

Multi-Sensor, Third-Party API, and Cross-Device Integrated Toolkit for Enriching Participatory Research

Hyoseok Yoon^{1*}

Abstract

Smart devices include various functions and hardware sensors for specific purposes. The built-in functions and sensors of these smart devices can be repurposed to suit various additional use cases. Qualitative research methods such as "photovoice" are used in community-based participatory research, where participants take photos regarding a particular research theme. These user-generated photos are later interpreted and discussed to discover individuals' needs and problems. In this study, we aim to create a photo-based application that enables smart devices like smartphones, smart glasses, and smartwatches to function as a digital toolkit for consumers and academics. By utilizing cameras, GPS, voice notes, internal sensors, and third-party APIs that can gather both qualitative and quantitative data at the time of capturing images, we develop a photo-based integrated toolkit.

Key Words: Smart Device, Multi-Sensor, Toolkit, Participatory Research.

I. INTRODUCTION

Nowadays, cameras and vision sensors, particularly those in smart devices and Internet-of-Things (IoT) are pervasive. 6.64 billion people use smartphones worldwide, which is more than 83% of the world's population, according to BankMyCell [1]. A digital camera is a versatile tool that captures meaningful moments in a photograph to explore various context, create values, and identify problematic challenges such as accessibility issues [2]. Many qualitative research methods, such as "photovoice" [3-4] use photos taken by participants or communities to observe and identify theme-oriented problems and solutions. Since researchers tackle different domains and specific themes, a varying set of questions or requests is given to participants to collect relevant data and elicit insights from discussions. To meet the demands and reduce the burden on researchers and end-users who use photographs in their studies and life-logging, we propose a general-purpose yet integrated digital toolkit to serve this purpose. There are many potential applications to consume these personal data if a smart device application can capture images as qualitative personal data and record quantitative personal data. Additionally, it will be easier to concentrate on only relevant data if researchers and end users are able to identify their objective

or select a group of data to be collected. Our contributions on this paper are as follows.

- We explore how a digital toolkit in a cross-device and multi-sensor environment can configure smart devices and sensors of interest for qualitative and quantitative research methods (i.e., photovoice and experience sampling methods [5]) by enriching photographs with in-situ user-provided descriptions and system-generated records.
- We investigate the viability and possibility of cross-device and multi-device data gathering and interpretation because many people use many devices at the same time. We require a method for quickly sharing personal data in an interoperable format after it has been obtained.
- We offer suggestions for data integrity, data sharing, and data exchange formats to accelerate interoperable data sharing between researchers and end users.

II. RELATED WORK

Digital photographs are an expressive and easy-to-produce medium for both researchers and end-users. Koch and Maaß investigated a digital diary application as a digital probes kit to document the day using various media, answer

Manuscript received December 01, 2022; Revised December 16, 2022; Accepted December 20, 2022. (ID No. JMIS-22M-12-050)

Corresponding Author (*): Hyoseok Yoon, +82-31-379-0645, hyoon@hs.ac.kr

¹Division of Computer Engineering, Hanshin University, Osan, Korea, hyoon@hs.ac.kr

questionnaires, create content, and take pictures [6]. Several digital diary approaches used geo-tagged multimedia content [7], emotion-based life-logging [8], and participant-generated photographs [9]. Tan et al. explored using a digital crowdsourcing strategy to involve the community in clinical trials [10]. As illustrated in Fig. 1 and Fig. 2, researchers use photographs as a tool to gather qualitative data, and users use photographs to concisely capture and share their meaningful moments.

Key themes in patient-generated images and films, such as contexts employed, values attained by patients and medical personnel, and difficulties encountered, have been discovered by Ploderer et al. [11]. Fig. 3 illustrates how patient-generated photos and videos are useful in research.

Digital technologies were investigated by Bruckermann et al. for citizen science during the stages of data collection and analysis [12]. Sharples et al. developed a sensor toolkit, the Sense-it app, to access embedded sensors on Android smartphones [13]. Our approach is similar to digital diary studies, where we use user-taken photos as the qualitative data source. However, we further encode and enrich these photos with sensors, as explored by [13], and third-party

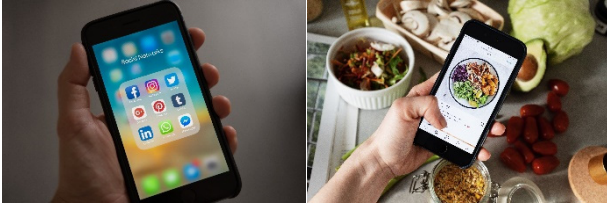


Fig. 1. Uses of photographs by users in social network services.

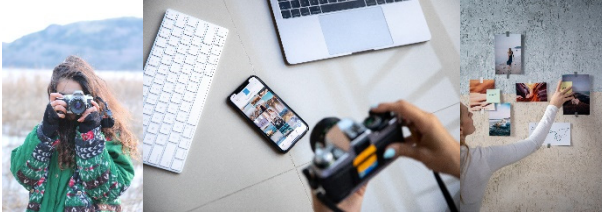


Fig. 2. Uses of photographs for capturing moments and remembering.

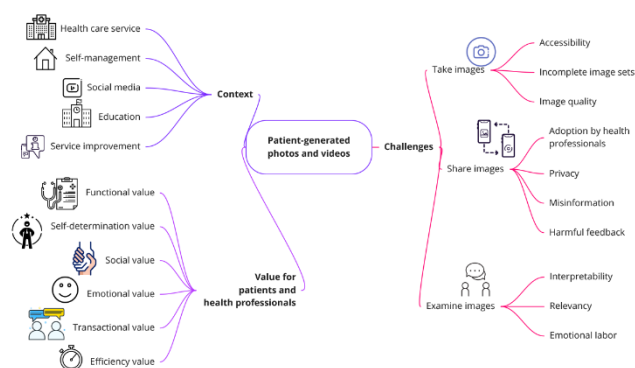


Fig. 3. Key themes identified in [11] (the image is from [11], an open access article).

Table 1. Comparative summaries of previous studies.

Studies	Brief description
[6]	Digital probes kit using questionnaires and pictures
[7]	Digital diary using geo-tagged multimedia
[8]	Digital diary exploring the life-logging concept
[9]	Mobile app for visual research using photographs
[10]	Digital crowdsourcing in the clinical trials context
[11]	Key themes and challenges in patient-generated images
[12]	Exploration in citizen science for data collection/analysis
[13]	Using embedded smartphone sensors for data collection
[14]	Outlining requirements for local civic participation tools

APIs for participatory research and civic participation tools [14]. Comparative summaries of previous studies are provided in Table 1.

III. DESIGN OF MULTI-SENSOR AND CROSS-DEVICE TOOLKIT

To gather data across several smart devices, both qualitative and quantitative, we created an integrated digital toolkit. Fig. 4 depicts the digital toolkit we propose. Smartphones, smartwatches, and smart glasses are just a few examples of smart devices that can run the proposed digital toolkit as an application.

3.1. Qualitative Data

When a user of this digital toolkit takes a picture, the user can leave text or voice messages on the picture. This produces qualitative data such as photo, text description, and voice memo. Qualitative data is categorized by how they are expressed, usually in descriptive language rather than numerical data. For example, photos can be labeled or described (i.e., a man is wearing a hat), and text and voice memos can be written and expressed in English.

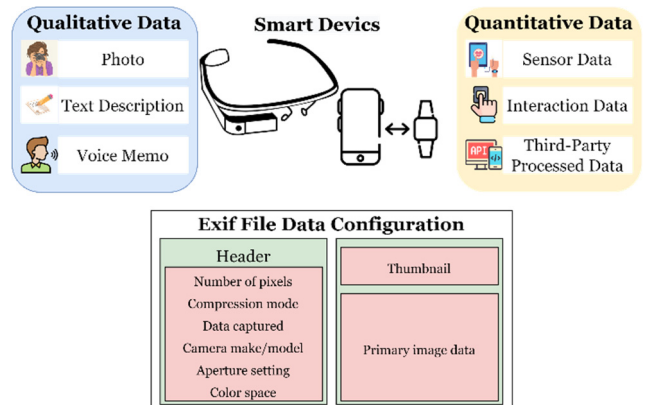


Fig. 4. Design of multi-sensor and cross-device toolkit.

3.2. Quantitative Data

For quantitative data, we collect multimodal data from embedded sensors [15], user interaction data, and third-party processed data. Data that can be measured and expressed as a numerical value is referred to as quantitative data, as opposed to qualitative data. An ambient light sensor, for instance, monitors ambient light in Lux units. In an application, the total number of user clicks can be counted. Image size is expressed as a number. Additionally, further interpretation of raw data is possible by using third-party APIs such as Google Vision AI (<https://cloud.google.com/vision/>) and Kakao Vision API (<https://vision-api.kakao.com/>). With the aid of these multimodal data [15], designers and service providers may better understand their target demographic and create systems that cater to their demands.

3.3. Exif File Data Configuration

This concept is already used in Exif (Exchangeable image file) where digital camera-related information (including geo-location) is stored. Our goal is to expand this concept so that it can be applied to smart devices and third-party APIs like vision APIs. The targeted embedded sensors include motion, environmental, positional, ambient light, proximity, accelerometer, gyroscope, barometer, and microphone data.

3.4. Smart Devices

It is important to consider cross-device and multi-device environments [16], as shown in Fig. 5 since many people carry more than one connected device with different levels of sensor accuracy, resolution, and sensing interval [17]. Our proposed digital toolkit supports Android-based smart devices (i.e., Google Glass Enterprise Edition 2, Samsung Galaxy Smartphone, and Samsung Galaxy Watch 4).

Wearable smart glasses such as Google Glass feature always-on, always-accessible, and always-connected characteristics. Various sensors and cameras in these wearable computers enable many engaging scenarios [17-20] including medication information provision [21] to provide essential information. For example, Apiratwarakul et al. explored using smart glasses as a tool for assessing people in mass-

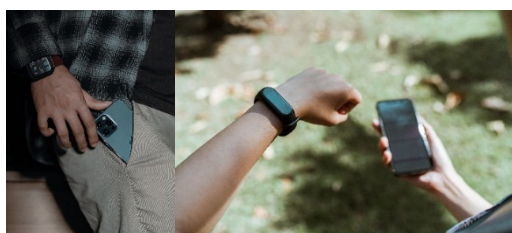


Fig. 5. Cross-device and multi-device environments.

casualty incidents [22]. We propose integrating third-party API, such as text recognition API, to automatically translate the acquired photo in order to support and develop a prospective killer application for wearable smart glasses in participatory research.

3.5. Data Sharing

Once these user-generated photographs with additional contextual information are collected, we must provide efficient and effective data-sharing and exporting methods. For example, we store integrity-checked photographs (i.e., MD5 checksum) in interoperable file formats such as JSON and XML. JSON and XML allow user-defined custom tags and values to be easily expanded to include user-annotated comments, ratings, and voice memos. These file formats can also be easily exported and refined by setting different conditions, such as filtering by dates and sensor value thresholds. These data are kept in a cloud database, a local SQLite database, or an internal program directory on an Android platform.

IV. PROTOTYPE DESIGN AND IMPLEMENTATION

We implemented a prototype to demonstrate our digital toolkit. The prototype can take pictures, add tags to the pictures, and export them as JSON files. Fig. 6 shows a user taking a photo of a ceiling fan and adding user-specified tags using an Android smartphone application. We developed our application using Java with the Android Studio (<https://developer.android.com/studio>) Bumblebee 2021.1.1 version. We have tested on android virtual device (AVD) as well as on two Android smartphones (Samsung Galaxy Note 9 running Android 10 OS and Samsung Galaxy Note 10 running Android 11 OS).

Fig. 7 and Fig. 8 shows a Google Glass application for automatically interpreting texts on a user's view by taking a photo and running through a text recognition API. This

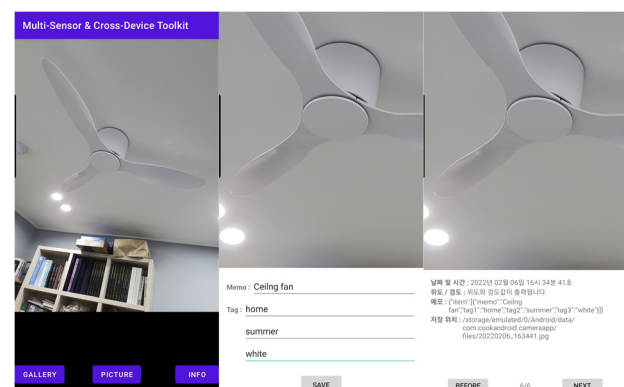


Fig. 6. A prototype Android app to tag and embed sensed data.

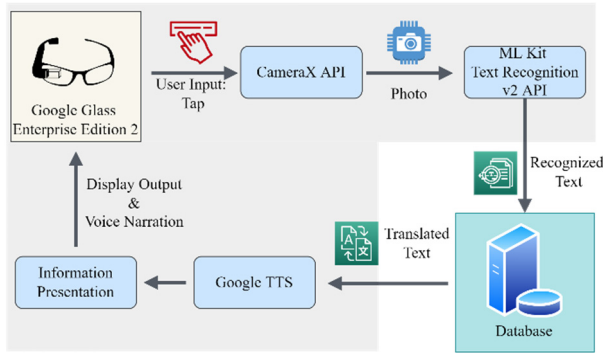


Fig. 7. Using Google Glass with OCR API.

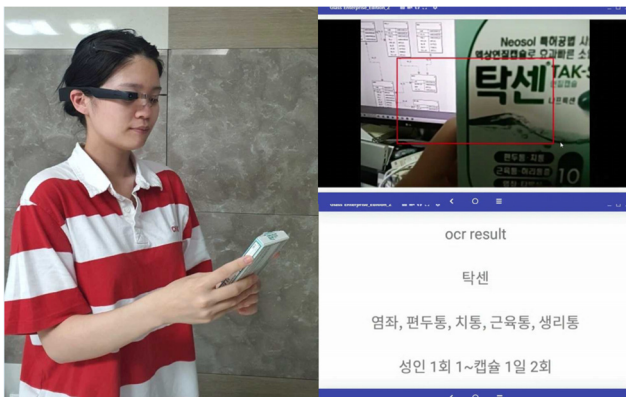


Fig. 8. A Google Glass wearer automatically recognizes texts.

photo is then analyzed by text recognition API (Application Programming Interface) and outputs a series of recognized texts. To implement and simulate this application, we designed a database to contain pre-registered textual information that can be matched and retrieved by text recognition results. Then automatically recognized information is concisely delivered using text-to-speech (TTS) or a comprehensible form of intuitive instruction layout.

Our wearable prototype was designed and implemented on Glass Enterprise Edition 2 (GEE2) by Google (<https://www.google.com/glass/tech-specs/>). The GEE2 runs on the Android Oreo operating system (system image version of OPM1.210425.001) and has Wi-Fi, Bluetooth, inertial sensors, a multi-touch gesture touchpad, a 640 by 360 display, and an 8-megapixel camera. We used a tap gesture on the touchpad to take a picture with the CameraX API (<https://developer.android.com/training/camerax>). For the text recognition API, we used the ML Kit Text Recognition V2 API (<https://developers.google.com/ml-kit/vision/text-recognition/v2>). When a region of interest (ROI) is identified in the taken picture, the image is analyzed by the ML Kit Text Recognition V2 API. The ML Kit supports text recognition in images or videos in various languages, including English, Korean, Chinese, and Japanese. The recognized text is output as a block of text. The recognized text is used to retrieve relevant data from a pre-registered database. For

the database, we used Firebase real-time database (<https://firebase.google.com/docs/database>). Using this database allows a cloud-hosted database to store JSON-based data.

V. CONCLUSION

In this paper, we proposed a design of a multi-sensor, third-party API, and cross-device integrated toolkit. The proposed integrated toolkit enhances applications for HCI researchers, self-tracking, and photo-based participatory research. We envision the proposed toolkit to support easy and efficient personal data acquisition and sharing for qualitative research methods such as ethnographic research, focus groups, record keeping, case study research, and interviews. We have yet to test our envisioned application with actual users and in actual scenarios. Research is still needed to examine the usability, satisfaction, and presentation layout [23] of the deployed input and output modalities on various smart devices. We plan to investigate systematic methods for compiling and analyzing personal multimodal data including physiological signals [15] collected from cross-device interaction.

ACKNOWLEDGMENT

This work was supported by Hanshin University Research Grant. We thank the members of the Hanshin HCI Lab. (Siyeon Kim, Nahyun Kim, and Hong Ji Lim) for application development. Fig. 1, Fig. 2, and Fig. 5 are from Pexels with the free license to use (<https://www.pexels.com/license>). Fig. 3 is from an open access article under CC BY 4.0 (Bernd Ploderer, Atae Rezaei Aghdam, Kara Burns. Originally published in the Journal of Medical Internet Research).

REFERENCES

- [1] A. Turner, How Many People Have Smartphones Worldwide, Dec. 2022. <https://www.bankmycell.com/blog/how-many-phones-are-in-the-world>.
- [2] Z. Sarsenbayeva, N. V. Berkel, E. Velloso, J. Goncalves, and V. Kostakos, "Methodological standards in accessibility research on motor impairments: A survey," *ACM Computing Surveys*, vol. 55, no. 7, p. 143, Dec. 2022.
- [3] C. Wang and M. A. Burris, "Photovoice: Concept, methodology, and use for participatory needs assessment," *Health Education & Behavior*, vol. 24, no. 3, pp. 369-387, Jun. 1997.
- [4] E. Smith, M. Carter, E. Walklet, and P. Hazell, "What are photovoice studies?" *Evidence Based Nursing*, vol.

- 25, pp. 6-7, 2022.
- [5] N. van Berkel, D. Ferreira, and V. Kostakos, "The experience sampling method on mobile devices," *ACM Computing Surveys*, vol. 50, no. 6, p. 93, Nov. 2018.
- [6] D. Koch and S. Maaß, "Digital probes kit: A concept for digital probes," *i-com*, vol. 17, no. 2, pp. 169-178, 2018.
- [7] A. Ahmad, I. Afyouni, A. Murad, Md. A. Rahman, and F. U. Rehman, et al., "ST-diary: A multimedia authoring environment for crowdsourced spatio-temporal events," in *Proceedings of the 8th ACM SIGSPATIAL International Workshop on Location-Based Social Networks*, 2015, p. 2.
- [8] Y. Park, B. Kang, and H. Choo, "A digital diary making system based on user life-log," in *Proceedings of the 2016 International Conference on Internet of Vehicles*, 2016, pp. 206-213.
- [9] S. Barriage and A. Hicks, "Mobile apps for visual research: Affordances and challenges for participant-generated photography," *Library & Information Science Research*, vol. 42, no. 3, 101033, 2020.
- [10] R. K. J. Tan, D. Wu, S. Day, Y. Zhao, H. J. Larson, and S. Sylvia, et al., "Digital approaches to enhancing community engagement in clinical trials," *NPJ Digital Medicine*, vol. 5, no. 37, Mar. 2022.
- [11] B. Ploderer, A. R. Aghdam, and K. Burns, "Patient-generated health photos and videos across health and well-being contexts: Scoping review," *Journal of Medical Internet Research*, vol. 24, no. 4, p. e28867, Apr. 2022.
- [12] T. Bruckermann, H. Greving, M. Stillfried, A. Schumann, M. Brandt, and U. Harms, "I'm fine with collecting data: Engagement profiles differ depending on scientific activities in an online community of a citizen science project," *PLoS ONE*, vol. 17, no. 10, p. e0275785, Oct. 2022.
- [13] M. Sharples, M. Aristeidou, E. Villasclaras-Fernández, C. Herodotou, and E. Scanlon, "The sense-it app: A smartphone sensor toolkit for citizen inquiry learning," *International Journal of Mobile and Blended Learning*, vol. 9, no. 2, pp. 16-38, Apr.-Jun. 2017.
- [14] F. Maas, S. Wolf, A. Hohm, and H. Jörn, "Citizen needs – to be considered: Requirements for local civic participation tools," *i-com*, vol. 20, no. 2, pp. 141-159, 2021.
- [15] M. N. Giannakos, K. Sharma, I. O. Pappas, V. Kostakos, and E. Velloso, "Multimodal data as a means to understand the learning experience," *International Journal of Information Management*, vol. 48, pp. 108-119, Oct. 2019.
- [16] H. Yoon and C. Shin, "Cross-device computation coordination for mobile collocated interactions with wearables," *Sensors*, vol. 19, no. 4, p. 796, 2019.
- [17] C. Luo, J. Goncalves, E. Velloso, and V. Kostakos, "A survey of context simulation for testing mobile context-aware applications," *ACM Computing Surveys*, vol. 53, no. 1, p. 21, Feb. 2020.
- [18] E. Zarepour, M. Hosseini, S. Kanhere, A. Sowmya, and H. Rabiee, "Applications and challenges of wearable visual lifeloggers," *Computer*, vol. 50, no. 3, pp. 60-69, Mar. 2017.
- [19] H. Yoon, S. K. Kim, Y. Lee, and J. Choi, "Google Glass-supported cooperative training for health professionals: A case study based on using remote desktop virtual support," *Journal of Multidisciplinary Healthcare*, vol. 14, pp. 1451-1462, Jun. 2021.
- [20] H. Yoon, S. B. Kim, and N. Kim, "Design and implementation of procedural self-instructional contents and application on smart glasses," *Journal of Multimedia Information System*, vol. 8, no. 4, pp. 243-250, Dec. 2021.
- [21] D. Roosan, Y. Li, A. Law, H. Truong, M. Karim, and J. Chok, et al., "Improving medication information presentation through interactive visualization in mobile apps: human factors design," *JMIR mHealth and uHealth*, vol. 7, no. 11, p. e15940, Nov. 2019.
- [22] K. Apiratwarakul, L. Cheung, S. Tiamkao, P. Phungoen, K. Tientanopajai, and W. Taweeprawadej, et al., "Smart glasses: A new tool for assessing the number of patients in mass-casualty incidents," *Prehospital and Disaster Medicine*, vol. 37, no. 4, pp. 480-484, June 2022.
- [23] S. K. Kim, Y. Lee, H. Yoon, and J. Choi, "Adaptation of extended reality smart glasses for core nursing skill training among undergraduate nursing students: Usability and feasibility study," *Journal of Medical Internet Research*, vol. 23, no. 3, p. e24313, Mar. 2021.

AUTHOR



Hyoseok Yoon received his B.S. degree in Computer Science from Soongsil University in 2005. He received his M.S. and Ph.D. degrees in Information and Communication (Computer Science and Engineering) from the Gwangju Institute of Science and Technology (GIST), in 2007 and 2012, respectively.

He was a researcher at the GIST Culture Technology Institute from 2012 to 2013 and was a research associate at the Korea Advanced Institute of Science and Technology, Culture Technology Research Institute in 2014. He was a senior researcher at Korea Electronics Technology Institute from 2014 to 2019. In September 2019, he joined the Division of Computer Engineering, Hanshin University where he is currently an assistant professor. His research interests include ubiquitous computing (context-awareness, wearable computing) and Human-Computer Interaction (mobile and wearable UI/UX, MR/AR/VR interaction).