# Indexing and Retrieving Voice Recordings by Instantly Tagging Mentioned Objects with Dots

van Bergen, Thibaut; Ishiyama, Rui; Makino, Kengo; Kudo, Yuta; Takahashi, Toru; Goosen, Hans

**DOI**
10.1109/WF-IoT.2019.8767339

**Publication date**
2019

**Document Version**
Accepted author manuscript

**Published in**
Proceedings of the IEEE 5th World Forum on Internet of Things (WF-IoT)

# Indexing and Retrieving Voice Recordings by Instantly Tagging Mentioned Objects with Dots

Thibaut van Bergen
Mechanical Engineering
Delft University of Technology
Delft, the Netherlands
T.P.M.vanBergen@student.tudelft.nl

Yuta Kudo
Data Science Research Laboratories
NEC Corporation
Kawasaki, Japan
y-kudo@bc.jp.nec.com

Rui Ishiyama
Data Science Research Laboratories
NEC Corporation
Kawasaki, Japan
r-ishiyama@bl.jp.nec.com

Toru Takahashi
Data Science Research Laboratories
NEC Corporation
Kawasaki, Japan
t-takahashi@hg.jp.nec.com

Kengo Makino
Data Science Research Laboratories
NEC Corporation
Kawasaki, Japan
k-makino@mb.jp.nec.com

Hans Goosen
Precision and Microsystems Engineering
Delft University of Technology
Delft, the Netherlands
j.f.l.goosen@tudelft.nl

***Abstract*—This paper presents a novel framework and its prototype tool for indexing and retrieving specific fragments of voice recordings obtained during discussions about physical objects such as text documents, pictures, or 3D models. When a specific part of an object is mentioned, it is tagged with an ink dot that is immediately registered in a database by capturing a microscopic image of the dot. Simultaneously, an index of the recording fragment is created and linked with the dot. After the recording, a dot can be scanned and identified by matching its microscopic image with the database to retrieve the linked recording fragment for playback. A handy tool was developed to facilitate these operations while the user concentrates on the ongoing discussion. Performance tests of the dot identification have shown genuine matches without error. In demonstrations of a realistic usage scenario, the tool successfully facilitated the creation of indexes with dots during a voice recording and correctly played back all the specific recording fragments linked to the dots.**

***Keywords—voice recording, indexing, retrieval, playback, pattern recognition, identification, image matching***

## I. INTRODUCTION

Meetings and interviews play an essential role in collecting and creating knowledge in business, research, and academia. Making recordings during the discussions is useful for memorising and organizing the great amount of information that is exchanged therein. Voice recorders are a popular tool for this purpose, as witnessed by the variety of consumer products available on the market. However, playing back the recordings for references can be troublesome. Particularly, retrieving a specific fragment from a long recording is a time-consuming and frustrating task. Therefore, a practical and useful framework for detailed indexing and targeted retrieval of recordings for playback is of great interest.

Video and audio indexing and retrieval has been a major research topic in multimedia information processing [1]. The most common retrieval method is to use automatic speech-recognition techniques [2][3] to produce a textual transcript, which can then be used to create a searchable index. These methods work well, especially for lectures and presentations [4] for which the topics or content are well-organized with preliminarily prepared related keywords. However, their performance is limited when exposed to creative or investigative speech, because of the out-of-vocabulary (OOV) problem [5] or the use of abstract or contextual words with non-verbal expressions. This is problematic, as non-verbal or contextual information can be critical indexes for retrieving relevant fragments. Specific parts of physical objects can be such indexes. Indeed, many meetings are conducted while showing objects, e.g. paper documents, maps, pictures, or prototyped three-dimensional (3D) models. Opinions, recommendations, or new ideas are given on the specific parts of the objects. In such cases, verbal pointers such as "here" or "this" accompanied by a gesture are more straightforward than cumbersome spoken descriptions. If such physical attributes can be immediately linked with the live on-going speech or video, and recordings can be played back simply by pointing to the part of the object as an index, a great amount of time and resources can be saved when proceedings need to be documented.

Therefore, we propose a novel framework in which specific parts of objects are marked with tags when each of these parts are mentioned during a recorded discussion. Any tag can later be used as an index to retrieve the relevant fragment of the recording. Similar concepts have been proposed as “paper-digital systems,” such as Anoto tools [6]; however such a system works only on specially printed paper in which tags are embedded beforehand. In contrast, our proposed concept is to create new labels on demand onto any standard paper, possibly any object, including 3D surfaces. The proposed framework rests on the following two key principles: (1) a basic method to create identifiable tags on demand and to link them to digital data, and (2) a useful tool to facilitate these processes without interrupting its user and speakers during recorded discussions.

The micro-sized Identifier Dot on Things (mIDoT) technology proposed in [7] is useful as the basic method. Previous studies offer the core algorithm and an implementation as a tool for industrial parts traceability in factory use [8]. However, no practical tool has been proposed for implementing our framework yet. Now, we have developed a functional prototype of such a tool. With simple but effective mechanisms using inexpensive hardware, a specific part of an object can be instantaneously tagged with a dot, which is then captured in a microscopic image and saved in the database for identification, all while carrying on a recorded discussion. After the voice recording ends, the fragment of the recording that mention a dotted part of the object can be easily retrieved by scanning any chosen dot.

IEEE 5th World Forum on Internet of Things, pp. 183—188, 2019

The dotting and capturing (i.e. retrieving) operation is accomplished with a simple one-handed action, enabled by the tool's design and mechanisms.

This paper is organized as follows. In section II, we review the mIDoT technology [7] and describe the requirements for the tool enabling its application to our framework. Section III explains the prototype's mechanisms and its operation. In section IV, we review the image-matching algorithm to identify the dots. Section V gives the experimental results. Finally, conclusions and future work are discussed in section VI.

## II. Functional Requirements for the Tool

### A. Micro-sized Identifier Dot on Things(mIDoT)

The mIDoT technology [7] provides instant tagging for identifying individual components and retrieving the data linked to them. An overview of the technology that is applied for our framework is given in Fig. 1. Tiny and unique ID tags are created on an object by marking it with a dot using an ink pen and by capturing the dot image with a microscope; the dot image serves as a "fingerprint" for identification. Tiny particles contained in the ink naturally form a unique pattern, enabling each dot to be identified by matching their images. In other words, when a dot is scanned, the pattern recognition algorithm extracts the image features, matches the dot with registered dots in the database to find the corresponding one, and retrieves the data linked to the dot.

The benefit of the mIDoT technology is that dots are easy to create; lightly touching an object with the pen's tip suffices. The minuscule size (approximately 2 mm) of the dots makes it possible to physically place tags on any objects, small to large. Further, every dot naturally creates a unique identifier on demand. This technology's implementation remains low-cost due to the use of an off-the-shelf $2 ink pen and an inexpensive microscope camera. The mIDoT has already shown it can be implemented successfully for identification of more than 10,000 tiny electric parts [8].

### B. Procedures for the Tagging Framework

For a successful practical implementation of the framework for indexing and retrieval using the mIDoT technology, the practical tool must facilitate first the indexing procedures, then the retrieval procedures. The fundamental steps for each procedure are given below.

*1) Indexing.* (1) When a new topic of discussion is initiated, choose the corresponding part of the mentioned object to be tagged, (2) tag the part with a dot, (3) align the microscope's centre of view with the dot while letting the ink of the dot dry (approximately 0.5 second), (4) capture an image of the dot (register), (5) extract the image features, (6) register it in the database, and repeat these steps for any additional index until the end of the recording.

*2) Retrieval:* (1) Choose the dot corresponding to the topic of interest, (2) align the microscope's centre of view with the dot, (3) capture an image of the dot (query), (4) extract the image features (5), identify the dot with the image-matching algorithm, (6) extract the data (in this case, fragment of the recording linked to the dot), and repeat these steps as many times as necessary.

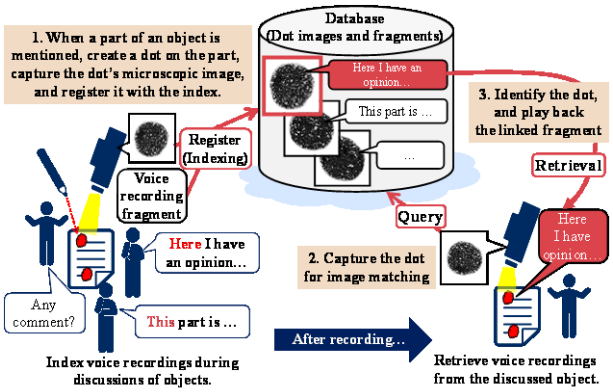



Fig. 1. Overview of our framework for indexing and retrieving voice recording using the mIDoT technology.

### C. Requirements for the Tool

In addition to the purely functional requirements of the tool to use the mIDoT technology for our framework, we determined a set of qualitative requirements in order to ensure our framework's practical usefulness.

*1) Easy one-handed operation.* We assume that the tool is used by a member of the discussion, an interviewer or a facilitator. Also, the user of the tool should not be a specialist; the user as well as the others, must be able to concentrate on the discussion while performing the indexing task. The difficulty is in capturing a tiny (approximately 2 mm) dot by using a microscope camera with a narrow field of view (approximately 5 mm), where precise alignment of the dot and microscope is crucial. The prototype should simplify this task such that is becomes a one-handed operation.

*2) Compactness for mobile use.* Often, meetings are held in different places in or even out of an office. A compact and light design makes it easier to carry the tool around between meetings and improves user-friendliness during its use.

*3) Low cost and easy maintenance.* Using inexpensive hardware makes the tool accessible to a wider public. Also, by incorporating the hardware in the assembly such that mounting and unmounting is quick and straightforward, maintenance of the tool remains simple.

*4) Robust and reliable performance.* The identification of the dots should yield one of two results; a genuine match such that only the correct data is retrieved, or no match at all when no data can be retrieved due to failing to find a match. Most importantly, there should be no instance in which incorrect data is retrieved. In other words, true positive, true negative and false negative matches are acceptable, but false positives are not.

## III. The Prototype: Components, Mechanisms, and Operation

### A. Hardware and Software

For the development of the prototype tool, the previously outlined requirements served as guidelines for the design of the hardware and software.

*1) Hardware.* The prototype relies on a $2 off-the-shelf ink pen (DryAce, Teranishi Chemical Industry, Co. Ltd.) and a pen-shaped USB digital microscope camera (A1-Microscope, MixMart Ltd.). Currently, the recording and speaker functionalities are fulfilled by a separated device (PHS002W, Jabra). However, ideally, these functionalities should all be embedded into the prototype hardware. Circuitry containing a controller board (Nano ATmega328, Arduino) and two Hall-effect sensors (US1881) in combination with magnets (4–2mm) are used for automation. Additionally, generic 3mm LEDs give feedback to the user to communicate the tool's status. All of these components are incorporated in a 3D-printed custom-made body, designed to be held in and operated with one hand.

*2) Software.* The main software for the tool is in a PC's Python program, which operates as the decision maker. That is, the program controls the microscope camera, recording (A/D conversion, data storage, and fragmentation with indexing), speaker, image-matching engine, controller board, and the LEDs. The controller board's Arduino program communicates the sensor values to the PC. With the help of the sensor data, the PC determines the action to undertake (e.g. capture an image) on the basis of a state-machine implementation. The communication between the camera, controller board, microphone-speaker system, and PC is established through USB cables. The image-matching engine developed in [9] was used for identification of the dots instead of the algorithm used in [7] and [8]. This engine provides rotationally invariant pattern matching processing at a high speed of more than 1,000 matching pairs within a second, running on a standard desktop PC.

### B. Mechanisms

The indexing procedures would be cumbersome if the pen and microscope were to be handled manually and separately. Finding a way to facilitate this task without disturbing an ongoing discussion is the key to implementing our proposed framework. The prototype tool was given the following four essential mechanisms for that purpose: (1) a sliding mechanism for consecutive dotting and capturing controlled with a finger grip, (2) built-in microscopic alignment of the pen and camera, (3) an automatic trigger for image capture after dotting, and (4) a push-button switch for control.

*1) Dotting and capturing.* To accommodate the consecutive dotting-and-capturing action, the pen and microscope are held by sliders. These sliders have pins which are guided along a pair of curved slots on the body of the prototype, as shown in Fig. 2. As such, alternating between the dotting and capturing configurations can be achieved by reciprocating the sliding movement of the two. This is done using a gear between two sliding racks; any vertical movement of one rack leads to the opposite movement of the other. In addition to the curved slots, the sliders are also slotted into the racks such that they follow the vertical movement of their respective rack. By adding a finger grip on one rack and attaching the other to the base with a pull spring, the pen and microscope can be moved reciprocatively along their respective slots and naturally return to the original configuration (Fig. 3). This mechanism makes it possible to accomplish the dotting-and-capturing procedure in one swift push and release of the finger grip.

*2) Microscopic alignment.* The pen and camera are precisely aligned by mounting identical conical caps on them, which precisely fit in a corresponding conical slot. As such, the microscope's centre of view and the contact point of the pen on the target precisely coincide when subsequently pressed down on the target (Fig. 4).

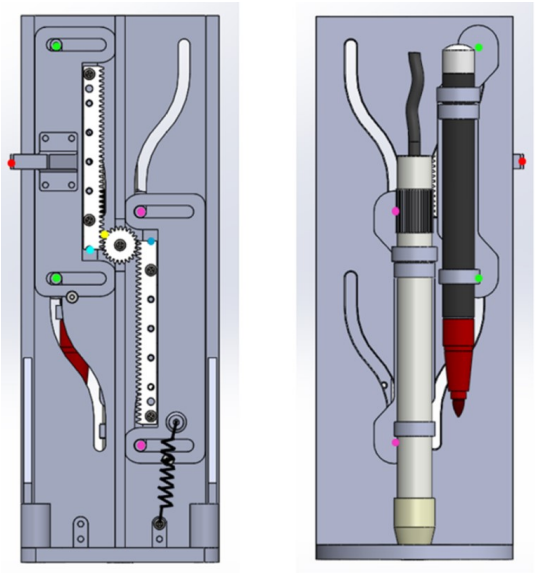
*3) Trigger for image capture.* The trigger mechanism uses two sets each containing one Hall effect sensor paired with a magnet. The sensors are fixed to the base, whereas the magnets are attached to the moving racks. These sensors and magnets are placed such that the pen or the microscope can be detected when either of them reaches the bottom position, corresponding to either the dotting or capturing configurations (Fig. 5). The state machine in the PC program tracks the sensory information and only triggers a capture once the dotting and capturing configurations are consecutively detected. This sequence corresponds to pushing the finger grip until it reaches the bottom and then releasing it back to its rest position. Only then is an image and timestamp registered into the PC database.

*4) Push-button switch.* A push-button switch is included in the design for switching from the indexing mode to the retrieval mode in which dots are captured to initiate playbacks. The switch is placed on the side of the tool below the finger grip for easy operation with a single finger.

### C. Operation

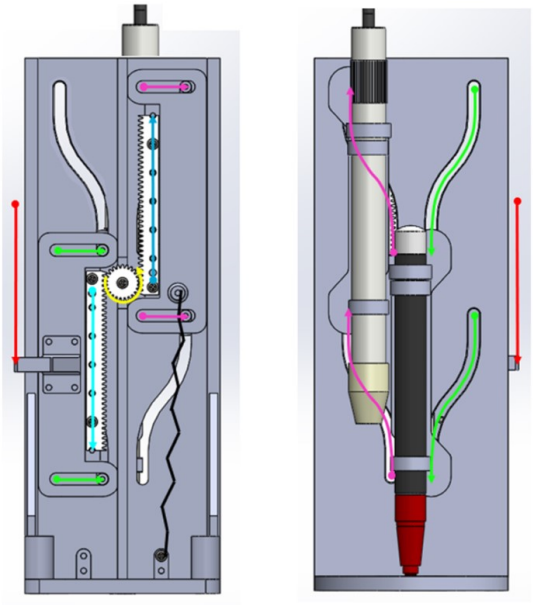
*1) Indexing: Tagging recording fragments onto physical objects.* When a specific topic is introduced, the user marks the corresponding part of the target object with a single push on the finger grip. Releasing the grip lowers the microscope back into the capture position, which automatically triggers an image registration. The user must hold the prototype still until the LEDs flash, indicating a successful capture, after which the image along with a timestamp is saved in the data base. The whole process is quick (approximately 1 sec) and easy to complete; thus, it does not disturb or interrupt the conversation or speech. Afterward, the recording can be ended by a single push of the push-button switch, which triggers the fragmentation of the recording on the basis of the timestamps. Each fragment is then linked to its corresponding dot image.

*2) Retrieval: Dot identification and recording playback.* Pressing on the push-button switch ends the recording and switches the prototype to retrieval mode. When a dot is selected, the prototype is placed such that the dot is fully inside the microscope's field of view. Then, pushing the push-button switch triggers a capture and initiates the image-recognition algorithm. When a match is found in the database for the query image, the linked recording fragment is played back. Otherwise, the LEDs give a signal when the identification has failed.



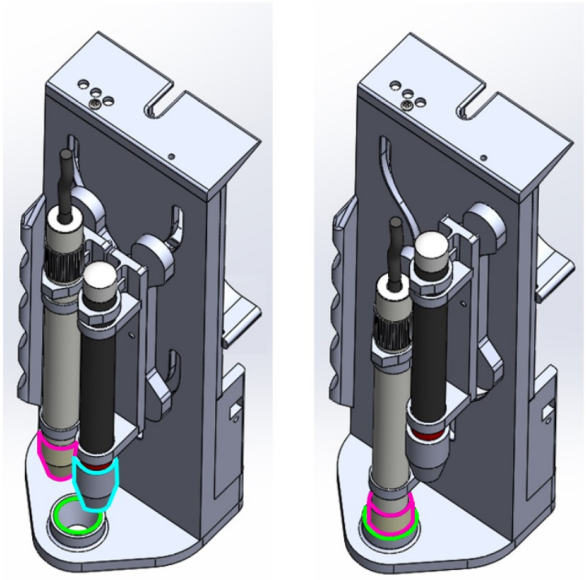
(a) (b)

Fig. 2. Back (left) and front (right) view of dotting and capturing mechanism. Colored dots indicate rest positions of moving components. (a) The button (red) is fixed on a rack (cyan), which is connected to a gear (yellow) such that the other rack (blue) reciprocates the vertical movement of the first one. (b) Pins guide the pen (green) and microscope (magenta) through the curved slots, while also being slotted to the racks in order to follow their vertical movement.



(a) (b)

Fig. 3. Back (left) and front (right) view of dotting-and-capturing mechanism. Colored arrows indicate movement of crucial components when actuated. (a) Pushing the finger grip (red) actuates the racks reciprocately (cyan and blue) such that (b) the pen (green) and microscope (magenta) alternate positions while moving up and down the slots.



(a) (b)

Fig. 4. Mechanism for accurate alignment of dot and microscope's field of view. Identical conical caps are mounted on the pen (cyan)

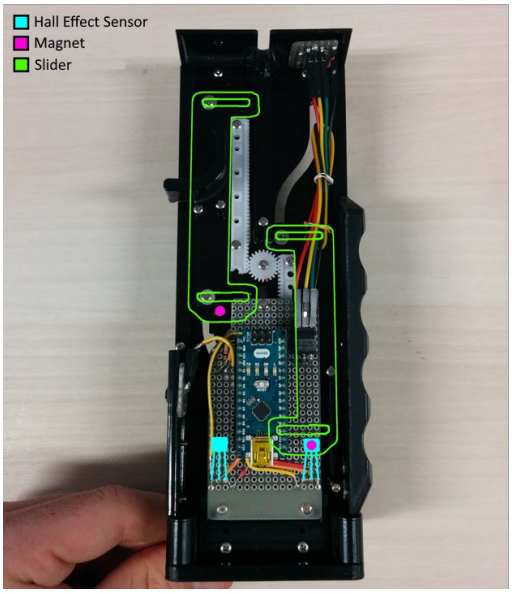



Fig. 5. Trigger mechanism for image capture. Hall effect sensors (cyan) detect the presence of magnets (magenta) attached to the pen- or microscope-slider (green). For example, if the pen reaches the bottom, i.e. if dotting occurs, the sensor sends a signal to the CPU to indicate a dotting action.

## IV. Image-Matching Algorithm

The tool described in Section III facilitates holding the microscope such that it is perpendicular to the target object's surface, with the created dot centered in its field of view. Thus, the database and query images are captured under a rigid transformation including any planar rotation and small amounts of translational and scale changes. Unlike what was done in [7] and [8], we use the Fourier-Mellin band-limited phase (FMBLP) matching algorithm proposed in [9]. This algorithm offers fast and accurate image-matching with robustness against rigid transformations. Here, we review the image-matching algorithm

### A. Geometric Invariant Feature Extraction

First, an input image $f(n_1,n_2)$ is converted into a feature (2D complex array) using the Fourier-Mellin transform (FMT), which gives rotation, scale, and translational invariant features. Let $F(k_1,k_2)$ denote the 2D discrete Fourier transform (DFT) of the input image. The FMT is implemented by applying the log-polar transform (LPT) to the amplitude spectrum $|F(k_1,k_2)|$. The algorithm uses $\log_{10}\{|F(k_1,k_2)|+1\}$ instead of $|F(k_1,k_2)|$. The FMT transforms changes in rotation and scale in an image into translational shifts of the resultant image. The feature we obtain by applying the LPT to the amplitude spectrum $\log_{10}\{|F(k_1,k_2)|+1\}$ is denoted by $F_{LP}(k_1,k_2)$.

Next, the 2D DFT is applied to $F_{LP}(k_1,k_2)$, and its amplitude spectra is normalized as follows:

$$V_f(l_1,l_2) = \frac{\mathcal{F}(F_{LP}(k_1,k_2)}{|\mathcal{F}(F_{LP}(k_1,k_2)|} = e^{j\theta_f(l_1,l_2)}, \quad (1)$$

where $\mathcal{F}(\cdot)$ means the 2D DFT operation and $e^{j\theta_f(l_1,l_2)}$ is a phase component of the FMT image $F_{LP}(k_1,k_2)$.

Finally, only the low frequency bands are selected from the $V_f(l_1,l_2)$. The selected bands of the FMP feature are called the FMBLP feature.

### B. Feature Matching

The correlation value between the query and the database FMBLP features is calculated as the matching score. Let $V_{db}(l_1,l_2)$ denote an FMBLP feature registered on a database and let $V_q(l_1,l_2)$ denote a query FMBLP feature. The cross-power spectrum of these features is described as

$$R(l_1,l_2) = V_{db}(l_1,l_2)\overline{V_q(l_1,l_2)} = e^{j\theta_{db}(l_1,l_2)}e^{-j\theta_q(l_1,l_2)}, \quad (2)$$

where $\overline{V_q(l_1,l_2)}$ means a complex conjugate of the query FMBLP feature $V_q(l_1,l_2)$. The correlation map $r(k_1,k_2)$ between the registered FMBLP feature $V_{db}(l_1,l_2)$ and the query FMBLP feature $V_q(l_1,l_2)$ is given as follows:

$$r(k_1,k_2) = \mathcal{F}^{-1}\{R(l_1,l_2)\} \quad (3)$$

where $\mathcal{F}^{-1}\{\cdot\}$ means 2D inverse DFT. We employ the peak value of the correlation map $r(k_1,k_2)$ as the matching score $s$ between the registered FMBLP feature $V_{db}(l_1,l_2)$ and the query FMBLP feature $V_q(l_1,l_2)$:

$$s = max\{r(k_1,k_2)\}. \quad (4)$$

Because each dot's microscopic image has unique features, the shape of the correlation map $r(k_1,k_2)$ has a sharp peak similar to a delta function if the database FMBLP feature $V_{db}(l_1,l_2)$ and the query FMBLP feature $V_q(l_1,l_2)$ are captured from the same dot. Otherwise (i.e. if they are captured from different dots), the correlation map $r(k_1,k_2)$ does not have a sharp peak. Therefore, we can identify individual dots by using the peak value of the correlation map $r(k_1,k_2)$ as a measure of similarity.

## V. Experiments

We tested the prototype in two experiments. One tested test the accuracy of indexing and retrieval, i.e. tagging and identification of dots. The other tested the usefulness of the tool in a real-life usage scenario: an interview of three people using a map for collecting new information about local restaurants.

### A. Indexing and Retrieval Performance

The most fundamental functionality test of the prototype examines its accuracy in tagging and identifying dots. Fig. 6 shows the setup for the test, in which a sheet with a grid of 100 numbered boxes is used. First, a dot was marked and registered for each of the boxes, resulting in 100 numbered dots on the sheet and 100 registered images in the database (shown in Fig. 7) with corresponding index numbers. Next, the prototype was set to retrieval mode. For each dot a query image was captured and then processed by the image matching algorithm to retrieve the dot index. The query images are also shown in Fig. 7. In this experiment, all 100 dots were successfully identified, i.e. the test yielded only genuine matches. Thus, correct indexes were retrieved.

Fig. 6. Experimental setup to test accuracy of dot tagging and identification.

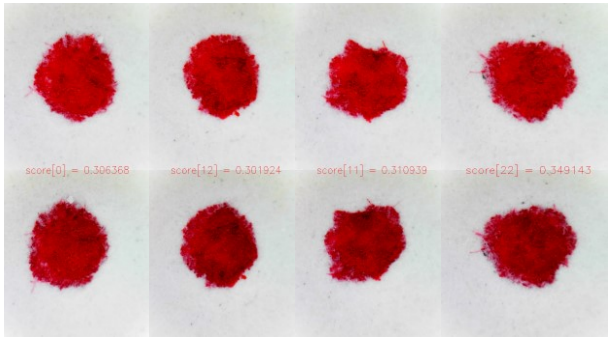



Fig. 7. Registered images (top row) and query images (bottom row) . Each column shows different images of an identical dot with corresponding match score. Note that the orientation of the images is not controlled by the user, but the algorithm is robust against rotations and still identifies the dots.

### B. *Demonstration of Possible Usage Scenario*

The prototype was also tested by enacting a practical usage scenario, as shown in Fig. 8. The chosen scenario was interviewing three people using a map to collect information about local restaurants. During the interview, the prototype was used to mark locations of mentioned restaurants on the map, for which the spoken recommendations were recorded. This scenario is a symbolic example of typical business tasks and designed to evaluate the prototype's previously discussed requirements, i.e. mobility, performance and ease of operation.

The prototype and a map of the area surrounding our office were brought out of our office for a meeting with the three participants. The interviews asked the participants about the names, locations and highlights of their favourite restaurants. Throughout the demonstration, our prototype fulfilled all the discussed requirements. The interviewer successfully marked the mentioned restaurants' locations on the map with dots without disturbing the interviewees. After finishing the meeting, each of the recorded recommendations were easily and successfully played back by capturing their respective dots on the map.

### C. *Discussion*

As of now, the prototype is not a standalone tool and relies on a PC, which limits the tool's mobility. Also, the USB cables connecting the prototype to the computer can be cumbersome during operation; sometimes, the cables obstructed the interviewer when attempting to make a dot. Embedding the CPU into the tool body would be an effective way to tackle this issue.

Additionally, the tip of the pen tends to dry when no dot has been created for a long period of time (beyond ten minutes). Consequently, the dot quality can deteriorate for long recording fragments, which can be problematic for the identification performance. However, keeping the pen pressed down on the target for longer than a second while making a dot temporarily recover the dotting ability.

## VI. Conclusion

We proposed a new framework for indexing and retrieving voice recordings of discussions, for which a new handy tool was developed as a proof of concept. The tool is novel in that it physically tags the mentioned part of the object during a recorded meeting. Tiny ink dots, called mIDoTs, are used as unique tags that can be instantly created and identified later to retrieve specific fragments of voice recordings. The handy and inexpensive tool was developed to conduct these operations with easy single-handed operation. During the experiments, the tool accurately identified all the dots. Further, the demonstration test has shown the tool's usefulness in a real-life usage scenario, suggesting this tool's usage could be extended to many other scenarios. The proposed prototype only represents the first iteration in the development of the tool, so it has room for improvement despite its satisfactory experimental results.

Further developments should consider embedding the CPU into the prototype body or working with a smartphone to obtain better mobility. The identification algorithm can already match more than 1,000 registered dots with a standard desktop PC's CPU within a second; consequently, a smartphone or a single-board PC suffices to run the algorithm because only a small fraction of that number of dots (i.e. indexes) is required. A more compact body for the tool would enable further applications, such as marking 3D surfaces with dots.

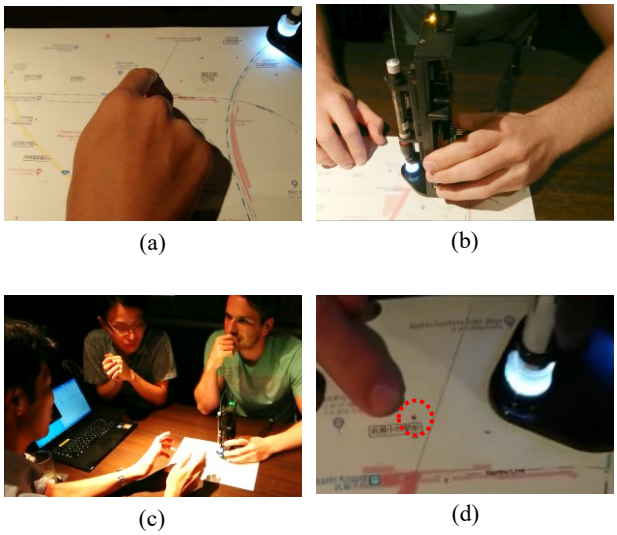



Fig. 8. Demonstration of usage scenario: discussing restaurants with a city map. (a) Interviewee points to the restaurant location and gives his recommendation. (b) Interviewer can instantly mark and register a dot at the indicated location without interrupting. (c) The natural course of conversation is maintained while being recorded and indexed with the single-hand operation. (d) The dot are captured as indexing to playback the corresponding speech.

## References

[1] G. Lu, "Indexing and Retrieval of Audio: A Survey," Multimedia Tools and Applications, 15, 269–290, 2001.

[2] A. Mathur et al. "On Robustness of Cloud Speech APIs: An Early Characterization," ACM Int. Joint Conf. and Int. Symp. on Pervasive and Ubiquitous Computing and Wearable Computers (UbiComp '18), pp. 1409-1413, 2018.

[3] V. Këpuska and G. Bohouta, "Comparing speech recognition systems (Microsoft API, Google API and CMU Sphinx)." Journal of Engineering Research and Application 7.3, pp. 20-24, 2017.

[4] W. Hurst, "Indexing, Searching, and Skimming of Multimedia Documents Containing Recorded Lectures and Live Presentations," ACM Int. Conf. on Multimedia (ACM MM'03), pp. 450-451, 2003.

[5] I. Bazzi, "Modelling Out-of-Vocabulary Words for Robust Speech Recognition" (Doctoral Dissertation). Retrieved from https://dspace.mit.edu/handle/1721.1/29241. 2002.

[6] B. Signer et al. "Advanced authoring of paper-digital systems." Multimedia Tools and Applications 70.2, pp. 1309-1332, 2014.

[7] R. Ishiyama, Y. Kudo, and T. Takahashi, "mIDoT: micro Identifier Dot on Things - An efficient alternative to barcodes, tags or serial marking for industrial parts traceability-," Proc. IEEE Int. Conf. on Industrial Technologies (ICIT2016), pp. 781--786, 2016.

[8] Y. Kudo, H. Zwaan, T. Takahashi, R. Ishiyama, and P. Jonker, "Tip-on-a-chip: Automatic Dotting with Glitter Ink Pen for Individual Identification of Tiny Parts," ACM Int. Conf. on Multimedia Systems (MMSys'18), pp. 502-505, 2018.

[9] R. Ishiyama, T. Takahashi, K. Makino, and Y. Kudo, "Fast Image Matching Based on Fourier-Mellin Phase Correlation for Tag-Less Identification of Mass-Produced Parts," IEEE Global Conference on Signal and Information Processing (GlobalSIP2018), pp. 380-384, 2018.