Occupancy-Driven Energy-Efficient Buildings Using Audio Processing with Background Sound Cancellation

Qian Huang

School of Architecture, Southern Illinois University Carbondale, Carbondale, IL 62901, USA; qhuang@siu.edu

Abstract: Demand-driven HVAC (heating, ventilation, and air conditioning) operation is essential in occupant-oriented smart buildings, where the levels of heating, cooling, and ventilation are intelligently regulated to avoid energy waste. Despite the great potential of building energy efficiency, one of the remaining technical challenges is how to accurately estimate building occupancy information in real time. In this paper, this design challenge is addressed. An advanced audio-processing technique is adopted that minimizes the impacts of environmental sounds on the recorded voice sounds of humans. Adopted mathematical modeling and signal processing procedures are elaborated in this work. Experimental studies show that our proposed audio processing with background sound cancellation algorithm improves the estimation accuracy of room occupancy quantity by approximately 11–12%, which results in an averaged ventilation energy reduction of 3.54% compared to the case of not applying background sound cancellation. The proposed audio-processing technique is promising to achieve non-intrusive, cost-effective, robust, and accurate solutions for building occupancy estimation.

Keywords: energy efficient building; background noise cancellation; occupancy detection; acoustic

1. Introduction

According to U.S. Energy Information Administration (EIA) statistics, more than 39% of carbon dioxide and 70% of electricity in the United States are consumed by buildings. Among various energy sources of energy usage in buildings, HVAC (heating, ventilation, and air conditioning) equipment accounts for up to 50% [1]. In fact, HVAC systems are typically sized to meet design full-loaded heating and cooling conditions that historically occur only 1% to 2.5% of the time [2]. Thus, HVAC systems are intentionally oversized most of the time. Heating and cooling equipment often operates at their respective part-load efficiencies. In traditional buildings, occupants basically have no control over building operations. Air-conditioning switches, temperature set points, and weekly schedules of HVAC operation are usually pre-set by property management personnel. Regardless of the behaviors and preferences of building occupants, this simple HVAC control method reduces the occupant comfort and energy efficiency of the building system [3,4]. As the occupants lose control of their indoor environment, their feelings of comfort are also degraded [5,6]. Consequently, this method has great potential to realize significant energy savings and comfort enhancements by improving the control of HVAC operations [3–6].

With rapid advances in smart cities [7], the Internet of Things (IoT) [8], and Li-Fi communication [9], next-generation smart buildings are supposed dynamically to sense the number of occupants in each room or thermal zone, then to adjust HVAC equipment accordingly. Moreover, these smart buildings provide daily operational data for performance analysis and visualization. To enable these attractive features, embedded and miniature environmental sensors are indispensable, such as motion sensors, indoor air-quality sensors, surveillance cameras, and security sensors. According to a report from the

U.S. Department of Energy (DOE) in 2017 [10], existing occupancy recognition and counting sensors are still far from meeting the requirements of next-generation smart buildings. These requirements in [10] include user-transparency, high accuracy, low failure rate, easy maintenance, low complexity, good privacy protection, and low price. It is expected in [10] that the development of future occupancy recognition and counting sensors will lead to drastic improvements in the way that HVAC systems operate in buildings. For example, it was reported in [11] that 10–15% of building energy can be saved using occupancy-driven HVAC operations. The researchers in [12] summarized existing occupancy detection approaches and results in demand-driven commercial office buildings. Later, the study in [13] revealed that energy savings can be 20–30% for occupancy-aware cooling in university residence halls. In [14], the impacts of building occupancy on HVAC energy efficiency are analyzed from occupancy transitions, variations, and heterogeneity. In [15], occupancy patterns of HVAC zones was incorporated into the building management system in commercial buildings. When retaining the indoor thermal comfort, a 38% reduction in heating energy was achieved in [15]. In order to calculate the number of people in a thermal zone, several detection mechanisms are presented in the literature. Table 1 summarizes the advantages and drawbacks of each existing occupancy-detection mechanism.

Table 1. Advantages and drawbacks of each existing occupancy-detection mechanism.

Occupancy Detection Mechanism	Advantages	Drawbacks
Passive infrared (PIR)	Good privacy protection + occupancy presence (and location) detection	No capability of occupancy counting + limitation of line of sight
Ultrasonic	Good privacy protection + no limitation of line of sight	No capability of occupancy counting
Radio-frequency identification (RFID)	Low cost + easy installation + fine-grained occupancy detection	Poor privacy protection
Image/Video camera	Fine-grained occupancy detection	Poor privacy protection + limitation of the line of sight + higher cost
Wi-Fi probe request	Good privacy protection + fine-grained occupancy detection	Need to carry a mobile device for Wi-Fi communication
CO_2 level	Good privacy protection	Influenced by environmental air flow and HVAC settings
Acoustic recognition	Low cost + intermediate occupancy detection in quiet environments	False counting due to interference from environmental sounds
Multiple hybrid types (e.g., PIR + CO_2 + acoustic)	Potential to further improve the accuracy of occupancy detection through multi-model data fusion	Complex data processing algorithms + higher cost + larger system size

Next, these existing occupancy-detection mechanisms are discussed individually. Since infrared (IR) radiation emitted by human movement is collected and identified by passive infrared (PIR) sensors [16,17], they are good at occupancy presence or absence detection in an area. In [16], multiple PIR sensors worked with machine-learning algorithms to estimate the occupancy number in a space. This system was implemented and tested in real office environments. In [17], to address the challenge that PIR sensors cannot detect stationary objects, the researchers presented a new chopped PIR sensor. The operating mechanism and experimental testing results were provided. Yet, these PIR sensors cannot count the number of occupants, so they are incapable of performing occupancy recognition and counting. Similarly, an ultrasonic sensor detects the presence of a building occupant by sending ultrasonic waves into space and measuring its return speed [18,19]. In [18], an ultrasonic system was created to estimate the occupancy status of rooms. The measurement results show that the ultrasonic signal is significantly attenuated with the number of occupants in a space. In [19], a broadband ultrasonic occupancy sensing system was presented with energy efficiency and scalability. It can detect the occupancy presence or quantity using proper data training efforts.

A direct line of sight is required between PIR sensors and building occupants, while ultrasonic sensors are also suitable for situations where it is impossible to keep a line of sight. Radio-frequency identification (RFID) tags are small, low-cost, and wearable devices to attach to building occupants [20,21]. In [20], in addition to adopting RFID technology to comfort building occupants, the authors also proposed a conflict-resolution architecture (CRA) to avoid conflicts of occupants' preferences. In [21], an RFID system was tested for occupancy information monitoring towards demand-driven HVAC operations. The average detection accuracy of the number of real-time occupants was located between 62–88%. Although with the use of RFID sensors it is very easy to achieve fine-grained occupancy counting, occupants are often concerned about personal privacy. The resultant poor privacy protection hinders its wide adoption in practice. With the help of computer vision algorithms, video cameras lead to fine-grained occupancy information [22]. In [22], a camera-based people detection and behavior classification system was developed and tested. Eleven classification models were built to analyze the behaviors of building occupants. In [23], a vision-based system using static cameras was built. Through video content analysis and multiple cascades of classifiers, the building occupancy count, location and activities were detected. Yet, the drawbacks of using image/video camera include poor privacy protection, limitation of the line of sight and higher cost. Wi-Fi probe request signals have been also studied to predict indoor occupancy information [24,25]. In [24], a Wi-Fi-based adaptive occupancy counting and tracking algorithm was proposed. Measured good occupancy tracking was reported. In [25], the design and implementation of Wi-Fi-enabled mobile devices were studied for fine-grained occupancy detection, tracking and counting. Despite the benefits of good privacy protection and fine-grained occupancy detection, this approach needs each occupant to carry a Wi-Fi device such as mobile phone or iPad. Furthermore, as the level of carbon oxide indirectly reflects the number of occupants, many studies have been conducted to extract the number of occupants [26,27]. In [26], the researchers measured carbon dioxide concentrations in 10 hospital patient rooms. The combination of multiple CO_2 sensors in different locations improves the accuracy of occupancy estimate. In [27], CO_2 and light sensors were selected and incorporated with a wireless sensor network for room-occupancy detection. The light sensor can be mounted on a door frame, and the prediction can be refined using CO_2 sensors. The main drawback is that the level of carbon dioxide fluctuates with HVAC operation and building status, such as unpredictable opening of doors and windows, locations of CO_2 sensors. Therefore, an accurate relationship between CO_2 level and occupancy number is not explicit. Acoustic-based occupancy estimation is another option. Yet, when occupants in an HVAC area do not make sounds, or when indoor vocal sound mixes with outdoor loud noise, the acoustic method causes the detection to fail.

In [28], acoustic energy calculation (i.e., short-time energy (STE)) was used to estimate the number of people inside a room. The proposed STE approach is non-intrusive and protects the privacy of building occupants. Yet, this work did not consider the interference of background noise. In [29], energy mode and babble speaker count methods were proposed for crowd size estimation in a party-mode room setting. Moreover, the impacts of distance between speakers and microphones were studied. In [30], measurements were conducted in six churches to study the effects of occupancy on speech transmission index values. The potential of energy savings based on occupancy-driven building operations was not involved in [30]. Based on Gaussian mixtures and hidden Markov models, an audio-based room occupancy analysis algorithm was developed in [31]. In [32], the researchers developed a networked embedded acoustic-processing system, which includes acoustic event detection, feature extraction, occupancy level models, etc. in order to estimate the occupancy level in buildings. In [33], the authors discussed three on-going research projects, which try to use sound to reduce energy consumption in buildings.

In order to take advantage of each single detection mechanism, researchers also perform multiple-sensor deployments and multi-model signal processing [34–40]. In [34], PIR sensors, CO_2 sensors, temperature sensors, acoustic sensors, volatile organic compounds (VOC) sensors, and infrared cameras were deployed in a test building. Then, the mathematical description of features was investigated

and the validity of the occupancy level estimate was demonstrated. In this example, the accuracy of occupancy detection was 84.59%. In [35], utilizing the spatial and temporal dependence of multiple sensor points, a low computational complexity sensor-fusion algorithm was developed to predict the occupancy status. Although this algorithm shows high-precision estimates of the presence or absence of occupants, it cannot be used to calculate the number of occupants. In [36], the researchers proposed an occupancy monitoring system using temperature, carbon dioxide concentration, door status, light, sound, motion, and humidity sensors. Artificial neural network (ANN) algorithms were used for multi-modal data fusion. In [37], a multi-sensor occupant detection system was developed with data analytics and fusion capabilities. In [38], in order to realize the recognition of human occupancy, multivariate sensors with a proposed feature extraction method and the most dominant sensor were presented and discussed. In [39], the researchers presented a prototype of multi-functional wireless sensor that includes five heterogeneous low-cost sensors and their system integration. The weakness of this work is the lack of multi-modal data-fusion algorithms. In [40], various emerging information technologies were reviewed and discussed. The authors pointed out the necessity and importance of studying the interaction and co-optimization of smart buildings and information technologies. Yet, this work did not present any specific multi-modal data fusion algorithm or case study. The inherent flexibility of a hybrid solution creates ample opportunities for customization in different buildings scenarios. However, the overheads of system cost, size, and design complexity are non-negligible. From a research perspective, it is still necessary to continue to analyze and optimize each individual occupancy-detection mechanism.

Even though previous studies in audio processing [28–33] have shown good prospects, these works do not consider the impacts of environmental noise on the occupancy estimation performance. Most audio-processing techniques for building occupancy estimation are suitable for outdoor quiet places, such as office buildings or research laboratories. Environmental noises from nearby traffic streets or farmers' markets may overwhelm the interior sounds made by building occupants. In these scenarios, it is necessary to further improve audio-processing algorithms to suppress outdoor noises and to maintain indoor human sounds as the main acoustic signal for occupancy extraction. This is the focus of this research. To deal with this challenge of background sound interference, a background sound-cancellation algorithm is studied and adopted in this work to enhance the impacts of human speech during acoustic-driven occupancy estimation. As there is no speech recognition or identification computations involved in our flowchart, user privacy is well protected in this work. Experimental results show that the proposed algorithm increases the average detection accuracy by approximately 11–12% in 10 typical noise environments, which results in a reduction of 3.54% in ventilation energy in a case study of building energy simulation.

2. Proposed Audio-Processing Algorithms for Building Occupancy Estimation

2.1. Audio-Processing Algorithms without Considering Outdoor Sound Interference

Two assumptions have been made in our previous work [28]: (1) indoor sound recordings are mainly human speech (excluding sounds from televisions, computers, music players, etc.); (2) the outdoor sound level is much weaker than the indoor speech level. Based on the above two assumptions, the noise from outside is considered as additive white Gaussian noise (AWGN) with a negligible magnitude and small temporary variation. Then, dedicated acoustic-based room occupancy estimation algorithms were developed for two distinct scenarios: meeting mode and party mode. In the meeting mode, where meeting participants are assumed to speak one by one, so voice sound is not coincident or mixed with each other, each speaker's voice is first recognized through acoustic signal processing and then summed up to obtain the total number of occupants. While the human voices are mixed together in the party mode, it is extremely difficult to clearly identify each occupant's voice. Instead, a feature of STE is used to estimate the total number of occupants. STE is an important feature of signal

energy within a short interval of time. The details of the audio-processing algorithm for party-mode occupancy number calculation are elaborated in our paper [28].

2.2. Audio-Processing Algorithms with Consideration of Outdoor/Background Sound Interference

The presence of loud background noise is inevitable in some places, such as busy restaurants or shopping malls. Therefore, the performance of the algorithm proposed in [26] is questionable. The algorithm in [28] takes into account of the collected background noise as a part of human sounds. As a result, the estimated number of occupants exceeds the actual number of occupants in these places. In order to solve this shortcoming, background sound cancellation algorithms are studied and adopted. Hence, clean acoustic signals with attenuated background noise are generated from raw noisy acoustic signals. Consequently, this study only assumes that the indoor sound recordings are primarily human speech (excluding sounds from televisions, computers, music players, etc.), rather than making two assumptions as in [28].

Figure 1 shows an overview of our proposed audio-processing flow. Raw acoustic signals are collected and recorded by microphones. Then, the raw acoustic signals are processed by the proposed background sound cancellation algorithm (i.e., speech-enhancement algorithm) to obtain clean acoustic signals. The clean acoustic signals are given for the STE analysis and to determine the estimated number of occupants. Since the STE analysis does not identify or interpret human speech, only the time-dependent acoustic energy spectrum is used for occupancy estimation, so the proposed work helps to protect the privacy of building occupants. In this section, the background sound-cancellation algorithm, which is also called the speech-enhancement algorithm, is derived and introduced.
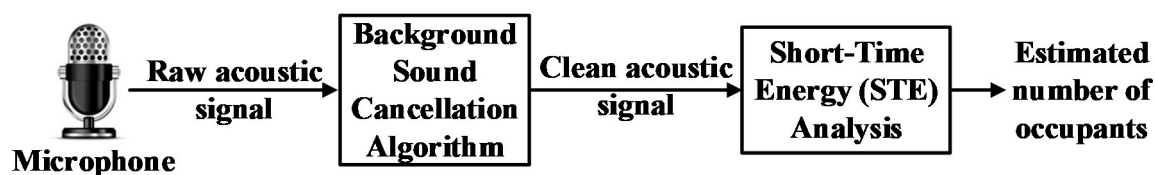

Figure 1. Overview of audio-processing algorithms with background sound-cancellation algorithm.

As been mentioned earlier, background noise severely overwhelms and corrupts human speech, so the occupancy quantity obtained using the STE approach is overestimated. Therefore, researchers investigate algorithms to detect the present noise level and to eradicate it efficiently. Due to their non-stationary nature, high-level background noises are hard to accurately describe and model. Although time-domain statistical models of probability distributions of speech and noise are attractive [41–43], one of the major limitations of these statistical models is the need for a priori knowledge of speech or noise [44]. Moreover, these statistical models mainly describe the long-term characteristics of speech or noise, which do not accurately characterize and reflect short-term features. In the literature [45], the researchers have found that Teager energy operator (TEO) could detect and model speech in an analytical approach [46], so it does not depend on a priori knowledge of speech or noise. To date, there have been a number of researchers to adopt the TEO method in human speech processing. On the other hand, in the area of acoustic signal processing, wavelet packet transform is also found to be a useful technique. For example, in [46], a speech-enhancement method was presented considering both the time and scale dependency of wavelet thresholds. In [47], the researchers presented a speech-enhancement algorithm using TEO and adaptive thresholds in the wavelet packet domain.

In this work, wavelet packet transform (WPT) and Teager energy operator (TEO) are used to reduce speech distortion from high background-noise environments [48]. The presented algorithm in [48] is based on two-dimensional TEO in the wavelet packet transform domain, where the human sound is treated as amplitude or frequency modulated signals by noise signals. To overcome the challenge of effective signal separation between human speech and noise, a state-of-the-art speech-and-noise separation algorithm [48] is selected and adopted in this study, where an improved speech presence

probability (SPP) estimator is established accordingly. Even though both independent and intersectional 2D TEOs have been developed in [48], for computational simplicity only intersectional ones are adopted. The intersectional 2D TEO kernels with respect to the horizontal–vertical direction are modeled in [48] as,

$$TH\{w(k,t)\} = \left\{\frac{\partial w}{\partial k}\right\}^2 + \left\{\frac{\partial w}{\partial t}\right\}^2 - w\left\{\frac{\partial^2 w}{\partial k^2} + \frac{\partial^2 w}{\partial t^2}\right\}^2$$

The intersectional 2D TEO kernels with respect to the diagonal direction are modeled as,

$$TD\{w(k,t)\} = 2\left\{\frac{\partial w}{\partial k}\right\}\left\{\frac{\partial w}{\partial t}\right\} - w\left\{\frac{\partial^2 w}{\partial k \partial t} + \frac{\partial^2 w}{\partial t \partial k}\right\}$$

Here $w(k,t)$ is the wavelet packet transform coefficient. Frequency and time are represented as k and t, respectively. The use of a contrast parameter s introduces the discrete form of nonlinear 2D versions:

$$T^{2,H}(k, t, s) = 2w(k,t)^{\frac{2}{s}} - (w(k-\Delta k, t)w(k+\Delta k), t)^{\frac{1}{s}} - (w(k, t-\Delta t)w(k, t+\Delta t))^{\frac{1}{s}}$$

$$T^{2,D}(k, t, s) = 2w(k,t)^{\frac{2}{s}} - (w(k-\Delta k, t+\Delta t)w(k+\Delta k), t-\Delta t)^{\frac{1}{s}} - (w(k-\Delta k, t-\Delta t)w(k+\Delta k, t+\Delta t))^{\frac{1}{s}}$$

Here Δt and Δk are the time and frequency lag window parameters. Then, the outlines of the energy distribution of 2D intersectional TEOs are modeled in [48] as,

$$S^{2,1}(k,t,s) = \frac{\left|H(k,t) * T^{2,H}(k,t,s)\right|}{\max(|H(k,t) * T^{2,H}(k,t,s)|)}$$

$$S^{2,2}(k,t,s) = \frac{\left|H(k,t) * T^{2,D}(k,t,s)\right|}{\max(|H(k,t) * T^{2,D}(k,t,s)|)}$$

Here $H(k,t)$ is a low-pass filter and the operator * indicates a convolution operation. As harmonic signals are represented as higher energy density and random noise is represented as lower-level energy density in 2D TEOs, the energy density obtained from TEOs generally reveals whether speech components exist or not [48]. In [48], the normalized outline of energy distribution for intersectional TEOs is applied as SPP estimators. The introduction of the proposed 2D TEOs enables the detection of speech components. Note this SPP estimation is computed without prior knowledge of speech and background noise. Therefore, it is preferred for short-term acoustic signal processing for occupancy count estimation. These 2D intersectional TEO-based SPP estimators are very sensitive to background noise. To avoid the over-than-enough sensitivity for SPP estimation, two lag window parameters Δk and Δt are used to derive the SPP values. $SPPT_l$ represents local SPP and $SPPT_g$ represents global SPP. Therefore, a new SPP estimator is modeled in [48] as

$$SPP(k,t,s) = SPPT_l(k,t,\Delta k_1, \Delta t_1, s) \cdot SPPT_g(k,t,\Delta k_2, \Delta t_2, s)$$

Here Δk_1 and Δt_1 are selected as unit values to represent the high resolution of a lag window, while Δk_2 and Δt_2 are selected as larger values to represent the low resolution of a lag window.

An advanced speech estimator was presented in [48], which is based on a generalized speech model in the WPT domain. In [48], a signal model is constructed of $w_y(k,t) = w_x(k,t) + w_r(k,t)$, where $w_y(k,t)$, $w_x(k,t)$, $w_r(k,t)$ are WPT coefficients in k-th sub-band at time t extracted from noisy speech, clean speech, and noise signal, respectively. Assuming $w_x(k,t)$ and $w_r(k,t)$ are independent on time and frequency from a statistical point of view, the minimum mean-square error (MMSE) estimator is modeled in [48] as

$$E(X|Y) = \frac{\int_{-\infty}^{+\infty} X p(Y|X) p(X) dX}{\int_{-\infty}^{+\infty} p(Y|X) p(X) dX} = \frac{\int_{-\infty}^{+\infty} X p_r(Y-X) p_x(X) dX}{\int_{-\infty}^{+\infty} p_r(Y-X) p_x(X) dX}$$

Here, X and Y represent the coefficients, $p_x(X)$ is assumed to follow the generalized gamma distribution described in [41] and $p_r(Y - X)$ is assumed to follow the Gaussian distribution described in [41]. According to [41,48], the SPP estimator can be further modeled as,

$$E(X|Y) = \sigma_r v \frac{\left[\exp\left(\frac{1}{4}Y_-^2\right)D_{-(v+1)}(Y_-)\right] - \left[\exp\left(\frac{1}{4}Y_+^2\right)D_{-(v+1)}(Y_+)\right]}{\left[\exp\left(\frac{1}{4}Y_-^2\right)D_{-v}(Y_-)\right] + \left[\exp\left(\frac{1}{4}Y_+^2\right)D_{-v}(Y_+)\right]}$$

$$Y_\pm = \beta\sigma_r \pm \frac{Y}{\sigma_r}$$

Here $D_{-v}(\cdot)$ is a parabolic cylinder function of order v, and σ_r is the estimated noise variance. Then, the results of the MMSE estimator goes through an inverse WPT computation and finally generates clean human speech.

2.3. Occupancy-Counting Algorithm Implementation with Background Noise-Cancellation Feature

In this work, these models of two-dimensional Teager energy operator (TEO) and wavelet packet transform (WPT) are implemented in MATLAB codes. The flowchart in Figure 2 illustrates the details of entire acoustic signal processing for building occupancy count estimation.

First, noisy speech is collected using microphones and is processed using the wavelet packet transform technique. Then, two-dimensional intersectional Teager energy operators are calculated, and the results are provided for both global and local SPP estimation. Then, the minimum mean-square error (MMSE) estimation is performed to effectively separate noise signals and clean human speech. Next, the clean speech signals are processed using the short-time energy calculation in our previous publication [28]. Finally, the building occupancy number is estimated accurately. Every step in this flowchart is implemented and run in MATLAB codes. From Figure 2, we can see that the flowchart does not involve speech recognition or identification. Therefore, the user privacy issue is eliminated in this study.
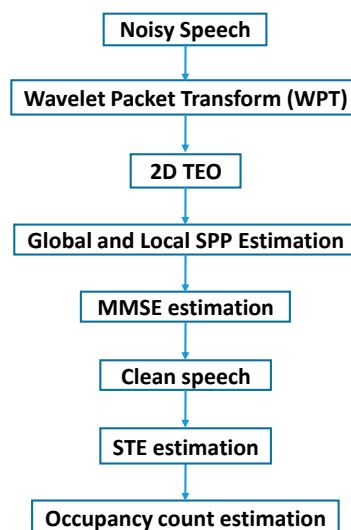

Figure 2. Flowchart of overall acoustic signal processing for occupancy count estimation.

3. Experimental Results

In this section, 100 clean speech files with an individual duration of 25 s are listened to by researchers in order to identify the exact number of speakers for each speech file. Next, as shown in Figure 3, these clean speech files are mixed with added noise files, either containing Gaussian noise or measured noise recorded from several noisy places, including airports, cafeterias, construction sites,

factories, streets, restaurants, subways, trains, flights, and exhibitions. Then, these mixed sound files are processed by the acoustic signal-processing algorithm in [28] and the proposed acoustic signal processing in this work, respectively. Finally, the exact occupancy number and two estimated occupant numbers are compared and discussed in this section.
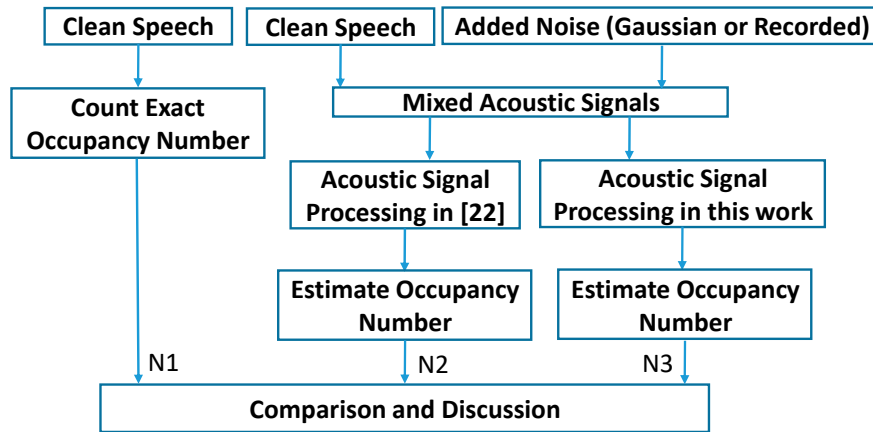

Figure 3. Comparison among occupancy number identification and estimation results.

3.1. Occupancy-Counting Results for Gaussian White Noise Mixed Human Speech

In a real environment, noise is often not caused by a single source, but a complex of many different sources. Assume that real noise is the addition of random variables with a very large number of different probability distributions, and that each random variable is independent. According to the central limit theorem, their normalized sum increases with the number of noise sources and is close to a Gaussian distribution. As a typical acoustic noise type, the probability density function of a Gaussian white noise follows a normal distribution. The room occupancy estimation performance is firstly evaluated when the background sound is assumed as Gaussian white noise.

Figure 4 plots the experimental results of using the STE feature in estimating the number of speakers. Figure 4a shows the STE-based acoustic processing results in a high accuracy of room occupancy estimation, when there is no background noise. It is clear that after processing 25 s of the recorded acoustic signals, the estimation accuracy is close to 1, which indicates a very small error. When there is a strong background white Gaussian noise source (e.g., 70 dB), the background noise is louder than human speech, thus, the estimation accuracy is drastically decreased as shown in Figure 4b. Especially for the cases of 10 speakers and 20 speakers, the estimation performance is very bad. With the proposed background sound-cancellation algorithm introduced in Section 2, the estimation accuracy is significantly improved and recovered as shown in Figure 4c. Comparing Figure 4b,c, at the time instance of 25 s, the estimation accuracy with our proposed background noise enhancement algorithm is boosted by at least 30%. Therefore, this proves the efficacy of using an appropriate speech enhancement algorithm in the overall signal processing of occupancy estimation.
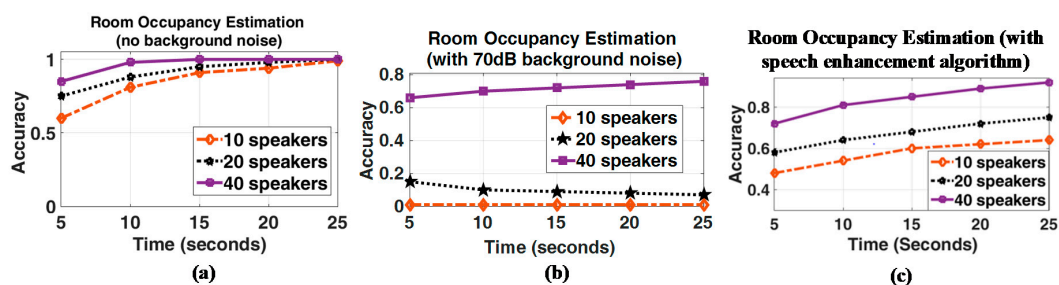

Figure 4. Speaker number estimation results assuming background sound is white Gaussian noise.

3.2. Occupancy-Counting Results for Gaussian White Noise Mixed Human Speech

In addition to the investigation of the noise-cancellation performance for Gaussian white noise, background sounds from 10 typical noisy places are recorded, including airports, cafeterias, construction sites, factories, streets, restaurants, subways, trains, flights, and exhibitions. These noise files are available to download from (https://sites.google.com/site/qianhuangshomesite/). Assuming that human speech is 60 dB, Tables 2 and 3 show the comparison of occupancy-estimation results before and after applying the proposed background sound cancellation algorithm to the 65 dB and 55 dB background sounds, respectively. It is observed that these 10 noisy locations lead to an average improvement in occupancy estimation by approximately 11~12%, which is lower than the performance enhancement in a Gaussian white noise environment. This is because a Gaussian white noise is a random signal with equal intensity at different frequencies, so its power spectral density is constant and relatively easy to remove. In contrast, actually recorded noise from 10 typical locations includes significant unpredictable variations in power spectral density. Therefore, the proposed background sound cancellation algorithm exhibits better performance in processing speech signals that are mixed with Gaussian white noise.

Table 2. Comparison of occupancy estimation accuracy before and after applying background sound cancellation algorithm, assuming 65 dB background sound and 60 dB human speech.

Noise Environments	Average Occupancy Estimation Accuracy without Background Sound Cancellation	Average Occupancy Estimation Accuracy with Background Sound Cancellation	Accuracy Improvement (%)
Airport	0.71	0.82	15.96
Cafeteria	0.71	0.81	13.08
Construction site	0.73	0.80	9.55
Factory	0.70	0.80	13.27
Street	0.74	0.82	11.76
Restaurant	0.72	0.82	13.36
Subway	0.72	0.8	10.65
Train	0.75	0.8	7.59
Flight	0.72	0.81	12.5
Exhibition	0.76	0.8	5.26
Average results	0.726	0.808	11.3

Table 3. Comparison of occupancy estimation accuracy before and after applying background sound cancellation algorithm, assuming 55 dB background sound and 60 dB human speech.

Noise Environments	Average Occupancy Estimation Accuracy without Background Sound Cancellation	Average Occupancy Estimation Accuracy with Background Sound Cancellation	Accuracy Improvement (%)
Airport	0.74	0.8	9.05
Cafeteria	0.68	0.71	5.42
Construction site	0.72	0.81	12.04
Factory	0.71	0.83	16.43
Street	0.66	0.82	25.38
Restaurant	0.73	0.83	13.24
Subway	0.71	0.81	14.08
Train	0.76	0.8	5.73
Flight	0.74	0.81	8.97
Exhibition	0.73	0.8	9.59
Average results	0.718	0.808	12.0

4. Building Energy Simulation Using EnergyPlus

This study is conducted using EnergyPlus software [49], which has been developed and released by the U.S. Department of Energy. The university bookstore in the Student Service Building of Southern Illinois University Carbondale campus is chosen, and its energy consumption is used as a baseline.

As depicted in Figure 5, the university bookstore is surrounded by a billiards room, bowling room, student dining area, McDonald's, and McDonald's seating area. The sounds made from activities in these nearby rooms are viewed as background noise to the university bookstore, and the measured worse case of background sound level is no higher than 65 dB.

The prerequisite knowledge for baseline modeling includes blueprints for original construction, historical energy bills, and current operating data in the building automation system. For example, the exterior wall consists of four layers, which are made of material M01 100 mm brick, M15 200 mm heavyweight concrete, I02 50 mm insulation board, and G01a 19 mm gypsum board, respectively. The interior wall is made of G01a 19 mm gypsum board. In our EnergyPlus simulations, the occupancy schedule is based on the percentage of occupants that occupy the bookstore on weekdays. According to the daily occupancy information from the building manager, a dedicated occupancy schedule is created and used in the EnergyPlus software. All useful data and statements provided by the physical plant engineers will be imported into an input file for EnergyPlus, which takes into account building envelope, windows, lighting, HVAC equipment, and weather. The output variables include the fan electric energy, zone air temperature, heating-coil electric energy, cooling-coil electric energy, site wind speed, site wind direction, site outdoor air humidity, zone exterior and interior windows total transmitted beam solar radiation rates, etc. These output variables are recorded for a whole year hourly, daily, and monthly, respectively.
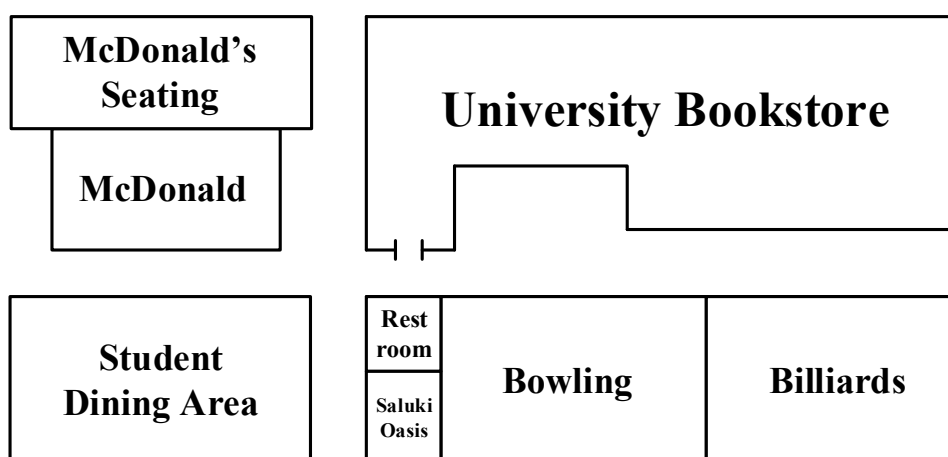

Figure 5. Floor plan of simulated university bookstore and its neighboring stores.

Assuming the average occupancy-detection accuracy is 70% (without background sound cancellation) and 80% (with background sound cancellation), Figure 6 shows the required monthly ventilation electricity for default maximum occupancy, real occupancy, estimated occupancy with and without background sound cancellation, respectively. Compared with the default maximum occupancy set points for HVAC equipment, results of real occupancy and the previously estimated occupancy estimation in [28] in Figure 6 achieve an average energy reduction of 14.2% and 8.6% in ventilation electricity, respectively. Then, in contrast with the previously developed estimation in [28], the adoption of our proposed background sound-cancellation algorithm further reduces ventilation energy consumption by 3.54%. This result validates the necessity of developing the background sound-cancellation algorithm. As shown in Figure 7, simulation results of cooling and heating electricity are not sensitive to occupancy number. This is because occupancy number is not the main factor to control HVAC equipment for cooling and heating in this building.

Figures 8 and 9 show the simulation results of daily ventilation electricity for two weeks in January and July, respectively. From Figure 8, it is clear that for typical winter days, the difference with or without using the adopted background sound cancellation is consistent. It indicates an average of 2.22% ventilation energy reduction when the background sound-cancellation algorithm is adopted.

The weekdays are from 5 January to 9 January, when this bookstore requires more ventilation energy than weekend days. From Figure 9, it is observed that for typical summer days, the average ventilation energy reduction is 5.67%, when the background sound cancellation algorithm is used. The weekdays are from 3 July to 7 July, when this bookstore requires more ventilation energy than weekend days.
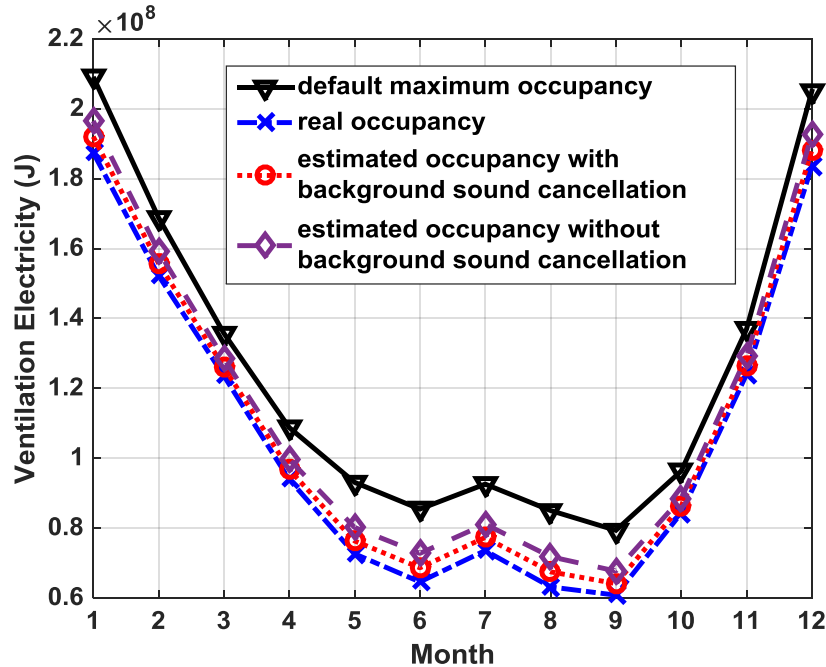
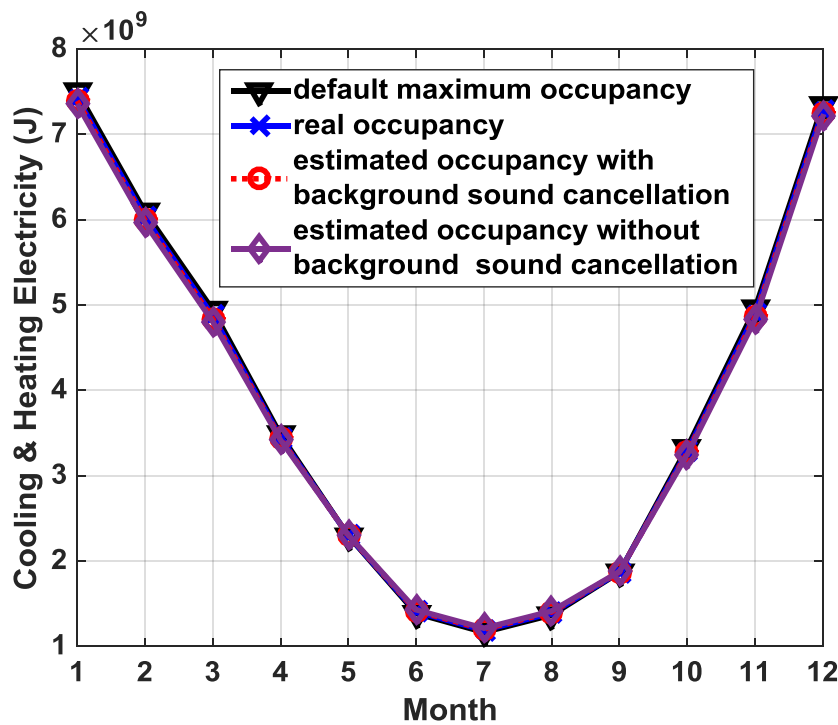**Figure 6.** Simulation results of monthly ventilation electricity for a one-year period.

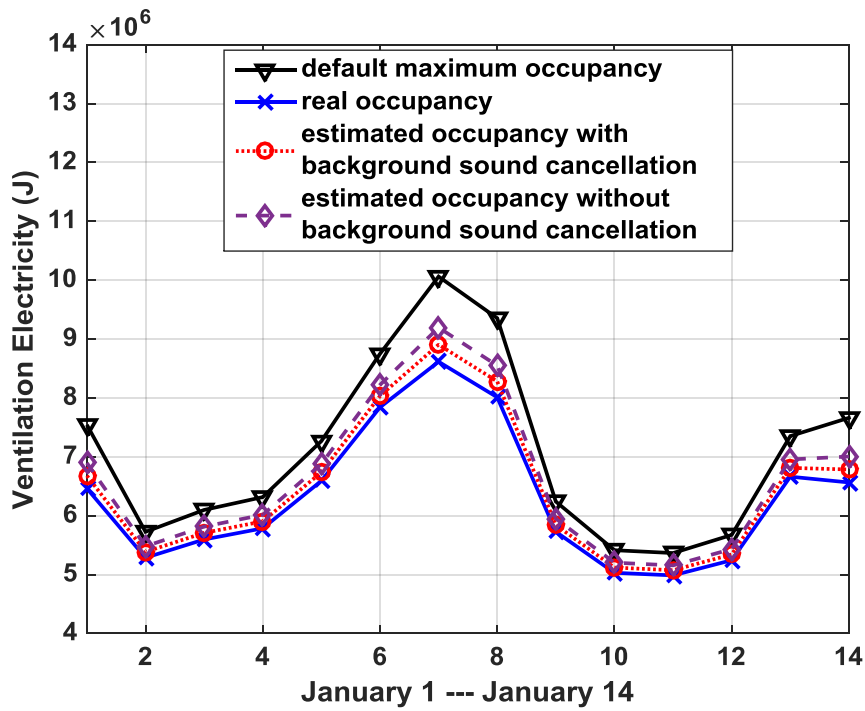**Figure 7.** Simulation results of monthly heating and cooling electricity for a one-year period.

Figure 8. Simulation results of daily ventilation electricity for two weeks in January.
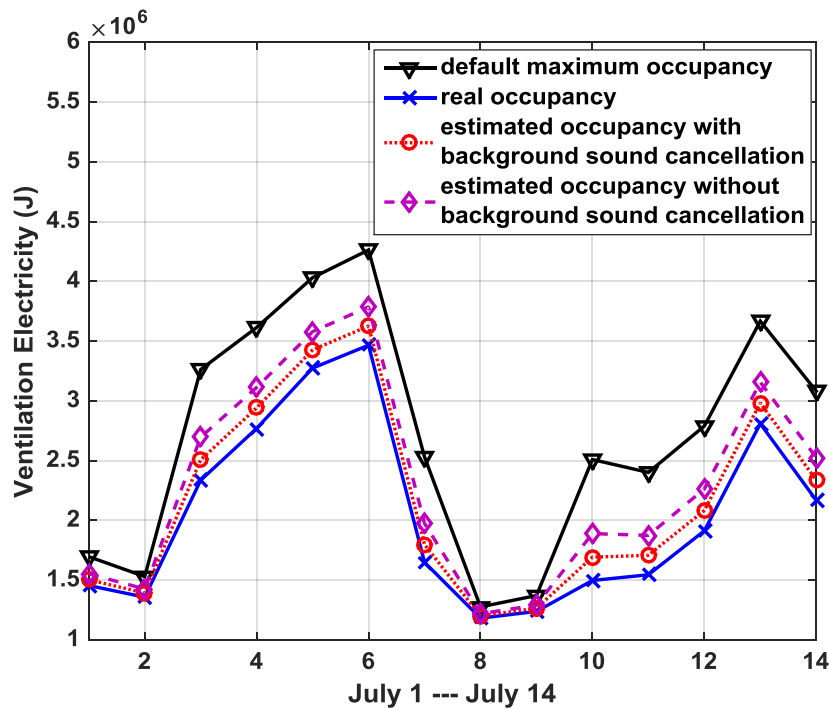

Figure 9. Simulation results of daily ventilation electricity for two weeks in July.

Figures 10 and 11 show the simulation results of hourly ventilation electricity for one day in January and July, respectively. From Figure 10, it is easy to see that for typical winter days, the difference with or without the adopted background sound cancellation algorithm is consistent, and the average ventilation energy reduction is 3.14%. Moreover, it was observed that the ventilation electricity at night is above 3×10^5 J, while it is below 3×10^5 J at daytime. This is because the outdoor temperature on winter nights

is much lower than at daytime, and therefore the HVAC equipment consumes more energy at night. From Figure 11, this is shown for typical summer days, the background noise cancellation algorithm helps to reduce the ventilation energy by 3.74%. As shown in Figure 11, most of this ventilation energy reduction is achieved from 9 a.m. to 6 p.m.
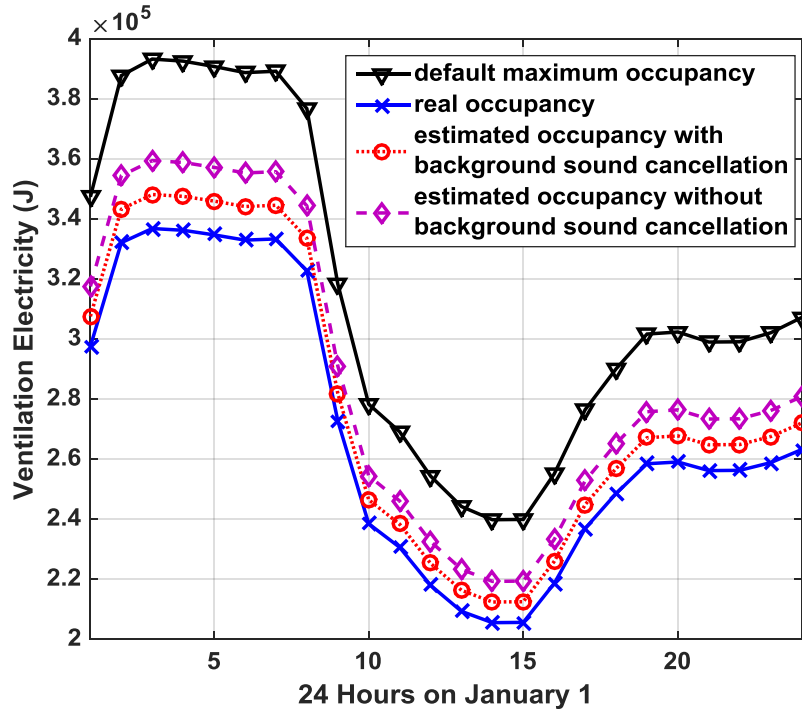
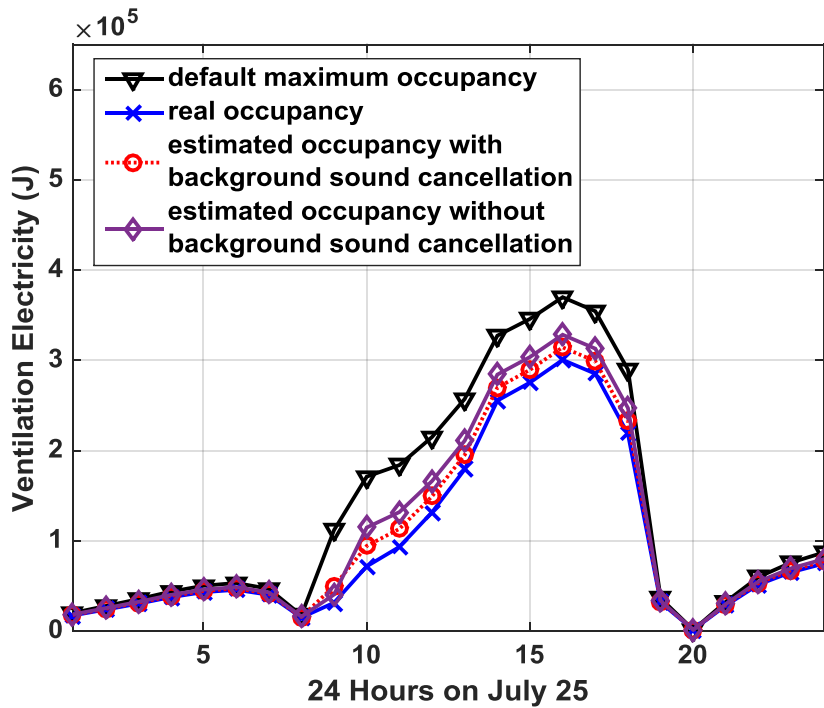**Figure 10.** Simulation results of hourly ventilation electricity for 1 January.

Figure 11. Simulation results of hourly ventilation electricity for 25 July.

5. Conclusions

With the adoption of various information technologies in next-generation smart buildings, demand-driven building operation is very attractive for reducing energy consumption in buildings. Therefore, it is imperative to investigate and develop occupancy recognition and counting techniques. While several promising occupancy-estimation techniques based on carbon dioxide sensors, RFID sensors, etc. are being explored, each of them has significant issues that need to be addressed. Researchers envision that future building occupancy counting techniques will have user-transparency, high accuracy, a low failure rate, easy maintenance, low complexity, good privacy protection, and low price. Occupancy estimation based on the acoustic processing of sound recorded in a room or thermal zone is low cost, non-intrusive, and has good detection accuracy in quiet environments. However, background noise in some noisy locations (such as restaurants, trains, and factory) mixes together or even overwhelms indoor human voice, thus degrading the occupancy estimation accuracy. To deal with this challenge of background sound interference, a background sound cancellation algorithm is adopted to enhance the impacts of human speech during acoustic-driven occupancy estimation. As there is no speech recognition or identification computations involved in our flowchart, user privacy is well protected in this work. Experimental results show that the proposed algorithm increases the average detection accuracy by approximately 11–12% in 10 typical noise environments, which results in a reduction of 3.54% in ventilation energy in a case study of building energy simulation. In this study, the motivation is not to prove that the proposed acoustic-based method using background noise-cancellation algorithm is more accurate or superior than other existing occupancy detection methods. The purpose is to investigate and evaluate the performance of combining STE calculation and background noise cancellation, and to show its potential to save building operation energy and costs.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Lombard, L.; Ortiz, J.; Pout, C. A review on buildings energy consumption information. *Energy Build.* **2008**, *40*, 394–398. [CrossRef]
2. Graham, C. *High-Performance HVAC*; Technical Reports; Viridian Energy & Environmental Inc: Norwalk, CT, USA, 2018; Available online: https://www.wbdg.org/resources/high-performance-hvac (accessed on 1 June 2018).
3. Klein, L.; Kwak, J.; Kavulya, G.; Jazizadeh, F.; Gerber, B.; Varakantham, P.; Tambe, M. Coordinating occupant behavior for building energy and comfort management using multi-agent systems. *Autom. Constr.* **2012**, *22*, 525–536. [CrossRef]
4. Yan, D.; Brien, W.; Hong, T.; Feng, X.; Gunay, H.; Tahmasebi, F.; Mahdavi, A. Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy Build.* **2015**, *107*, 264–278. [CrossRef]
5. Levin, H. Designing for people: What do building occupants really want? In Proceedings of the Healthy Buildings, Singapore, 7–11 December 2003; pp. 1–18.
6. Rasheed, E.; Byrd, H.; Money, B.; Mbachu, J.; Egbelakin, T. Why are naturally ventilated office spaces not popular in New Zealand? *Sustainability* **2017**, *9*, 902. [CrossRef]
7. Jin, J.; Gubbi, J.; Marusic, S.; Palaniswami, M. An information framework for creating a smart city through Internet of Things. *IEEE Internet Things J.* **2014**, *1*, 112–121. [CrossRef]
8. Zanella, A.; Bui, N.; Castellani, A.; Vangelista, L.; Zorzi, M. Internet of Things for Smart Cities. *IEEE Internet Things J.* **2014**, *1*, 22–32. [CrossRef]

9.   Man, K.; Yue, Y.; Lu, C.; Huang, Q. System design, analysis, and optimization of Li-Fi based energy harvesting systems for "Internet of Things" Applications. In Proceedings of the International Conference on Internet of Things and Convergence, Seoul, Korea, 26–28 October 2015; pp. 189–192.

10.   DOE APRA-E Funding Announcement DE-FOA-0001737, Saving Energy Nationwide in Structures with Occupancy Recognition (Sensor). Available online: https://arpa-e-foa.energy.gov/Default.aspx?Search=DE-FOA-000173 (accessed on 1 June 2018).

11.   Agarwal, Y.; Balaji, B.; Gupta, R.; Lyles, J.; Wei, M.; Weng, T. Occupancy-driven energy management for smart building automation. In Proceedings of the ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building, Zurich, Switzerland, 3–5 November 2010; pp. 1–6.

12.   Labeodan, T.; Zeiler, W.; Boxem, G.; Zhao, Y. Occupancy measurement in commercial office buildings for demand-driven control applications-a survey and detection system evaluation. *Energy Build.* **2015**, *93*, 303–314. [CrossRef]

13.   Pritoni, M.; Woolley, J.; Modera, M. Do occupancy-responsive learning thermostats save energy? A field study in university residence halls. *Energy Build.* **2016**, *127*, 469–478. [CrossRef]

14.   Yang, Z.; Gerber, B. How does building occupancy influence energy efficiency of HVAC systems? *Energy Procedia* **2016**, *88*, 775–780. [CrossRef]

15.   Ardakanian, O.; Bhattacharya, A.; Culler, D. Non-intrusive techniques for establishing occupancy related energy savings in commercial buildings. In Proceedings of the ACM International Conference on Systems for Energy-Efficient Built Environments, Palo Alto, CA, USA, 16–17 November 2016; pp. 21–30.

16.   Raykov, Y.; Ozer, E.; Dasika, G. Predicting room occupancy with a single passive infrared (PIR) sensor through behavior extraction. In Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, 12–16 September 2016; pp. 1016–1027.

17.   Liu, H.; Lin, H.; Wang, K.; Wang, Y. A novel chopped pyroelectric infrared sensor for detecting the presence of stationary and moving occupants. In Proceedings of the ASME Conference on Smart Materials, Adaptive Structures and Intelligent Systems, Snowbird, UT, USA, 18–20 September 2017; p. V002T05A006.

18.   Shih, O.; Rowe, A. Occupancy estimation using ultrasonic chirps. In Proceedings of the International Conference on Cyber-Physical Systems, Seattle, WA, USA, 14–16 April 2015; pp. 149–158.

19.   Shih, O.; Lazik, P.; Rowe, A. AURES: A wide-band ultrasonic occupancy sensing platform. In Proceedings of the International Conference on Systems for Energy-Efficient Built Environments, Palo Alto, CA, USA, 16–17 November 2016.

20.   Lee, H.; Jae, S.C. A conflict resolution architecture for the comfort of occupants in intelligent office. In Proceedings of the IEEE International Conference on Intelligent Environments, Seattle, WA, USA, 21–22 July 2008; pp. 1–8.

21.   Li, N.; Carlis, G. Measuring and monitoring occupancy with an RFID based system for demand-driven HVAC operations. *Autom. Constr.* **2012**, *24*, 89–99. [CrossRef]

22.   Ahmed, H.; Faouzi, B.; Caelen, J. Detection and classification of the behavior of people in an intelligent building by camera. *Int. J. Smart Sens. Intell. Syst.* **2013**, *6*, 1317–1342. [CrossRef]

23.   Benezeth, Y.; Laurent, H. Towards a sensor for detecting human presence and characterizing activity. *Energy Build.* **2011**, *43*, 305–314. [CrossRef]

24.   Ciftler, B.; Dikmese, S.; Guvenc, I.; Akkaya, K.; Kadri, A. Occupancy Counting with Burst and Intermittent Signals in Smart Buildings. *IEEE Internet Things J.* **2018**, *5*, 724–735. [CrossRef]

25.   Zou, H.; Jiang, H.; Yang, J.; Xie, L.; Spanos, C. Non-Intrusive Occupancy Sensing in Commercial Buildings. *Energy Build.* **2017**, *154*, 633–643. [CrossRef]

26.   Dedesko, S.; Stephens, B.; Gilbert, J.; Siegel, J. Methods to assess human occupancy and occupant activity in hospital patient rooms. *Build. Environ.* **2015**, *90*, 136–145. [CrossRef]

27.   Mao, C.; Huang, Q. Occupancy estimation in smart building using hybrid CO_2/light wireless sensor network. *J. Appl. Sci. Arts* **2016**, *1*, 5.

28.   Huang, Q.; Ge, Z.; Lu, C. Occupancy estimation in smart buildings using audio-processing techniques. In Proceedings of the International Conference on Computing in Civil and Building Engineering, Osaka, Japan, 6–8 July 2016; pp. 1413–1420.

29.   Chen, S.; Epps, J.; Ambikairajah, E.; Le, P. An investigation of crowd speech for room occupancy estimation. In Proceedings of the Interspeech, Stockholm, Sweden, 20–24 August 2017; pp. 324–328.

30. Desarnaulds, V.; Carvalho, A.; Monay, G. Church acoustics and the influence of occupancy. *Build. Acoust.* **2002**, *9*, 29–47. [CrossRef]

31. Valle, R. ABROA: Audio-based room-occupancy analysis using Gaussian mixtures and hidden Markov models. In Proceedings of the Future Technologies Conference, San Francisco, CA, USA, 6–7 December 2016; pp. 1270–1274.

32. Uziel, S.; Elste, T.; Kattanek, W.; Hollosi, D.; Gerlach, S.; Goetze, S. Networked embedded acoustic processing system for smart building applications. In Proceedings of the Design and Architectures for Signal and Image Processing, Cagliari, Italy, 8–10 October 2013; pp. 349–350.

33. Kelly, B.; Hollosi, D.; Cousin, P.; Leal, S.; Iglar, B.; Cavallaro, A. Application of acoustic sensing technology for improving building energy efficiency. *Procedia Comput. Sci.* **2014**, *32*, 661–664. [CrossRef]

34. Ekwevugbe, T.; Brown, N.; Pakka, V. Real-time building occupancy sensing for supporting demand driven HVAC operations. In Proceedings of the International Conference for Enhanced Building Operations, Montreal, QC, Canada, 8–11 October 2013; pp. 1–58.

35. Hsiao, R.; Lin, D.; Lin, H. Room occupancy determination using multimodal sensor fusion. *Sens. Mater.* **2015**, *27*, 605–610.

36. Yang, Z.; Li, N.; Gerber, B. A non-intrusive occupancy monitoring system for demand driven HVAC operations. In Proceedings of the Construction Research Congress, West Lafayette, IN, USA, 21–23 May 2012; pp. 828–837.

37. Nasir, N.; Palani, K.; Chugh, A.; Prakash, V. Fusing sensors for occupancy sensing in smart buildings. In *International Conference on Distributed Computing and Internet Technology*; Springer: Cham, Switzerland, 2015; pp. 73–92.

38. Ang, I.; Salim, F.; Hamilton, M. Human occupancy recognition with multivariate ambient sensors. In Proceedings of the IEEE International Conference on Pervasive Computing and Communication Workshops, Sydney, Australia, 14–18 March 2016; pp. 1–6.

39. Huang, Q.; Mao, C.; Chen, Y. A compact and versatile wireless sensor prototype for affordable intelligent sensing and monitoring in smart buildings. In Proceedings of the ASCE International Workshop on Computing in Civil Engineering, Seattle, WA, USA, June 25–27 2017; pp. 155–161.

40. Huang, Q.; Lu, C.; Chen, K. Smart building applications and information systems hardware co-design. *Big Data Anal. Sens. Netw. Collected Intell.* **2017**, 225–240. [CrossRef]

41. Erkelens, J.S.; Hendriks, R.C.; Heusdens, R.; Jensen, J. Minimum mean-square error estimation of discrete Fourier coefficients with generalized gamma priors. *IEEE Trans. Audio, Speech Lang. Process.* **2007**, *15*, 1741–1752. [CrossRef]

42. Ephraim, Y.; Malah, D. Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator. *IEEE Trans. Audio Speech Lang. Process.* **1984**, *32*, 1109–1121. [CrossRef]

43. Park, J.; Kim, J.; Chang, J.; Jin, Y.; Kim, N. Estimation of speech absence uncertainty based on multiple linear regression analysis for speech enhancement. *Appl. Acoust.* **2015**, *87*, 205–2011. [CrossRef]

44. Cohen, I. Speech enhancement using a non-causal a priori SNR estimator. *IEEE Signal. Process. Lett.* **2004**, *11*, 725–728. [CrossRef]

45. Kandia, V.; Stylianou, Y. Detection of sperm whale clicks based on the teager-kaiser energy operator. *Appl. Acoust.* **2006**, *11*, 1144–1163. [CrossRef]

46. Bahoura, M.; Rouat, J. Wavelet speech enhancement based on time-scale adaptions. *Speech Commun.* **2006**, *48*, 1620–1637. [CrossRef]

47. Sanam, T.; Shahnaz, C. Noisy speech enhancement based on an adaptive threshold and a modified hard thresholding function in wavelet packet domain. *Digit. Signal. Process.* **2013**, *23*, 941–951. [CrossRef]

48. Sun, P.; Qin, J. Wavelet packet transform based speech enhancement via two-dimensional SPP estimator with generalized gamma priors. *Achiev. Acoust.* **2016**, *41*, 579–590. [CrossRef]

49. EnergyPlus Energy Simulation Software. Available online: https://energyplus.net/ (accessed on 1 June 2018).