ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume **25**, Number 1, 2022, 80–91

LifeTags++: A Multi-User, Multi-Device, and Multi-Perspective System for Recording and Abstracting Visual Life with Tag Clouds

Adrian AIORDĂCHIOAE and Radu-Daniel VATAVU*

MintViz Lab | MANSiD Research Center, Ştefan cel Mare University of Suceava, Romania
E-mail: `radu.vatavu@usm.ro`*, `adrian.aiordachioae2@student.usv.ro`
* Corresponding author

Abstract. We present LifeTags++, a multi-user multi-perspective lifelogging and life abstraction system designed to automatically record, extract, process, and store visual concepts from a wide range of video camera devices and systems, including personal mobile and wearable devices, video cameras worn on the body, and public IP video cameras from the environment. Other features of LifeTags++ include integration with mobile crowdsensing to accept queries and deliver notifications about specific concepts of interest entered by users. We present technical details regarding the design and engineering of LifeTags++ as a distributed, cloud-based software architecture accommodating mobile and wearable devices and public ambient video cameras. Also, we report the results of an evaluation study with a dataset of 18,780 visual concepts extracted from a number of nearly 1,000 snapshots captured from three video cameras simultaneously recording the life of a user from three different perspectives: egocentric eye-level provided by a pair of video camera glasses, egocentric back view with a wearable clip-on camera, and exocentric ambient from a public IP camera. Our results show the benefits of employing multi-perspective life abstraction for creating rich lifelogs of visual concepts. Based on our results, we present opportunities for future lifelogging systems that employ mobile devices, wearable cameras, and personal robots in the form of drones in conjunction with ambient sensing and computing infrastructure..

Key-words: Mobile, distributed systems, cloud computing; Lifelogging; Smartglasses; Camera glasses; Image processing; Smart wearables; Assistive vision.

1. Introduction

We are witnessing an increase in the availability, affordability, and diversity of wearable computer devices and systems, from smartwatches [8, 10] to *Augmented* and *Mixed Reality* (AR/MR) smartglasses and head-mounted displays [16], smart bracelets, fitness trackers, and smart jewelry for health monitoring [20], smart rings for the control of remote devices and rapid POS payments [30], and personal wearable cameras for lifelogging applications [14]. Among these, smartglasses are expected to gain considerable adoption in the foreseeable future, as per a January 2021 Gartner Report [13], mainly due to the numerous applications and services they enable [25] for a variety of user categories, including workplace and lucrative activities [27], assistive technology for users with visual impairments [22, 32], and lifelogging [2, 4], to mention just a few examples. Regarding the latter, wearable devices with built-in video cameras are gaining more attention as consumers have embraced lifelogging as a way of life in order to recollect, reminisce, retrieve, reflect, and remember relevant aspects of their lives [24]. However, current systems and applications designed to support lifelogging are limited by the single perspective they provide for recording life. Consequently, lifeloggers have access to photos and video records captured from one single point of view, *e.g.*, that of the clip-on video camera, which limits the richness of recollecting and reflecting on their life experiences. Moreover, most commercial wearable cameras are designed as clip-ons, such as the Narrative Clip 2, which leads to a perspective of recording visual life different from the one originally experienced at eye level by the lifelogger.

We introduce in this paper LifeTags++, a multi-user, multi-device, multi-perspective video camera-based system for recording and abstracting visual life. LifeTags++ extends the functionality and scope of our prior work represented by LifeTags [4], a single-user, single-device, and single-perspective lifelogging system. Moving beyond this prior work, LifeTags++ enables integration of multiple points of view in the lifelog and, thus, presents users with a richer perspective on their visual life. Our paper is organized as follows: Section 2 positions our contributions in the context of lifelogging and mobile crowdsensing research; Section 3 presents the technical details of the design and engineering of LifeTags++, our distributed, cloud-based software architecture accommodating mobile devices, wearables, and public IP video cameras to enable three different perspectives, *i.e.*, egocentric eye-level, egocentric back view, and exocentric ambient, for lifeloggers; Section 4 presents an evaluation of LifeTags++ on a dataset of 939 snapshots captured from three video sources, from which a total number of 18,780 concepts were extracted and analyzed; we conclude in Section 5 by presenting other application opportunities for lifelogging systems.

2. Related Work

Lifelogging has been described as a phenomenon, whereby people record various aspects of their daily life [14] with the lifelog serving as "a form of pervasive computing consisting of a unified digital record of the totality of an individual's experiences, captured multi-modally through digital sensors and stored permanently as a personal multimedia archive" [11] (p. 431). A wealth of data can be recorded about people's daily activities to create lifelogs for various applications [3, 14, 29], such as food logging [17], monitoring sleep patterns [19], quality of life [33], and travel logs [5]. Some of these applications fuse information from multiple sensors [5, 19, 33] to create rich records for the lifelog and capture a diversity of data, *e.g.*, visual, aural, physiological, etc., while others use just one source, mainly video cameras embedded in various devices or worn directly on the body [6, 7, 17]. Aiordchioae *et al.* [3] discussed lifelogging applications at the intersection with cross-reality [23] and multimedia alternate realities [9], and proposed a design space for broadcasting personal visual realities captured by wearable devices with built-in video cameras to remote audiences.

From the scientific literature on lifelogging, the most relevant to our work is LifeTags [4], a lifelogging and life abstracting wearable system that presents users with word clouds representing summaries of their visual lives. This specific way to abstract life instead of simply recording it enables better privacy for passers-by, an important aspect frequently highlighted in the scientific literature for wearable video cameras [6, 18]. However, LifeTags [4] was designed for single users and offered a single point of view, *i.e.*, that of the video camera integrated in the smartglasses. A follow-up version [2] extended LifeTags with a mobile application and a software architecture based on cloud platforms and services, where multiple users could participate to mobile crowdsensing goals, *e.g.*, to identify missing objects or points of interest. Mobile crowdsensing [12] describes the approach where a large group of individuals with mobile devices collectively share data and extract information for a common goal. In this context, the approach from [2] was to repurpose lifelogging video cameras and systems, such as LifeTags [4], with a shift of perspective in terms of privacy concerns about these systems. Passers-by, who are usually impacted negatively by wearable cameras in public places, may find themselves among the beneficiaries of the data that is collected. However, one of the limitations of the opportunistic crowdsensing system from [2] is that the user does not have access to multiple perspectives in their lifelog, nor does the software architecture enable simple addition of new perspectives provided by other devices.

3. LifeTags++

We introduce LifeTags++, a software architecture and system implementation that employ computer vision services available on the web to extract visual concepts from snapshots captured by video cameras and store those concepts for future consultation (*e.g.*, to support reminiscing and reflection [24]), but also for informing users about specific objects or concepts of interest located in their proximity. LifeTags++ is a multi-user, multi-device, multi-perspective lifelogging system enabling data collection from multiple video sources simultaneously, as follows:

- Glasses with built-in video cameras provide an egocentric, eye-level perspective;
- Clip-on wearable cameras provide an egocentric, body-level perspective, *e.g.*, the Narrative Clip 2 camera can be attached to the user's shirt, hat, backpack, etc.;
- Exocentric, aerial perspective provided by video cameras built into drones. This category includes both small drones for indoor operation or large drones designed to operate outdoors in public places;
- Exocentric, ambient perspective provided by video cameras installed in public places, either indoors or outdoors, such as IP cameras;
- Exocentric, other-people perspective provided by video cameras worn by other people, e.g., passers-by from the proximity of the LifeTags++ user.

LifeTags++ draws inspiration from narration multiperspectivity [15], where the audience of a story is presented with multiple, different perspectives on a given topic, *i.e.*, "a basic aspect of narration or as a mode of storytelling in which multiple and often discrepant viewpoints are employed for the presentation and evaluation of a story and its storyworld" [15] (p. 353). Another inspiration source is mosaic novels, which report individual stories that share a common setting or characters to present the reader with a linear story encompassing a plurality of viewpoints and styles.

Figure 1 presents the block diagram of the LifeTags++ software architecture, highlighting devices, software components, cloud services, and dataflows that enable its operation. Wearable devices with built-in video cameras, such as smartglasses and clip-on cameras, connect to the hotspot of the user's smartphone or, optionally, to a wireless router to which the smartphone is also connected. These devices deliver snapshots following HTTP requests performed by the smartphone application; see dataflow "a" (acquisition) in Figure 1. The snapshots are captured at a frequency of a few seconds apart, which was found sufficient according to [4] to generate a rich lifelog. The smartphone application, implemented in our case for the Android operating system, converts the snapshots in the base64 format and stores the timestamp, user and device IDs, and GPS coordinates, when available. The data is transmitted to a node.js server – see dataflow q (queuing) in Figure 1 – that employs a computer vision service from Clarifai, a Computer Vision, NLP, and Machine Learning platform on the web, – see dataflow "p" (processing) in Figure 1 – to obtain a list of concepts describing the snapshot. Then, the server performs the following tasks:

- Stores the snapshot to the Firebase Storage (FS) and the additional information to the Firebase Realtime Database (FRD), which are cloud-hosted NoSQL databases synchronized in real-time to all of the connected clients, for future consultation of the lifelog; see dataflow "s" (storage) represented in Figure 1.

- Checks for matches of the identified concept(s) in a list of concepts of interest specified by the users of LifeTags++. In the case of a match being found, a message is sent to the smartphone application that delivers a notification to the user, according to the dataflow n (notification) from Figure 1. The notification contains the snapshot with the identified concept as well as the GPS coordinates of the location where the snapshot was originally taken.

- Following its predecessor [2], LifeTags++ implements mobile crowdsensing by delivering notifications to users when a snapshot contains a concept of interest. When the server detects such a match, the user is notified on the smartphone application with a message containing the image of the respective concept.

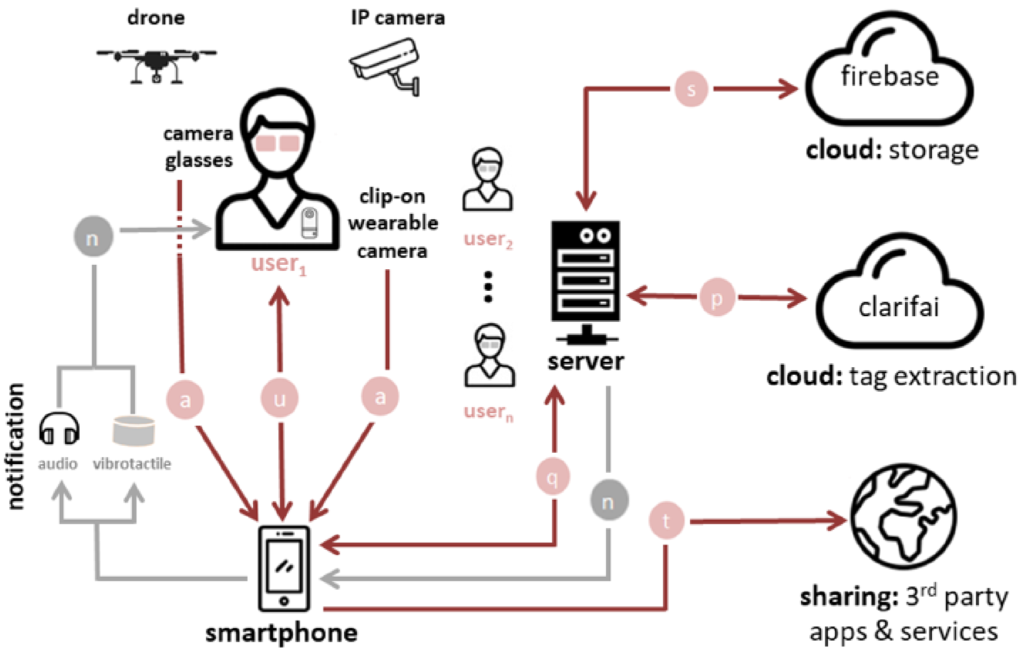

Fig. 1. System architecture of LifeTags++ showing mobile and wearable devices, software components, and corresponding dataflows. For example, a snapshot captured by a pair of glasses with built-in camera is processed by a specialized cloud service and the result stored in the lifelog.

The lifelog is presented to the user in the form of a cloud of tags, a concept introduced with the first version of LifeTags [4] in order to summarize, by means of abstraction, the visual information collected for the lifelog. LifeTags++ builds on the design requirements of the previous versions [2, 4] regarding passive and automated lifelogging, inconspicuous and privacy-orientated lifelogging, and support for memory access provided at various levels of granularity. A cloud of tags CT is formally defined as a set composed of concepts c_i represented by individual words w_i extracted from snapshots by employing a computer vision model, such as that used by the LifeTags++ software architecture illustrated in Figure 1, for which associated frequency information f_i can be computed and context information i acquired and stored, respectively, as follows:

$$CT = \{c_i = (w_i, f_i, \Omega_i) \mid i = 1..n\} \tag{1}$$

where n denotes the size of the cloud. The context information Ω_i represents a set of additional data, for which the structure and size depend on the capabilities of the sensors used by the lifelogging application, *e.g.*, the GPS information can be stored to associate the identified concepts with physical locations, but also by the actual usage intended for the application employing the tag cloud, *e.g.*, lifelogging, crowdsensing, or both. Figures 2 and 3 present the abstraction process for several snapshots captured by three devices corresponding to three different perspectives for recording the visual life of a LifeTags++ user. The first word cloud CT_1 was generated from images recorded from an eye-level perspective (camera glasses), the second cloud CT_2 from images recorded from a body-level perspective (clip-on video camera placed on the user's back), and the third cloud CT_3 from an exocentric ambient perspective (public IP camera from the surroundings of the user). By representing tag clouds as sets, specific operations can be applied when multiple perspectives are available, such as union or intersection of tag clouds. For example, the tag cloud from the right of Figure 3 was obtained by combining the eye-level, back-view, and ambient perspectives and can be formalized mathematically as the union $CT_1 \cup CT_2 \cup CT_3$. More generally, the tag cloud CT_+ obtained from a number of m different perspectives independently recorded from m different video sources that generated tag clouds CT_j is:

$$CT_+ = \bigcup_{j=1}^{m} CT_j = \left\{ c_i = \left(w_i, \sum_{\substack{(w,f,\Omega)\,\in\, CT_j \\ w=w_i,\, j=1..m}} f, \bigcup_{\substack{(w,f,\Omega)\,\in\, CT_j \\ w=w_i,\, j=1..m}} \Omega \right) \mid i = 1.. \left| \bigcup_{j=1}^{m} CT_j \right| \right\} \quad (2)$$

where $\bigcup$ denotes the union operation and $|\cdot|$ the cardinality of a set. Correspondingly, we can highlight the difference CT_- between the perspectives reflected by two tag clouds computed from two video sources, *e.g.*, the concepts CT_1 that were detected by the camera glasses but were missed by the ambient camera CT_3, as follows:

$$CT_- = \{c_i = (w_i, f_i, \Omega_i) \in CT_1 \mid w_i \notin \{w \mid (w, f, \Omega) \in CT_3\}, i = 1.. \mid CT_1 - CT_3 \mid\} \quad (3)$$

The difference of perspective caught by a specific tag cloud CT_* with respect to those provided by other m tag clouds CT_j computed from m video sources is formalized as:

$$CT_- = \big\{\{c_i = (w_i, f_i, \Omega_i) \in CT_* \mid w_i \notin \{w \mid (w, f, \Omega) \in CT_j, j = 1..m\}, i = 1..n_*\big\} \quad (4)$$

where $n_* = \left| CT_* - \bigcup_{j=1}^{m} CT_j \right|$. In our evaluation from Section 4, we present more examples of operations for tag clouds. Our implementations in the form of node.js, HTML/JavaScript, and Java for Android applications together with a sample of snapshots are available from http://www.eed.usv.ro/mintviz/projects/MintVizAwardingParticipationH2020.

4. Evaluation

We performed a technical evaluation of LifeTags++ to understand the impact on the richness of the lifelog when recording visual life from multiple perspectives represented by multiple video sources capturing snapshots simultaneously compared to the situation of the single perspective of LifeTags [2,4] addressed in our prior work.

4.1. Dataset

We asked a volunteer to wear the NorthVision Technologies glasses with built-in video camera (corresponding to the egocentric, eye-level perspective) as well as a clip-on video camera on the back (egocentric, back view perspective) and to walk in a delimited perimeter outdoors that was covered by a public IP camera from the top of a nearby building (corresponding to the exocentric, ambient perspective); see Figure 2 for examples of snapshots captured from these video sources. The camera glasses and the backside camera were Full HD (1920×1080 pixels) and the public video camera provided video over IP at a resolution of 640×480 pixels. According to the study of Alharbi *et al.* [6], the positioning of wearable video cameras and the shooting angle can influence user behavior because of the social discomfort caused by the surveillance of passers-by. For this reason, we have carefully chosen the video sources for our study as follows. We preferred a wearable video camera placed on the user's back as we felt it was less suspicious for passers-by than a frontal camera placed directly on the chest. We also preferred a fixed IP video camera from the environment to record an exocentric perspective in the lifelog, complementary to the egocentric ones provided by the two wearable camera devices.

A total number of 939 snapshots were captured from the three video sources, from which a total number of 18,780 concepts were extracted with a confidence level over 85% by the Clarifai computer vision service; see Figure 1 for the software architecture of LifeTags++. Of these, we found that a number of 311 concepts were distinct, representing approximately one new concept detected for every three and a half snapshots. The data were saved to the Firebase Storage and the Firebase Realtime Database during the experiment.

Fig. 2. Examples of snapshots captured by a pair of camera glasses (first column), a wearable video camera (second column), and a public IP camera (third column) for a LifeTags++ user walking in a pedestrian area. Note the differences in image quality and resolution for these video sources, but also the difference with respect to the video recording perspective, from egocentric to exocentric.

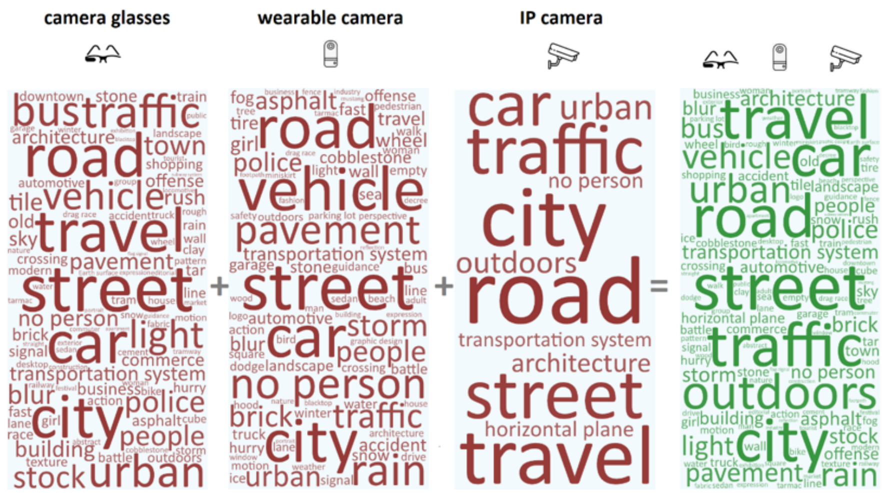

Fig. 3. Synchronized word clouds computed from the recordings of three video cameras according to the scenario exemplified in Figure 2. The word cloud from the right combines all the three perspectives.

4.2. Results

Figure 4, left presents the total number of concepts that we extracted from each video source and corresponding perspective for lifelogging, but also in combinations of two and three, respectively, corresponding to the union operation defined in Eq. 2. Figure 4, right presents the number of concepts that were distinct.

We found that the concepts extracted with a confidence level above 85% intersected to a large extent across all of the three perspectives, *e.g.*, a number of 138 of the 202 concepts extracted in the egocentric back view perspective were also found among the concepts detected with the camera glasses in the eye-level perspective. The number of distinct concepts extracted from the perspective of the IP ambient camera was smaller compared to the other two conditions of our experiment, a result that we explain by the fact that the IP camera was fixed and had a lower resolution. However, a number of 11 out of the 12 distinct concepts that were extracted with a confidence level above 95% from the IP camera perspective were also found among the concepts identified in the two egocentric, eye-level and back view perspectives. Overall, the concepts extracted in the two egocentric conditions had 57 distinct concepts in common out of a total of 99 (for the eye-level perspective of the camera glasses) and 89 (for the back view perspective of the back clip-on video camera), respectively. Had only the camera glasses been used as in the original version of LifeTags [4], a number of 33 distinct concepts would have been lost. Furthermore, by analyzing the lists of distinct concepts extracted in each of the three conditions considered in our experiment, we identified concepts unique for each perspective. This analysis corresponds to the difference operation for tag clouds formalized with Eqs. 3 and 4. For example, the camera glasses positioned at eye level captured snapshots in which concepts such as “sky”, “tile”, “shopping”, “commerce”, “desktop”, “rush”, “texture” were identified. The back view camera enabled detection of other concepts, such as “pedestrian”, “footpath”, “parking lot”, “tire”, “square”, “dodge”, etc., not visible from the front perspective of the eye-level video camera.

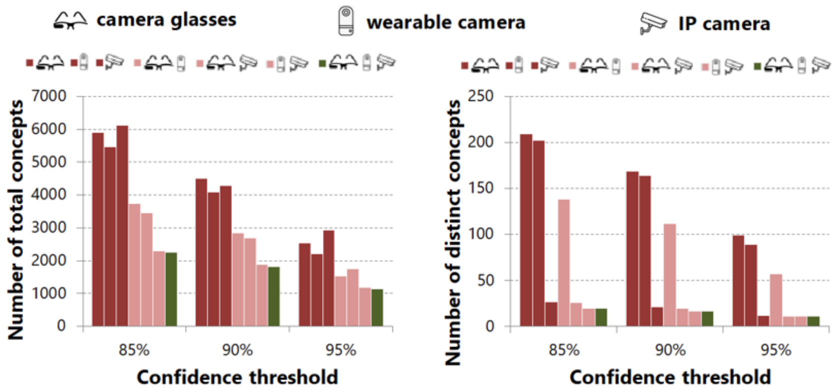

Fig. 4. Total number of concepts (left) and number of distinct concepts (right) extracted from the three video recording devices: camera glasses, wearable camera, and IP camera.

5. Conclusion

We presented LifeTags++, a multi-user, multi-device, and multi-perspective system for recording and abstracting visual life. Also, we presented results from an evaluation study conducted with LifeTags++ that involved three video cameras used to record visual life from three different perspectives, egocentric eye-level, egocentric back view, and exocentric ambient, respectively. Our empirical evaluation showed that incorporating data from multiple video sources into the lifelog leads to richer descriptions in terms of the number of distinct concepts stored in the lifelog. Based on these findings, we suggest several opportunities for future lifelogging systems integrating mobile, wearable, and ambient devices:

• Integration of other perspectives provided by other sensors to be added in the software architecture of LifeTags++. For example, personal drones can capture aerial perspectives by operating in an autonomous, "follow me" mode; see Figure 5 for a snapshot captured with the Mambo Fly drone. As drones become more and more popular and accessible, exocentric perspectives provided by their built-in cameras, but also the contextual information they can collect, *e.g.*, GPS, could be used to enrich the multiperspectivity of lifelog tag clouds;

• Integration of other interactive devices in the software architecture of LifeTags++ towards increased functionality. For instance, AR smartglasses can deliver visual notifications for situations of inattentional blindness, *e.g.*, by highlighting target objects with visual cues. Also, interaction techniques designed to control, activate, and deactivate different perspectives for LifeTags++ as well as switching between various recording perspectives are interesting to investigate as part of future work.

• Security and privacy aspects for lifelogging systems implemented with glasses with built-in video cameras [21, 26], clip-on video cameras, and cameras from the environment should be further examined. These include aspects of data confidentiality, integrity, availability, access control, etc. While the original version of LifeTags [2] was designed with explicit requirements about privacy and security, the extended functionality enabled by LifeTags++ demands revisiting the privacy and security implications of such prototypes for both users and passers-by.

• Examination of LifeTags++ for specific user categories is also an interesting direction for future work, such as for people with various disabilities [31] that could use smart wearables to access information and services in the context of ambient intelligence scenarios [1,28].

Fig. 5. An aerial perspective for LifeTags++ provided by a personal drone. In this scenario, the user is followed by an autonomous drone and videos are captured by the drone's built-in camera.

Acknowledgements. This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI-UEFISCDI, project number PN-III-P3-3.6-H2020-2020-0034 (12/2021). Original versions of the icons used in Figures 1, 3, and 4 were made by Freepik (http://www.freepik.com, "Miscellaneous Elements" pack) from Flaticon (http://www.flaticon.com), and licensed under Creative Commons BY 3.0 (http://creativecommons.org/licenses/by/3.0).

References

[1] ACAMPORA G., COOK D.J., RASHIDI P., VASILAKOS A.V., *A survey on ambient intelligence in healthcare*, Proceedings of the IEEE 2013 **101**(12), pp. 2470–2494, 2013.

[2] AIORDĂCHIOAE A., FURTUNĂ D., VATAVU R.-D., *Aggregating life tags for opportunistic crowd-sensing with mobile and smartglasses users*, Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good, ACM, New York, NY, USA, pp 66–71, 2020.

[3] AIORDĂCHIOAE A., PAMPARĂU C., VATAVU R.-D., *Lifelogging meets alternate and cross-realities: An investigation into broadcasting personal visual realities to remote audiences*, Multimedia Tools and Applications 2021.

[4] AIORDĂCHIOAE A., VATAVU R.-D., *Life-Tags: A smartglasses-based system for recording and abstracting life with tag clouds*, Proceedings of the ACM on Human-Computer Interaction 2019 **3**(EICS), pp. 15:1–15:22, 2019.

[5] AIORDĂCHIOAE A., VATAVU R.-D., POPOVICI D.-M., *A design space for vehicular lifelogging to support creation of digital content in connected cars, Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, ACM, New York, NY, USA, pp. 1–6, 2019.

[6] ALHARBI R., STUMP T., VAFAIE N., PFAMMATTER A., SPRING B., ALSHURAFA N., *I can't be myself: Effects of wearable cameras on the capture of authentic behavior in the wild*, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **2**(3), pp. 90:1–90:40, 2018.

[7] AL MAHMUD A., BRAUN J., MARTENS J.-B., *Designing to capture and share life experiences for persons with aphasia*, Proceedings of the 12th International Conference on Human-Computer Interaction with Mobile Devices and Services, ACM, New York, NY, USA, pp. 391–392, 2010.

[8] CECCHINATO M.E., COX A.L., BIRD J., *Smartwatches: the good, the bad and the ugly?* Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, ACM, New York, NY, USA, pp. 2133–2138, 2015.

[9] CHAMBEL T., KAISER R., NIAMUT O., OOI W.T., REDI J.A., *AltMM 2016: 1th International Workshop on Multimedia Alternate Realities*, Proceedings of the 24th ACM International Conference on Multimedia, ACM, New York, NY, USA, pp. 1481–1482, 2016.

[10] DAWI N.M., JALIL N.A., *Integrated model for smartwatch adoption*, Proceedings of the International Conference on Advanced Information Science and System, ACM, New York, NY, USA, pp. 1–7, 2019.

[11] DODGE M., KITCHIN R., *Outlines of a world coming into existence: Pervasive computing and the ethics of forgetting*, Environment and Planning B: Planning and Design **34** (3), 431–445, 2007.

[12] GANTI R.K., YE F., LEI H., *Mobile crowdsensing: Current state and future challenges*, IEEE Communications Magasinte **49**(11), 32–39, 2011.

[13] GARTNER, *Gartner forecasts global spending on wearable devices to total $81.5 billion in 2021* https://www.gartner.com/en/newsroom/press-releases/2021-01-11-gartner-forecasts-global-spending-on-wearable-devices-to-total-81-5-billion-in-2021, Accessed July 21, 2021.

[14] GURRIN C., SMEATON A.F., DOHERTY A.R., *LifeLogging: Personal big data*, Foundations and Trends in Information Retrieval **8**(1), pp. 1–125, 2014.

[15] HARTNER M., *Multiperspectivity*, in De Gruyter: *Handbook of Narratology*, pp. 353–363, 2014.

[16] ITOH Y., LANGLOTZ T., SUTTON J., PLOPSKI A., *Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays*, ACM Computing Surveys **54**(6), pp. 120:1–120:36, 2021.

[17] KITAMURA K., YAMASAKI T., AIZAWA K., *Food log by analyzing food images*, Proceedings of the 16th ACM International Conference on Multimedia, ACM, New York, NY, USA, pp. 999–1000, 2008.

[18] KOELLE M., KRANZ M., MLLER A., *Don't look at me that way! Understanding user attitudes towards data glasses usage*, Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, ACM, New York, NY, USA, pp. 362–372, 2015.

[19] METSIS V., KOSMOPOULOS D., ATHITSOS V., MAKEDON F., *Non-invasive analysis of sleep patterns via multimodal sensor input*, Personal and Ubiquitous Computing **18**(1), pp. 19–26, 2014. https://doi.org/10.1007/s00779-012-0623-1

[20] MOSCONI P., RADREZZA S., LETTIERI E., SANTORO E., *Use of health apps and wearable devices: Survey among Italian associations for patient advocacy*, JMIR mHealth and uHealth **7**(1), pp. e10242, 2019.

[21] OPASCHI O., VATAVU R.-D., *Uncovering practical security and privacy threats for connected glasses with embedded video cameras*, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **4**(4), pp. 167:1–167:26, 2020.

[22] PAMPARĂU C., VATAVU R.-D., *FlexiSee: Flexible configuration, customization, and control of mediated and augmented vision for users of smart eyewear devices*, Multimed Tools and Applications **80**(20), pp. 30943–30968, 2021.

[23] PARADISO J.A., LANDAY J.A., *Guest editors' introduction: Cross-reality environments*, IEEE Pervasive Computing, 8(3), pp. 14–15, 2009.

[24] SELLEN A.J., WHITTAKER S., *Beyond total capture: A constructive critique of lifelogging*, Communications of the ACM **53**(5), pp. 70–77, 2010.

[25] SHRESTHA P., SAXENA N., *An offensive and defensive exposition of wearable computing*, ACM Computing Surveys **50**(6), pp. 92:1–92:39, 2017.

[26] STEIL J., KOELLE M., HEUTEN W., BOLL S., BULLING A., *PrivacEye: Privacy-preserving head-mounted eye tracking using egocentric scene image and eye movement features*, Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications, ACM, New York, NY, USA, pp. 1–10, 2019.

[27] TERHOEVEN J., SCHIEFELBEIN F.-P., WISCHNIEWSKI S., *User expectations on smart glasses as work assistance in electronics manufacturing*, Procedia CIRP **18**, pp. 1028–1032, 2018.

[28] UNGUREAN O.-C., VATAVU R.-D., *Users with motor impairments' preferences for smart wearables to access and interact with ambient intelligence applications and services*, Proceedings of ISAmI **21** the 12^{th} International Symposium on Ambient Intelligence, Salamanca, Spain, pp. 1–10, 2021.

[29] VATAVU R.-D, *Connecting research from assistive vision and smart eyewear computing with crisis management and mitigation systems: A position paper*, Romanian Journal of Information Science and Technology **23**(S), pp. S29–S39, 2020.

[30] VATAVU R.-D., BILIUS L.-B., *GestuRING: A web-based tool for designing gesture input with rings, ring-like, and ring-ready devices*, Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology, ACM, New York, NY, USA, pp. 710–723, 2021.

[31] WATSON C., KIRKHAM R., KHARRUFA A., *PIP Kit: An exploratory investigation into using lifelogging to support disability benefit claimants*, Proceedings of the CHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, pp. 1–14, 2020.

[32] ZHAO Y., SZPIRO S., SHI L., AZENKOT S., *Designing and evaluating a customizable head-mounted vision enhancement system for people with low vision*, ACM Transactions on Accessible Computing **12**(4), pp. 15:1–15:46, 2019.

[33] ZINI F., REINSTADLER M., RICCI F., *Life-logs aggregation for quality of life monitoring*, Proceedings of the 5th International Conference on Digital Health, ACM, New York, NY, USA, pp. 131–132, 2015.