HAL
open science

A Real-Time Streaming and Detection System for Bio-acoustic Ecological Studies after the Fukushima Accident

Hill Hiroki Kobayashi, Hiromi Kudo, Hervé Glotin, Vincent Roger, Marion Poupard, Daisuké Shimotoku, Akio Fujiwara, Kazuhiko Nakamura, Kaoru Saito, Kaoru Sezaki

To cite this version:

Hill Hiroki Kobayashi, Hiromi Kudo, Hervé Glotin, Vincent Roger, Marion Poupard, et al.. A Real-Time Streaming and Detection System for Bio-acoustic Ecological Studies after the Fukushima Accident. Multimedia Tools and Applications for Environmental & Biodiversity Informatics, 2018. hal-01879592

HAL Id: hal-01879592

https://hal.science/hal-01879592

Submitted on 24 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Real-Time Streaming and Detection System for Bio-acoustic Ecological Studies after the Fukushima Accident

Hill Hiroki Kobayashi, Hiromi Kudo, Hervé Glotin, Vincent Roger, Marion Poupard, Daisuké Shimotoku, Akio Fujiwara, Kazuhiko Nakamura, Kaoru Saito and Kaoru Sezaki

Abstract Acoustic ecology data have been used for a broad range of soundscape investigations. Counting sounds in a given soundscape is considered an effective method in ecology studies that provides comparative data for evaluating the impact of human community on the environment. In 2016, Kobayashi and Kudo collected a particularly valuable dataset containing recordings from within the exclusion (i.e., difficult-to-return-to) zone located 10 km from the Fukushima Daiichi Nuclear Power Plant in the Omaru District (Namie, Fukushima, Japan). These audio samples were continuously transmitted as a live stream of sound data from an unmanned remote sensing station in the area. In 2016, the first portion of their collected audio samples covering the transmitted sound recordings from the station was made available. Such data cover the bioacoustics in the area. This paper describes the methodologies by which we processed these recordings, in extreme conditions, as preliminary eco-acoustic indexes for demonstrating possible correlations between biodiversity variation and preexisting radioecology observations. The variations in some of these vocalizations were also studied.

1 Introduction

According to a report describing the Chernobyl nuclear disaster penned and published by the International Atomic Energy [1], it is academically and socially important to conduct ecological studies focused on ascertaining the levels and effects radiation

Hill Hiroki Kobayashi, Hiromi Kudo, Daisuké Shimotoku, Akio Fujiwara, Kazuhiko Nakamura, Kaoru Saito and Kaoru Sezaki
CSIS, GSFS, GSALS, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan, e-mail: kobayashi@csis.u-tokyo.ac.jp

Hervé Glotin, Vincent Roger and Marion Poupard
AMU, University of Toulon, UMR CNRS LIS, 83957 La Garde, France, e-mail: herve.glotin@univ-tln.fr

exposure has had on wild animal populations over several generations. Although numerous studies and investigations have been conducted regarding the Chernobyl nuclear disaster, there were very few captured audio samples available. In 2012, over 25 years since the Chernobyl disaster occurred, Cusack published audio recordings captured from within the exclusion zone in the Ukraine [2]. To understand the impact a nuclear disaster or other such catastrophic event has on wildlife, we first need long-term and wide-range monitoring of the effects nuclear radiation has on animals because there is little evidence of the direct effects of radioactivity on wildlife in Fukushima [3].

Immediately following the Fukushima Daiichi Nuclear Power Plant disaster, Ishida, a research collaborator at the University of Tokyo, started conducting regular ecological studies of wild animals in the northern Abukuma Mountains where high levels of radiation were detected. In [3], Ishida noted that it is essential to place automatic recording devices (e.g., portable digital recorders) at over 500 locations to properly collect and analyze the vocalizations of the target wild animals . To monitor such species, an effective method is for experts to count the recorded voices of animals; here, acoustic communication is used by various animals, including mammals, birds, amphibians, fish, and insects [4], thus a broad range of species may be covered using this technique. This audial method, in conjunction with visual counts, is commonly used to investigate the habitat of birds and amphibians [4]. It is often surprisingly difficult to analyze this recorded data, which requires observers to manually count and list identified species via repeated playbacks.

Given the intensity of such work, it is also very difficult to continue these activities for long periods of time. Therefore, in this study, we aim to establish a long-term continuously operating ubiquitous system that delivers and analyzes, in real time, environmental information, including bio-acoustic information, for citizen scientists. More specifically, in this paper, we discuss the development and evaluation of an implementation of this system used as part of a bio-acoustic ecology study of the Fukushima accident. Based on related work, we first developed a real-time streaming and identification system for this study, then designed a new experimental human-computation system based on related studies and methodologies. We discuss the methodologies we use to process these recordings as ecoacoustic indexes that demonstrate the variations in biodiversity. Note that while this study is not intended to provide scientific insight into the Fukushima accident, it does provide a comparable dataset and multimedia system for further bio-acoustic studies in the area.

2 Background

As introduced above, in ecology studies, it is often desirable to develop a multimedia system that most effectively supports a study with minimal resources. Recently, the use of the Internet and cloud-sensing has engaged citizen scientists [5], i.e., members of the public that are motivated to work together with professionals interested in science-based conservation to expand the shared knowledge base and explore large-

scale solutions. While this has the potential to move science in a good direction, it is difficult to rely largely on citizen participation alone. More specifically, through the cooperation of citizen scientists, activities were also conducted to analyze data actually recorded in restricted area (not exclusion zone), but there is a problem of continuity [6]. Instead, of asking participants to work directly, those that use the knowledge that the participants entered by accident by keeping the system running all the time. We suggest here that this could solve the continuity problem, and we have conducted the research described herein to realize this.

Developing a system that can be operated for long periods of time

Ecological studies of the environment near urban areas are now being conducted using cell phones [7]; however, it is difficult to use such information devices within an exclusion zone because these areas do not tend to have reliable infrastructure and can be dangerous areas for individuals to enter. Therefore, we conclude that it is necessary to develop a remotely controlled monitoring and evaluation system capable of operating for multiple years to ensure long-term stability under unmanned operating conditions. Note that we previously researched and developed proprietary systems that deliver and record remote environmental sounds in real time for ecology studies [8]. Our previous system was continuously operational on Iriomote Island (Okinawa, Japan) from 1996 to 2010. To date, the fundamental research performed at Iriomote Island has expanded into the Cyberforest project that we are conducting at the University of Tokyo [9].

Observing user behavior via our developed system

We have a record of casual comments and analysis from over 2,000 users who wrote during experiments we conducted from 1996 to 2010 [8]. Among the conditions that keep environmental sounds running in real time, it was revealed that the user most consciously was the animal's voice. This implies that if we continue to stream real-time environmental sounds to users who are interested in environmental issues, these users will share the names of bark animals with others. Moreover, despite being in situations in which users do not know when an animal will make any noise, users continue to listen carefully until an animal makes sound, then carefully report it. Note that in most of these cases, we are not asking the users to do anything specific; they are taking these actions on their own accord. Given these behaviors, we felt that if we could evaluate these activities performed by citizen scientists, we could solve the aforementioned continuity problem.

Developing an interface that makes it easy to obtain comments from users

The development of a wearable forest [10] and a tele echo tube [10] was meant as a work of art that demonstrates its effect. Here we placed a speaker next to a microphone already installed in the forest, then observed the reaction by adding cue sounds to environmental sounds and user actions. From the exhibition of this artwork, it became clear that the concentration of the user's sound is high. Based on these findings, we also conducted research on platform for Citizen Science [5]. More specifically, we are working with ecology scientists to develop a new type of bird

census method, i.e., an audio census, that uses our live streaming audio and social media systems (e.g., Internet Relay Chat and Twitter) [9]. When ornithologists in separate locations used our developed sound system to remotely conduct a woodland birds census with the cue sounds, more species were identified than from a field-based spot census (i.e., 36 identifications versus only 28, respectively). Given the above issues, we summarize our goals via the problem statements listed below.

1. Social Problem: It is academically and socially important to conduct ecological studies focused on ascertaining the levels and effects radiation exposure has had on wild animal populations over several generations in Fukushima. Understanding of research activities from society is important.
2. Technical Problem: Since there are limitations on the working hours and abilities of researchers, including both professionals and citizen scientists, it is necessary to improve the efficiency of work to the extent possible by utilizing artificial intelligence (AI) techniques.
3. Computational Problem: Since it requires a certain level of expertise and time to create proper training data, it is necessary for anyone to be able to make training data as efficient as possible. It is necessary to clarify the design theory to obtain highly accurate data when using unsatisfactory training data.

Fig. 1 (a) microphone, (b) sync node station and (c) website project site in exclusion zone, which is 10 km from Fukushima Daiichi Nuclear Power Plant

3 Developed system

For this study, we installed the first transmitter station [12] within the exclusion zone area shown in Figure 1; more specifically, this location was 10 km from the Fukushima Daiichi Nuclear Power Plant. The transmitter station is located within the Oamaru district in the town of Namie, Fukushima (i.e., 37°28′4.3″ N 140°55′27.5″ E). We selected this site within the exclusion zone because it is one of the most difficult areas for long-term continuous investigations. Here, the exclusion zone is the most radioactively polluted zone in Fukushima. Further, no remote sensing methods are available on the surface due to the lack of power, information, and traffic infrastructures. Given these restrictions, although field surveys are required, the number of workable hours is extremely limited due to radiation exposure concerns. Finally, frequently used portable recorders require regular replacement given their limited

memory and battery capacities, which is impractical for long-term continuous investigations.This project development began after a year of literature searching from 2011 to 2014. A transmitter station with satellite Internet was installed in 2015. We then conducted an intensive feasibility survey of all transmission facilities at the location. Finally, after official preparations and approvals, we signed a service contract with a local electricity company in 2016. The final construction was completed at the end of March, 2016 [11].

In this project, we set out to collect, share, and analyze soundscape data within the exclusion zone. At the first, we developed both a Live Sound System and a Streaming/Archiving System that enabled us to distribute sound data from the exclusion zone to the public via the Internet to make such data publicly available for listening in real time via `http://radioactivelivesoundscape.net/` and for asynchronous listening. The Live Sound System was composed of separate subsystems, i.e., a Field Encoding System to digitize live sounds from the forest and a Streaming/Archiving System to deliver the live sound data via the Internet and archive the sound data in a recorded file. Note that the technical operational testing notes of the Live Sound System were discussed previously in [11].

The Field Encoding System was composed of two key components, i.e., an audio block and a transmission block. Microphones (omnidirectional SONY F-115B) were individually connected to an amplifier (XENYX 802, Behringer) of the audio block, and their outputs served as input to an audio encoder (instreamer100, Barix) that converted sounds captured by a microphone into MP3, which was the format used for subsequent digital sound delivery. These characteristics of our Internet service plan were important considerations since research funds required to conduct such a long-term ecological study are likely to fluctuate over time. As there was no prior Internet connection at the exclusion zone site, we used a satellite Internet service, which was provided by IPSTAR in April 2016.

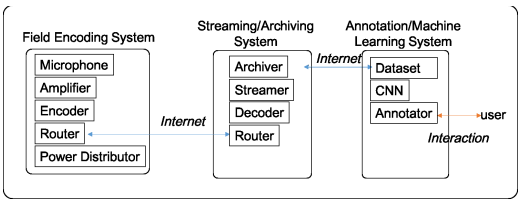

Fig. 2 System diagram of Live Sound System: Field Encoding and Streaming/Archiving System.

The Streaming/Archiving System is located in the server room of our laboratory and uses a normal bandwidth Internet connection, allowing simultaneous public access to transmissions in Figure 2. We employed two servers; one for streaming, the other for archiving. The processed audio signal sent from the microphone was encoded into an MP3 live stream in the Field Encoding System. After transfer to the Streaming/Archiving System, the MP3 live stream can be simultaneously played on MP3-based audio software worldwide. The operating system was the standard single package of Linux Fedora 14, and the sound delivery service was implemented in Icecast 2 software.The servers were established in our laboratory rather than at

the contaminated forest site. This setup avoids the technical difficulties in providing power and adequate data download at the remote location.

4 Data Processing

The recording was conducted from June 2016 to June 2017 in Japan Standard Time in the exclusion zone of Fukushima, Japan. The summer and winter seasons in this zone last from June to September and from December to March, respectively. The average monthly temperature is highest in August (23.0 °C and lowest in January (2.0 °C). The average annual rainfall is 1,511 mm (Namie Meteorological Station; JMA 2017) [12].

The audio files were processed by peak normalization high-pass filtering with a 500-Hz cutoff and 20 dB/decade attenuation. These calculations are made with the software *sox* version 14.4.2.

To process the (24 hours × 365 days) sound stream recordings of the environment surrounding the station, the human computation must be augmented with automated analysis. Both are presented below.

Manual detection

In this first strategy of manual annotation sampling, as the acoustic activity of birds is highest near sunrise, we analyzed the live streaming audio data between 10 min before sunrise and 60 min after sunrise [12]. Mainly following the procedure in [13] and [15], we also studied the sunset sounds. Audio files containing the sunrise and sunset times were annotated. The recordings were started at 06 min of every hour and lasted for 1 hour (i.e., streaming events were recorded at 00:06 – 01:05, 01:06 – 02:05, and successive hours throughout the day). Denoting the sunrise or sun-set time by n:m, if $6 \leq m < 36$, we take the first 30 minutes of the sound file containing the sunrise or sunset time. Otherwise (i.e., if $0 \leq m < 6$ or $36 \leq m < 60$), we take the last 30 minutes of the file. To alleviate the workload of the listeners, these 30-minute audio files were separated into two parts.

In this experiment, 21 students were recruited to index the selected audio stream. The students were instructed to identify four events: the songs and calls of the target birds, rain, and wind. They were also instructed to subjectively identify the signal levels of each event (strong or weak). The selected target bird was the Japanese Nightingale (*Horornis diphone*) for comparison with Ishida et al. [14].

Moreover, the songs of nightingales are unique and common, so their directions are more easily detected than songs of other bird species. The *ho ho ke kyo* songs of male Japanese Nightingales attract females in Figure 3 (a), and the *chi chi chi* songs of both sexes warn against predators or herald the presence of one bird to others. For this reason, calls are harder to localize and identify than songs in Figure 3 (b).

To minimize the human error factor, each audio file was scanned at least three times by different listeners. Listeners dedicated four hours to this task, and were directed to index as many files as possible. Forty-eight audio files were allocated to

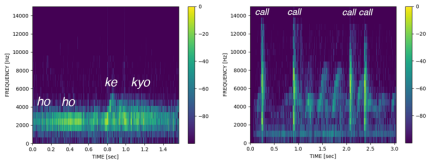

Fig. 3 Vocal activity of bush warblers: colors describes signal strength in dB Full Scale. (a) warblers song. (b) warblers call

21 listeners (11 listeners for the former part, 10 listeners for the latter part). This human listening experiment accumulated 2,225 events, including 711 calls, 572 songs, 628 winds and 314 winds. From these human-based inputs, 8,006 MP3 files were computed. The CNN detector depicted in next section has been trained on these annotations.

A second strategy of human annotation has been conducted. To build efficiently the Received Operation Characteristics (ROC), specialists made binary annotation for 2,500 files, 4 per day at sunrise, sunset, mid day and mid night. For this purpose the DYNI team designed a web-based application for collaborative audio annotation based on a front-end developed by M. Cartwright et Al. and described in [1]. This application makes the annotation of large amount of audio files easier and more robust by allowing the (customizable) integration of annotations from an unlimited number of users. This makes it a useful interface to build large ground truth and train machine learning algorithms, within a citizen science scope. A demo of the system can be found at `http://sabiod.org/EADM/crowdannot` and is proposed with open licence for academic research. The interface is presented in Figure 4. We then produced in 5 hours an expert annotation of 2,500 chunks of 10 seconds. In Toulon 4 professionals labeled for existence of sound sourced by a bird, and in Tokyo 2 professionals labeled for sound by a warbler. Regarding these data as a ground truth, the detector model performance was evaluated.

Automatic detection

For the automated analysis we tried two methods : (1) a raw signal processing and (2) a deep neural net approach. The signal processing approach is based on energy and spectral flatness [19] thresholding. The threshold for local decision were guided by the annotated data and Receiving Operating Characteristics analysis.

With the second automatic method, the neural net, the audio files were splitted and differentiated by windows of 0.252s with 50 % overlap. To increase the robustness of the model, the converted raw files were augmented by added noise, then fast-Fourier transformed was computed as the input of a 8-layers Convolutional Neural Network (CNN). A denoising auto-encoder was firstly trained to initialize the convolutional

[1] M. Cartwright, A. Seals, J. Salamon, A. Williams, S. Mikloska, D. MacConnell, E. Law, J. Bello, and O. Nov. "Seeing sound: Investigating the effects of visualizations and complexity on crowdsourced audio annotations." In Proceedings of the ACM on Human-Computer Interaction, 1(1), 2017

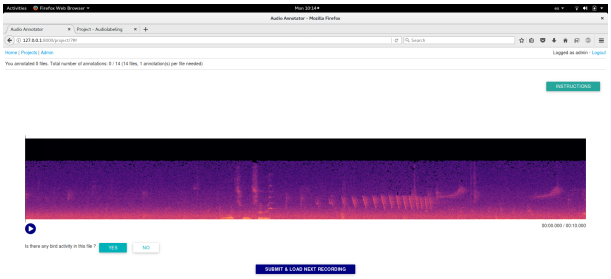

Fig. 4 Illustration of the online collaborative annotation system we developed to annotate 40 seconds of recordings per day with annotator based in Japan and in France, in order to compute the ROC of two detectors and other statistics.

layers. The number of connections in this architecture was 32 – 512 [16]. The parameters of each layer (weights and biases) were L1/L2 regularized (multiple combinations and hyper-parameters were tried). We tried multiple combinations of conduction for the convolution layers (maximum pooling, average pooling and batch normalization).

5 Results

The recordings worked fine, yielding to a full year, 7/24 soundscape recording, and nearly 2 To of data. We trained the CNN on the first annotation strategy, and ran it at a scale of 0.5 sec. on the whole year. The Accuracy of the CNN model was poor (under 30%). The reason could the hudge amount of rain noise, and the high variability of the used annotations for the training stage.

We then also score the automatic detector based on the maximum of the spectral flatness on the 10 sec. sections. The ROCs of this detector are given Figure 5 for Bird and for Warbler detectors. Even if it is a very fast annotator running in few hours the whole year, the area under the curve fird Bird detector is 0.63 and 0.66 for Warbler, which is allowing preliminary analyses depcited below.

At 0.5 second intervals, the approach outputted whether the sound included a call or a song (a warble), but did not distinguish between calls and songs. The model counted 6,520 hours of warbling. The vocal activity of warblers was most variable at dawn and dusk in Figure 6 (a).

Corresponding to temperature and season, two groups of warbler detection can be classified, demonstrating that the model allow to produce ecological features that could be correlated at long term with the variation of the radioactivity rate.

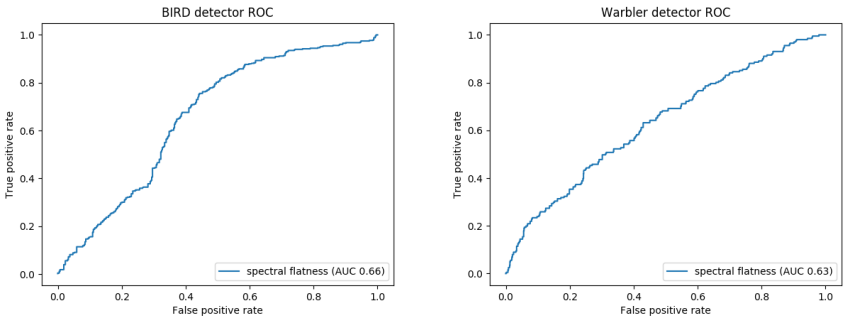

Fig. 5 ROC curve of (a): Bird Detector, and (b): Warbler Detector based on spectral flatness

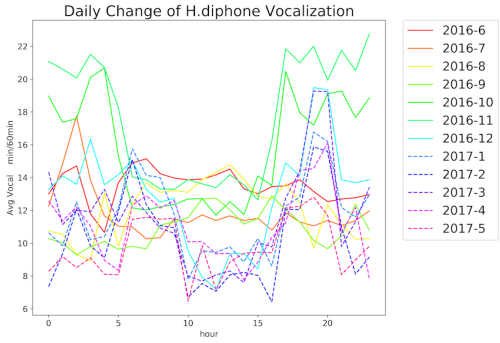

Fig. 6 Monthly vocalization change of warblers: hourly-average is illustrated corresponding to each o'clock. Legends describes years-months.

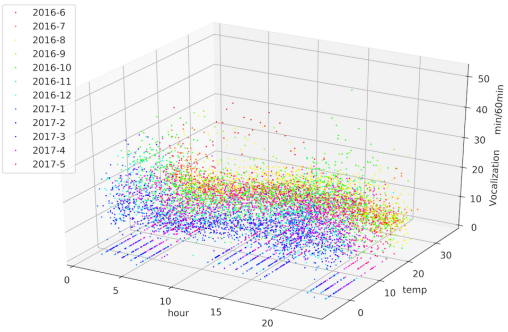

Fig. 7 Numbers of warbling vs. hours and temperature: warbling count is shown with o'clock-time of a day and temperature. Legends describe years-months.

6 Discussion

The raining conditions (weak signal) and difficulties on the field to increase the recording capacity yield to difficulties into the automatic detections with trained CNN. However, we demonstrated that we get reasonable eco-acoustic results all over the year.

Vocal activity was identified during periods of heavy rain (≥ 20 mm/h), when warbling is known to cease. As evidenced by the lack of target sounds in their assigned audio files, humans attention lapses, the signal can be unrecognized by the human. When simultaneous signals arrive from multiple directions, they will overlap, which complicates the computation. In contrast, humans can focus on a specific direction (the so-called cocktail party effect), if multiple microphones are installed at the same site. If a bird is present but does not sing, it will be unrecognized by both humans and the model. This problem could be overcome by a visual observation method; for instance, the Cyberforest Project [9] is installing a monitoring camera at their observation site.

The difficulty in the automatic method yields in the weak signal of the target (under rain). Recently the Bird activity detection challenge confirmed the superiority of CNN approaches as the model of Pelligrin [16]. Also CNN is the best model for Bird species classification [18]. These CNN-based systems required supervisors annotated by professional ornithologists, however creating supervisors is very time consuming task that it is not possible to rely completely on experts [6]. Asking amateurs with soundscape mania can help this problem. This study demonstrated limitations and possibilities of connecting human-based annotation to machine learning systems. Future work will focus to obtain highly accurate annotated samples using the produced interface, or potentially recruiting massive international crowd sourcing as the interface allows cooperative annotation. Detection of humans at the training stage are not always suitable as computer inputs. Multiple calls are always heard at any one time. Whereas some listeners record these calls as separate events, others count them as a group in Figure 3(b). Other than that factor, the limited proficiency of listeners in using the listening software, which degraded listening precision 1/1000 second to 1/10 second. To control and overcome such problems and to accelerate the experiment, united interface between users and CNN is desired. To overcome these limitations, we augmented the raw data and applied filtering thresholds on the features to maximize the area under the curve with labeled data. Machine learning can over-estimate the number of vocal activities of bush warblers, whereas human listeners cannot always distinguish between bird warbling and the calls of frogs, and may combine both sounds into the learner.

As previously mentioned, it is necessary to improve the efficiency of work to the extent possible by utilizing semi-supervised automatic algorithm for the professionals and citizen scientists.

We however demonstrated the presence of nightingales near the station varied on a monthly basis between June 2016 and July 2017. The acoustic index was elevated during the breeding season in spring, when male birds sing frequently to attract females. This study commenced in 2016, and because no comparative record exists before the Nuclear Accident, the effect of radioactive contamination cannot be assessed. The dataset collected at Namie, Fukushima, is the largest dataset available for ecoacoustic research. Using this dataset, we developed a methodology for tracking environmental effects on biodiversity. The dataset is also useful for measuring

and comparing future long-term changes. As the recording has continued every 24 hours since June 2017, we can analyze the yearly changes from this time onward. The Cyberforest project operates eight other observational open-microphone sites located throughout Japan. All the sites are connected to the on-line archive system at `http://www.cyberforest.jp/`. Applying this method at the other sites, we could simultaneously compare the bio-acoustic data at multiple sites. This technology could monitor climate change through birdsong analyses.

There is little known in advance about the introduction and long-term operation of consumer electronics products in high-dose zones. Long-term operation for decades in this project has many technical problems. Especially, although the report is informal, it has been reported that the operation time with batteries is shortened when consumer electrical appliances are operated in a high dose zone. In other words, it means that the brought-in equipment can not operate for the prescribed time at first. However, since it is not permitted to bring radioactive equipment out of the same zone by operating in a high dose zone, research is necessary for actual verification. Since the equipment used in this study is already highly radioactive, it becomes a sample for such experiments.

As previously mentioned, it is academically and socially important to conduct ecological studies focused on effects radiation exposure has had on wild animal populations over several generations in Fukushima. (Social Problem)

Although 6.5 years has passed since the Great East Japan Earthquake, the situation in the difficult-to-return area has remained static. Electricity remains in disrepair, and many patches of road are still collapsed. As decontamination activities are rarely performed, it is extremely difficult for researchers to enter the site. Building a multimedia system at such a point and using it for a long time is not only technically but also a socially big challenge. Since the power supply and road collapsed due to the earthquake, it can not be expected to be a base station of mobile phones or a stable power supply. Special permission is required for the surveyor to conduct the ground survey directly and the staying time is extremely limited due to the high dose. However, due to the large environmental problem of nuclear disaster, there are high social concerns about the animals left here. In order to make full use of this limited resource, this research is aimed at constructing a multimedia system capable of long-term operation, taking advantage of environmental concerns in the general society, It is to realize support. The system constructed for this research has operated continuously for 1.5 years. The system is designed to operate for 24 hours over 365 days each year. Prior to this project, we operated a similar system for more than 10 years [8]. Therefore, the present project is expected to continue until approximately 2030. By special arrangement with the power company, it also receives a stable power supply which is unlikely to disconnect in future operation. After finishing the operation of the mp3 file, we will commence operations on the uncompressed recordings.

The National Institute of Radiological Sciences have confirmed plant malformations and other abnormalities at the site [17], we are also considering a camera-based study. Plant malformations are known to be short-term events, but could be recorded

by a camera system running for 24 hours. Moreover, if individual animals can be discriminated in the camera data and the relevance of deformed plants assessed, we could begin to research the linkage between animals and plants.

Citizens can contribute to ecological surveys by participating in identification activities and research at dedicated events. Experience programs and exhibitions are open to the public at museums, science museums, and similar institutes. Asking participants directly for such work would certainly assist science, but listening to mass data is mentally demanding and probably unsustainable. The present research takes an indirect approach. Our method processes a large amount of data by AI without compromising the creation of training data. The following questions remain to be addressed:

1. How can we request a user to create training data that are easily handled by AI? The training data must be appropriately selected for the total data size. For instance, birdsongs are often sampled in the morning and evening, when vocal activity is most intense.
2. How can we handle differences among individuals in human-generated training data for AI processing? When preparing training data, it is necessary to resolve the timing deviations among users. The timing of a bird chirp varies from user to user. Whereas one person records the moment the bird starts singing, others might record a moment during the singing or when the song has finished.
3. How can we sustain 1 and 2 in a sustainability-oriented society that is increasingly aware of environmental issues? Instead of unilaterally requesting users to create training data, we must develop an interface that promotes environmental consciousness through enjoyable activities and games.

7 Conclusion

This paper discussed the acquisition and analysis of environmental sounds in a difficult-to-return area 10 km from the Fukushima Daiichi Nuclear Power Station. After obtaining official permission, we established a real-time acquisition and distribution system and acquired over 8,000 hours of continuous data. To process these data, we used a signal processing approach and tried a CNN model with human-computation. Prior to our study, no samples in this area had been continuously collected over the long term. We believe that by future data acquisition and analysis, we can investigate the influence of radiation on wild animals.

Acknowledgements Special thanks go to Fumikazu Watanabe and Laboratory Members. This study was supported by JSPS KAKENHI Grants 26700015 and 16K12666, MIC SCOPE Grants 142103015 and 162103107, JST PRESTO 11012, the Telecommunications Advancement Foundation, the Moritani Scholarship Foundation, the Tateisi Science and Technology Foundation, and the Mitsubishi Foundation. We thank the `http://sabiod.org/EADM` GDR CNRS MADICS that support development of the annotator platform. We thank Pascale Giraudet for annotation participation and other volonteers in Toulon and in Japan.

References

1. Chernobyl Forum. Expert Group Environment., International Atomic Energy Agency. (2001) Environmental consequences of the Chernobyl accident and their remediation: twenty years of experience; report of the Chernobyl Forum Expert Group 'Environment'. Radiological assessment reports series. International Atomic Energy Agency, Vienna
2. Cusack P (2012) Sounds from Dangerous Places. ReR Megacorp, Thornton Heath, Surrey,
3. Ishida K (2013) Contamination of Wild Animals: Effects on Wildlife in High Radioactivity Areas of the Agricultural and Forest Landscape. In: Nakanishi TM, Tanoi K (ed) Agricultural Implications of the Fukushima Nuclear Accident. Springer, Japan
4. Begone M, Harper JL, Townsend, CR (1990) Ecology: individuals, populations, and communities. 2nd. edn. Blackwell Scientific, Boston
5. Louv R, Fitzpatrick JW (2012) Citizen Science Public Participation in Environmental Research. 1st. edn. Cornell University Press, Ithaca
6. Keita F, Yoshio M, Akira Y, Nao K, Kumiko T, Takeshi O (2016) Mammal assemblages recorded by camera traps inside and outside the evacuation zone of the Fukushima Daiichi Nuclear Power Plant accident. Ecological Research 31 (4):493-493. doi:10.1007/s11284-016-1366-7
7. Slabbekoorn H, Peet M (2003) Ecology: Birds sing at a higher pitch in urban noise. Nature 424 (6946):267-267
8. Kobayashi H (2010) Basic Research in Human-Computer-Biosphere Interaction. The University of Tokyo, Tokyo, Japan
9. Saito K, Nakamura K, Ueta M, Kurosawa R, Fujiwara A, Kobayashi HH, Nakayama M, Toko A, Nagahama K (2015) Utilizing the Cyberforest live sound system with social media to remotely conduct woodland bird censuses in Central Japan. Ambio 44 (Suppl 4):572-583. doi:10.1007/s13280-015-0708-y
10. Kobayashi H, Hirose M, Fujiwara A, Nakamura K, Sezaki K, Saito K (2013) Tele echo tube: beyond cultural and imaginable boundaries. In: 21st ACM international conference on Multimedia, Barcelona, Spain, pp 173-182. doi:10.1145/2502081.2502125
11. Kobayashi HH, Kudo H (2017) Acoustic Ecology Data Transmitter in Exclusion Zone, 10 km from Fukushima Daiichi Nuclear Power Plant. Leonardo 50 (2):188-189. doi:10.1162/LEON_a_01416
12. Japan Meteorological Agency – Namie Meteorological Station. Japan Meteorological Agency. http://www.jma.go.jp. 2013
13. Ueta M, Hirano T, Kurosawa R (2012) Optimal time of the day to record bird songs for detecting changes of their breeding periods. Bird Research 8:T1-T6. doi:10.11211/birdresearch.8.T1
14. Mark C, Ayanna S, Justin S, Oded N, Stefanie M, Duncan M, Edith L, Juan B, J. Bello and Oded N (2017) Seeing sound: Investigating the effects of visualizations and complexity on crowdsourced audio annotations. In Proceedings of the ACM on Human-Computer Interaction, 1(1).
15. Ishida K, Tanoi K, Nakanishi TM (2015) Monitoring free-living Japanese Bush Warblers (Cettia diphone) in a most highly radiocontaminated area of Fukushima Prefecture, Japan. Journal of Radiation Research 56 (Suppl 1):i24-i28. doi:10.1093/jrr/rrv087
16. Stowell D, Wood M, Stylianou Y, Glotin H Bird detection in audio: A survey and a challenge. In: 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), 13-16 Sept. 2016 2016. pp 1-6. doi:10.1109/MLSP.2016.7738875
17. Watanabe Y, Ichikawa Se, Kubota M, Hoshino J, Kubota Y, Maruyama K, Fuma S, Kawaguchi I, Yoschenko VI, Yoshida S (2015) Morphological defects in native Japanese fir trees around the Fukushima Daiichi Nuclear Power Plant. Scientific Reports 5:13232. doi:10.1038/srep13232
18. Sevilla, A, and Glotin H (2017) Audio bird classification with inception-v4 extended with time and time-frequency attention mechanisms. In: 2017 CLEF Working Notes 1866, Dublin
19. Ricard, J. and Glotin H (2016) Bird song identification and monitoring system. LSIS internal research report, University of Toulon (ed)