y. Lee and J. W. Choi. Remote-controlled home automation system via Bluetooth home network. In SICE 2003 Annual Conference, volume 3, pages 2824–2829. IEEE, 2003.