HONEY: A Multimodality Fall Detection and Telecare System

Quan Zhang[1], Lingmei Ren[1], and Weisong Shi[1, 2]
[1]Tongji University, China
[2] Wayne State University, USA
Contact author: weisong@wayne.edu

ABSTRACT— The increasing cost of finance and healthcare resources is driving healthcare providers to provide home-based telecare instead of institutionalized healthcare. Falling is one of the most common and dangerous accidents for elderly group and a significant factor affecting the living quality of the elderly. Many efforts have been put towards providing a robust method to detect falls accurately and timely. To facilitate a reliable, safe and real-time home-based healthcare environment, we propose the HONEY system to detect falls for elderly people in the home telecare environment. The basic idea of HONEY is a three-step detection scheme which consists of multimodality signal sources, including an accelerometer sensor, audio, images and video clips via speech recognition and on-demand video techniques. The magnitude acceleration, corresponding to a user's movements, triggers fall detection combining speech recognition and on-demand video. If a fall occurs, an alarm email is delivered to a medical staff or caregivers at once, containing the fall information so that caregivers could make a primary diagnosis based on it. This paper also describes the implementation of the prototype of HONEY. A comprehensive evaluation with 10 volunteers shows that HONEY has high accuracy of 94% for fall detection, 18% higher than *Advanced Magnitude Algorithm* (AMA), which is a wearable sensor-based method, while the false positive rate and false negative rate are 3% and 10%, 19% and 16% lower than AMA, respectively. The average response time for a detected fall is 46.2 seconds, which is also short enough for first aid.

Index Terms—Fall detection, Health information management, Home telecare, Medical information systems

I. INTRODUTION

The impending influx of elderly citizens aged over 65 will soon be a much larger ratio to all population around the world. For example, in 15.07% of 23 million population of Shanghai is more than 60 years old [1]. Falling is a common and dangerous accident for elderly group and an important factor affecting the living quality of the elderly. It is estimated that more than one in three elder people aged over 65 living at home fall at least once each year. The risk of falling will also rise with the increasing age. Falling also leads to decreased mobility, fear of fall, and even deaths [2, 3, 4]. Furthermore, a research among Medicare (the largest healthcare insurer in USA) beneficiaries shows that the total medical care costs each year for elderly adults who fall once a year, are 29 % higher than the ones for other elderly adults who never fall in one year. Similarly, the costs for elderly adults who fall more than two times a year are 79 % higher than ones for those who report no fall [5]. However, the limited funds for public healthcare service and an aging population are driving factors for reducing the institutionalized healthcare based on entity service and moving to home healthcare based on wireless sensor networks.

Accelerometers with low-cost and low-power features make the wearable and reliable fall detection into a reality. Multiple sensors with accelerometers placed at various bodily sites are used for real time human movement detection [6, 7, 8]. Many systems [2, 9, 10, 11, 12] employ triaxial accelerometers to detect fall according to the acceleration of body motion and posture angle. To achieve better accuracy, later systems [13, 14, 15] detect fall combining accelerometer with barometric pressure sensor, image process, and gyroscope etc. *Information Technology for Assisted Living at Home* (ITALH) is a project using new technologies to help the older citizens live more comfortable [16, 17]. The ITALH includes two items: the IVY project concerns detecting fall at home or in office environments, and the SensorNet project concerns developing an integrated, safe and wireless sensor to monitor the user. However, those previous systems have several restrictions: **(1)** the methods are device-centric, not user-centric; **(2)** the devices are expensive and complicated; **(3)** the information received by the doctor is insufficient to make an accurate diagnose in a timely fashion. In most of the systems, the final decisions are based on the data collected from the sensors and the user cannot express his/her ideas on his/her own initiative and just passively accept the decision. In addition, some of the previous systems use acoustic or vibration sensors, and image processing software etc. Most of them are high cost and not universally accessible. Ordinary users cannot control them as their will. Few systems send an SMS message as a simple alarm. However, the text message is not enough for describing a patient's symptoms, and caregiver cannot take accurate treatments to rescue the user through this simple text message. Therefore, we propose a three-step fall detection scheme: *First*, we employ multimodality resources, i.e., movement, audio, images and video, for precise fall detection. *Second*, HONEY informs caregivers that a fall has happened by sending an alert email once a fall is detected. *Thirdly*, we provide the video clip of the fall scene by uploading the video to network storage for further investigation.

In this paper, we propose a ***HO**me healthcare senti**NE**l s**Y**stem* (**HONEY**) for home-based fall detection and response by leveraging multiple sensors and home-area network, which will be a common setting for most houses apartments. In addition, HONEY is an innovative multimodality detection scheme by leveraging a triaxial accelerometer, speech recognition, and on-demand video. We implement a prototype that implements the core functions of HONEY and perform a comprehensive evaluation of the system. Our results show that (1) HONEY can distinguish fall and non fall actions with high accuracy; (2) the speech recognition of HONEY can work well on a reasonable level in quiet and noisy environment; (3) we compare HONEY with AMA presented in [12], which is a wearable sensor-based method, and the accuracy rate of HONEY to detect fall is higher than AMA. (4) the average response time of HONEY is 46.2 seconds, and this is a short period for an emergency first aid.

In the remaining of this paper, we describe the architecture, implementation, and evaluation of HONEY. We will introduce more details about why we need a home-based healthcare system for elder people in Section □. Next, we present a detailed description of the HONEY architecture in Section □. Section □ provides the implementation of a HONEY prototype. The evaluation of the HONEY prototype is presented in Section □. The final sections summarize the work related to fall detection for elder people in Section □, and conclude in Section □.

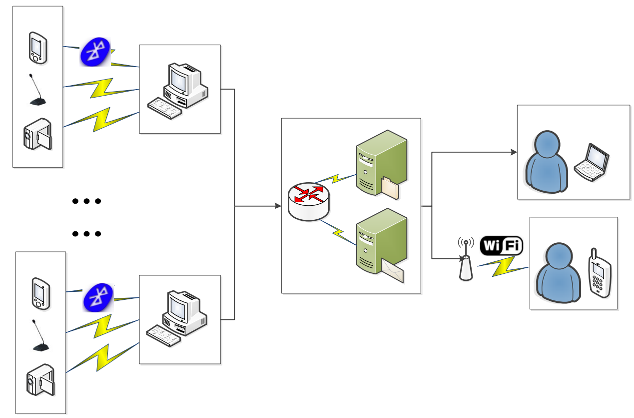
II. BACKGROUND

Due to the growing ratio of elder people, most of elderly people live on their own in a home-dwelling environment. In China, the population over 60 is about 167 million by 2010, 13.2 % of all. By 2050, the ratio will be over 30 %. The persons aged over 80 will be more than 90 million by 2050 [18]. The trends are also found in Australia and other developed countries [19, 20]. The high costs for healthcare becomes a significant issue that every nation has to address [21]. An emerging method is the use of wireless sensors to detect problems as early as possible and to prevent incidents. Among these incidents, falling is the leading cause of nonfatal or fatal injuries for elderly group. Fifty-five percent of falling injuries happen inside the home, and an additional 23 % is near the home [22]. Fall often causes many physiological and psychological problems, such as restricted activity, fear of falling, fear of living alone and even death. To avoid the situation that many individuals are forced to leave the comfort and privacy of their home to live in a nursing home; it is a significant issue to develop a home-based fall detection system to protect elderly people by using an accurate, automatic, non-intrusive, and socially acceptable method.

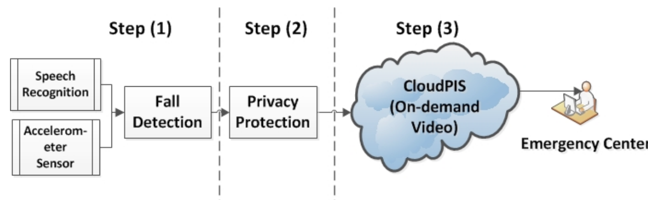
Recent developments in wireless communication technologies and wireless sensor network provide a foundation for the remote supervision of physiological monitoring. Most of these systems only give the judgments about whether the subject falls. A few of them send a SMS message. However, the information is too less for caregivers to give an accurate diagnosis. To resolve the problems aforementioned, we propose HONEY system that is a home-based healthcare system. Our system not only provides instantaneous and continuous fall detection, but also provides more information for caregivers. The time between the occurrence of a fall event and the sending of an alarm email for caregivers is also short enough for first aid.

III. HONEY SYSTEM ARCHITECTURE

Figure 1(a) provides the high-level overview of HONEY by showing the key components. HONEY consists of *FallSensor*, *Homeserver and Cloud-based Personal Information Storage* (CloudPIS). FallSensor includes an accelerometer, microphones and cameras. The accelerometer is fixed on the user's waist, and the microphones and cameras are distributed in every room of a house in pairs. Homeserver is the control center, which determines whether the user falls or not by utilizing the data collected from FallSensor. CloudPIS is an information sharing platform for patients and caregivers, and serves as a permanent storage for a patient, including all the information about the patient's medical history and real-time data, such as video clips. The caregivers can obtain the video information of a user's fall scene through CloudPIS in a secure way. Figure 1(b) shows the three-step fall detection at the functional level. First, an accelerometer monitors a user's acceleration and posture, if triggering conditions are met, the user's audio message can be used for speech recognition function to confirm or cancel the alarm. When a fall has been



(a). The basic operation diagram of HONEY.



(b). The key function components in HONEY.

Figure 1. A high-level overview of HONEY.

detected, an alarm email is sent to caregivers. Second, due to the fact that the alarm email contains images which may involve user's privacy, thus we have to take measures to protect his or her privacy information during the mail delivery. Similarly, during the data transmission between Homeserver and CloudPIS, the video clip data also needs strict protection to avoid privacy disclosure. Thirdly, when caregivers receive an alarm email, they can review the fall scene video through CloudPIS, and the caregivers can make an accurate diagnosis based on the symptoms presented in video clip and the medical history stored in CloudPIS.

A. Fall Detection

A sensor is fixed on the waist of a subject, and it monitors the acceleration and the tilt angle of the subject in real time by processing the data collected from accelerometer. If FallSensor detects magnitude acceleration, it will trigger Homeserver for further detection, which includes speech recognition and kinescope recording using a pair of microphone and camera preset. Later, FallSensor may give the posture information. However, this process depends on the result of speech recognition. Homeserver will combine the results of speech recognition with the tilt angle sent by FallSensor to make a final decision whether the subject falls or not. An alarm email will be sent to a doctor and his or her relatives if HONEY gives a positive decision. The following step is uploading the video recorded by HONEY to the user's CloudPIS. The caregivers can review the video and take appropriate first aid measures in the light of the user's symptom shown in the video.

When HONEY is working properly, the triaxial accelerometer on FallSensor will sample three axes acceleration at a frequency of 45 Hz. Then the total acceleration of body movement is calculated at every sampling point. If the total acceleration is over than the threshold preset, FallSensor detects a possible fall. Then *Early Warning Information* is sent from FallSensor to Homeserver.

FallSensor keeps monitoring the tilt angle of body movement using those sampling values in following interval (about 20 seconds). Now Homeserver receives Early Warning Information and locates the position where the subject falls. The existing localization methods based on Radio Frequency- (RF-) or Electromagnetic [23] are good enough and can be leveraged in our localization algorithm. HONEY leverages either RF- or Electromagnetic-based developed indoor localization algorithm, which uses a magnetic sensor, an accelerometer, and a moving vehicle with servo motors to determine the location of the fall event. Details are beyond the scope of this paper. The pair of microphone and camera closest to the fall position is turned on, and then a further detection begins immediately. The camera will not stop recording on the falling scene until HONEY finishes current detection. This action continues for 12 seconds. During the recording, the camera also captures images every 2 seconds and there are totally six images. The video file and images are stored in the local storage temporally. At the same time, Homeserver begins speech recognition with the microphone. Homeserver will identify whether the subject says the specified keywords. According to the cases after the patient falls, HONEY's working states will be divided into the following two categories:

1. If the patient is conscious, he or she could speak out the specified keywords clearly. The patient could speak out positive keyword such as "Help me", which means he or she cannot take care of himself/herself and needs a first aid right now, or negative keyword such as "I am fine", which means he or she is not injured and could recover himself/herself. If a positive keyword is detected, Homeserver sends *Abort Information* to FallSensor. This information will stop FallSensor watching the tilt angle changes, and drive FallSensor into next detection round. All those actions should be completed in 20 seconds. Otherwise, HONEY will work in the other process. In most cases, the speech recognition finishes earlier than the kinescope recording. Consequently, when the video recording is completed, Homeserver sends an alarm email to the target mailboxes. The prior six images captured will be the attachments of the alarm email. After the email is sent out, the video file will be uploaded to CloudPIS. A doctor could get the patient's state based on the images and take first aid measures immediately. If image information is insufficient to decide the patient's state, the doctor can review the video file for more details.

On the contrary, a negative keyword stops current detection. The Abort Information is sent and FallSensor steps into next detection round. The kinescope recording and image capturing will finish normally. Then the whole HONEY enters into the standby state. Afterwards, an investigation of this uninjured fall could help to improve the living environment according to those video file and images.

2. If the subject is seriously injured and unconscious, he or she cannot speak anything in this situation. The camera is still on duty as usual to kinescope record on the scene and capture images, but the speech recognition function becomes invalid. Therefore, Homeserver will not send the Abort Information. FallSensor still keeps monitoring the tilt angel of body movement after Early Warning Information is sent. Now it is FallSensor's turn to continue the detection. When the 20 seconds interval ends, according to the current tilt angle of the patient, FallSensor sends *Alert Information* to Homeserver. This information includes the judgment made by FallSensor based on the tilt angle. The analysis result of Alert Information provides the foundation for Homeserver to make the final decision. If it is a fall, HONEY will send an alarm email which is the same as the one previously described. Otherwise, HONEY terminates current detection and enters into the standby state.

In all categories, FallSensor or speech recognition may draw a wrong conclusion. In this situation, the images and video clip could be served as a remedy measure. A doctor can correct the false positive cases by scanning the images and the video, which also reduces unnecessary waste of medical resources.

B. Privacy Protection

Due to the fact that HONEY is used in a home environment and related to human beings, the privacy protection and the data security are important issues in reality. We assume the patients are willing to share their information with caregivers. In HONEY, the cameras are located in several different rooms of a house, such as bedroom, living room and dining room. When HONEY detects falls, the system has to capture some images and record a video clip. In those images and clips, some personal or individual privacy information may be included, and a privacy disclosure may occur during the transmission. A more serious case is that revealing the privacy information may affect a user's personal safety. Thus, HONEY provides a robust protection on privacy protection and data security. HONEY takes the following three approaches to achieve the purpose:

1). We choose Bluetooth as the communication protocol between FallSensor and Homeserver. Up to now, the effective signal range of Bluetooth is about 10 meters. This distance could cover most ranges of a house, not much larger than a house. It is hard for the attacker out of the house to receive the Bluetooth signal, which means that it is unlikely and difficult to intercept sensitive information. An 8-bit decimal *Personal Identification Number* (PIN) code is also used for establishing the connection in the safe mode. The connection mode is *point-to-point*, not spreading. A short message length is necessary to reduce the risk of data loss, which could also extend the sensor's lifetime.

2). We choose the *Secure Socket Layer* (SSL) [24] technology as the privacy protection tool on the Internet. A large number of applications prove that SSL is safe and reliable in an open Internet environment. When HONEY sends the alarm email or uploads the video clip, the sensitive information will be transferred over an open Internet environment. The attacker may capture the packets that contain the privacy information. SSL is built on a reliable transport protocol, such as Transmission Control Protocol (TCP), and provides data encapsulation, compression, encryption and other basic functions to ensure the data and communication secure.

3). When the video is uploaded, a time-limited link address will be sent to a doctor. This link address is the one for the doctor to access a user's CloudPIS. If there is a need to review the video, the doctor could get the video through the address. It is certain that the link address forbids downloading, deleting and modifying. If the time exceeds the limitation, the address becomes invalid. The doctor also has to promise not to reveal the images in the alarm email.

Under the protections of above measures, the privacy information is well protected, without the worry of privacy disclosure. Because the privacy protection is beyond the scopes we focus on, the protections may be faulty. In the future, we can employ other more robust measures to make HONEY safer.

C. CloudPIS

Nowadays, multiple terminals owned by individual make the synchronization of file or other information difficult and troublesome. Thus, cloud-based personal storage, such as Google Docs, Dropbox, and so on, has come into being and increasingly become part of our human life. In addition to providing basic storage services, these cloud-based storages also provide a nice easy-to-use sharing function among friends without copying and pasting. The data stored in those cloud-based storages is encrypted, and the date transmission is also protected by various methods, such as authentication, authorization and so on. CloudPIS will leverage existing privacy protection techniques for CloudPIS's personal information protection and control. Details are beyond the scope of this paper. In practice, researchers have designed such a common storage layer called Wukong [25]. This system provides a user-friendly and highly-available facilitative data access method for mobile devices in cloud settings, and supports heterogeneous storage services.

Privacy protection on CloudPIS is an important component. The reality is that many countries have corresponding privacy law or act, which has a mandatory requirement on the privacy issue, such as Health Insurance Portability and Accountability Act (HIPAA) in US [26]. In addition, the corporations of those commercial storage products, such as 115.com, Dropbox, Google Docs, also have their specified privacy policy. It is worth noting that different countries and corporations have different privacy requirements, so we will follow the policy of corresponding storage service provider for their content, and the laws of corresponding country to decide to deploy HONEY to the country.

In HONEY, the fall scene video clip is uploaded to CloudPIS, and we hope that caregivers can obtain this video clip automatically and immediately, without complex and difficult file sharing client and server. Cloud-based personal storage provides a good foundation for us to build our CloudPIS. By leveraging these services for storing and sharing personal information, such as medical history and other medical information, caregivers have an access to those information using a PC or PDA. At the same time, the aims of informing caregivers of a fall that has happened and providing them more video information in CloudPIS are achieved. In HONEY, CloudPIS is designed to employ any of this type of service by providing a general storage interface for different cloud storage services, and the user can use it without caring the different features between various services. Homeserver can upload the video to CloudPIS, and the user can manage his/her own medical history and other medical information with cloud platform. Doctors can also use the video and user's medical history stored in CloudPIS to give some detailed suggestions or treatments for the user to improve living quality or rescue him/her in emergency.

To this point, we describe the three-step detection scheme of HONEY, and all the aforementioned methods make HONEY a reliable, safe and responsive system. HONEY provides reliable fall detection by using triaxial accelerometer, speech recognition and on-demand video. HONEY also sets the user's heart at rest by using various privacy protection measures. The response time is also short enough to avoid the serious consequence caused by not having any medical assistance in time.

IV. IMPLEMENTATION

We have implemented a prototype of HONEY. The program running on FallSensor is written in Java for the Android environment and the one running on Homeserver is written in C and C# for the Windows environment. The prototype implements the most core functions of HONEY. The devices and standards used in our prototype are described in Table 1.

A. FallSensor

A triaxial accelerometer is integrated in FallSensor, and FallSensor sends Early Warning Information if the trigger conditions are met with handling the three axes' sample values at a frequency of 45 Hz. In our prototype, an HTC G3 Hero smartphone acts as FallSensor. G3 has a triaxial accelerometer and the frequency of this accelerometer is up to 70 Hz. In addition, the Bluetooth module and a high performance processor on G3 satisfy FallSensor's requirements very well.

It is known and verified that a sensor based on a triaxial accelerometer can distinguish the body movements more precisely when it is fixed on the patient's waist [27]. In HONEY, the G3 is worn on the waist of the body (see Figure 2). The triaxial accelerometer will output three acceleration values of x-, y- and z-axis at every sampling point, and the units is m/s^2. When the body is stationary, the total acceleration of the body is *g* (the gravity of Earth, 9.8 m/s^2), vertical down. When the body is moving, the acceleration changes along with the movement intensity. Our algorithm running on FallSensor is based on the assumption that a fall is always associated with a magnitude impact. An estimation of the degree of body movement intensity can be obtained from the *signal magnitude vector* (SVM). Define SVM by the relation:

$$SVM = \sqrt{x_i^2 + y_i^2 + z_i^2} \tag{1}$$

where x_i is the i-th sample value of the x-axis signal (similarly for y_i and z_i). Therefore, comparing the SVM to a preset *SVM threshold* (*SVM_th*) allows detecting the associated fall.

Similarly, when the body falls, the space relationship between body and ground also changes significantly. In order to determine the space posture of the body, we define *Tilt Angle* (TA) as the angle between positive z-axis and SVM by the relation:

$$TA = \arccos(\frac{z}{SVM}) \tag{2}$$

where *z* is the sample value of z-axis signal. TA refers to the relative tilt of the body in space. We also need an angle distinction between the upright postures of sitting and standing, as well as the lying in various conditions. Karantonis's works [8] provide the range of TA corresponding to the different body postures: if the patient's TA is from 0 to 20°, it is classified as *standing*, whereas values from 20 to 60° indicates a *sitting* posture; if TA is between 60 and 120°, it is regarded as a *lying* posture. In most cases, a fall starts from a standing posture, and directly ends with lying on the floor. However, no fall would be predicted if the user falls in such a way that he or she was not parallel with the ground. This is important in various cases during a fall. A user will try to grasp a wall, chair, or other objects and end up slumping next to the object, such as sitting on a chair, rather than lying on the floor. Therefore, a sitting posture following a magnitude SVM is regarded as a fall by our application too. Moreover, if a patient is intact, a short recovery interval (around 20 seconds) should be provided for the

Table 1. Devices/Standards Description.

Device/Standard	Device/Standard Parameters
HTC G3 Smartphone	Android 2.1 OS. Accelerometer at 45Hz, Controller and Bluetooth module.
Sony Laptop	Windows 7 OS Professional SP1 and Bluetooth module.
Microphone	Sampling Frequency: 96,000Hz and Bit-depth: 24bits.
Camera	CMOS, 640×480, 2 megapixels, 30FPS
Speech Recognition Engine	Microsoft Speech Engine With Speech SDK 5.3.
Bluetooth	Bluetooth Version 2.1.
SSL	SSL Version 3.0. Port: 587

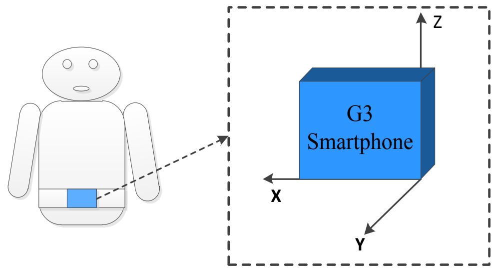


Figure 2: FallSensor is fixed on the waist. In the upright position, the orientation of x-axis, y-axis and z-axis in relation to the sensor is also illustrated.

patient to recover. The algorithm that is simplified running on FallSensor is developed from the AMA and illustrated in Figure 3. We add the Bluetooth communication process which provides simple information exchange between FallSensor and Homeserver. FallSensor just needs to send the triggering information (Early Warning Information), and the posture information (Alert Information), so we also subtract the second posture determination process following the first delay interval of AMA. The delay interval of HONEY is 20 seconds, while the one of AMA is 12 seconds.

When the current SVM exceeds the *SVM_th*, Early Warning Information is sent out. In the following time interval, FallSensor will wait for Abort Information. If no Abort Information is received, FallSensor predicts the body's posture according to the current TA at the end of the recovery interval. Then, the positive or negative Alert Information will be sent to Homeserver. Finally, FallSensor comes back to the initial state.

B. *Homeserver*

Homeserver is the control center of HONEY. It connects FallSensor and external network. The speech recognition, kinescope recording, image capturing, email sending and video uploading are under the control of Homeserver. Speech recognition and image capturing are important means to detect fall. Therefore, the camera should have a high resolution and a widely view. In this way, the images and video are clear for later review. The microphone for speech recognition should also have a high sampling rate, thus reducing the negative effect on accuracy caused by audio quality.

In our prototype, a laptop running Windows 7 operating system acts as the cornerstone of Homeserver. The camera and the microphone are peripherals connected to the laptop. A Bluetooth

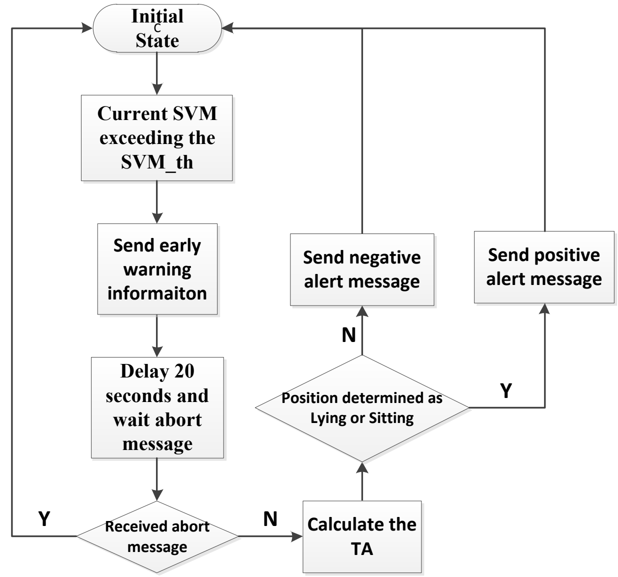


Figure 3. The algorithm running on FallSensor.

module and Internet access are also available on the laptop. All those devices work together as Homeserver. Figure 4 provides the algorithm running on Homeserver. In idle period, Homeserver listens on whether a message comes in. If Early Warning Information is received, Homeserver starts image capturing and speech recognition in the following interval (around 20 seconds). If a keyword is detected, Homeserver sends Abort Information to FallSensor, and then determines whether the alarm email should be sent or not, depending on the detected keyword. If no keyword is detected, Homeserver waits for the FallSensor's Alert Information. If positive Alert Information is received, Homeserver sends an Alarm Email. Otherwise, Homeserver goes back to the initial state.

a) *Speech Recognition*

In some situations, a patient falls without injury or light injury, and he or she can ask for help or cancel the alert by speaking out some words. In HONEY, a user is asked to speak out some keywords if the user wants a help, or wants to cancel the alert. Those keywords are pre-specified. The user can also modify them to make it easy to remember or customized. But they are better not to conflict with daily words.

Nowadays, several companies provide free speech recognition engines. Microsoft's newest operating system, Windows 7, also integrates an engine for free. It is easy to access by using *Microsoft's Speech API* [28] tools. This engine can identify multiple languages including English and Chinese. Training is also available to improve the engine's performance. We do not train the engine because HONEY should work well for anyone rather than an individual patient. The user could train the engine for a particular object. We also define two types of keywords: *Positive Type* and *Negative Type*, and they have the opposite meanings. Positive Type keywords including "I fall", "Help me" and "Call 120" (China) or "Call 911" (US), are used for affirming the alert. Negative Type keywords including "I'm fine" and "Cancel alarm" are used for canceling the alert. In addition, the accuracy of speech recognition relies on the microphone's location and the quality of voice sample. In order to reduce the distance between the microphone and the patient, we place the microphone in the middle of the room. We also adjust the microphone's sampling

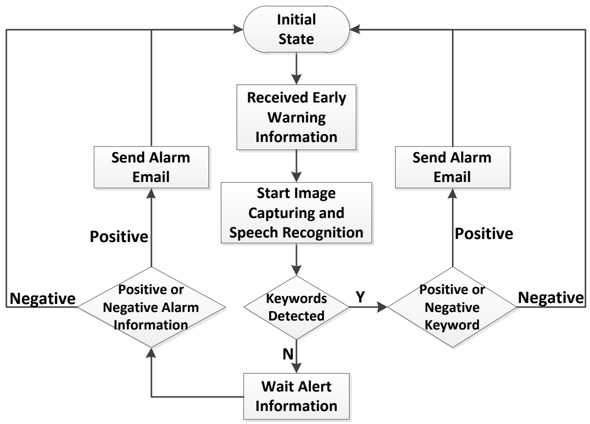


Figure 4. The algorithm running on Homeserver.

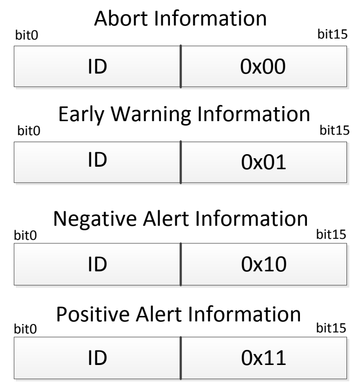


Figure 5. Four types of message format. Bit0 is the highest bit of the message and bit 15 is the lowest.

frequency and quality to the highest level.

b) Kinescope Recording, Image Capturing and Email Sending

Vision information can provide information directly and in details. In the design of HONEY, Homeserver should upload the video clip to the user's CloudPIS, and a doctor could acquire the patient's symptoms by reviewing the fall video immersive. Then targeted aid measures could be made immediately. Because the video transmission is time-consuming, we run this task in the background. When the alert email sending is finished, the video transmission begins.

In our prototype, we use a camera to capture images of the fall scene and send those images to a doctor through an email. The camera is two megapixels with a highest resolution of 640×480 pixels. The horizontal view angle is sixty degree and the vertical view angle is forty-four degree. The camera is installed on the wall, 2 meters high from the ground, where most areas of the experimental chamber can be included in the video. The mat used for experiment is located in front of the camera on the ground. The images are saved in JPG format. The image size is 640×480 pixels and those images are compressed before they are sent.

When a fall is detected, HONEY sends an alarm email, which encloses six images, to corresponding caregivers' mailboxes, such as a doctor's or relatives'. A mailbox account is needed for sending the alert email. The user has to specify the target mailboxes in advance. In our prototype, we use two Gmail accounts as the source and the destination. The alarm email contains fixed text and the images captured beforehand.

c) Video Transmission

As designed, HONEY will transmit a 12-second fall video to CloudPIS once a fall is detected. In order to simulate the real application environment, we transmit the fall video to a commercial cloud-based storage in China, 115.com. 115.com provides the similar storage service as Dropbox does, has 30 million users in China. The time-limited link pointing to this video is also sent to the caregivers and emergency center along with the alert email. The caregivers can review the video through this link. The video transmission is time consuming and mainly depends on the bandwidth of network. In a home environment, a wireless LAN with ADSL or 3G cellphone network may be common network accesses. The bandwidth is relatively fixed, compared to the video size or video quality which we can adapt according to our requirements. In medical service, "Golden Hour" is the highest likelihood that prompt emergency treatment will prevent death [29]. There is also another phrase, "Emergency Platinum Ten Minutes", which means the survival rate is much higher in the first ten minutes after the injury occurred. So the video transmission is better to be completed in a short period after the fall occurred. In practical, we can adapt the resolution according to the severity of user's physical condition. If the user has heart disease, we select a lower resolution for reducing the transmission time in this situation, in which an immediately first aid is more important than a clear video. Otherwise even the transmission time is a little longer, but a high resolution video is more useful to determine the user's condition. Another factor affecting the video quality is the illumination condition. When the light intensity is low, the video may be indistinct. This may has an adverse influence on the caregiver's determination. As our system is used in a home environment, the illumination is sufficient at most conditions.

C. Privacy Protection

The disclosure of privacy usually happens when the sensitive information is transported on the Internet. There are three data transmission processes in HONEY: the Bluetooth communication between FallSensor and Homeserver, the video transmission and the images uploading. We will discuss the privacy protection concretely in the following paragraphs.

a) Bluetooth Communication

Privacy protection on Bluetooth communication includes two parts: one is the configuration of Bluetooth properties, and the other is the definition of special messages format. The following approaches are adopted to configure Bluetooth: First, we make the Bluetooth module work in a safe mode. FallSensor should be set to undiscoverable for other Bluetooth devices. Thus other irrelevant devices cannot launch an initiative connection with FallSensor. Second, an 8-bit decimal *Personal Identification Number* (PIN) code is used for establishing a safe channel. The PIN code is preset in HONEY. After the channel is established, Bluetooth provides a service level security, supporting authentication, encryption and authorization. Third, we use the *Media Access Control* (MAC) address to establish the channel directly and make this channel in point-to-point mode. Thus, Homeserver could find FallSensor without device searching even if FallSensor is undiscoverable. The specific messages' size is two bytes: the first byte is an *identity* (ID) that presents the current detection; the second byte is the message that presents specific information transferred between FallSensor and Homeserver. Figure 5 shows the format of four messages. The ID is generated randomly within 0 to 255, and FallSensor and Homeserver do not save it until current detection ends. It means that the messages own the same ID in the same detection. If the ID is different from the one that

FallSensor and Homeserver save, we believe that HONEY fails or is attacked, and we should resume HONEY.

b) SSL Communication

When HONEY determines to send an alarm email, six images will be sent as the attachments of the email. Those images will be uploaded to the mail server before delivery. In addition, the video will also be transmitted after a fall occurs. Thus, a privacy disclosure may happen during the images uploading and video transmission periods. Therefore, we establish a secure connection channel between HONEY and corresponding servers using SSL (Version 3.0). The data transferred through SSL is encrypted, and both endpoints are authenticated. Most common mail servers support SSL and it is easy to use SSL. In HONEY, we choose Gmail, as the mail service provider for the reasons that Gmail only provides SSL connection and it is common and widely used. The video storage server provided by 115.com also support SSL transmission between users and servers. We implement those functions with the protection of SSL.

V. EVALUATION

In this section, we design four experiments for evaluating the HONEY performance, includes Determination of SVM Threshold, Speech Recognition Accuracy, Video Transmission and Overall HONEY Performance. We invited 10 volunteers to do the following tests, include two females and eight males.

A. Determination of SVM Threshold

In HONEY, the fact that value of SVM exceeds *SVM_th* is the trigger condition of fall detection. Thus, a suitable *SVM_th* should distinguish the daily activities and falls. In order to select a reasonable value, we invited 3 volunteers to do this test. The average age, weight and height are 24 years old, 65 kg and 168 cm. The G3 is fixed on the waist of the volunteers. In the test, we ask volunteers to wear G3 and do the following tasks described in Table 2: walking, running, ascending stairs, descending stairs, sitting down, standing up, squatting, rising, and fall. All those activities are done in natural status without any restriction. When a subject performs those actions, the G3 will record SVM in real-time at 45 Hz. After all tests, we choose 500 sample values of each action except fall for further process. Due to the duration of a fall is too short, we only choose 200 sample values. Then we calculate the maximum, the minimum, the upper limit of 95% confidence interval and the lower limit of 95% confidence interval for each action statistically. Figure 6 provides all actions' results and the *SVM_th* (1.9g) we select.

From Figure 6 we can see that the *SVM_th* we select could distinguish most daily activities from fall except running, ascending stairs and descending stairs. The maximum values of walking, sitting down, standing up, squatting, and rising are lower than the *SVM_th*, so those actions will not trigger the detection. All data of fall and some data of running, ascending stairs, and descending stairs are above the *SVM_th*. Therefore, if the elder fall down, the event of SVM exceeding *SVM_th* will certainly trigger the detection of HONEY. Besides fall, running, ascend stairs and descend stairs also may be regarded as fall by mistake. On this situation, we can correct it in the following detection.

In this test, the acceleration data is acquired from the activities performed by those young volunteers. Though those data certainly differs from the data got from real elderly people's activities, in the following tests, all the volunteers we invited are in the similar physical state, including age, height and other aspects. Therefore, the *SVM_th* selected here is appropriate in the following test.

Table 2. Tasks Description.

Task	Description
Walking	The subject walks at their normal pace for 15 seconds at least.
Running	The subject runs at their normal speed for 15 seconds at least.
Ascending stairs	The subject ascends stairs continuously for 15 seconds at least.
Descending stairs	The subject descends stairs continuously for 15 seconds at least.
Siting down	From an initially standing position, the subject sits down and remains seated. Do this at least ten times.
Standing up	From an initially seated position, the subject stands up and remains standing. Do this at least ten times.
Squatting	From an initially standing position, the subject went down slowly and gets in to a squat. Do this at least ten times.
Rising	From an initially squate position, the subject stands up and remains standing. Do this at least ten times.
Simple Fall	Simple fall: Initially standing, the subject falls onto the mattress directly in an unspecified manner and remains lying on the mattress. Do this simple fall as least five times.
Complex Fall	Complex fall: Initially standing, the subject falls into a chair, or falls with grasping the wall and remains inactive. Do this complex fall at least five times.

B. Speech Recognition Accuracy

The speech recognition accuracy is an important factor affecting HONEY's performance. In practice, we cannot ensure the fall environment is quiet, and other voice sources, such as a TV or a radio, may exist in the room. Therefore, two sets of experiments are done to evaluate the speech recognition accuracy. One is the *Quiet Environment* test where the background volume is 25 dB。It is a typical quiet home environment. The other is the *Noisy Environment* test where the background volume is 65 dB. Sixty-five dB also is the normal speech volume. By utilizing playing music in a chamber, we keep the background volume of the test environment in 65 dB. In addition, the sampling quality of microphone decides the result of speech recognition. We preset the sample frequency at 96,000 Hz, and the bit-depth is 24. Then we do the test on different distance away from the microphone. At each distance, we ask three volunteers to speak out a fifty-word paragraph, which consists of five keywords mentioned in *sub-subsection* ***Speech Recognition***，**Section** . The average age, weight and height of those volunteers are 24 years old, 65 kg and 168 cm. The normal speech volumes of those volunteers are 62 dB, 64dB and 68 dB. Those keywords are repeated by the participants in disorder, and the volunteer speaks in his or her normal speech volume at the distance of 1m, 2m, 3m, 4m and 5m, which separate the microphone and the speaker. The microphone is installed on the floor, center of the chamber.

The average accuracy on each point is illustrated in Figure 7 for both environments. Along with the increasing distance, the

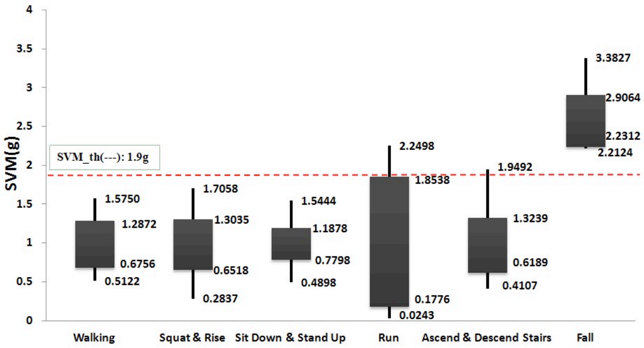


Figure 6. Ninety-five percent confidence interval of SVM. The dextral four data is on behalf of maximal value, upper limit of confidence interval, lower limit of confidence interval and minimal value in order. Selected SVM_th is marked (---).

accuracy decreases. It is also obvious that the accuracy in quiet environment is higher than that in noisy environment. In an ordinary family, the area of most rooms is smaller than 6×6 square meters. If we install the microphone in the middle of a room, the longest distance away from microphone in the room is within 4.3 meters except the corner areas. On this condition, the accuracy rate in quiet environment is above 90 % and 77 % for the noisy environment. The low accuracy in noisy environment is mainly caused by following three reasons: First, the microphone we used is sensitive to all directions, which is needed in our system, while this is adverse to the recognition accuracy. A professional microphone with a high unidirectional sensitivity may be the best choice for speech recognition, while this is not suitable in our application. Second, the noisy volume also has an influence on the accuracy. In our test, the noisy volume is roughly equal to the volunteers' voice volume. When the speaker is far away from the microphone, the actual speaker's voice volume detected by the microphone is lower than the speaker's normal speech volume showed aforementioned. So the noisy volume will be much higher than the speaker's voice volume. Of cause, the noisy volume may be higher than we assumed in our tests, this may leads to lower accuracy. Third, the speech recognition engine is not professional. This may also affect the accuracy. But this engine provided by Windows 7 OS is easy to use and access in home. Overall, the accuracy level can satisfy the practical application of HONEY in home environment. Therefore, in the overall HONEY performance test, we do all experiments within a certain range of 3 meters.

In this test, limitations also exist. When an elderly person falls, his or her voice may be weak or speech unclearly. This affects the performance of speech recognition. But this may not affect the performance of HONEY, as the accelerometer on FallSensor will continue to detect the fall if the speech recognition fails. Another limitation is the noisy voice volume. In the real scenario, the noisy volume changes according to the elderly persons' living environment. Here we assume that a TV or a radio's volume is equal to the normal speech volume. So we set the noisy volume is 65 dB. Sometime, the elderly person may need a larger volume to enjoy the TV or radio show, which is caused by weak auditory. Here we assume the auditory of elderly people is strong enough to heart the TV or radio's voice, even if with a hearing aid. Ideally closer the microphone located to the user, higher the recognition accuracy is. In this situation, the distance between the user and microphone is less than 1 meter, and the recognition accuracy closes to 100 % according to our experimental result showed

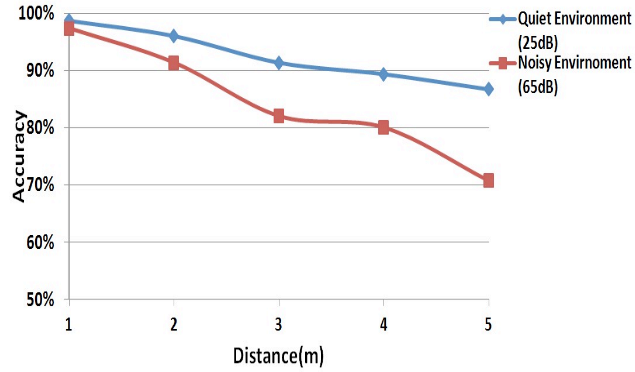


Figure 7. The average accuracy of speech recognition at different distances in quiet and noisy environment.

in Figure 7. Consequently what we plan to do is to integrate the microphone into FallSensor, and make HONEY more practical.

C. *Video Transmission*

This function has been implemented in our prototype. In our system, the video has constant time length and *frames per seconds* (FPS), so the video file size is mainly depend on the resolution. Thus a separated test evaluated the performance of video transmission function. Four kinds of resolution and two types of network access are considered in this test. As we all know, the bandwidth as well as the data length, affects the data transmission time for both upload and download. In China, the bandwidth of connection to ADSL is normally 2-Mbps in a home environment. The download speed is 256 kilobyte (KB) per second and the upper limit speed of upload is 56 KB per second. Another network access is 3G cell phone network. With the spreading among people, one can connect to network with a 3G cell phone. Hence the video transmission performance under 3G network is also evaluated. A computer uses a 3G network adapter to get the access to 3G cell phone network. The actual network bandwidth of 3G is dynamically changing based on the 3G signal quality in different areas.

Without other network load, we upload a 25.91 megabytes (MB) file to network storage, and repeat this operation five times in two different network conditions. The same to the download test. Then we conclude that the average upload speed is 44 KB per second, and the download is 256 KB per second with an ADSL access. And the average upload speed of 3G network is 94 KB per second, and the download speed is 59 KB per second.

Next, we kinescope 12-second video in different resolutions, and the FPS is fixed to 30. We did this with a proper illumination level, in which condition the video is clear enough to distinguish the objects. We estimate the upload and download time based on the actual network speed. Table 3 provides the resolution, the related file size and the estimated times in different network conditions. The *Estimated upload time* and *Estimated download time* are estimated by following equations:

$$\text{Estimated upload time} = \text{Size} \times 1024 \,/\, \text{Upload Speed}. \quad (3)$$

$$\text{Estimated download time} = \text{Size} \times 1024 \,/\, \text{Download Speed}. \quad (4)$$

According to the results showed in Table 3, it is obvious that no matter what the network condition is, the video file size and the estimated time decrease along with the resolution reducing. The total time varies from 5 minutes to 45 minutes. When the video is in the lowest resolution, the total time is about 335 seconds, less

Table 3. Estimated time for different network types.

Resolution (pixels)	Size (MB)	Network Condition	Estimated upload time(s)	Estimated download time(s)
640×480	97.4	ADSL	2,267	389
		3G	1,062	1,691
352×288	40.6	ADSL	945	162
		3G	443	705
320×240	30.8	ADSL	717	123
		3G	336	535
176×144	12.3	ADSL	286	49
		3G	134	214

than 6 minutes. It is important to make the total time within 10 minutes (Emergency Platinum Ten Minutes), which means the emergency aid will be opportune. But the poor quality video may affect the doctor's diagnosis. If we choose the video with a resolution of 320×240 or 352×288, the total time is less than 20 minutes. This exceeds 10 minutes, but compared to 60 minutes (Golden Hour), it is also appropriate in emergency service, and the video quality is more convenient for the doctor. With the highest resolution video, the total transmission time is over 45 minutes. Even though the video quality is quite well, the transmission time is too long for the emergency aid. From the perspective of time consumption, there is no significant timing difference between ADSL and 3G network. Because HONEY mainly works in home environment, the affection on the system caused by network type is not as apparent as in a public environment, where 3G network is easier to find and access. As HONEY's design, an alert message, the alert email will be sent to the caregivers once HONEY detects a fall before the video is transmitted. Due to the fact that the time consumption on email sending is much shorter than the video transmission, so we run the transmission task on Homeserver in the background, while the alert email is sent immediately.

D. Overall HONEY Performance

This test of the HONEY prototype is in a controlled lab environment. Figure 8 shows the real experimental scene. According to the result of speech recognition accuracy, all tests are performed within 3 meters. We also define *Response Time* -- the time between detecting the magnitude SVM and sending the alarm email. The delivery time of an alarm email depends on the network condition that is not controlled by us, so Response Time dose not includes this period. There are two groups of experiments based on different detection methods. One group is wearable sensor-based method as the comparison. The other is based on the HONEY system. We implement AMA on G3 as the wearable sensor-based method. In each group, we ask 10 volunteers to do the test. The average age, weight and height are 24 years old, 73 kg and 173 cm. Each volunteer does 12 actions including seven daily activities and five falls. Seven daily activities include walk-to-sit, run-to-lie, squat-to-sit, jump-to-sit, stand-to-sit, stair ascending and stair descending. Five falls include three simple falls and two complex falls. The simple falls and complex falls all are defined in Table 2. In all tests, the G3 is fixed on the waist of the subject during whole activity. In contrast, the tests for HONEY are different from AMA. The volunteers should speak out the keywords for speech recognition during the fall test. The test environment includes quiet environment and noisy environment.

Figure 8. The real experimental scene. The volunteer is lying on the mat. The camera is capturing images and the image and some text information are displayed on the UI.

Table 4. AMA and HONEY Statistics.

Detect Methods	Fall	Non Fall	Overall Correct	Overall Incorrect	False Positive	False Negative
AMA	50	70	76%	24%	22%	26%
HONEY	50	70	94%	6%	3%	10%

Other experiment conditions are the same. When the subject does those actions and a fall is detected by HONEY, the Response Time will be recorded.

Table 4 shows the results. The total accuracy of AMA is 76 %. The false positive rate is 22 %, and the false negative rate is 26 %. The reason for the low accuracy is the similarity between daily activities and falls. In those seven non-fall actions, run-to-lie and jump-to-sit are most similar to fall. The SVM of running and jump are easier to exceed the *SVM_th* than other movements, and this event will trigger the detection. The ending postures of those two activities are also similar to the simple fall and complex fall's ending postures. Therefore, those two activities are easy to be detected as falls according to AMA in our test. In all 15 false positive cases, 13 of them are caused by run-to-lie and jump-to-sit. The two remaining cases are squat-to-sit. This may be caused by the fact that the subject squats down and rises up rapidly. In the fall tests, AMA gives the wrong results for 13 times. Five of them occur in simple falls, while the remaining cases occur in complex falls. Most of those false negative cases are caused by the ending posture. In some tests, the subject did not lie on his or her back but sideways, and the G3 moves a lot caused by magnitude impact, which leads to a wrong posture prediction and an increase on wrong judgments.

When we use HONEY to do the same tests, the performance is better than AMA. The total accuracy of HONEY is 94 %, 18 % higher than AMA. The false positive rate is 3 % and the false negative rate is 10 %. In non-fall tests, HONEY misses two cases, which are caused by the run-to-lie. In non-fall test, HONEY may start detection without informing the users. Therefore, the user does no stopping operation to cancel detection, and then a false positive occurs. In fall tests, there are five cases missed by HONEY. One of them is caused by not meeting the trigger condition. The reason for the remaining cases is speech recognition failure and the system fault. Speech recognition causes three mistakes because the positive keyword recognition misses. Failure to send the alarm email also makes false detection. However, the total performance is much higher than AMA.

The responding time of HONEY is fast and reasonable. Previous system, such as iFall [30] has implemented a smartphone based fall detector, which will send a SMS to the pre-specified cell phone number. But the response time or detection time this system need is not evaluated. If we assume that the SMS sending and receiving can be completed in a short time which is negligible, those systems' response time is up to the detection algorithm. The AMA method does not send any alarm message out, so the response time is decided by the constant algorithm, which sets the response time of AMA between 12 seconds and 24 seconds. While in all tests for HONEY, total 47 emails including 45 true alarms and 2 false alarms emails are sent by HONEY. The average response time is 46.2 seconds, less than one minute. If we add the video transmission time into the response time, according to the results in subsection Video Transmission, Section V, the total time is less than 7 minutes and 21 minutes, under low and high resolution condition correspondingly. As a result, the patient can get the first aid from caregivers quickly.

In this test, all data is acquired through simulating the fall events performed by ten young volunteers. Actually, it is dangerous and difficult to obtain the real fall data in a real fall event. So there must be some differences between the simulating fall data we tested and the real fall data. First, the fall environment is different from the real fall condition, which may have no other protection measures to ensure the safety of the experimenters. Second, the data will also be different caused by the physical state of young volunteers and elderly people. Third, the instinctive reaction of young volunteers when a fall occurs is also more dexterous than elderly people. So we do not use the real fall data, which leads to a limitation in our experiment.

After all, HONEY provides a robust fall detection method for home-based environment. It reduces false positive and false negative significantly in accuracy, and makes the alert process fast enough for first aid. The consideration of privacy protection leads to a safe environment for applying our system.

E. Real-life Trial

Next, we apply our HONEY system into a realistic environment to verify the system performance and applicability. Eleven elder people living in Guangdong province, China, participated in our experiment voluntarily, including four males and sever females. The average age of them is 58 years old. All of them are living independently with their caregivers. For the safety consideration, it is not possible so that the elderly really falls in the experiment, and the children of these volunteers are also unwilling to allow us to do the fall test. So we only require the elderly to do the actions of daily life. In the test, those volunteers wearing our device can do free activities, and our system monitors the actions of the elderly continually. At the end of the experiments, each participant will do a short survey about their physical facts and suggestions or comments to our system.

During the test, the elderly did some daily life actions, mainly including walking, sitting down and standing up, ascending stairs and descending stairs, picking up things from the ground and lying in bed, and so on. The results show that a total of 46 times detection processes were triggered, and 4 false positive alarms were generated. One of these false positive alarms was detected by the TA judgment, and the rest of false positive alarms were caused by speech recognition failure. The false positive rate is 9%, higher than the 3% which we got in the lab environment experiment. In the aspect of speech recognition, our system conducted 14 times recognition processes, and succeeded 10 times, failed 4 times. The accuracy is 71%. The average response time is 49.7 seconds, and this is almost the same with what we acquired in the lab trial.

After the test, we analyzed the data, and the following several reasons may explain why the results are not good as we got in the lab trial. First, one of the false positive alarms causing by the TA checking is triggered by the action of lying in bed. Due to the acceleration of lying in bed acted by one volunteer is larger than the *SVM_th*, our system began the detection. However the lying in bed is very similar to falling on the ground, so the TA value is larger than 60° which means a fall on the ground. Then a false positive occurs. There are also few other actions that trigger the fall detection, such as ascending stairs and descending stairs, while these actions will not bring out a false positive because the TA values are less than the threshold of fall determination. Secondly, in aspect of speech recognition, during the whole test, only 14 times recognition processes are conducted. Although we trained the volunteers to use this function, the elderly rarely remember to use it. Another fact is that the accented dialect has a significant influence on recognition accuracy. As a result the recognition error rate is 29%, higher than what we got in the speech recognition experiment. The mainly reason we think is that in our lab trial, all the young volunteers can speak out the keywords using a nearly standard Mandarin. However part of the elder volunteers cannot speak standard Mandarin fluently, and our system perform well with a standard Mandarin. Those two reasons cause the high error rate in the real life trial. Thirdly, the response time is approximately the same with the lab result. The little difference is caused by the image and email processes.

Obviously, the biggest limitation of this test is no real fall detection. As we all know, fall is very dangerous to elderly and no elderly would like to fall intentionally. So we cannot verify the performance of our system for real fall detection. This is a realistic problem that rare similar systems have been applied in real life usage. Another limitation of our device is the battery lifespan. At present, our prototype device only can works continuously for four hours. This also makes it impossible to achieve a 24×7 hour monitoring. But we have already started to create a miniaturized device, which consumes much less energy and has a smaller volume. In the future, we will use this new device to implement a real all-day monitoring, and then the fall detection performance can be verified.

VI. RELATED WORK

Fall detection is an important application for elder healthcare. Fall can be associated with intrinsic or extrinsic factors. The fear of falling can lead to the deterioration of one's mental health, physical health, autism and the general degradation of his or her living quality [31]. Currently, there are mainly three approaches for fall detection in the existing systems:

1). Acoustic or vibration recognition: Alwan et al [32] achieve fall detection by using vibration recognition. It is based on the facts that human movements can cause measurable vibrations on the floor and a fall can be detected by monitoring the vibration patterns in the floor. Alwan's method can also distinguish a human fall from objects falling. Popescu et al [33] provide an acoustic fall detector equipped with two microphones. A linear array of acoustic sensors is used in the detector. When a sound is detected, the features are extracted from this sound for pattern recognition. If those features match the fall's features, an alarm is generated.

2). Video based detector: By installing a camera at a fixed location, one can track a user and learn movement patterns. The systems [34, 35] detect an event of a fall based on image processing that is designed to indentify unusual inactive. *Smart Inactivity Monitor using Array-Based Detectors* (SIMBAD) [36] uses an infrared camera to capture blurry images of the user. A fall will be confirmed by analyzing the images later. Nevertheless, the user experience is not good due to the bad feeling of being-watched. Adam Williams et al [14] use a distributed camera network to detect falling. To avoid the clumsy hardware and extra wearable device, a network of camera is proposed by Williams et al. In this solution, the energy consumption of each camera is low，and a number of cameras work together in a single room to complete the detection. Due to the multiple cameras they used, this application also provides 2D world coordinates of a fall.

3). Wearable sensor based method: A number of fall detection methods require the user to wear an external sensor. Usually these sensors are the triaxial accelerometer, the gyroscope or other devices. These systems [2, 9] develop different fall detection algorithms based on a triaxial accelerometer. A smartphone and wireless network are used in [37, 30]. Garrett Brown [12] gives a series comparison of fall detection algorithms. A clear evolution of those algorithms is provided according to the different scenarios. Other applications [3, 38] used for classifying the movements of human body could also detect fall among various body movements. ITALH project [39] implements several prototypes about home-based healthcare systems. Their works are similar to us, but the methods are different. ITALH provides a full-time video surveillance that may be annoying to a patient. In HONEY, the video is short, and only when the doctor needs it, the video is accessible. Moreover, the speech recognition is an important way to detect fall in HONEY, but ITLAH does not support. Another important fact is that HONEY is an on-line data process system, while ITLAH is off-line and the data should be collected first, and then an analysis could carry on.

Our HONEY has significant improvements and advantages compared to aforementioned classes. First, our system is easier to distribute and operate than acoustic and vibration recognition. This kind of method needs many acoustic or vibration sensors, which should be installed in a specified arrangement. However the microphones and cameras in our system can installed more freely. Moreover, in our future work, the microphone will be integrated in FallSensor, which make an easier installation. Second, our system is more robust comparing with other video based fall detector. All HONEY and video based method use camera as a significant tool to detect falls. When the illumination is weak, the camera may be disabled, and is not proper functioning. But the FallSensor of HONEY can work properly in a weak illumination environment, as well as speech recognition. So our system is more robust in some specific situations. Third, compared to the wearable sensor based method, HONEY adds speech recognition and video support for the user and caregiver, which makes HONEY more accurate and functional. Another important advantage is easy to use, which is a key factor affecting the user's choice.

The wireless protocol between FallSensor and Homeserver has multiple options, such as Bluetooth, Zigbee, and *Wireless Internet Service Provider* (WISP) [40] etc. But we decide to choose Bluetooth because of its wide deployment. The effective range of Zigbee is around 100 meters, much larger than a house's range, which stands a chance to disclosure personal information. WISP could be a good choice since it does not need power, but it is not widely available, and need a dedicated server. In addition, the effective range of WISP is up to 10ft with harvested RF power, which is not enough for covering the house.

In our application, we take some measures to protect the user's privacy. In Malasri's [41] opinion, implantable medical devices will play a major role in pervasive healthcare, enabling applications ranging from patient identification to remote administration of drug treatments. However, the relay and physic attacks in *internal identifications* (IIDs), *deny of service* (DoS) in *Intermediate Drivers* (IMDs) and the wireless reprogramming ability of *International code of Diseases* (ICDs) have not been well addressed in the literature. They propose some countermeasures, such as Data Encryption, Data Access Control etc. The data security is beyond the scopes we focus on. In the future, we can employ more protection measures in our system with more professional technology.

VII. REMAINING ISSUES AND CHALLENGES

Through the experimental study, we conclude that the following three remaining issues need further investigation. First, speech recognition is an important part to detect fall. But the results of lab trial and real life application show that dialectal voice recognition accuracy is not as good as the standard mandarin. Although Lu *et al.* have given a hybrid model of Hidden Markov Model and BP Neural Network to detect dialectal small-vocabulary, the corpus and training process could be serious problems [42]. Thus, it is hard to really use in a real-time recognition process, like fall detection. So a recognition system which is compatible with a variety of dialects is needed in a real setting. Another problem is where the microphone locates. According to our research results, the closer the microphone is located to the user, the higher recognition accuracy is. So we intend to integrate the microphone into FallSensor. The following result could be higher complexity and power consumption.

Second, the power management is another issue we are focusing on. The power consumption directly influences the lifespan of continuous monitoring without battery recharge. In addition, the results of a questionnaire following the real life trial show that a **half year lifespan** at least with no battery recharge is acceptable. At the same time, miniaturization reduces the battery capacity which aggravates the power supply. Based on our research, the following methods may relieve the power consumption. First, develop an algorithm using less acceleration data will reduce the energy consumption of the accelerometer and processor, because the sample frequency could be decreased. Secondly, if we leave the data to Homeserver to handle, the wireless transmission will consume most energy of whole device. Thirdly, if we integrate the microphone into FallSensor for enhancing speech recognition accuracy, extra energy consumption will be added. So how to balance those issues is also a tricky work.

Third, an individual SVM threshold for each user could make a system more adaptable. In our system, the threshold is fixed before real application, and this threshold could not satisfy everyone's requirements, because the physic condition and activity pattern are not all the same in most situation. An adaptive and autonomous learning mechanism may resolve this problem. Adding some algorithms such as BP Neural Network and Generic Algorithm, for an adaptive threshold determination can satisfy different cases of all users. But the algorithm running time and the computing resource requirement may be a big problem for the processor capability and also bring in extra power consumption. So a simple adaptive and autonomous learning algorithm will be the next target.

VIII. Conclusions and Future Work

In this paper, we present HONEY for fall detection in home-based environment. We use a tri-axial accelerometer to trigger the whole detection. In order to improve the performance and accuracy, we deploy speech recognition and images to reduce the false positive rate and false negative rate. The privacy protection is also considered in HONEY. Finally, we perform a comprehensive evaluation of HONEY. Our evaluation shows that HONEY provides reliable, safe, and responsive fall detection in a home environment.

As the next step, we intend to extend HONEY by leveraging *Body Area Networks* (BANs) [43] and online diagnose. BANs are made feasible by novel advances on lightweight, small-size and intelligent monitoring sensors, such as blood glucose, blood pressure, electrocardiogram (ECG), electroencephalogram (EEG), electromyography (EMG) etc. The patient's identification can be stored in the *Radio Frequency Identification* (RFID) tag, and the healthcare provider can retrieve not only the medical history, but also the physiological information in real-time. Thus, a medical staff can make a more accurate and timely diagnosis using those information, which can be utilized by data mining to make a diagnosis of potential disease.

IX. Acknowledgements

This work is in part supported by Tongji University, Introduction of Innovative R&D team program of Guangdong Province (NO. 201001D0104726115) and Wayne State University Wireless Health Initiative fund. We thank the anonymous reviewers for their valuable comments and Zhiying Lin for discussions during the early stages of the project. We are also grateful to the volunteer, especially the 11 senior volunteers, in our experiments.

References

[1] The Sixth National Population Census Data Bulletin of Shanghai 2010. *Shanghai Statistics Bureau*, May 2010. URL: http://www.stats-sh.gov.cn/2005shtj/tjgb/node94/userobject1ai14673.html.

[2] K. Doughty, R. Lewis, and A. McIntosh. The design of a practical and reliable fall detector for community and institutional telecare. *J.Telemed. Telecare*, vol. 6, Suppl. 1, pp. S150-154, 2000.

[3] A. Ozcan, H. Donat, N. Gelecek, M. Ozdirenc and D. Karadibak. The relationship between risk factors for falling and quality of life in older adults. *BMC Public Health*, vol. 5, pp. 90, Aug 26, 2005.

[4] R. W. Sattin, D. A. Lanbert Huber, C. A. DeVito, J. G. Rodriguez, A. Ros, S. Bacchelli, J. A. Stevens and R. J. Waxweiler. The incidence of fall injury events among the elderly in a defined population. *Am. J. Epidemilo.*, vol. 131, pp. 1028-1037, Jun 1990.

[5] A. Shumway-Cook, M. A. Ciol, J. Hoffman, B, J. Dudgeon, K. Yorkston and L. Chan. Falls in the medicare population: Incidence, associated factors, and impact on health care. *Phys. Therapy*, vol. 89, pp. 324-332, 2009.

[6] J. Fahrenberg, F. Foerster, M. Smeja and W. Müller. Assessment of posture and motion by multichannel piezoresistive accelerometer recordings. *Psychophysiol.*, vol. 34, pp. 607-612, 1997.

[7] F. Foerster and J. Fahrenberg. Motion pattern and posture: Correctly assessed by calibrated accelerometers. *Behav. Res. Meth. Instrum. Comput.*, vol. 32, pp. 450-457, 2000.

[8] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell and B. G. Celler. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Transaction on Information Technology in Biomedicine*, vol. 10, no. 1, Jan 2006.

[9] G. Williams, K. Doughty, K. Cameron and D. A. Bradley. A smart fall and activity monitor for telecare applications. *Proc. 20th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, 1998.

[10] M. Mathie, J. Basilakis and B. G. Celler. A system for monitoring posture and physical activity using accelerometers. *23rd Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society,* 2001.

[11] R. Sallen, D. MacKenzie, M. Mathie and B. G. Celler. Low power tri-axial ambulatory falls monitor. In *Proc. 10th Int. Conf. on Biomedical Engineering*, 2000.

[12] G. Brown. An accelerometer based fall detector: Development, Experimentation and Analysis. URL:http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.128.6988&rep=rep1&type=pdf.

[13] F. Bianchi, S. J. Redmond, M. R. Narayanan, S. Cerutti and N. H. Lovell. Barometric pressure and triaxial accelerometer-based falls event detection. *IEEE Transaction on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 6, Dec 2010.

[14] A. Willams, D. Ganesan and A. Hanson. Aging in place: Fall detection and localization in a distributed smart camera network. In Proceedings of the 15th International Conference on Multimedia, pp. 892-901, 2007.

[15] Q. Li, J. A. Stankovic, M. A. Hanson, A. T. Barth, J. Lach and G. Zhou. Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information. 2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks, pp. 138-143, Jun 2009.

[16] M. Eklund, T. Hansen and S. Sastry. Information technology for assisted living at home: building a wireless infrastructure for assisted living. The 27th Annual International Conference of the EMBS, pp. 3931-3934, 2005.

[17] M. Eklund. Executive Summary: Information Technology for Assisted Living at Home (ITALH). URL: https://buffy.eecs.berkeley.edu/PHP/resabs/resabs.php?f_year=2006&f_submit=one&f_absid=101094.

[18] Development trend forecast of China's aging population. *China National Committee On Ageing* URL: http://www.cncaprc.gov.cn/info/1408.html

[19] L. D. Gillespie, M. C. Robertson, W. J. Gillespie, S. E. Lamb, S. Gates, R. G. Cumming and B. H. Rowe. Intervention for preventing falls in older people living in the community. *Cochrane Database Syst. Rev.*, vol. 2, 2009.

[20] P. A. Stalenhoef, H. F. J. M. Crebolder, J. A. Knotternus and F. G. Van der Horst. Incidence, risk, factors and consequences of falls among elderly subjects living in the community: a criteria-based analysis. *Eur. J. Public Health*, vol. 7, No. 3, pp. 328-334, 1997.

[21] UN. World population ageing 1950-2050. Technical report, United Nations New York, 2001. URL: http://www.un.org/esa/population/publications/worldageing

19502050.
[22] A. Kochera. Falls among older persons and the role of the home: an analysis of cost, incidence, and potential savings from home modification. *AARP Public Policy Institute*, 2002. URL: http://www.aarp.org/home-garden/home-improvement/info-2002/aresearch-import-788-IB56.html.
[23] Y. Gu, A. Lo and I. Niemegeers. A survey of Indoor Positioning Systems for Wireless Personal Networks. In *IEEE Communication Surveys & Tutorials,* vol. 11, no. 1, First Quarter, 2009.
[24] Secure Socket Layer (SSL). URL: http://www.rfc-editor.org/cgi-bin/rfcsearch.pl.
[25] H. Mao, N. Xiao, W. Shi and Y. Lu. Wukong: toward a cloud-oriented file service for mobile devices. In *Proc. IEEE SCC 2010*, pp.498-505, Jul 2010.
[26] HIPAA. *Center for Medicare & Medicaid Services*. URL: http://www.cms.gov/HIPAAGenInfo/.
[27] J. Chen, K. Kwong, D. Cahng, J. Luk and R. Bajcsy. Wearable sensors for reliable fall detection. In *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 3551-3553, Sep 2005.
[28] Microsoft Speech API. URL: http://msdn.microsoft.com/en-us/library/ms723627(v=VS.85).aspx.
[29] Golden Hour. Cowley R. A. A new concept for the delivery of critical care. *J Med Soc N Jersey*, 1977.
[30] F. Sposaro and G. Tyson. iFall: an android application for fall monitoring and response. *31st Annual International Conference of the IEEE EMBS*, pp. 6119-6122, Sep 2009.
[31] E. W. Peterson, C. C. Cho and M. L. Finlayson. Fear of falling and associated activity curtailment among middle aged and older adults with multiple sclerosis. *Mult Scler*, vol.13, no. 9, pp. 1168-1175, 2007.
[32] M. Alwan, P. J. Rajendran, S. Kell, D. Mask, S. Dalal, M. Wolfe and R. Felder. A smart and passive floor-vibration based fall detector for elderly. *Information and Communication Technologies, ICTA'06. 2nd*, pp. 1003-1007, 2006.
[33] M. Popescu, Y. Li, M. Skubic and M. Rantz. An acoustic fall detector system that uses sound height information to reduce the false alarm rate. *30th Annual International IEEE EMBS Conference*, pp. 4628-4631, 2008.
[34] T. Lee and A. Mihailidis. An intelligent emergency response system: preliminary development and testing of automated fall detection. *Journal of Telemedicine and Telecare*, 11(4): 194-198, 2005.
[35] Shaou-Gang Miaou, Pei-Hsu Sung and Chia-Yuan Huang. A customized human fall detection system using omni-camera images and personal information. In *Proceedings of the 1st Distributed Diagnosis and Home Healthcare Conference*, pp. 39-41, 2006.
[36] A. Sixsmith and N. Johnson. A smart sensor to detect the falls of the elderly. *Pervasive Computing, IEEE*, 3(2):42-47, Jul 2004.
[37] T. Zhang, J, Wang, P. Liu and J. Hou. Fall detection by embedding an accelerometer in a cellphone and using kfd algorithm. *IJCSNS International Journal of Computer Science and Network Security*, 6(10): 277-284, Oct 2006.
[38] M. J. Mathie, A. C. F. Coster, B. G. Celler and N. H. Lovell. Classification of basic daily movements using a triaxial accelerometer. *Med. Biol. Eng. Comput.*, vol. 42, pp. 670-687, 2004.
[39] T. R. Hansen. Information technology for assisted living at home. *Progress report, EESC, University of California, Berkeley and Aarhus University*. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.116.3793&rep=rep1&type=pdf.
[40] Wireless Identification and Sensing Platform (WISP). *Intel Labs Seattle*. URL: http://www.seattle.intel-research.net/wisp/.
[41] K. Malasri and L. Wang. Securing wireless implantable devices for Healthcare: Ideas and Challenges. *IEEE Communication Magazine,* pp. 74-80, July 2009.
[42] Ling Lu, Xiangyang Zeng, Xiaobin Cheng, and Zhaoli Yan. Dialectal alarm words recognition based on a hybrid model of Hidden Markov Models and the BP Neural Network. *Journal of the Acoustical Society of America*, vol. 131, issue 4, pp. 3351-3351, 2012.
[43] M. Chen, S. Gonzalez, A. Vasilakos, Huasong Gao and C. M. Leung. Body Area Networks: A survey. *Mobile Netw Appl*, 16: 171-193, 2011.