A Theoretical Study on The Placement of Microphone Arrays for Improving The Localization Accuracy of A Fall

Yun Li, *IEEE student member*, K. C. Ho, *IEEE fellow*, Mihail Popescu, *IEEE senior member*, Marjorie Skubic, *IEEE senior member*

***Abstract*— Falling is a common health problem for elderly. To address the problem, we are currently developing an acoustic fall detection system, FADE, which automatically detects a fall and reports to the caregiver. Of great importance of the fall detection system is a low false alarm rate that can be achieved by knowing where the acoustic signal comes from. The previous work showed the sound source localization can be determined by using an 8-microphone circular array, but the accuracy varies when placing the array at different positions. To further improve the localization accuracy, a second array can be added. In this paper we investigate the variations of localization accuracy of a fall signal when one or two arrays are placed at different positions in a room. The accuracy is evaluated by the Cramér-Rao Lower Bound (CRLB). The CRLB aids the determination of the best theoretical placement of one or two arrays in a room for locating the sound source.**

I. Introduction

It is reported by CDC [1] that more than one third of about 38 million adults of 65 and older fall each year in the United States. A fall can cause serious health problems such as head injuries and hip fractures [1]. The total medical cost for all fall injuries exceeded $19 billion in 2000 [2]. Older adults who live alone are more likely to be unable to get help immediately after a fall occurs [3]. The annual risk for a person living alone of being found helpless or dead at home by paramedics is about 3.2% [4]. Other studies have shown that the longer the people lie on the floor, the poorer is the outcome of the medical intervention [4]. To address the problem of medical intervention delay, we need to develop fall detection techniques which detect the fall as soon as possible so that the immediate assistance can be provided.

A variety of fall detection methods have been published in the recent scientific literature. There are two types of fall detection devices: wearable and non-wearable. Wearable devices, like accelerometer-based ones, detect falls by measuring the applied acceleration along the vertical axis. However, we are more interested in non-wearable devices since such as floor vibration sensors [5], video cameras [6], radars [7] and smart carpets. Floor vibration sensors have the advantages of low cost and privacy preserving but their efficiency depends on the floor type. Video cameras, radars and smart carpets are promising technologies that are still trying to address the challenges related to cost, field of view and privacy. The acoustic sensors, proposed in early work [8-9], handle well with the light variation; however, they are susceptible to false alarms.

The early development of the acoustic fall detection system (FADE) [8-9] was based on a linear array of microphones. The advantage of using an array of microphones instead of just one is that the array can determine the position of the acoustic source which can be utilized to improve the fall detection performance. Since a fall sound mostly comes from the ground, [9] use the height estimate of a sound source to reduce the false alarm rate, and processes only the sound signals coming from near the ground (e.g. below 1m) for fall vs non-fall classification. This approach not only reduces the false alarm rate but also improves the computation efficiency. However, the height (location) estimation using a linear array is not reliable in the presence of strong interference and reverberations because it is a 1-D array, and 2-D or 3-D array geometries like a circle or cube can provide better performance. Another advantage of 2-D or 3-D array is that they can provide the point position of the source in 3-D, which is not possible for a linear array. As a result, [10-11] proposed an 8-microphone circular array for footstep tracking and fall detection. [10] shows that the footstep tracking errors are acceptable for signal-to-noise ratio (SNR) about 10dB. In addition to using one microphone array, utilizing multiple microphone arrays increases the effective array baseline which can further improve the performance.

The previous work has not examined how the localization accuracy varies with respect to the positioning of the microphone array. The spatial distribution of the sound source location accuracy obtained when placing the array at a certain position could help us determine the best placement of the microphone arrays before deploying the acoustic fall detection system in real-world. For this purpose, we evaluate the spatial distribution of localization accuracy from a microphone array using the Cramér-Rao Lower Bound (CRLB) [12]. The CRLB is a performance bound that tells the minimum possible mean-square estimation error of any unbiased estimator, when a set of data measurements at a certain SNR is given.

The structure of the paper is as follows: Section II briefly describes the CRLB technique. Section III presents the architecture of FADE. Section IV shows the results and Section V concludes the paper.

Y.L. was with Elect. and Comp. Eng. Dept., University of Missouri, Columbia, MO 65211, USA. (corresponding author's email: yl874@mail.missouri.edu).

M.P. is with the Health Mngmt. & Inform. Department, University of Missouri, Columbia, MO 65211, USA (email: popescum@missouri.edu).

K.C.H. and M.S are with Elect. and Comp. Eng. Dept., University of Missouri, Columbia, MO 65211, USA (email: hod@missouri.edu and SkubicM@missouri.edu).

II. DESCRIPTION OF CRLB

Cramér-Rao Lower Bound (CRLB) [13] places a lower bound on the variance of any unbiased estimator, which is the location estimator in our case. It is physically impossible to find an unbiased estimator with variance less than the bound. The bound is often achievable when using the Maximum Likelihood Estimator (MLE).

Let us assume that we wish to estimate a vector of parameters $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_p]^T$. The estimator $\widehat{\boldsymbol{\theta}}$ is assumed to be unbiased, that is $E(\widehat{\boldsymbol{\theta}}) = \boldsymbol{\theta}$. The CRLB gives the parameter σ_i^2, the i^{th} diagonal element of the inverse of Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$, that bounds the accuracy of $\widehat{\boldsymbol{\theta}}_i$:

$$\mathrm{var}(\widehat{\boldsymbol{\theta}}_i) \geq [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{ii} = \sigma_i^2. \quad (1)$$

$\mathbf{I}(\boldsymbol{\theta})$ is the $p \times p$ matrix defined by

$$[\mathbf{I}(\boldsymbol{\theta})]_{\mathrm{ij}} = -E\left[\frac{\partial^2 \ln p(\boldsymbol{X};\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right] \quad i,j = 1,2,\ldots,p \quad (2)$$

where $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N]^T$ is a sequence of N observations and $p(\boldsymbol{X};\boldsymbol{\theta})$ is the probability density function of $\boldsymbol{X}$ parameterized on $\boldsymbol{\theta}$. Note that each observation $\boldsymbol{x}$ could be described by a vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_M]^T$ of length M.

In our case, the unknown parameter of interest is $\boldsymbol{\theta} = [\theta_x,\ \theta_y,\ \theta_z]^T$, the 3D position of the unknown source. The observation is the time difference of arrival (TDOA) vector, obtained by cross-correlating the received signals from different elements in the microphone array. Given the positions of the M microphones, for each unknown source position, we can generate the corresponding TDOA's with respect to the first microphone, $t_{k,1} = t_k - t_1$ for $k = 2,3,\ldots,M$ and giving the TDOA vector $\boldsymbol{t} = [t_{2,1}, t_{3,1}, \ldots, t_{M,1}]^T$. As discussed by [12], if the spatial noise is incoherent and the noise power received by each sensor is identical, the covariance matrix $\mathbf{C}$ of the TDOA vector is equal to

$$\mathbf{C} = \rho^2 \begin{bmatrix} 1 & 0.5 & 0.5 & \cdots & 0.5 \\ 0.5 & 1 & 0.5 & \cdots & 0.5 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 0.5 & 0.5 & 0.5 & \cdots & 1 \end{bmatrix} \quad (3)$$

where ρ^2 is a constant related to the power spectral densities of the signal and noise. We define the true delay in distance for the signal arriving at each microphone with respect to the first one as $\boldsymbol{d} = [d_{2,1}, d_{3,1}, \ldots, d_{M,1}]^T$, where

$$d_{k,1} = \sqrt{(\theta_x - x_k)^2 + (\theta_y - y_k)^2 + (\theta_z - z_k)^2} - \sqrt{(\theta_x - x_1)^2 + (\theta_y - y_1)^2 + (\theta_z - z_1)^2},\ k = 2,3,\ldots,M \quad (4)$$

and (x_i, y_i, z_i) is the position of the i^{th} microphone. The TDOA vector is approximately Gaussian distributed with mean $\boldsymbol{d}/c$ and covariance matrix $\mathbf{C}$ [*], where c is the sound speed. Thus the likelihood function of the unknown parameter $\boldsymbol{\theta}$ given the observation $\boldsymbol{t}$ is

$$p(\boldsymbol{t};\boldsymbol{\theta}) = \frac{1}{(2\pi)^{(M-1)/2}|\mathbf{C}|^{1/2}} \times exp\left\{-\frac{1}{2}\left(\boldsymbol{t} - \frac{\boldsymbol{d}}{c}\right)^T \mathbf{C}^{-1}\left(\boldsymbol{t} - \frac{\boldsymbol{d}}{c}\right)\right\} \quad (5)$$

Since $\boldsymbol{d}$ is function of $\boldsymbol{\theta}$, putting (4) into (2) and taking its inverse, we can evaluate $[\mathbf{I}^{-1}(\boldsymbol{\theta})]$ and

$$\begin{aligned} \mathrm{var}(\hat{\theta}_x) &\geq [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{11} \\ \mathrm{var}(\hat{\theta}_y) &\geq [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{22} \\ \mathrm{var}(\hat{\theta}_z) &\geq [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{33} \end{aligned} \quad (6)$$

To form a single CRLB for the source position estimate $\widehat{\boldsymbol{\theta}}$, we take the trace of $\mathbf{I}^{-1}(\boldsymbol{\theta})$ to combine the variance information along the three coordinates:

$$\mathrm{var}(\hat{\theta}_x) + \mathrm{var}(\hat{\theta}_y) + \mathrm{var}(\hat{\theta}_z) \geq \mathrm{tr}[\mathbf{I}^{-1}(\boldsymbol{\theta})] \quad (7)$$

Thus, we calculate CRLB($\widehat{\boldsymbol{\theta}}$)= $\mathrm{tr}[\mathbf{I}^{-1}(\boldsymbol{\theta})]$ for each possible source location and generate a contour plot showing the distribution of the maximum possible localization accuracy over the area of interests.

III. SYSTEM ARCHITECTURE

FADE consists of one or two 8-microphone circular arrays. Each microphone has a mini amplifier and is mounted on a Cana Kit UK009 board. The microphones are installed on a plywood board in a circular pattern with a 25-cm radius. While a greater radius would improve the detection performance, it also limits the deployment options in an apartment setting. Also, to avoid spatial aliasing, the microphones of the array cannot be placed too far from each other. We choose a radius of 25cm based on the studies presented in [10]. The microphone array was hanged vertically on a wall facing the room. The working hypothesis for FADE is that the person is alone in an apartment (room) hence only one moving person has to be tracked. If motion is detected during a given interval (one minute) after a detected fall event, the caregiver alert is not issued. Instead, the event that provoked the alarm is recorded as a false alarm and used to retrain the classifier(s). In order to preserve the privacy of the resident, the sound will be internally processed on a microprocessor board and only the detection and classification results will be sent to the caregiver. Currently, we are mainly investigating the sound localization methods for fall detection. That is, we do not consider the motion detector and the communication processing with the caregiver. We are also developing a multiple-array system to locate a sound source. We have built two identical arrays (each has 8 microphones and 25cm radius). To aviod the synchronization issue between the two arrays, we record the data using a 16-channel analog input data acquisition card (NI-6212 DAC) for both arrays simultaneously (instead of one 8-channel NI-9201 DAC for each array).

In this paper, we are interested in studying the impact of the placement of one or two arrays on the localization accuracy over an area of interest. The study is helpful for determining the best placement of the arrays before conducting any real-world experiment. We prefer the placement with nearly uniform localization accuracy coverage over the entire room, with as high accuracy as possible.

The study emulates the deployment of the array(s) in a living room at the TigerPlace [14], an assisted living facility in Columbia, MO designed for aging in place, that has a size of 6.4×4.4×4 (Length×Width×Height) meters. We designate the 6.4 m wall, as the "long wall" and the 4.4m one as the "short wall". The arrays are placed on the walls without inclination. The center of each array has its height fixed at

2m except for the case when the array is placed on the ceiling. Thus, for comparison on the localization accuracy between one array and two arrays, also considering the symmetric property of the room space, we create a total of 9 scenarios of placement. Each scenario describes the array positions by specifying the 3D coordinates (x, y, z) of the array center(s) in meters, with respect to the origin at one corner on the ground of the room as detailed below

— *One Array*

Scenario 1: One array on long wall, (3.2, 0, 2)

Scenario 2: One array on short wall, (0, 2.2, 2)

— *Two Arrays*

Scenario 3: Both arrays on long wall, 3 meters apart with each other, (1.7, 0, 2), (4.7, 0, 2)

Scenario 4: Both arrays on long wall, 1 meter apart of each other, (2.7, 0, 2), (3.7, 0, 2)

Scenario 5: One array on long wall, the other on opposite, (3.2, 0, 2), (3.2, 4.4, 2)

Scenario 6: One array on short wall, the other on opposite, (0, 2.2, 2), (6.4, 2.2, 2)

Scenario 7: One array on long wall, the other on short wall, (3.2, 0, 2), (0, 2.2, 2)

Scenario 8: One array on long wall, the other on ceiling, (3.2, 0, 2), (3.2, 2.2, 4)

Scenario 9: One array on short wall, the other on ceiling, (0, 2.2, 2), (3.2, 2.2, 4)

For fall localization, we are more interested in the sound sources on the ground. Thus we only show the 2D CRLB contour plots on the ground when the height parameter $\theta_z = 0$.

IV. Results and Analysis

The CRLB contour plots of the 9 scenarios are in order of the scenario number. The measurement noise covariance matrix **C** is set according to equation (3). The localization error is normalized by the TDOA error standard deviation ρ, i.e. $\sqrt{\mathrm{CRLB}(\hat{\boldsymbol{\theta}})/(c\rho)}$. Only the localization errors in the x-y plane are shown since we are interested in the signals coming from the ground. The solid red circles in each figure denote the microphones. Also shown is the corresponding intensity image of each scenario. The localization accuracy is evaluated at a position resolution of 0.01m in both x and y coordinates. For each case, we can compare each scenario based not only on the contour values but also on the intensity by visual assessment (black means high-accuracy and white means low-accuracy).

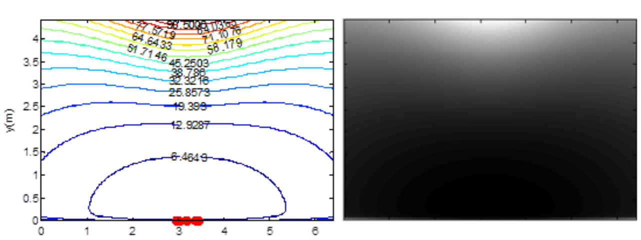

Fig 1. CRLB contour plot and intensity image for Scenario 1.

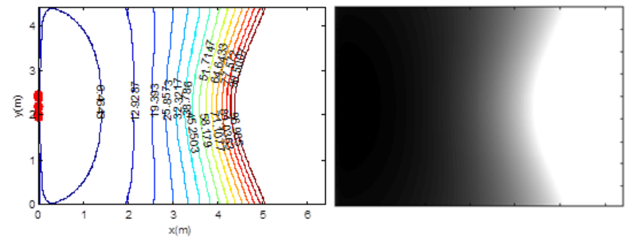

Fig 2. CRLB contour plot and intensity image for Scenario 2.

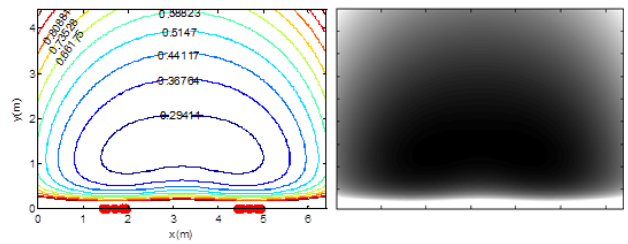

Fig 3. CRLB contour plot and intensity image for Scenario 3.

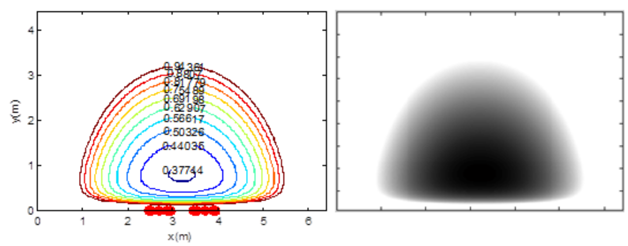

Fig 4. CRLB contour plot and intensity image for Scenario 4.

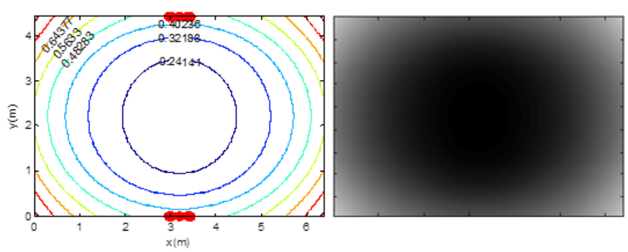

Fig 5. CRLB contour plot and intensity image for Scenario 5.

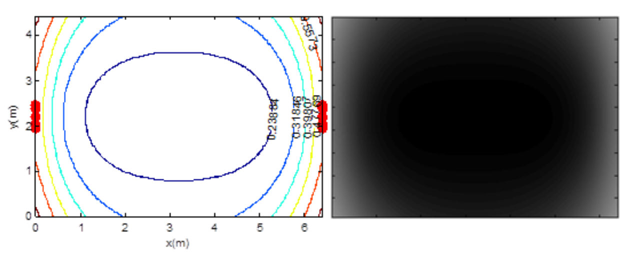

Fig 6. CRLB contour plot and intensity image for Scenario 6.

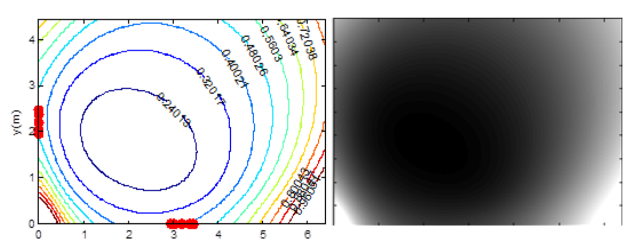

Fig 7. CRLB contour plot and intensity image for Scenario 7.

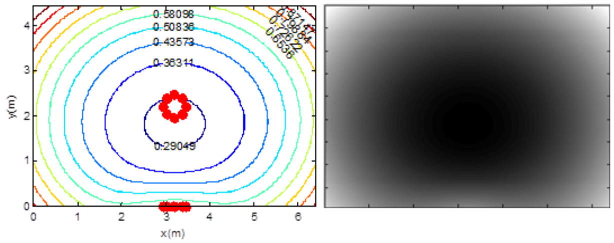

Fig 8. CRLB contour plot and intensity image for Scenario 8.

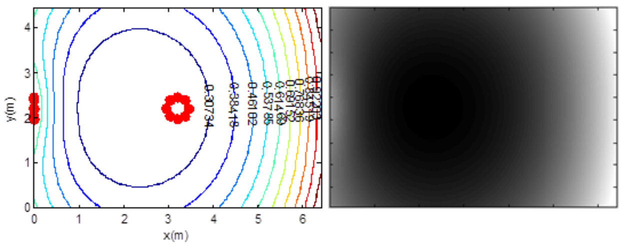

Fig 9. CRLB contour plot and intensity image for Scenario 9.

Figure 1 and Figure 2 show the distribution of localization accuracy using one array. The accuracy decreases as the source moves away from the array. It decreases fastest along the direction through the center of the array and perpendicular to the plane of the array. Obviously, Scenario 1 has larger coverage of high-accuracy area than Scenario 2.

Figure 3-9 show the spatial distribution of the localization accuracy using two arrays. Figure 3 and Figure 4 show that if both arrays are put on the same wall, the high-accuracy area significantly shrinks to a small part when the distance between of the arrays decreases. Figure 5 and Figure 6 show that if two arrays are separately placed on the opposite short walls, the accuracy is higher larger than when they are put on the opposite long walls. Figure 7 shows that if two arrays are put on the adjacent walls, the accuracy is significantly reduced at the corner areas although the accuracy is the highest near the center area. Figure 8 and 9 illustrate that if one of the arrays is on the ceiling, the accuracy in the corner areas is low and the accuracy in the center area is not as good as those of other scenarios. Thus, it appears scenario 8 should be the best choice. In addition, from the contour values of each case, the overall accuracy in the two-array case is much higher than that of the one-array case.

We have done some preliminary laboratory measurements and the initial results correlate very well with the theoretical study using CRLB.

V. Conclusions

In this paper, we investigated the effect of the placement of microphone array on the localization accuracy of an acoustic fall signal. To evaluate the localization accuracy we employ the CRLB, which provides the lower limit on the estimation variance of any unbiased estimator. Ideally, the microphone array(s) should be placed to achieve the height localization accuracy over as large area as possible over the entire room. The study indicates that a fall detection system with only one microphone array should be placed on the long wall (Scenario 1), while a system having two microphone arrays should be placed on the two opposite short walls (Scenario 7). We also observe that using two microphone arrays can increase the localization accuracy considerably compared with one microphone array.

The current study is idealized by assuming an empty room, direct line of sight propagation, ignoring the reverberation effects. We plan to develop a theoretical room reverberation model and include it in this study to account for more realistic acoustic environment such as different room size, room shape, noise type and presence of obstacles. The outcomes of the study on microphone arrays with different geometries such as the Microsoft Kinect device will be investigated as well. Finally, the fall detection performance using real-world dataset and the optimal array placement will be compared with the one without considering the optimal array placement.

Acknowledgements

This work has been supported in part by the NSF grant CNS-0931607.

References

[1] Center for Diseas Contro (CDC). http://www.cdc.gov/HomeandRecreationalSafety/Falls/adultfalls.html

[2] J. A. Stevens, "Fatalities and injuries from falls among older adults – United States, 1993–2003 and 2001–2005," *MMWR* 2006; 55(45).

[3] D. A. Sterling, J. A O'Connor, and J. Bonadies, "Geriatric falls: injury severity is high and disproportionate to mechanism," *Journal of Trauma–Injury*, Infection and Critical Care, 50(1): pp. 116–9, 2001.

[4] R. J. Gurley, N. Lum, M. Sande, B. Lo, and M. H. Katz, "Persons found in their homes helpless or dead," *N Engl J Med*.; 334(26): pp. 1710-6, 1996.

[5] M. Alwan, P. J. Rajendran, S. Kell, D. Mack, S. Dalal, M. Wolfe, and R.Felder, "A smart and passive floor-vibration based fall detector for elderly," in *Proc. 2nd IEEE Int. Conf. on Inf. & Comm. Tech.: from Theory to Applications*-ICTTA'06, April 24-28, 2006, Damascus, Syria.

[6] D. Anderson, R. H. Luke, J. Keller, M. Skubic, M. Rantz, and M. Aud, "Linguistic summarization of activities from video for fall detction using voxel person and fuzzy logic," *Computer Vision and Image Understanding*, vol. 113 (1), pp. 80-89, 2009.

[7] L. Liu, M. Popescu, K. C. Ho, M. Skubic, and M. Rantz, "Doppler radar sensor positioning in a fall detection system," in *Proc.* in *Proc. 34th Int. IEEE EMBS Conf.*, San Diego, CA, 2012, pp 256-259.

[8] M. Popescu and S. Coupland, "A fuzzy logic system for fall detection," in *Proc. AAAI Fall Symp.*, Washington DC, Nov 7-9, 2008, pp. 78-83.

[9] 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," in *Proc. 30th Int. IEEE EMBS Conf.*, Vancouver, BC, Aug. 20-24, 2008, pp. 4628-4631.

[10] Y. Li, Z. L. Zeng, M. Popescu, and K. C. Ho, "Acoustic fall Detection Using a Circular Microphone Array," in *Proc. 32th Int. IEEE EMBS Conf.*, Buenos Aires, Argentina, 2010, pp. 2242-2245.

[11] Y. Li, K. C. Ho, and M. Popescu, "A microphone array system for automatic fall detection," *IEEE Trans. Biomedical Engineering*, vol. 59, pp. 1291-1301, 2012.

[12] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Process.*, vol. 42, pp. 1905-1915, Aug. 1994.

[13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, 1993.

[14] Center for Eldercare and Rehabilitation Technology, Columbia MO. http://eldertech.missouri.edu/