

Article

Evaluation of a Sensor System for Detecting Humans Trapped under Rubble: A Pilot Study

Di Zhang ¹ , Salvatore Sessa ² , Ritaro Kasai ¹, Sarah Cosentino ¹, Cimarelli Giacomo ², Yasuaki Mochida ², Hiroya Yamada ², Michele Guarnieri ² and Atsuo Takanishi ^{3,4,*}

¹ Graduate School of Advanced Science and Engineering, Waseda University, Tokyo 169-8555, Japan; gwlrzd@fuji.waseda.jp (D.Z.); ritaro.kasai@gmail.com (R.K.); sarah.cosentino@aoni.waseda.jp (S.C.)

² Hibot Corporation, Watanabe Corporation Building 4F, 5-9-15 Kitashinagawa, Shinagawa-ku, Tokyo 141-0001, Japan; sessa@hibot.co.jp (S.S.); cimarelli@hibot.co.jp (C.G.); mochida@hibot.co.jp (Y.M.); yamada@hibot.co.jp (H.Y.); guarnieri@hibot.co.jp (M.G.)

³ Department of Modern Mechanical Engineering, Waseda University, Tokyo 169-8555, Japan

⁴ Humanoid Robotics Institute (HRI), Waseda University, Tokyo 162-0044, Japan

* Correspondence: contact@takanishi.mech.waseda.ac.jp; Tel.: +81-3-5369-7329

Received: 6 February 2018; Accepted: 10 March 2018; Published: 13 March 2018

Abstract: Rapid localization of injured survivors by rescue teams to prevent death is a major issue. In this paper, a sensor system for human rescue including three different types of sensors, a CO₂ sensor, a thermal camera, and a microphone, is proposed. The performance of this system in detecting living victims under the rubble has been tested in a high-fidelity simulated disaster area. Results show that the CO₂ sensor is useful to effectively reduce the possible concerned area, while the thermal camera can confirm the correct position of the victim. Moreover, it is believed that the use of microphones in connection with other sensors would be of great benefit for the detection of casualties. In this work, an algorithm to recognize voices or suspected human noise under rubble has also been developed and tested.

Keywords: life detection; earthquake rescue; gas sensor; voice recognition; thermal vision camera

1. Introduction

During the 21st century, more than 522 significant earthquakes happened [1], with a death toll of more than 430,000 worldwide [2]. The majority of deaths are caused by buildings collapsing and trapping occupants under the rubble. In fact, if the casualty is an uninjured, healthy adult with a supply of fresh air, then they can survive for about 72 h. Eighty percent of survivors can be rescued alive within 48 h of a collapse, but after 72 h the survival rate reduces exponentially [3]. This time limit can be much shorter due to air supply shortage, environmental temperature, the health condition of the casualty, etc. Therefore, to reduce mortality after a natural disaster, the rapid detection of survivors inside collapsed structures is of the utmost importance. The current searching method is based on survivors' testimony to establish the possible presence of casualties under the rubble. Rescue operations are generally carried out in subsequent steps. First, the rescue team accesses the area with dogs to search for casualties on the surface. Then, the rescue team uses video cameras to check the situation under the rubble. Finally, the rescue team tries to verify the presence of people trapped under the rubble [4]. However, the first objective of the rescue team is to assess two essential characteristics of the searching area: the existence of a sufficient number of survival spaces, and the stability of volume of the ruins [5]. This assessment is subjective and prone to change due to structural instability and the unknown situation under the rubble. Accessing collapsed structures is extremely dangerous for rescue teams because subsequent aftershocks might furthermore undermine the stability of structures. Moreover, rescue workers are at great risk for the development of physical, cognitive,

emotional, or behavioral symptoms of stress [6]. Hence, rapid localization of survivors under the rubble, avoiding direct access and exploration of the affected area, is essential for rescue teams.

To reduce the risks of rescue operations and accelerate the localization of casualties, several methods based on the use of sensor technologies have been proposed. Currently, rescue teams use life detection systems mainly based on microphones, optical/thermal cameras, and Doppler radar [7]. Audio signal analysis is an effective method to detect humans trapped under rubble, and some systems are already commercially available, such as the Acoustic Life Detector, which is based on audio signal processing to identify victims' low-frequency sounds. Moreover, several refined audio processing algorithms have been developed to detect human presence [8–10]. However, microphones become less accurate in the case of high background noise such as pneumatic drills, breakers, vehicles, wind, power cables, and water flows that can be present in a real scenario. Another limitation of audio detection systems is that they cannot locate unconscious victims.

Cameras are also widely used in rescue operations. Cameras are often mounted on mobile robots to explore dangerous and inaccessible areas because they are an efficient interface for human rescue [11–14]. Some researchers proposed thermal cameras to detect trapped humans to overcome the problems of limited visibility under the rubble [15,16]. However, even though cameras are an efficient method to detect casualties, their effectiveness is limited by their inherent reduced angle of view, the presence of obstacles, and the generally limited visibility under the rubble. In a real scenario, rapid localization and accurate estimation of the person's position are fundamental for an efficient rescue operation, and images alone do not provide enough information.

Doppler radar has been widely used in disaster rescue operations due to its efficiency in detecting motion behind obstacles [17]. In fact, frequency or phase shift in a reflected radar signal can be used to detect motions of only a few millimeters such as heartbeat or breathing [18]. However, Doppler radar requires accurate calibration and even small environmental changes due to aftershocks and structural instability have a negative impact on the performance of this kind of system [19]. Moreover, due to its narrow-angle view, this system is not suitable for wide disaster areas.

The use of gas sensors for human detection via analysis of changes in carbon dioxide (CO_2) and oxygen (O_2) in the environment due to human breath has also been proved feasible [3]. However, this system and several other experimental sensor systems for life detection have only been tested in controlled laboratory settings [20–25].

The objective of this study is to evaluate the performance of a system based on three different sensors in detecting live human presence under the rubble in a high-fidelity simulated disaster area in the open.

The system was composed of these three types of sensors:

1. Gas sensors (O_2 and CO_2) for the detection of human breath and quality of air.
2. Microphones for the detection of voices, human-produced sounds, or environmental noise.
3. Thermal vision camera for a direct view of the environment, localized temperature patterns.

The only a priori information during the experiment was that one person, and only one, was present in the area.

The article follows this structure: Section 2 introduces the sensors being tested, the specific sound recognition algorithm used, the data analysis method, and the experimental protocol. Sections 3 and 4 present the results and performance evaluation for each sensor. The last section summarizes the results and proposes future work.

2. Materials and Methods

In this section, we describe the sensors being tested, the sound recognition algorithm, the data analysis method, and the experimental protocol.

2.1. Gas Sensors (CO_2 and O_2 Sensors)

The FIGARO TGS4161 CO_2 sensor was chosen for its high sensitivity. This sensor can detect CO_2 in a range of 350~10,000 ppm. Moreover, this sensor exhibits a linear relationship between the change in electromotive force and CO_2 gas concentration on a logarithmic scale and shows excellent durability against the effects of high humidity.

The FIGARO SK-25F O_2 sensor was chosen. The advantage of this sensor is that it is not influenced by other gases such as CO_2 , CO , and H_2S that can be present in the environment. It shows a good linearity up to 30% O_2 , inside the measurement range in the real disaster area, and has chemical durability.

These two sensors were connected to a Waspmove motherboard from Libelium Comunicaciones Distribuidas S.L. (Zaragoza, Spain). The motherboard transmits the data stream via USB to a PC for data storage and analysis every 10 s. The CO_2 sensor needs 10 min of warm-up time to stabilize its data output. The O_2 sensor does not need an initial warm-up. The Waspmove board was mounted on a long telescopic pole and the pole was introduced in the gaps in the rubble for more than two minutes, then the collected CO_2 and O_2 data were analyzed.

2.2. Thermal Vision Camera

The LEPTON thermal camera from FLIR (Wilsonville, OR, USA), which is a complete long-wave infrared (LWIR) camera, was chosen. Its size is $8.5 \times 11.7 \times 5.6$ mm (without socket). The lens horizontal range is 56 degrees, the diagonal range is 71 degrees, and the resolution is 160×120 active pixels. The images are sent in streaming to PC via LAN communication using a Hi-Bot Corp. TITech M4 Controller as grabber. Dedicated software visualized the image data automatically, adapting the temperature range to a red-blue color map. The software also estimated the highest and lowest temperature in the image (Figure 1).

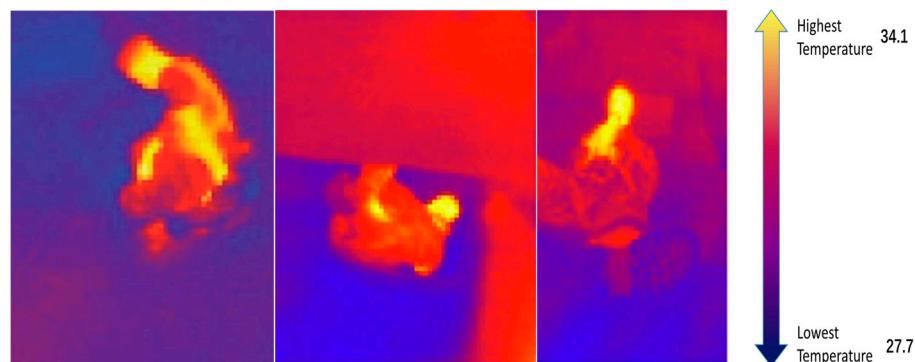


Figure 1. Three images from thermal camera from different directions.

The thermal camera was mounted on another telescopic pole; the pole was introduced in the gaps in the rubble and manually rotated to check the surrounding environment under the rubble. In Figure 1, a thermal image of an object with an outline similar to a human is shown. When a human-like thermal outline is detected, the affected area is tested from different angles and directions to verify if it really is a human victim.

2.3. Microphone

2.3.1. Hardware and Audio Signal Process

The low-energy Bluetooth SONY ECM-AW4 microphone was chosen. This is a non-directional microphone with a frequency response in the range of 300–9000 Hz.

To discriminate human voice from environmental noise, six voice features, usually used for voice detection, have been computed with MatLab.

- Energy Entropy

Entropy is a measure of state unpredictability. The definition of entropy H of a discrete random variable X with possible values x_i and probability mass function $P(X)$ is:

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i). \quad (1)$$

- Signal Energy

The energy E_s of a continuous-time signal $x(t)$ is defined as:

$$E_s = \int_{-\infty}^{\infty} |x(t)|^2 dt. \quad (2)$$

- Zero Crossing Rate

The rate of sign changes of a signal, a useful parameter of Voice Activity Detection (VAD) [26]:

$$ZCR = \frac{1}{T-1} \sum_{t=1}^{T-1} 1_{R<0}(s_t s_{t-1}), \quad (3)$$

where s is a voice single of length T and $1_{R<0}$ is an indicator function.

- Spectral Roll-Off

The roll-off frequency is defined as the frequency under which a percentage (85% cutoff) of the total energy of the signal spectrum is contained.

$$\sum_{n=1}^{R_t} M_t[n] = 0.85 \times \sum_{n=1}^N M_t[n], \quad (4)$$

where $M_t[n]$ is the magnitude of the Fourier transform at frame t and frequency bin n , and R_t is the frequency.

- Spectral Centroid

The Spectral Centroid C is calculated as the weighted mean of the frequencies present in the signal, determined using an FFT with their magnitudes as the weights [27]. If $x(n)$ represents the weighted frequency value, or magnitude, of bin number n , and $f(n)$ represents the center frequency of that bin, the Spectral Centroid C is:

$$C = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}. \quad (5)$$

- Spectral Flux

Spectral Flux is a measure of how fast the power spectrum of the signal is changing, comparing the power spectrum of one frame with the power spectrum of the previous frame:

$$F_t = \sum_{n=1}^N (N_t[n] - N_{t-1}[n])^2, \quad (6)$$

where $N_t[n]$ and $N_{t-1}[n]$ are the normalized magnitude of Fourier transform at frames t and $t-1$.

The audio signal was divided into non-overlapping frames of 10 ms and for each frame the above six features and their statistical deviation are calculated. In particular, for Energy Entropy, Zero Crossing Rate, Spectral Roll-off, and Spectral Centroid, the Standard Deviation has been computed, while for Signal Energy and Spectral Flux, the Standard Deviation by Mean Ratio has been computed. These six statistical values are the final feature values that characterize the audio signal.

2.3.2. Human Voice Detection Algorithm

The human voice detection algorithm was based on Support Vector Machine (SVM) from MatLab and consisted of a training phase and a classification phase. The Hard-margin SVM [28] classifies data identifying the best hyperplane that divides all data points into two groups [29,30].

- Training Phase

A database composed of 1588 samples of speech voice files, including male and female voices speaking in several languages, and 1687 samples of environment noise files including different types of environmental noise was created. All the sound samples were pre-processed with a bandpass filter (50 Hz~3000 Hz). The six statistical audio features were computed for each sound sample, and arranged in two matrices, a 6×1588 matrix for human voice samples and a 6×1687 matrix for environmental noise samples. These matrices were used in the SVM based algorithm as training data.

- Classification Phase

The flow chart of the classification phase is shown in Figure 2. Its fundamental steps are:

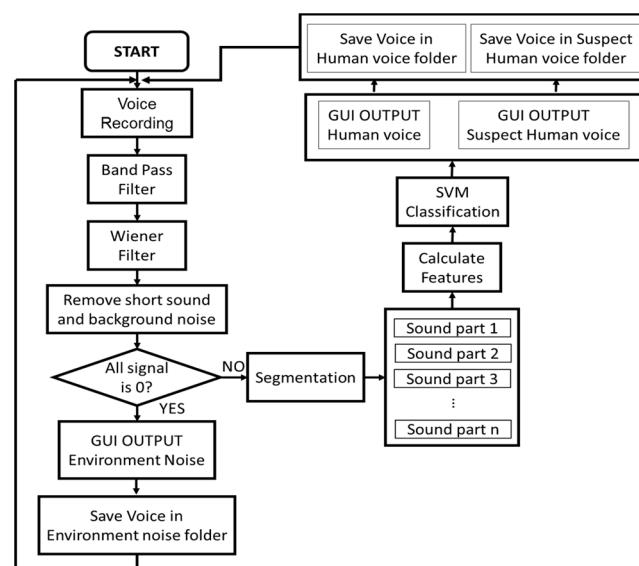


Figure 2. SVM classification for human voice and environment noise.

1. Voice recording phase: the system records voice at 5-s intervals.
2. Recorded data are bandpass filtered (50 Hz~3000 Hz)
3. Data are filtered with a Wiener filter. The Wiener filter minimizes the Mean Square Error (MSE) between the estimated random process and the desired operation. This filter is generally used to remove noise from a recorded voice.
4. Short sounds and background noise are removed. First, an adaptive threshold to remove background noise has been used. The reference level of environmental noise must be calculated.

As the noise in the disaster area is high and highly variable, an adaptive background noise reference has been defined according to the equation:

$$Ref_{noise} = \alpha Vol_t + (1 - \alpha) Vol_{t-1} \quad (7)$$

where α is the smoothing factor of Ref_{noise} change, Vol_t is the average volume [dB] of current 5 s voice data, Vol_{t-1} is the volume of previous 5 s voice data. It has been empirically found that $\alpha = 30\%$ yields the best performance. Then, if the volume of the sound sample is lower than 1.3 times Ref_{noise} , the algorithm identifies the sound sample as environmental noise and discards it. Sound signals that are 1.3 times higher than Ref_{noise} are suspect sounds. Then, the algorithm checks the length of this suspect sound. As human voice sound is assumed to last more than 300 ms, sounds shorter than 300 ms are removed. After removing short sounds, this suspect sound is processed with SVM to identify possible human noise.

5. Segmentation. The 5-s audio signal, after removing short sounds and background noise, is broken into shorter audio samples of 10 ms.
6. Audio statistical features, as described in Section 2.3.1, are computed for these shorter 10-ms audio samples.
7. SVM Classification. Sounds are differentiated in human voice or noise.

2.4. Experiments

2.4.1. Experimental Environment

The tests were conducted at a site at the Singapore Civil Defence Force (SCDF) facilities, Singapore. It is a high-fidelity disaster area meant to simulate collapsed buildings after a massive earthquake. Figure 3 shows the test area, which is approximately $8\text{ m} \times 24\text{ m}$ (192 m^2) organized as a grid of cells of $2\text{ m} \times 2\text{ m}$. This area is composed of two parts, a simulated two floors building partially collapsed (rows 6–13), and a simulated total collapse (rows 1–5). In rows 6–13 there are some accessible and stable paths for rescuing operations, while rows 1–5 represent a totally collapsed area with no accessible rescue paths.

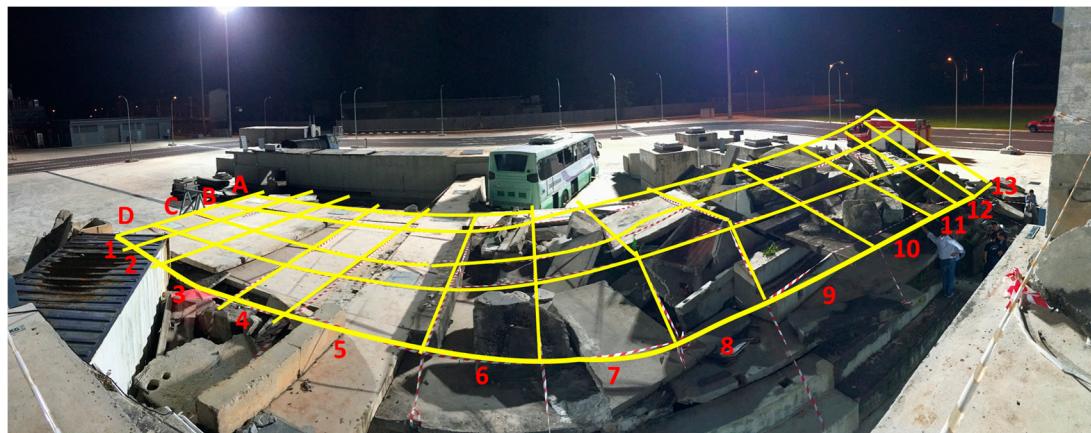


Figure 3. Experiment environment (panorama image).

2.4.2. Experimental Protocol

No environmental and structural information about the simulated disaster area was available before starting the experiment. At least 30 min before starting the sensor-based rescue experiment, a person entered the area and randomly hid inside the rubble, simulating an unconscious earthquake casualty. The casualty position had to be estimated within a 2-h time limit, without directly accessing

the rubble. However, tools could be inserted through the gaps to acquire data under the rubble. After scanning the entire areas, the position of the casualty had to be estimated. The acceptable identification area consisted of a square of $4\text{ m} \times 4\text{ m}$, a 4 cell square. The entire experimental session lasted three days and consisted of three trials per day (a morning, an afternoon, and an evening trial), or nine trials in total.

3. Results and Discussion

3.1. Experimental Results

In Table 1, the time needed to detect the casualty in each trial is shown, about one hour on average. We successfully detected the casualty in eight out of nine trials performed. Being fast and precise in casualty detection is a key factor because 80% of survivors are recovered alive if rescued within 48 h.

Table 1. Global results of the tests.

TEST	Execution Time	Result
Day 1 morning	1 h 35 min	Success
Day 1 afternoon	56 min	Success
Day 1 evening	1 h 25 min	Success
Day 2 morning	33 min	Success
Day 2 afternoon	50 min	Success
Day 2 evening	1 h 12 min	Failed
Day 3 morning	2 h 13 min	Success
Day 3 afternoon	20 min	Success
Day 3 evening	31 min	Success

The results of each trial are shown Figures 4–6 and described and commented on in the rest of this section. O_2 is measured as concentration, while CO_2 is in parts-per-million (ppm). Because the CO_2 sensor is not calibrated, the CO_2 data do not represent the real concentration and the absolute measured values in each trial vary widely depending on the time the measurement was taken and the environment around the site. For this reason, relative variations of CO_2 during trials were considered, and further confirmation from a rescuer or other sensors was required to verify the presence of the casualty in that specific area. The areas with relatively high levels of CO_2 are indicated in yellow. Areas manually checked with a thermal camera are circled in purple.

Figure 4 shows the results of the first day's trials.

Day 1, morning trial: The gas sensor located several possible locations for the casualty. The reason for those abnormal concentrations is that the person reached the center of the site through tunnels in the test sites (C5, A5, A8, A10, and A11 are sections of the same tunnel). The thermal camera images confirmed the presence of the casualty in the estimated area indicated by the red square, in the square composed by cells B9, C9, B10, and C10.

Day 1, afternoon trial: The gas sensor identified an area with a peak CO_2 concentration and the thermal camera confirmed the presence of the casualty in the area indicated by the gas sensor data, in the square composed by cells B7, C7, B8, and C8.

Day 1, evening trial: In this test, the casualty was located in the square composed by cells B11, C11, B12, and C12, using only the thermal camera. The gas sensor did not work properly because the affected area is a large area in which the wind could easily change the CO_2 concentrations, so the presence of a casualty did not significantly change the CO_2 concentration in this situation. This test was useful to analyze the factors that can lead to localization failures when using a gas sensor. However, this kind of area can be easily searched by a rescue team or a rescue dog because it is near the boundaries of the disaster area, outside the collapsed structure.

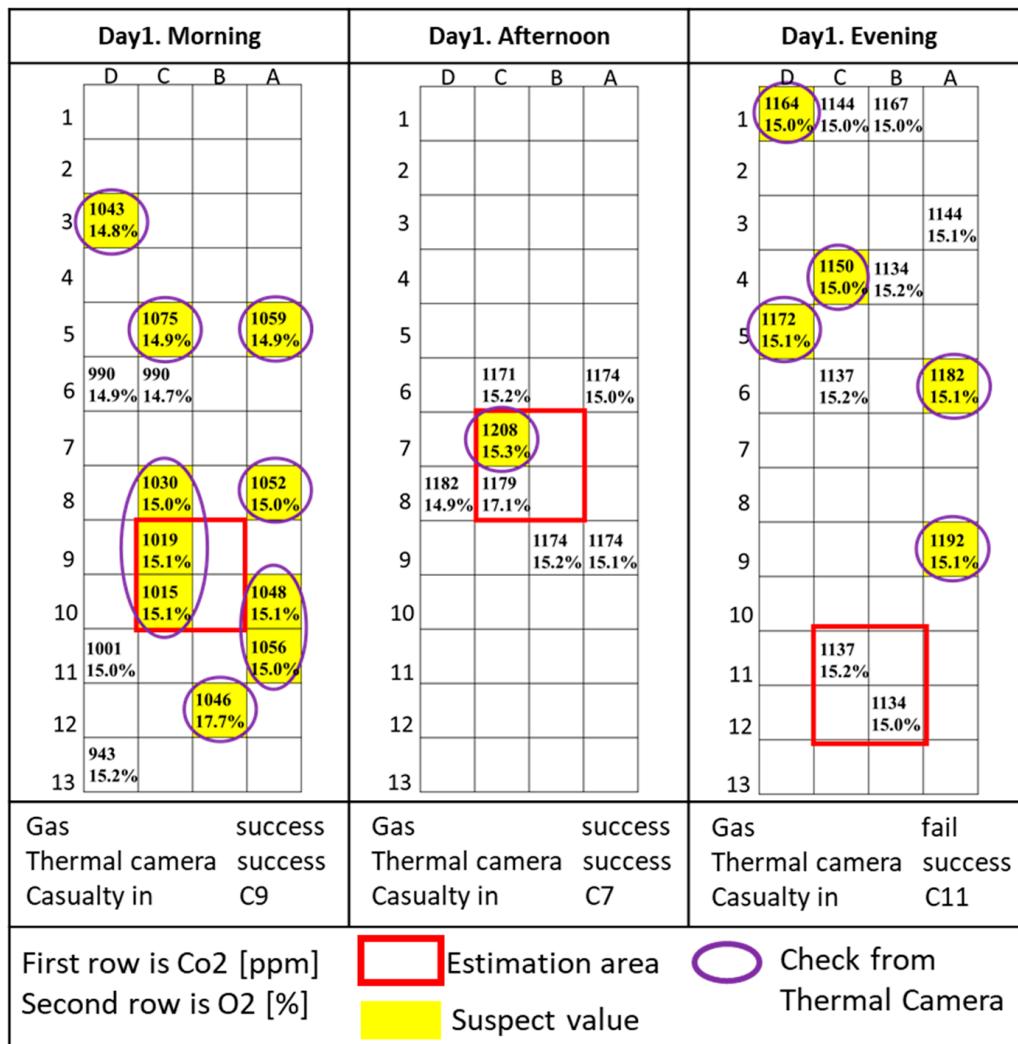
**Figure 4.** Day 1 results.

Figure 5 shows the results of the second day trials.

Day 2, morning: Both the gas sensor and the thermal camera located the casualty. The C11, D12, C12, D12 area is part of a corner in which the gas concentration was unusually high, and the camera could be inserted through a hole in the rubble to verify the presence of the casualty.

Day 2, afternoon: Both the gas sensor and the thermal camera located the casualty in the square composed by cells C6, D6, C7, and D7 that was beside a wall in a corridor where the gas sensors and the camera could be placed. It is important to note that, in this case, the gas sensor detected a high concentration of CO₂ in the whole corridor, so the exact position of the casualty could only be confirmed with a thermal camera.

Day 2, evening: This was the only trial in which the sensor system failed to locate the casualty. A high concentration of CO₂ was found in the area around B2, C2, B3, and C3, but the presence of many obstacles obstructing the view made verification via thermal camera impossible. This area is a maze of corridors in a semi-closed area with low air circulation, with the possible presence of grass and animals that might raise the concentration of CO₂. Moreover, the corridor in C2 was not reachable by gas sensors on the telescopic pole.

Day2. Morning				Day2. Afternoon				Day2. Evening			
D	C	B	A	D	C	B	A	D	C	B	A
1 639 14.1%				1				1 451 14.5%			
2				2				2 654 14.7%	683 14.7%		
3				3				3 451 14.8%	647 14.8%		
4				4				4			
5 636 14.0%			646 14.1%	5				5			
6				6 867 15.2%	875 15.3%	879 15.3%		6 664 14.7%			
7 649 14.1%				7 854 15.1%	840 15.2%			7			
8				8				8			
9			651 14.1%	9		847 15.2%		9			
10				10 849 15.0%	849 15.2%			10 631 14.7%	620 14.7%		
11 659 14.1%	664 14.0%			11				11			
12 669 14.0%	672 14.2%			12				12			
13				13				13			
Gas success	Thermal camera success	Casualty C11	Gas success	Thermal camera success	Casualty C6	Gas fail	Thermal camera fail	Casualty C7			
First row is CO ₂ [ppm]				Estimation area				Check from Thermal Camera			
Second row is O ₂ [%]					Suspect value						

Figure 5. Day 2 results.

Figure 6 shows the results of the last day's trials.

Day 3, morning: The gas sensor found a high CO₂ concentration very close to the casualty. However, the presence of many obstacles obstructing the view made verification via thermal camera impossible, so the casualty was located in the square composed by cells B3, C3, B4, and C4 based only on the gas sensor data.

Day 3, afternoon: The casualty was located very fast because the gas sensor measured a relatively high level of CO₂ in the square composed by cells B11, C11, B12, and C12 and the thermal camera confirmed the presence of the casualty through a hole in the corridor.

Day 3, evening: The casualty was located in the square composed by cells B9, C9, B10, and C10 using only the thermal camera. The data from the gas sensor were corrupted because of hardware problems on the gas sensor board.

Day3. Morning				Day3. Afternoon				Day3. Evening															
	D	C	B	A		D	C	B	A		D	C	B	A									
1	921 14.6%									1													
2	951 14.4%					2				2													
3	965 14.9%					3				3													
4	939 14.6%	976 14.8%	1103 14.8%			4				4													
5	946 14.6%					5				5													
6	956 14.5%	929 14.4%	931 14.5%			6				6	488 14.5%	456 14.4%											
7	931 14.5%					7				7													
8			977 14.5%			8				8	455 14.5%												
9			989 14.6%			9				9		438 14.5%											
10						10				10													
11	938 14.5%					11	832 14.8%			11													
12						12	825 14.8%			12													
13			951 14.4%			13	795 14.7%			13													
Gas success				Gas success				Gas fail															
Thermal camera success				Thermal camera success				Thermal camera success															
Casualty B4				Casualty D11				Casualty B8															
First row is CO ₂ [ppm]								Estimation area															
Second row is O ₂ [%]								Check from Thermal Camera															
								Suspect value															

Figure 6. Day 3 results.

3.2. Evaluation of the Gas Sensor and Thermal Camera

O₂ measurements were not useful to determine the presence of life under the rubble. CO₂ measurements were highly correlated with the possible position of the casualties; however, the CO₂ sensor failed to locate the casualty in three trials out of nine, one time due to hardware problems and the other times due to environmental conditions. The thermal camera failed to locate the casualty in two trials out of nine, confirming that, although visual analysis is useful, a multi-sensor system is more robust due to sensor redundancy and complementarity. Figure 7 shows the relationship between high casualty localization rate and high casualty presence exclusion rate depending on the CO₂ threshold. Areas with a high casualty localization rate indicate that the possible presence of the casualty is high, while a high casualty presence exclusion rate indicates areas in which the possibility of presence of the casualty can be reasonably excluded, and so do not need to be cross-checked with the thermal camera. From these empirical data, a method can be devised to estimate a reasonable CO₂ absolute threshold, correlated with a high casualty localization rate but also with a high casualty presence exclusion rate.

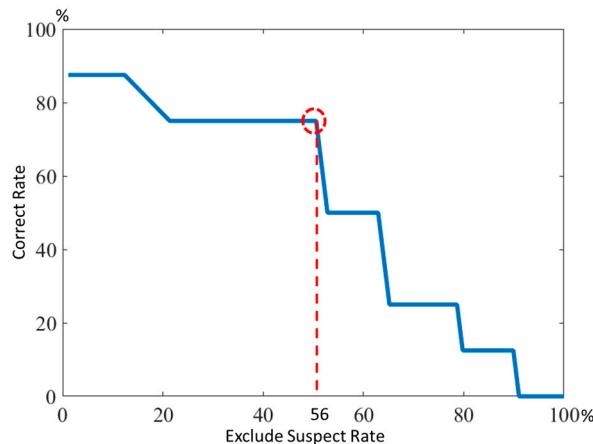


Figure 7. Correct rate and exclude suspect rate.

The closest point to (100%, 100%) was found with a CO₂ threshold of 27 ppm, leading to a reduction of the possible casualty presence area to 44% of the total area and significantly shortening the search and rescue operations. The sensitivity of the CO₂ sensor is 75% and specificity is 53.1%, as shown in Table 2.

Table 2. Evaluation of gas sensor system.

	Predicted Condition Positive	Predicted Condition Negative
Condition positive	6	2
Condition negative	38	43

3.3. Evaluation of Microphone and Audio Processing Algorithm

In all the experimental trials, the casualty was supposedly unconscious. Therefore, the person did not speak or produce other sounds such as scratching during the whole trial. However, in real disaster scenarios, there are cases in which the casualty is not unconscious and can produce sounds. For this reason, an algorithm for the detection of sounds that might be related to the presence of a casualty was designed and tested. The hardest problem was to make the algorithm less sensitive to background noise. A disaster site is often a noisy environment, with people searching for victims, vehicles, and various natural and artificial sounds. A dynamic threshold for the classification between a possible sign of life and background noise, based on the average level of sound in the area, was proposed. Of course, this method implies that in extremely noisy environments the detection of feeble sounds will not be possible. However, in this way the system is more robust and automatically rejects sounds that are not linked with the presence of casualties, reducing the number of sounds that must be listened for to check the presence of casualty in a specific area. In particular, speech has characteristic features that were used to separate it from other suspect noises. Figure 8 shows the results of the Day 3 afternoon test, in which we spoke directly to the casualty after locating them to test the audio recognition system. The microphone was placed on a telescopic pole and inserted in a hole in the same corridor where the person was detected by using the gas sensor and the camera. Then, we asked the casualty to perform three different tests: to not move and stay in silence while we talked outside, to call for help at a low volume inaudible by the human ear from outside, and to simply scratch on the ground. Audio detection results are shown in Figure 8. Moreover, we detected an unwanted cough, confirming the presence of the casualty, in the area with a high level of CO₂ during the Day 1 afternoon trial, and another suspect noise during the Day 2 afternoon trial, when the person moved into the corridor.

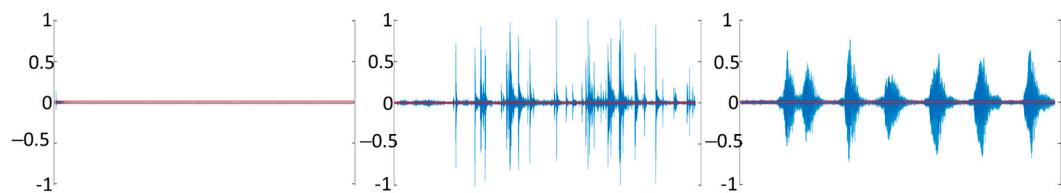


Figure 8. GUI of sound recognition.

The result of audio detection performance evaluation is shown in Table 3. The proposed algorithm can automatically differentiate the sound data and save it in different folders. In Table 3, the first row α/β represents the correct sound identification rate, where β is the total number of automatically classified sound files present in each category folder and α is the correctly classified number of sound files, which were validated manually.

Table 3. Evaluation of microphone.

TEST	Human Voice	Suspect Noise	Noise
Test Day 1 afternoon	87.5%	89.36%	100%
Test Day 2 afternoon	89.4%	91.21%	100%
Test Day 3 afternoon	90.6%	98.18%	100%
Average	89.36%	93.95%	100%

The correct voice recognition rate is 89.36% in a noisy environment. The correct classification rate for human-related suspect noise, including scratching and coughing, is 93.85%. Therefore, using a microphone in connection with other sensors would be beneficial for the detection of casualties.

4. Conclusions

In this study, a new sensor system for detecting human presence under rubble was proposed and tested. The effectiveness of each sensor was evaluated and confirmed. A CO₂ sensor can provide useful information to locate a casualty, but an O₂ sensor does not. A voice recognition algorithm based on SVM was also tested and from the results obtained it was confirmed that using the microphone would be of great benefit in the detection of casualties. This system has some limitations; for example, the gas sensor is difficult to use in open spaces due to stronger airflow affecting the CO₂ concentration. A sensor system using only a thermal camera is not robust because some areas cannot be directly accessed using a telescopic pole or directly observed due to the presence of obstacles.

In future work, a sensor system should be developed that includes multiple sensors, such as microphones and gas sensors, to be distributed in the area by the rescue team to alert them if one measures signs of a casualty under the rubble.

This kind of distributed sensor system can also be used in search and rescue operations with robotic aids that can release such sensors in areas inaccessible to or very risky for human rescue teams.

Acknowledgments: This study was partially supported by the Research Institute of Science and Engineering, Waseda University. This research has been supported by the HiBot Corporation and the Consolidated Research Institute for Advanced Science and Medical Care, Waseda University. We thank the team from the Singapore Civil Defence Force, in particular the Assistant Commissioner Ling Young Ern (Director Operations Department Singapore Civil Defence Force) and Captain Clara Toh (Commander, Banyan Fire Station Singapore Civil Defence Force), who provided great insight and expertise.

Author Contributions: Ling Young Ern and Clara Toh conceived and designed the experiments; Di Zhang, Ritaro Kasai, and Sarah Cosentino performed the experiments and analyzed the data; Cimarelli Giacomo, Yasuaki Mochida, Hiroya Yamada, Michele Guarneri, and Atsuo Takanishi contributed materials and analysis tools; Di Zhang wrote the paper, with contributions from Salvatore Sessa and Sarah Cosentino.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Significant Earthquakes—2017. Available online: <http://earthquake.usgs.gov/earthquakes/browse/significant.php> (accessed on 23 February 2017).
- Earthquakes. Available online: <http://earthquake.usgs.gov/earthquakes/> (accessed on 23 February 2017).
- Huo, R.; Agapiou, A.; Bocos-Bintintan, V.; Brown, L.J.; Burns, C.; Creaser, C.S.; Devenport, N.A.; Gao-Lau, B.; Guallar-Hoyas, C.; Hildebrand, L.; et al. The trapped human experiment. *J. Breath Res.* **2011**, *5*, 046006. [CrossRef] [PubMed]
- Vaswani, K. Nepal Earthquake: How Does the Search and Rescue Operation Work? 2015. BBC News. Available online: <http://www.bbc.com/news/world-asia-32490242> (accessed on 13 March 2018).
- Kiriazis, E.; Zisiadis, A. *Technical Handbook for Search & Rescue Operations in Earthquakes*, 2nd ed.; Zoi, V., Dandoulaki, M., Eds.; Access Soft Limited: Athens, Greece, 1999; pp. 1–48.
- Berger, W.; Coutinho, E.S.F.; Figueira, I.; Marques-Portella, C.; Luz, M.P.; Neylan, T.C.; Marmar, C.R.; Mendlowicz, M.V. Rescuers at risk: A systematic review and meta-regression analysis of the worldwide current prevalence and correlates of PTSD in rescue workers. *Soc. Psychiatry Psychiatr. Epidemiol.* **2012**, *47*, 1001–1011. [CrossRef] [PubMed]
- Younis, M.; Akkaya, K. Strategies and techniques for node placement in wireless sensor networks: A survey. *Ad Hoc Netw.* **2008**, *6*, 621–655. [CrossRef]
- Sun, H.; Yang, P.; Liu, Z.; Zu, L.; Xu, Q. Microphone array based auditory localization for rescue robot. In Proceedings of the 2011 Chinese Control and Decision Conference (CCDC), Mianyang, China, 23–25 May 2011; pp. 606–609.
- Latif, T.; Whitmire, E.; Novak, T.; Bozkurt, A. Sound Localization Sensors for Search and Rescue Biobots. *IEEE Sens. J.* **2016**, *16*, 3444–3453. [CrossRef]
- Yang, P.; Sun, H.; Zu, L. An acoustic localization system using microphone array for mobile robot. *Int. J. Intell. Eng. Syst.* **2007**, *2*, 18–26. [CrossRef]
- Rudol, P.; Doherty, P. Human Body Detection and Geolocalization for UAV Search and Rescue Missions Using Color and Thermal Imagery. In Proceedings of the 2008 IEEE Aerospace Conference, Big Sky, MT, USA, 1–8 March 2008; pp. 1–8.
- Kadous, M.W.; Sheh, R.K.-M.; Sammut, C. Effective User Interface Design for Rescue Robotics. In Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction, Salt Lake City, UT, USA, 2–3 March 2006; pp. 250–257.
- Murphy, R.R. Human-robot interaction in rescue robotics. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2004**, *34*, 138–153. [CrossRef]
- Fenwick, J.W.; Newman, P.M.; Leonard, J.J. Cooperative concurrent mapping and localization. In Proceedings of the IEEE International Conference on Robotics and Automation, Washington, DC, USA, 11–15 May 2002; Volume 2, pp. 1810–1817.
- Baker, M.; Casey, R.; Keyes, B.; Yanco, H.A. Improved interfaces for human-robot interaction in urban search and rescue. In Proceedings of the 2004 IEEE International Conference on Systems, Man and Cybernetics, The Hague, The Netherlands, 10–13 October 2004; Volume 3, pp. 2960–2965.
- Nourbakhsh, I.R.; Sycara, K.; Koes, M.; Yong, M.; Lewis, M.; Burion, S. Human-robot teaming for search and rescue. *IEEE Pervasive Comput.* **2005**, *4*, 72–79. [CrossRef]
- Chen, K.-M.; Huang, Y.; Zhang, J.; Norman, A. Microwave life-detection systems for searching human subjects under earthquake rubble or behind barrier. *IEEE Trans. Biomed. Eng.* **2000**, *47*, 105–114. [CrossRef] [PubMed]
- Garg, P.; Srivastava, S.K. Life Detection System during Natural Calamity. In Proceedings of the 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, India, 12–13 February 2016; pp. 602–604.
- Li, C.; Lubecke, V.M.; Boric-Lubecke, O.; Lin, J. A review on recent advances in Doppler radar sensors for noncontact healthcare monitoring. *IEEE Trans. Microw. Theory Tech.* **2013**, *61*, 2046–2060. [CrossRef]
- Suzuki, T.; Kawabata, K.; Hada, Y.; Tobe, Y. Deployment of wireless sensor network using mobile robots to construct an intelligent environment in a multi-robot sensor network. In *Advances in Service Robotics*; InTech Open Access Publisher: Rijeka, Croatia, 2008.

21. Wang, Y.; Wu, C.-H. Robot-assisted sensor network deployment and data collection. In Proceedings of the International Symposium on Computational Intelligence in Robotics and Automation, Jacksonville, FL, USA, 20–23 June 2007; pp. 467–472.
22. Bahl, P.; Padmanabhan, V.N. RADAR: An in-building RF-based user location and tracking system. In Proceedings of the Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, Tel Aviv, Israel, 26–30 March 2000; Volume 2, pp. 775–784.
23. Thrun, S.; Liu, Y.; Koller, D.; Ng, A.Y.; Ghahramani, Z.; Durrant-Whyte, H. Simultaneous localization and mapping with sparse extended information filters. *Int. J. Robot. Res.* **2004**, *23*, 693–716. [CrossRef]
24. Corke, P.; Hrabar, S.; Peterson, R.; Rus, D.; Saripalli, S.; Sukhatme, G. Autonomous deployment and repair of a sensor network using an unmanned aerial vehicle. In Proceedings of the 2004 IEEE International Conference on Robotics and Automation, New Orleans, LA, USA, 26 April–1 May 2004; Volume 4, pp. 3602–3608.
25. Tuna, G.; Gungor, V.C.; Gulez, K. An autonomous wireless sensor network deployment system using mobile robots for human existence detection in case of disasters. *Ad Hoc Netw.* **2014**, *13*, 54–68. [CrossRef]
26. Ramirez, J.; Górriz, J.M.; Segura, J.C. *Voice Activity Detection. Fundamentals and Speech Recognition System Robustness*; InTech Open Access Publisher: NewYork, NY, USA, 2007.
27. Peeters, G. A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project. 2004. IRCAM Web Site. Available online: http://recherche.ircam.fr/anasyn/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf (accessed on 13 March 2018).
28. Yu, H.; Kim, S. Svm tutorial—Classification, regression and ranking. In *Handbook of Natural Computing*; Springer: New York, NY, USA, 2012; pp. 479–506.
29. Friedman, J.; Hastie, T.; Tibshirani, R. *The Elements of Statistical Learning*; Volume 1 Springer Series in Statistics; Springer: Berlin, Germany, 2001.
30. Andrew, A.M. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods* by Nello Christianini and John Shawe-Taylor; Cambridge University Press: Cambridge, UK, 2000; ISBN 0-521-78019-5.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).