Bathroom Activity Monitoring Based on Sound

Jianfeng Chen, Alvin Harvey Kam, Jianmin Zhang, Ning Liu, and Louis Shue

Institute for Infocomm Research,
21 Heng Mui Keng Terrace, Singapore, 119613
jfchen@i2r.a-star.edu.sg

Abstract. In this paper an automated bathroom activity monitoring system based on acoustics is described. The system is designed to recognize and classify major activities occurring within a bathroom based on sound. Carefully designed HMM parameters using MFCC features are used for accurate and robust bathroom sound event classification. Experiments to validate the utility of the system were performed firstly in a constrained setting as a proof-of-concept and later in an actual trial involving real people using their bathroom in the normal course of their daily lives. Preliminary results are encouraging with the accuracy rate for most sound categories being above 84%. We sincerely believe that the system contributes towards increased understanding of personal hygiene behavioral problems that significantly affect both informal care-giving and clinical care of dementia patients.

1 Introduction

The world is rapidly graying. Older adults already constitute one-fifth of the population of much of Western Europe and Japan [1]. Here in Singapore, it is estimated that one in five persons will be over 65 years old in 30 years' time. Older people are known to be the most expensive demographic group to be treated as longevity gives rise to costly age-related disabilities and chronic diseases.

A major challenge in ensuring sustainable healthcare costs and maintaining the quality of life for the elderly is to allow senior citizens to live independently in their own homes for as long as possible while facilitating informal care-giving. A critical element in ensuring the well being of the elderly, especially those afflicted by dementia and other forms of cognitive decline, is an accurate account of the subject's physical, behavioral and psychosocial functioning. This is achieved in some ways through a detailed understanding of their activities of daily living (ADL).

One class of ADL of utmost concern to both caregivers and clinicians is personal hygiene activities occurring within the private confines of the bathroom. It is quite common for people with dementia to forget about, or lose interest in, bathing and other personal hygiene activities. There were various reasons for this [2]: a) washing and cleaning oneself are intimate, private activities. People with dementia may feel particularly embarrassed if they are incontinent, and may refuse to bathe in denial of the problem; b) the person may feel uncomfortable. The room may be too hot or cold, or produce feelings of claustrophobia which confuses them; c) getting undressed, having a wash and brushing teeth can be overwhelmingly complex tasks because of

H.W. Gellersen et al. (Eds.): PERVASIVE 2005, LNCS 3468, pp. 47 – 61, 2005.

the many steps involved; d) some people with dementia may have a changed sense of perception of hot and cold water, caused by damage to the hypothalamus. They may also feel a different sensation from water; e) fear of falling may be another problem. Feeling out of control and powerless may add to a person shunning all personal hygiene activities within the bathroom altogether.

The understanding of personal hygiene behavioral patterns and problems thus significantly affect both informal care-giving and clinical care. The current practice in the field of elderly care is to obtain behavioral information through observer or self reporting. Direct behavioral observations of bathroom and toileting activities by clinicians or even family caregivers are embarrassing and even humiliating due to privacy reasons. Due to similar privacy reasons especially during the testing phase, the use of video sensors is inappropriate. Self-reporting meanwhile is not reliable for dementia patients suffering from cognitive decline.

The automated bathroom monitoring system we proposed in this paper is developed with the motivation to address the above unmet needs. The system is able to objectively capture behavioral patterns within the bathroom accurately and free from manpower constraints. Further benefits lie in the need for nighttime observations, such as in clinical studies of nocturia, nocturnal polyuria or nocturnal incontinence where self-reporting for even cognitively normal subjects is troublesome and highly inaccurate. Although less intrusive than using video, we acknowledge that the system may still result in some loss of privacy for the individual being studied. We will address this issue in depth further on.

1.1 Review of Related Work

Automated activity monitoring within bedrooms and bathrooms is not new. For example, motion sensors were used for activity detection within the confines of the bedroom and the bathroom [3]. These sensors could however only acquire limited information pertaining mainly to movement trajectories of the subject without being able to provide useful behavioral information relating to personal hygiene activities. In [4], various ‘off-the-shelf’ sensors were used to collect data on four different behavioural domains: medical adherence, movements throughout the house, bathroom use and meal preparation. While the above system permit continuous, unobstructive monitoring of certain ADLs, it is really the objects in the environment, e.g. pill bottle, the refrigerator door, a kitchen cabinet, that are electronically monitored, not the individual her/himself.

Activity recognition based on body worn sensors, in particular accelerators, has been demonstrated in [5]. This approach is unfortunately inappropriate for cognitively challenged dementia patients who will remove these sensors at will or is it suitable for monitoring activities such as bathing. In [6], the COACH system, which monitors progress and provides assistance in the washroom during hand washing, requires the wearing of a bracelet, which facilitates the tracking of user’s movements. Besides being found to be obtrusive and bothersome, the system generated a fair degree of errors and false alarms.

More recently, a system which uses radio-frequency-identification (RFID) technology to infer the undergoing activities was proposed in [7]. Low cost RFID tags were embedded into representative objects that are highly relevant to the activities of interest. An RFID-detecting glove was designed to detect nearby objects and a probabilistic engine would infer activities from detected object-level interactions. This system could provide detailed information on the exact steps taken in performing a certain activity but the need to wear a glove probably makes it unsuitable for tracking personal hygiene activities such as bathing.

A vision of a Smart Home, which facilitates aging-in-place, was described in [8]. Under this vision, a wireless network of sensor motes will be installed ubiquitously within the home (including the bathroom) to monitor the activities of the subject. The type of sensor modalities that would be useful was however not discussed. Simple motion detecting, pressure-based or even water flow sensors may not be sufficient to delineate the subtly different personal hygiene and toileting activities. For example, a subject may be detected as sitting on the toilet seat but no meaningful conclusion could be inferred on exactly what the subject is really doing. Or a tap may be turned on but we will not know why. The gap between 'things happening within the home' and actual human activities taking place and specific behaviors being exhibited need to be bridged.

In spaces where the use of video surveillance is not socially acceptable, sounds may provide the alternative source of information regarding activities that are occurring and behaviors that are exhibited. This is especially true for personal hygiene and toileting activities within the private confines of the bathroom, each of which are typically associated with distinct sounds. Microphones that capture sounds with sufficient fidelity for automated processing are also much cheaper in comparison with other sensors. Finally, acoustics-based behavioral understanding systems work from a distance: they do not constrain subjects by requiring them to wear special devices; a prerequisite for individuals with dementia. It is therefore no surprise that computational auditory scene analysis (CASA), the understanding of events through processing environmental sounds and human vocalizations, has become an increasing important research area [9]-[12].

1.2 A Discussion on Privacy

One key question pertaining to the issue of privacy remains: will sound surveillance be socially acceptable in private places like the bathroom where the use of video is not? In attempting to answer this question, we quickly realized that we do not have a ready framework to address this issue. We felt that it would be necessary to return to basics and first attempt to understand what privacy really is. We will not be so bold as to define privacy, but we will attempt to qualify, within the scope of this work, the phrase personal privacy.

A useful term that can make this discussion more concrete is Palen and Dourish's [13] *genre of disclosures*, which are socially constructed patterns of privacy management involving recognizable, socially meaningful patterns of information disclosure and use. Amidst a given genre, people expect each other to disclose *this* information but not *that*, under *these* conditions but not *those*, to *this* but not *that* person, and to use information in *this* but not *that* way. Within this framework, the degree of perceived 'privacy loss' caused by the introduction by a new technological construct is related to its non-conformity to these expectations of disclosures within a given genre. With this, we can rephrase our original question more meaningfully, i.e.: does sound surveillance within the bathroom conform to the expectations of our societal genre of disclosures?

We do not have a clear answer to this question but will be using this framework to address this issue once again towards the end of this paper.

1.3 System Summary

In this paper, we describe an acoustics-based system that is able to detect, identify and selectively record activities occurring within the bathroom with the ultimate aim of automatically generating customized personal hygiene behavioral reports for the benefit of caregivers and geriatric clinicians. Personal hygiene activities that are studied and modeled include showering, brushing of teeth, washing hands and urination. The system is designed to present a list of detected and classified bathroom activities with associated details such the time of occurrence, duration and sequence of occurrence for each bathroom visit. All these information are also automatically condensed into a summarized daily report, with adjustable triggers for automatic alert notification based on customized definition of 'abnormal' behaviors.

Fig. 1. Photo of a typical Singapore Housing Development Board flat's bathroom (the very bathroom used for our full blown trial as will be described later on)

The targeted bathrooms for our system are those within Singapore Housing Development Board's flats, a typical structure of which is shown in Fig. 1. It will not be appropriate to install advanced sensor systems that take up precious space and unnecessarily increase the complexity of a system that needs to be as simple as possible.

For the remainder of the paper, we shall assume that, without the loss of generality, the system will be detecting activities performed by a single individual. For bathrooms frequented by more than one individual, we shall assume that there is an available identification system, based on RFID or other technologies, which helps resolve the identify of the subject being monitored.

The rest of the paper is organized as follows. In section 2, the sensor setup, system training and feature extraction methodologies, as well as the full operation of the system will be described in detail. Experimental results and performance evaluation of the system will be shown in section 3. We conclude and discuss further work in section 4.

2 System Description

The main sounds associated with activities occurring within in a bathroom include (a) showering, (b) washing-hands, (c) brushing-teeth, (d) flushing, (e) urination, (f) defecation, (g) human vocalizations (cough, laugh, sigh, etc), and various miscellaneous sounds such as footsteps, sounds of dressing/undressing, combing of hair, etc. In this paper, we shall describe the detection and automatic classification of sounds (a) to (e), as our work on defecation detection and human vocalization classification is still on-going.

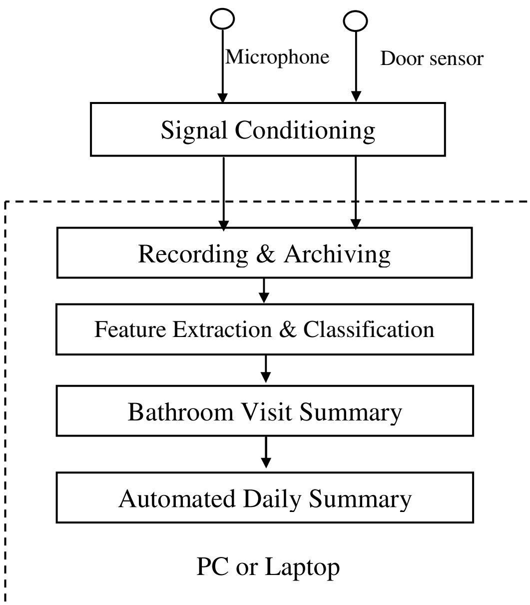

Fig. 2. Block diagram of system

2.1 Sensor Setup

The block diagram of our system is shown in Fig. 2 with the main sensor being a single miniature omni-directional microphone. A multiple microphone system as in [12] would unnecessarily complicate the sound-processing task due to the presence of strong reverberations within the targeted small bathrooms. Signal from the microphone is pre-amplified and routed into one of the two channels of the stereo *line-in* port of a PC sound card. To maximize overall signal quality of all major bathroom activities, we have found that the microphone is best installed near the washing basin to ensure that the weak sounds associated with brushing teeth and washing hands can be picked up clearly while in no way sacrificing the quality of the resultant more distant sounds of flushing, showering etc. which are typically loud.

An infrared door sensor is set up outside the bathroom to detect the subject entering the bathroom; its output being routed to the other *line-in* stereo port of the sound card, making this channel an 'entering/leaving' indicator. A simple energy-based detection scheme on this channel will be sufficient to notify the system every time the subject is entering or leaving the bathroom.

2.2 Methodology of Collection of Sound Samples for System Training

The exact nature of sounds arising from activities occurring within a bathroom are obviously a function of a bathroom's size and layout, material of the floor, type of shower and flushing system used etc. Proper training of the system using sound samples acquired from the same bathroom in which the system is to be installed is essential, analogous to the practice adopted for speech recognition whereby the system is individually trained on each user for speaker dependent recognition.

For clear audio-temporal delineation during system training, the sound capture for each activity of interest was carried out separately. A number of male and female subjects were used to produce the sounds of interest; each subject would typically go into the bathroom, generated the sounds of interest and leave. These steps would be repeated numerous times for each individual of the subject pool and for each activity of interest. It is important to note that in the generation of these sounds, associated 'background' sounds such as the shuffling of feet, undressing, application of soap, etc., are being simultaneously recorded. The variability in the captured sounds of the each activity provide realistic input for system training, and increase the robustness and predictive power of the resultant classifier.

Flushing sounds are generally loud and fairly consistent. Very few samples are typically needed to sufficiently train the classification model. Hands washing, on the contrary, exhibited a high degree of variability even for the same individual. The duration of the hand washing episodes varied significantly; sometimes the test subject applied soap and sometimes they did not. This required us to collect many more samples for hand washing sounds to capture the diversity of the sound of this activity.

2.3 General Observations of Bathroom Sounds

The typical waveforms of the sounds of four different bathroom activities are shown in Fig. 3. As can be seen, the flushing sound as depicted in the waveform of Fig. 3(a) is of short duration (about 5 seconds) and of rather high amplitude. It is almost unchanged every time. On the other hand, showering sounds will last for different lengths from a few seconds to more than an hour. Urination sounds (especially of a male subject) will be intermittent, lasting a couple of seconds long, while hand washing sounds are typically continuous but last for different lengths. Signals of human sighing are not displayed as it is difficult to find a 'typical' sigh.

The spectral of these typical sounds are analyzed and shown in Fig. 4. We can see that each sound has its distinct frequency distribution, corresponding to their distinct resonant frequency which is a function of the sound generation mechanism. For example, the flushing sound has a resonant frequency in the range of 265-285Hz, while for washing hand sound, the resonant frequency lies between 200-210Hz. Urination sounds have a strong frequency component at 600Hz while showering sounds shows a rather wide bandwidth of up to 6000Hz. For showering sounds, there are some strong frequency components between 100-300Hz, which can be explained as the low frequency splashing sound and their reverberations within the small shower enclosure of our bathroom. On the other hand, the high frequency components are contributed by sounds of water directly striking the floor. Generally, the major energy of these sounds is distributed below 5500 Hz and therefore a sampling rate of 11,025 Hz will be enough to preserve most features of these sounds.

2.4 Feature Extraction and Classification Methodology

It was obvious that simple frequency characterization would not be robust enough to produce good classification results. To find representative features, in [14], Cowling carried out an extensive comparative study on various transformation schemes, including the Fourier Transform (FT), Homomorphic Cepstral Coefficients (HCC), Short Time Fourier Transform (STFT), Fast Wavelet Transform (FWT), Continuous Wavelet Transform (CWT) and Mel-Frequency Cepstral Coefficient (MFCC). It was concluded that MFCC may be the best transformation for non-speech environmental sound recognition. A similar opinion was also articulated in [9]-[10]. These finding provide the essential motivation for us to use MFCC in extracting features for bathroom sound classification.

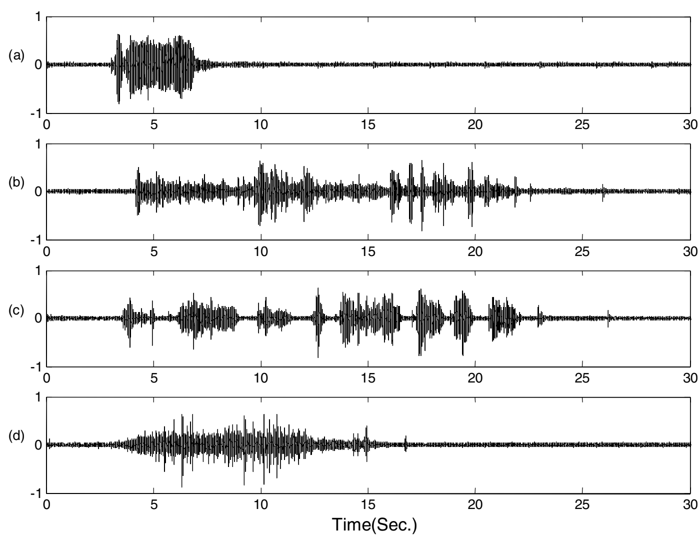

Fig. 3. Typical waveforms of four different sounds: Flushing, (b) Showering, (c) Urination (man), (d) Washing hands

An accurate and robust sound classifier is critical to the overall performance of the system. There are however many classifier approaches in the literature, e.g. those based on Hidden Markov Models (HMMs), Artificial Neural Networks (ANN), Dynamic Time Warping (DTW), Gaussian Mixture Models (GMM), etc. From these options, we have chosen an approach based on HMM as the model has a proven track record for many sound classification applications [15]. Another advantage is that it can be easily implemented using the HMM Tool Kit (HTK) [16]. It should be noted that HTK was originally designed for speech recognition, which meant we needed to carefully adapt the approach when applying in for sounds of our interest.

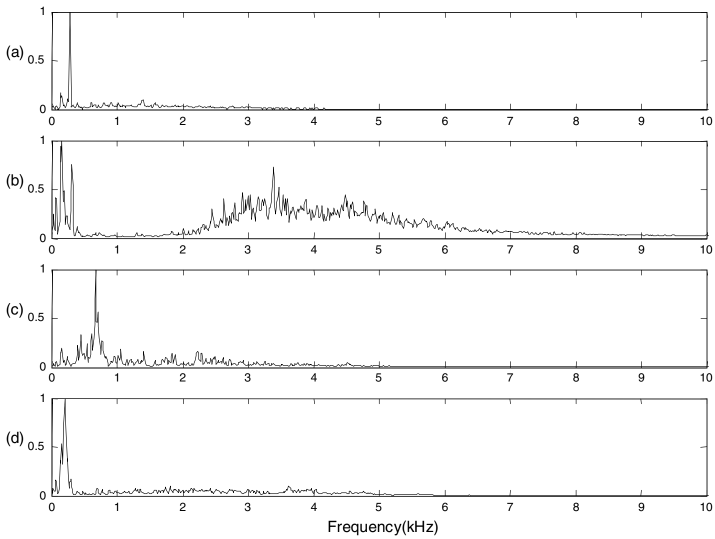

Fig. 4. Typical spectral of four different sounds: (a) Flushing, (b) Showering, (c) Urination (man), (d) Washing hands

Each sound file, corresponding to a sample of a sound event, was processed in frames pre-emphasized and windowed by a Hamming window (25 ms) with an overlap of 50%. A feature vector consisting of a 13-order MFCC characterized each frame. We modeled each sound using a left-to-right six-state continuous-density HMM without state skipping. Each HMM state was composed of two Gaussian mixture components. After a model initialization stage was done, all the HMM models were trained in three iterative cycles.

For classification, continuous HMM recognition is used. The grammar used is as follows: (<flush | shower | sigh | urination | wash-hand | silence >), which means that there is no predefined sequence for all the activities and each activity may be repeated for each bathroom visit. Special case was taken to avoid too many repetitive transcripts appearing in the final result for a single long event such as showering, while at the same time, ensuring that activities of short durations are recognized accurately.

As most bathroom sounds are intermittent, our classification system is unique that frequently, the resulting transcription number may not be equal to the number of predefined sound events. In other words, because sounds corresponding to urination or showering for example are intermittent in nature, there would be a number of ‘urination’ or ’shower’ transcriptions being generated by the classification engine for these occurrences of these events. A general rule that we used to resolve this problem is that if a sound stops, and then restarts within a short time interval, it is considered to be a continuation of the activity of the previous instance.

3 Experimental Set-Up and Results

Experiments were carried out in two phases: a simplified scenario for proof-of-concept and a real continuous trial. The two phases were carried out in different bathrooms and involved different subjects. Common system components used for both phases included a miniature Panasonic microphone WM-034CY and its preamplifier, a commercial infrared door sensor and our software residing on a Windows platform laptop. Data was captured in real time at 11,025Hz sampling rate and 16 bits data resolution. All the data were then saved in the hard disk for post processing.

3.1 Simplified Scenario for Proof-of-Concept

The purpose of this phase is to test the performance of our system in recognizing the major bathroom events. The system was trained and tested to recognize the following 5 bathroom activities: showering, flushing, washing hands, urination (male) and human sighing. Four subjects, two males and two females, were involved in generating sounds of each activity except for urination which only the two male adults participated.

As explained earlier, the sound recording for each activity was carried out separately. For example, for showering, each subject entered the bathroom to take a shower and then leave, with this repeated a number of times for the same individual. The other subjects followed the same protocol and the entire process was repeated for each activity being tested. The resultant sound database is summarized in Table I.

The training data set was formed utilizing a ‘*leave-one-out*’ strategy. That is, all the samples would be used for their corresponding models’ training except those included in the signal under testing. Hence, each time the models were trained respectively to ensure that the samples in the testing signal were not included in the training data set.

Table 1. Composition of the ‘simplified scenario’ database

Activities	Samples	Total length (sec.)
1. Showering	39	724
2. Urination	12	144
3. Flushing	13	68
4. Washing Hands	49	715
5. Sighing	19	22

Since each sound file contains only a single sound event, we developed a statistical testing methodology to evaluate the system's performance on continuous sound streams of various sound events. 100 test signals were generated, each of which contained a mixture of all five sound events randomly selected from their respective classes and concatenated together with short periods of recorded background noise. For testing, the recognition accuracy rate is calculated using the following rules: the length of each segment is factored into account so that an error in transcripting a very short segment is scaled proportionately and vice versa. As a result, the final evaluation result is an accurate reflection of the true operating performance of the system.

Sound recognition results are presented in Table 2. The recognition accuracy is encouraging with most being above than 87%. The correct classification of sighing sounds was found to be very challenging due to the sounds' shortness in duration and weakness in strength, hence the increased frequency for them to be wrongly classified as a 'non-recognizable' sound event.

Table 2. Sound recognition rate for 'simplified scenario'

Class	Correct Rate (%)
1. Showering	92.57
2. Urination	88.82
3. Flushing	91.22
4. Washing Hand	87.89
5. Sighing	72.95

3.2 Real Continuous Trial

In the second phase, a trial testing the system's performance on actual bathroom behavior was conducted. The same system setup was used but on a bathroom different from that used for the first phase. For this trial, we concentrated on the following six activities: (1) showering; (2) washing hands; (3) urination; (4) defecation[♣]; (5) flushing and (6) brushing teeth. Sighing was omitted because they occur too infrequently to provide meaningful statistical results. Four subjects were involved in the trial: a young man (32-year-old), a young lady (27-year-old), an old man and (62-year-old) an old lady (61-year-old). Bathroom visit activities were recorded for 10 full days with about **160** entries chalked up. Detailed number of bathroom activities records captured from each person and the number of samples used for training/testing are summarized in Table 3. The 'ground truth' of these activities was kept in a manual log-book against which our system will be benchmarked.

[♣]Defecation was inferred through a set of heuristic logic rules operating on the relationship between the duration of a subject sitting on the toilet seat, the occurrence of urination and a flushing sound. As a temporary solution to detect someone seating on the toilet seat, we installed a simple mechanical whistle, extracted from a child's toy, under the toilet seat, which will emit a squeaky sound when someone sits down or stands up thereafter. We are using this as a 'workable hack' while searching for a more acceptable solution.

Table 3. Composition of the Database for Each Activity (unit: number of records)

Class	M1	M2	F1	F2	Training	Testing
1. Showering	23	15	18	12	24	44
2. Urination	25	37	21	52	64	71
3. Flushing	34	47	35	63	74	105
4. Washing Hands	59	75	52	78	99	165
5. Defecation	10	8	9	6	13	20
6. Brushing teeth	20	22	20	24	33	53

Note: M1- young man, M2-old man, F1-young lady, F2-old lady, Training/ testing: number of records or entries used for training / testing. (Training +Testing = M1+M2+F1+F2)

3.2.1 Special Consideration for System Training and Activity Classification

Due to the complexity of the trial set-up (the four subjects chosen are members of an actual family using their bathroom in the normal course of their daily lives), we utilized a slightly different approach for system training. Sounds recorded from the normal course of bathroom use for the first four days were used as data for system training, based on which the system's performance is benchmarked for the remaining six days of the trial.

Upon reviewing the recorded sounds after the first 4 days, we found lots of sounds generated by secondary events such as coughing, spitting, floor washing, towel squeezing, water flowing down the drain pipe, etc. The features of some of these sounds overlapped among themselves and even with some primary sounds of interest. These sounds needed to be modeled explicitly. The following integrative steps were taken to handle this challenge:

1) **Forming of 2 New Sound Classes:** Besides the six primary sound classes discussed above, three more classes were added: a) those generated by the human vocal system, such as coughing, laughing, spiting, sneezing etc., collectively grouped under the *vocal interference class*; b) those generated all other environmental sounds, such as footsteps, object dropping, floor being washed, towel being squeezed, water flowing, door opening, noise from outside the toilet outside etc., collectively grouped under the *environmental interference class*.
2) **System Training:** Sounds from the first four days of normal bathroom use were utilized as the training data. Each sound event was manually categorized into one of eight sound classes to be modeled. The same features extraction methodology described in section 2.4 was used to determine the HMM parameters of each sound class.
3) **Classification:** The grammar for the classification is broadened to include the two new interference classes.

3.2.2 System Performance for Real Continuous Trial

The overall recognition rates achieved by the system during final 6 days of the trial are tabulated in Table 4. Accuracy rates for defecation, flushing and showering are high as the sounds involved are either distinct or occur sufficiently long in real life to ease recognition. Urination and teeth brushing can be recognized reasonably well, leaving hands washing being the most challenging activity to recognize due to its relatively weak sound amplitude and in many times overlapping with other sounds.

Compared with the ‘simplified scenario’, the main reason behind the degradation of the recognition accuracy is the increased complexity of a real sound stream (as opposed to an artificial one) with lots more unpredictable sound interferences.

Table 4. The accuracy of recognition

Class	Correct Rate (%)
1. Showering	87.45
2. Urination	77.23
3. Defecation	93.60
4. Flushing	90.10
5. Washing Hand	68.67
6. Brushing teeth	84.23

4 System Operations

4.1 Using the System

Our Windows based software controls all functions of the system: sound detection, selective recording, recognition, consolidation and printing. The graphical user interface (GUI) of our system is shown in Fig. 5 for the benefit of our readers.

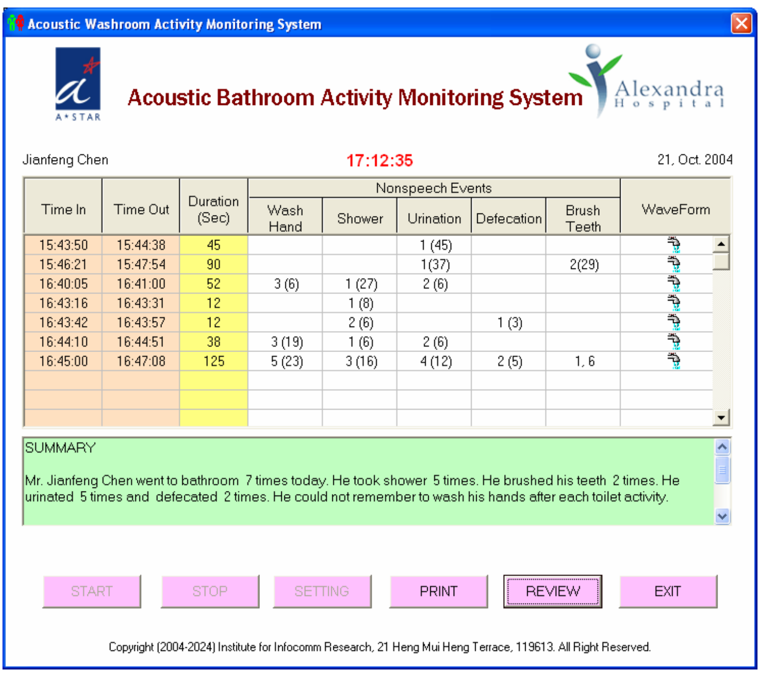

Fig. 5. Graphic User Interface of the System

Each time the subject enters the bathroom, the door sensor will be triggered and its audio output detected. The recording module is then immediately activated. Recording will continue until the subject leaves the bathroom as detected by the same door sensor. The entry time, leaving time and the duration of the bathroom visit will automatically be recorded in columns 1-3 of the GUI table of Fig. 5. The entire sound waveform captured during this period will be routed to the sound recognition and classification module which will detect and label each occurring sound event.

Detected activities will be denoted by a numerical digit indicating the sequence of its occurrence, 1 to indicate occurring first, 2 to indicate occurring second and so on. The duration of occurrence of each event is also captured and can be displayed as well if chosen to.

4.2 Automatically Generated Daily Summaries

On the stroke of midnight of each day, the whole day's activities are summarized and a daily behavioral report is automatically generated and archived in the hard drive for further reference. This report typically contains a consolidation of the frequency of occurrences of major activities of interest. Under the guidance of our geriatric clinician collaborators, normal baseline behaviors were defined, against which deviations are detected and highlighted in the 'comments' section. A series of pre-prepared words are used and intelligently strung together to form simple sentences that conform to the basic rules of English grammar. An example of this is shown below:

DAILY REPORT

Mr. Jianfeng Chen went to bathroom 7 times today. He took shower 5 times. He brushed his teeth 2 times. He urinated 5 times and defecated 2 times. He could not remember to wash his hands after each toilet activity.

Comments:

Normal but seems to have showered excessively. Recommend probing this possibly abnormal behavior.

The archival of these daily reports enables a caregiver or a doctor to review records very quickly and in the process, build a detailed understanding of the subject's bathroom behavioral patterns.

4.3 Revisiting the Issue of Privacy

We feel that it is appropriate if we return once again to the privacy framework introduced in section 1.2 and asked ourselves if it is socially acceptable for the type of information captured and the automated daily reports generated by our system be disclosed to care-givers and doctors who are caring for the monitored subject. Again we must acknowledge that we do not have a ready answer for this question. We can only hope that we have made some progress towards an affirmative answer by producing a solution that may actually be less intrusive than the alternative: the loss of independence that follows when a person is sent to a nursing home, partially due to a lack of understanding of the person's bathroom behavioral problems.

5 Conclusion

In this paper, we described a novel acoustic bathroom activity monitoring system that automatically detects and classifies major activities occurring within a bathroom. Carefully designed HMM parameters using MFCC features are used for accurate and robust bathroom sound event classification. Experiments to validate the utility of the system were performed firstly in a constrained setting as a proof-of-concept and later in an actual trial involving real people using their bathroom in the normal course of their daily lives. Preliminary results are encouraging with the accuracy rate for most sound categories being above 84%. We sincerely believe that the system contributes towards increased understanding of personal hygiene behavioral problems that significantly affect both informal care-giving and clinical care.

Besides further improving the recognition accuracy, we plan to enhance the capability of the system to identify different types of human vocalization, which provides useful information pertaining to the mental well-being of the subject. Urine flow sensors will also be integrated into the system to enable clinicians acquire better understanding in battling incontinence and lower urinary tract syndromes. The enhanced system will be shortly tested in a full-blown trial on the most needy dementia patients residing within the wards of a national hospital before evaluating its suitability as a benevolent behavior understanding system within the homes of these patients.

References

1. Cohen, J.E.: Human Population: The Next Half Century, Science, Nov 14, (2003), 1172-1175
2. Andresen, G.: Caring for People with Alzheimer's Disease, Health Professions Press, Baltimore, MD., (1995), 52
3. Chan, M., Campo, E., Esteve, D.: Monitoring Elderly People Using a Multisensor System, Proc. of 2nd International Conference On Smart Homes and Health Telematic, Singapore, (2004), 162-169
4. Glascock, A., Kutzik, D.: Behavioral Telemedicine: A New Approach to the Continuous Nonintrusive Monitoring of Activities of Daily Living, Telemedicine J., vol. 6, no. 1, (2000), 33–44
5. Lukowicz, P., Ward, J., Junker, H., Stäger M., Tröster G., Atrash A., Starner, T.: Recognizing Workshop Activity Using Body-Worn Microphones and Accelerometers, Proc. of Pervasive, Vienna, Austria (2004), 18-32
6. Alex, M.: Intelligent Environment for Older Adults with Dementia, http://www.agingtech.org/item.aspx?id=18&CAT=4&CA=1 (2004)
7. Philipose, M., Fishkin, K. P., Perkowitz, M., Patterson, D. J., Fox, D., Kautz, H., Hähnel, D.: Inferring Activities from Interactions with Objects, IEEE Pervasive Computing, Oct. (2004), 50-57
8. Philip, E. R.: Managing Care Through The air, IEEE Spectrum, vol. 12, (2004), 14-19

9. Okuno, H.G., Ogata, T., Komatani, K., Nakadai, K.: Computational Auditory Scene Analysis and Its Application to Robot Audition, International Conference on Informatics Research for Development of Knowledge Society Infrastructure (ICKS), (2004), 73–80
10. Eronen, A., Tuomi, J., Klapuri, A., Fagerlund, S., Sorsa, T., Lorho, G., Huopaniemi, J.: Audio-based Context Awareness - Acoustic Modeling and Perceptual Evaluation, Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03), vol. 5, (2003), V-529-532
11. Peltonen, V., Tuomi, J., Klapuri, A., Huopaniemi, J., Sorsa, T.: Computational auditory scene recognition, Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '02), vol. 2, (2002), 1941-1944
12. Drake, L., Katsaggelos, A.K., Rutledge, J.C., Zhang, J.: Sound Source Separation via Computational Auditory Scene Analysis-enhanced Beamforming, Workshop Proceedings on Sensor Array and Multi-channel Signal Processing, (2002), 259–263
13. Palen L., Dourish P.: Unpacking "Privacy" for a Networked World, Proc. of CHI 2003 Conference on Human Factors in Computing Systems, Florida (2003), 129-136
14. Cowling, M.: Non-Speech Environmental Sound Recognition System for Autonomous Surveillance, Ph.D. Thesis, Griffith University, Gold Coast Campus (2004)
15. Rabiner, L. R., Juang, B. H.: Fundamentals of Speech Recognition, PTR Prentice-Hall Inc., New Jersey (1993)
16. Young, S.: The HTK Book, User Manual, Cambridge University Engineering Department (1995)