Indoor Localization, Tracking and Fall Detection for Assistive Healthcare Based on Spatial Sparsity and Wireless Sensor Network

Mohammad Pourhomayoun, Department of Electrical and Computer Engineering, Binghamton University, State University of New York (SUNY), Binghamton, NY, USA

Zhanpeng Jin, Department of Electrical and Computer Engineering, Binghamton University, State University of New York (SUNY), Binghamton, NY, USA

Mark Fowler, Department of Electrical and Computer Engineering, Binghamton University, State University of New York (SUNY), Binghamton, NY, USA

ABSTRACT

Indoor localization and fall detection are two of the most paramount topics in assistive healthcare, where tracking the positions and actions of the patient or elderly is required for medical observation or accident prevention. Most of existing indoor localization methods are based on estimating one or more location-dependent signal parameters. However, some challenges caused by the complex scenarios within a closed space significantly limit the applicability of those existing approaches in an indoor environment, such as the severe multipath effect. In this paper, the authors propose a new one-stage, three-dimensional localization method based on the spatial sparsity in the x-y-z space. The proposed method is not only able to estimate and track the accurate positions of the patient, but also capable to detect the falls of the patient. In this method, the authors directly estimate the location of the emitter without going through the intermediate stage of TOA or signal strength estimation. The authors evaluate the performance of the proposed method using various Monte Carlo simulation settings. The results show that the proposed method is (i) very accurate even with a small number of sensors and (ii) very effective in addressing the multi-path issues.

Keywords: Compressive Sensing (CS), Fall Detection, Indoor Localization, Received Signal Strength (RSS), Sparsity, Time of Arrival (TOA)

DOI: 10.4018/ijmstr.2013040105
1. INTRODUCTION AND BACKGROUND

Indoor localization has been a long-standing and important issue in the areas of signal processing and sensor networks that has raised increasing attention recently (Pahlavan, Krishnamurthy, & Beneat, 1998; Pahlavan, Li, & Makela, 2002; Humphrey & Hedley, 2008; Li & Pahlavan, 2004; Becker et al., 2008; Chen & Kobayashi, 2002; Záruba, 2007; Shinno et al., 2008; Cheng et al., 2011; Hatami, Pahlavan, Heidari, & Akgul, 2006). As the number of elderly people grows rather quickly over the past few decades and continues to do so (He, Sengupta, Velkoff & DeBarros, 2005), it is imperative to seek alternative and innovative ways to provide affordable health care to the aging population (United Nations, 2009). A compelling solution is to enable pervasive healthcare for the elderly and people with disabilities at their own homes, while reducing the use and dependency of healthcare facilities. To this aim new technology and infrastructure must be developed for an in-home assistive living environment. One of the key demands in such an assistive environment is to promptly and accurately determine the state and activities of an inhabitant subject. The three-dimensional indoor localization provides an effective means in tracking the positions, motions, and reactions and detecting the falls of a patient, the elderly or any person with special needs for medical observation or accident prevention.

In assistive healthcare applications, the individual may wear a small device that could emit a radio frequency (RF) signal for localization. This emitter(s) propagates a signal that could be received and captured by several pre-mounted wireless sensors located in known positions. The sensors can estimate the location of the emitter after sharing some data and performing some processing.

The classic approach for localization is to first estimate one or more location-dependent signal parameters, such as time-of-arrival (TOA), angle-of-arrival (AOA) or received-signal-strength (RSS). Then in a second step, the collection of estimated parameters is used to determine an estimate of the emitter’s location. However, the AOA-based systems need multiple antennas or a scannable antenna that is usually very expensive (Humphrey & Hedley, 2008). The RSS-based methods require a costly training procedure and complex matching algorithms, while the positioning accuracy of these methods is largely limited by the large variance in indoor environments (Li & Pahlavan, 2004; Hatami, Pahlavan, Heidari, & Akgul, 2006). The TOA-based methods are usually very accurate. However, the accuracy of the classic TOA based methods often suffer from massive multipath conditions for indoor localization, which is caused by the reflection and diffraction of the RF signals from objects (e.g., interior walls, doors or furniture) in the environment (Pahlavan, Krishnamurthy, & Beneat, 1998).

In Comsa et al. (2011), the authors suggested a two-stage source localization method based on time-difference-of-arrival (TDOA) in a multipath channel exploiting the sparsity of the multipath channel for estimation of the line-of-sight component. In this method, the sensors don’t need to know the information on the specific transmitted symbols but, they require knowledge of the pulse shape of the transmitted signal. In Cevher et al. (2008), the authors presented a compressive sensing based distributed target localization. In this method, each sensor approximates the transmitted signal through its own received signal mapped to each one of the grid points. This idea helps to reduce the amount of data transmission in the sense of distributed localization but it lowers the quality of the estimation since each sensor estimates the transmitted signal just using its own received signal. Also, each sensor computes its own estimation of the emitter location that is not necessarily consistent with other sensors’ estimations.

In Pourhomayoun, Jin, and Fowler (2012), we proposed a 2-dimensional positioning approach based on the spatial sparsity and reported preliminary results for a stationary human target. However, that preliminary study came...
with several limitations, given the fact that various assumptions were made for the purpose of simplification and ease of implementation.

In this paper, we seek to exploit the spatial sparsity of the emitter (e.g., a device worn by the human subject) in the X-Y-Z space, and use the convex optimization theory to estimate the location of the emitter directly without going through the intermediate stage of TOA-based estimation. We will see that this method is very robust and effective in dealing with the multipath condition, which is a very serious problem in indoor localization due to the many reflections from furniture and walls. In our method, we don’t need any time synchronization between the emitter and the receivers, since the method is implicitly based on time difference of arrival (TDOA) between receivers. Moreover, the proposed approach can also estimate the height of the worn emitter above the floor, and thus immediately detect a fall of the patient, indicated as a sudden change of vertical position of the emitter.

We evaluate the performance of the proposed method using a set of Monte Carlo computer simulations for different cases. The simulation results show the accurate and fast localization performance of this method even in the conditions with multipath inference, low SNRs and small number of sensors; a superior advantage over conventional TOA, RSS or other single-stage methods.

2. PROBLEM FORMULATION

2.1. Spatial Sparsity Based Approach

As mentioned above, in this paper we exploit spatial sparsity of the emitter in the x-y-z space to estimate the location of the emitter directly. Assume that we divide up the X-Y-Z space into fine enough grids. It is manifest that in emitter localization problems, the number of emitters is much smaller than the number of all grid points in the X-Y-Z space. Thus, by assigning a positive number to each grid containing an emitter and zeros to all the rest of grid cells, we will have a very sparse 3-dimensional grid matrix that can be reformed as a sparse vector. In this context, a sparse vector is a vector containing only a small number of non-zero elements (Baraniuk, 2007). Since each element of this grid vector corresponds to one grid point in the X-Y-Z space, we can estimate the location of emitters by extracting the position of non-zero elements in the sparse vector. To this end, we have to estimate the sparsest vector that minimizes the cost between the predicted received signals and the actual received signals with respect to the signal model and delay relationship between the transmitted signals and the received signals.

In principle, sparsity of the grid vector can be enforced by minimizing its $\ell_0$-norm which is defined as the number of non-zero elements in the vector. However, since the $\ell_0$-norm minimization is an NP-hard non-convex optimization problem, it is very common (e.g., in compressive sensing research) to approximate it with the $\ell_1$-norm minimization, which is a convex optimization problem and can also achieve a sparse solution very well (Baraniuk, 2007). Thus, after formulating the problem in terms of the sparse grid vector, we can estimate this vector by pushing sparsity using $\ell_1$-norm minimization on the grid vector, and at the same time minimizing the cost between the predicted received signals and the actual received signals of the sensors.

2.2. Signal Model and Parameters

Suppose that an emitter transmits a signal and $L$ sensors receive that signal. The complex baseband signal observed by the $l$th sensor is:

$$r_i(t) = \alpha_l s(t - \tau_i) + w_i(t)$$

where $s(t)$ is the transmitted signal, $\alpha_l$ is the complex path attenuation, $\tau_i$ is the signal delay and $w_i(t)$ is a white, zero mean, complex Gaussian noise. Assume that each sensor collects $N_s$ signal samples at sampling frequency $F_s = 1 / T_s$. Then we have:
\( \mathbf{r}_i = \alpha_i \mathbf{D}_i \mathbf{s} + \mathbf{w}_i \)  \hspace{1cm} (2)

\( \mathbf{s} = [s(t_1), s(t_2), \ldots, s(t_{N_s})]^T \)

\( \mathbf{r}_i = [r_i(t_1), r_i(t_2), \ldots, r_i(t_{N_s})]^T \)

\( \mathbf{w}_i = [w_i(t_1), w_i(t_2), \ldots, w_i(t_{N_s})]^T \)

where \( \mathbf{r}_i \) is the vector containing \( N_i \) samples of the received signal by \( l \)th sensor, \( \mathbf{s} \) is \( N_s \) samples of the transmitted signal and \( \mathbf{D}_i \) is the time sample shift operator by \( n = (\tau / T_s) \) samples. We can write \( \mathbf{D} \) as an \( N_s \times N_s \) permutation matrix defined as

\[ [\mathbf{D}]_{ij} = \begin{cases} 1 & \text{if } i = j + 1, \end{cases} \quad [\mathbf{D}]_{i0,N-1} = 1 \quad \text{and} \quad [\mathbf{D}]_{ij} = 0 \text{ otherwise}. \]

Now, we assign a number \( \delta_{x,y,z} \) to each one of the grid points \((x,y,z)\). Assume that \( \delta_{x,y,z} \) is one for the grid points containing an emitter and zero for the rest of the grid points. Thus, the signal vector received by \( l \)th sensor will be:

\( \mathbf{r}_i = \sum_x \sum_y \sum_z \delta_{x,y,z} \alpha_{l,x,y,z} \mathbf{D}_{l,x,y,z} \mathbf{s} + \mathbf{w}_i \)  \hspace{1cm} (3)

where \( \mathbf{D}_{l,x,y,z} \) is the time sample shift operator with respect to sensor \( l \) assuming that the emitter is located in the grid point \((x,y,z)\) and the summations are over all grid points in the desired \((x,y,z)\) range. Now, if we reform all of the grid points in a column vector and re-arrange the indices, we will have:

\( \mathbf{r}_i = \sum_{n=1}^{N} \alpha_{l,n} \mathbf{D}_{l,n} \mathbf{s} + \mathbf{w}_i \)  \hspace{1cm} (4)

where \( N \) is the total number of grid points in the \((x,y,z)\) range of interest.

In assistive healthcare application, we can reasonably assume that the transmitted signal \( \mathbf{s} \) is known by the receiver sensors. However, in other applications when the signal is not known for receivers, we can consider the transmitted signal \( \mathbf{s} \) as a deterministic unknown signal. Then, for each grid point, we estimate the transmitted signal using the Minimum Variance Unbiased estimator (MVU) as:

\( \hat{\mathbf{s}}_n = \frac{1}{L} \sum_{l=1}^{L} \mathbf{D}_{l,n}^{-1} \mathbf{r}_i \)  \hspace{1cm} (5)

where \( \hat{\mathbf{s}}_n \) is the MVU estimate for the transmitted signal from grid point \( n \).

We define the matrix \( \Gamma_n \) as the delay operator with respect to all \( L \) sensors, assuming that the received signal comes from the grid point \( n \) (i.e., an emitter exists at the grid point \( n \)):

\( \Gamma_n = \begin{bmatrix} \alpha_{1,n} \mathbf{D}_{1,n} \\ \alpha_{2,n} \mathbf{D}_{2,n} \\ \vdots \\ \alpha_{L,n} \mathbf{D}_{L,n} \end{bmatrix} \) as an \( LN_s \times 1 \) vector containing all signals received by all \( L \) sensors when the emitter is at the grid point \( n \) as:

\( \theta_n = \Gamma_n \times \hat{\mathbf{s}}_n \)  \hspace{1cm} (7)

If we arrange all vectors \( \theta_n \) for \( n = 1, 2, \ldots, N \) as the columns of a matrix \( \Theta \) as:

\( \Theta = [\theta_1 \ \theta_2 \ \ldots \ \theta_N]_{LN_s \times N} \)  \hspace{1cm} (8)

then, the received signals are given by:

\( \mathbf{r} = \Theta \times \mathbf{v} + \mathbf{w} \)  \hspace{1cm} (9)

\( \mathbf{r} = [\mathbf{r}_1^T \ \mathbf{r}_2^T \ \ldots \ \mathbf{r}_L^T]^T \) as a \( LN_s \times 1 \) vector containing all signals received by all \( L \) sensors when the emitter is at the grid point \( n \) as:
\[ \mathbf{v} = [\delta_1 \delta_2 \ldots \delta_N]^T_{N \times 1} \]

where \( \mathbf{r} \) is the vector of all \( L \) received signals, \( \mathbf{v} \) is the sparse vector of \( \delta \)-values assigned to each grid point, and \( \mathbf{w} \) is the noise. Now, we can solve our problem by forming a BPIC (Basis Pursuit with Inequality Constraints) problem (Pourhomayoun, Jin, & Fowler, 2012; Pourhomayoun, Fowler, & Jin, 2012; Pourhomayoun & Fowler, 2012; Duarte & Eldar, 2011) as following:

\[
\begin{aligned}
\hat{\mathbf{z}} &= \arg \min \left\| \mathbf{v} \right\| \\
\text{s.t.} & \quad \left\| \Theta \times \mathbf{v} - \mathbf{r} \right\|_\infty \leq \varepsilon
\end{aligned} \tag{10}
\]

or regularized BPDN (Basis Pursuit Denoising) problem (Pourhomayoun, Jin, & Fowler, 2012; Pourhomayoun, Fowler, & Jin, 2012; Pourhomayoun & Fowler, 2012; Duarte & Eldar, 2011) as:

\[
\hat{\mathbf{v}} = \arg \min \left\| \Theta \times \mathbf{v} - \mathbf{r} \right\|_1 + \lambda \left\| \mathbf{v} \right\|_p \tag{11}
\]

where \( \left\| \cdot \right\|_p \) is the \( \ell_p \)-norm defined as \( \left\| \mathbf{v} \right\|_p = \left\{ \sum_i |v_i|^p \right\}^{1/p} \), \( \varepsilon \) is an appropriately chosen bound on the noise magnitude and \( \lambda \) is the regularization parameter balancing the sparsity versus estimation cost.

2.3. Reducing Computational Complexity for 3-Dimensional Localization

In this section we propose an alternative approach to reduce the computational complexity of the problem. This is very beneficial in cases when we have a large number of grid points, such as implementing 3-dimensional high-resolution localization in a large area. In this method, we can split the area of interest into several sub-areas and then find the best grid point that meets the minimization problem for each sub-area according to Equation 10 or 11. In a second step, we will only address the minimization problem among the selected grid points from each sub-area to find the final result. Note that the computational complexity of the minimization problem in (10) or (11) using traditional convex optimization methods for a vector \( \mathbf{v} \) of length \( N \), is about \( O(N^3) \). Thus, the computational complexity decreases significantly by reducing the size of the sparse vector.

Assume that we aim to estimate the position in a \( N_x \times N_y \times N_z \) grid space. Thus, the vector \( \mathbf{v} \) has \( N_x N_y N_z \) elements and the computational complexity of the minimization problem is \( O((N_x N_y N_z)^3) \). Now, if we split the grid space into \( N_z \) grid planes with size of \( N_x \times N_y \) each (as shown in Figure 1), we will have the complexity of \( O((N_x N_y)^3) \) for each plane.

![Figure 1. Splitting the 3D grid space into 2D grid planes](image-url)
Consequently, the total computational complexity will be \( O(N_z (N_x N_y)^3) \) for all \( N_z \) grid planes in addition to \( O(N_z^3) \) for final step. This is almost \( O(N_z^2) \) times smaller than the complexity of the original problem.

Another method that can be applied to reduce the computational load is to start with a coarse grid, and then increase the resolution of the grid iteratively. In other words, after estimating the approximate location of the target in a coarse grid space, we generate a finer grid in a smaller area of interest around the approximately estimated position, and then re-calculate the location in this new area. We continue this process until we achieve the desired positioning accuracy.

Furthermore, to reduce the computational complexity and achieve more accurate results in real-time patient tracking, we can exploit the prior trajectory information, to estimate the new position of the patient. In this case, we only need to start the localization process over the entire grid space once to estimate the initial position. After that, we can always restrict the area of interest to only a small vicinity around the past estimated position, considering the maximum possible movement of the subject. This technique leads to a significant reduction in size of the sparse vector and computational complexity, and also helps to achieve more accurate results by choosing a finer grid in the specific area of interest.

3. SIMULATION RESULTS

We examined the performance of the proposed method using multiple computer simulations in different scenarios. In the first scenario, we ran the Monte Carlo simulation with 500 runs each time for different numbers of sensors (from 3 to 8 sensors). We simulated the multipath conditions in a typical apartment layout as shown in Figure 2a. The sensors are mounted on the floor at \( x-y \) locations \((0,0), (0,10), (10,0), (10,10), (0,4), (4,10), (10,6), (6,0)\) respectively and the location of the target has been chosen randomly. In this simulation, we used a BPSK signal with carrier frequency of 1 GHz. The sampling frequency is 200 KHz and the number of samples is equal to 256. We run this simulation one time for \( \text{SNR} = 0 \) dB and another time for \( \text{SNR} = 10 \) dB.

Figure 2b shows the same apartment with four of the receiver sensors (i.e., red dots) mounted at the corners. Figure 2c illustrates a simple example for multipath scenario. In this figure, the solid (blue) lines present the direct paths (line-of-sight) and the dashed (red) lines indicate the reflected paths. However, given the extremely complex nature of the reflections within such a closed environment and the tremendous difference in the reflection rates for different building materials, it is impossible to conclude a rather perfect multi-path reflection model for the indoor circumstance. However, it is well agreed that the strength of reflected signals deteriorates after each reflection. Moreover, the TOA based localization systems usually suffer from first-order reflections since they generate the side-lobes very close to the main peak in the correlation stage used in traditional TOA based methods. Thus, the models like in Figure 2c seem reasonable for the purpose of research.

Figure 3 shows the RMS Error versus the number of sensors for estimating the location of the target. As we expected, the accuracy gets better by increasing the number of receiving sensors. The results demonstrate that the proposed method has a very good performance even for a small number of sensors (e.g., 3 sensors). This finding enables the possibility of using small number of sensors to reduce the complexity and cost of the system.

In the second scenario, we aim to track a moving target (e.g., a patient walking in the apartment). Again, we simulated the multipath conditions, and used only 4 sensors mounted at the corners of a room as shown in Figure 4a for the purpose of localization. In this simulation, we used a BPSK signal with the carrier frequency of 40 MHz. The sampling frequency is 80 KHz, the number of samples is equal to 256, and \( \text{SNR} = 10 \) dB. Figure 4a shows the ac-
Figure 2. (a) A typical apartment layout; (b) Four sensors mounted in the corners; (c) A simple case for multipath scenario. The solid line is direct path and dashed lines are reflected paths.
Figure 3. RMS Error for X and Y (meter) versus number of sensors
Figure 4. (a) True position of the patient (in blue) and the estimated position (in red); (b) Error in positioning for each location in part (a)
tual trajectory (blue line) of the patient walking around in the room, and the estimated path (red line) by the proposed system. Figure 4b shows the error in positioning defined as:

\[ e = \sqrt{e_x^2 + e_y^2 + e_z^2} \]  

(12)

where \( e_x \), \( e_y \) and \( e_z \) are root-mean-square (RMS) errors for positioning in the X, Y and Z dimensions. Note that the position of the human subject is not necessarily located at the grid points. However, the algorithm always finds the nearest grid point to the patient’s location. Thus, a significant portion of the errors in Figure 4b is resulted from the distance between the target and the determined grid point that is close to the target most. However, this kind of error can be reduced by employing finer grids.

In the third scenario, we used the proposed localization technique to detect the person’s fall, which is one of the most serious sources of injury for elderly people (Lan, Nahapetian, Vahdatpour, Au, Kaiser, & Sarrafzadeh, 2009; Degen, Jaeckel, Rufer, & Wyss, 2003; Noury, Fleury, Rumeau, Bourke, Laighin, Rialle, & Lundy, 2007). As mentioned before, we are always able to detect falls using the proposed 3-dimensional localization technique. Specifically, we obtain the accelerations in all three directions by computing the second derivative of the emitter position versus time, and then compare the norm of the accelerations against the predefined thresholds. For example, the second derivative of the height above the floor (versus time) can represent the acceleration in the Z-direction. The acceleration thresholding technique is an effective method for fall detection with a low false-positive rate (Degen, Jaeckel, Rufer, & Wyss, 2003).

Furthermore, the recognized location of the human subject in the X-Y plane can help to reduce the rate of missed detections (false-negatives). For example, the observation of a person lying down on the floor in the kitchen area is not expected. Therefore, by properly combining and interpreting the estimated location in the X-Y plane and the detected suspicious change in the Z-direction, our proposed approach can accurately recognize a life-threatening fall of the subject. This is the advantage of this method over traditional acceleration based fall detectors.

Figure 5 shows the actual path (blue line) of the human subject as well as the estimated path.
(red line) in a 3-dimensional space. It is shown that the system can track the exact position of the emitter (that can be conveniently worn by the people in their life) in the 3-dimensional space, and also detect fall activities.

4. CONCLUSION

Indoor localization is a very user-friendly tool in assistive healthcare environment to help track the locations, behaviors and reactions of the human subject (e.g., the patient, the elderly, or the people with disabilities) for medical observation, symptoms identification or accident prevention. On the other hand, timely fall detection can minimize the adverse impacts of unconscious falls to the patient or elderly. In this paper, we propose a novel method to address the increasing needs for both indoor localization and fall detection, based on spatial sparsity and wireless sensor networks.

Existing indoor localization methods are susceptible to performance degradation due to the likely occurrence of multipath reflections in an indoor setting. To combat this degradation, we develop a one-stage localization method based on spatial sparsity of the target(s) in the grid space. In this method, we assign a non-zero number to each one of the grid points containing an emitter (target) and zeros to all the rest of the grid points, all of which thus form a sparse, unknown vector. Since each element of this vector corresponds to one grid point in the grid space, we can estimate the location of the emitters by extracting the positions of non-zero elements of the sparsest vector that minimizes the cost between the predicted received signals and the actual received signals by the sensor network. We evaluate the performance of the proposed method using various Monte-Carlo simulations. The simulation results show that the proposed method has very good performance even with small number of sensors. The results also indicate that, in contrary to the classic TOA-based methods, the proposed approach is a very effective and robust tool to overcome multipath issues. Furthermore, the system works very well in noisy environments with low SNRs. It implies that, even with low transmitted power (to keep the worn device small with long battery life), we can still achieve a high localization accuracy.

REFERENCES


