# EV-Sounding: A Visual Assisted Electronic Channel Sounding System

Gang Li†, Jin Teng†, Fan Yang†, Adam C. Champion†, Dong Xuan†, Hong Luan* and Yuan F. Zheng‡
†Department of Computer Science and Engineering, ‡Department of Electrical and Computer Engineering,
The Ohio State University, Columbus, Ohio, USA, 43210
*School of Information Science and Engineering, University of Jinan, P.R. China
{lgang,tengj,yanfan,champion,xuan}@cse.ohio-state.edu, zheng@ece.osu.edu, luanhong.163@163.com

***Abstract*—Electronic channel sounding plays a vital role in developing wireless communication systems. It is critical for transceivers' equalization and filtering operations. However, current pure electronic channel sounding techniques are not well-suited for emerging scenarios such as opportunistic spectrum access, channel impulse response (CIR) based wireless positioning, and wireless security applications, which demand *rapid*, *high-resolution*, and *spectrum agile* channel measurements on *commercial off-the-shelf (COTS) devices*. To address these critical issues, this paper proposes EV-Sounding, a novel methodology for visual assisted electronic channel sounding. Based on frequency domain channel sounding, EV-Sounding leverages cameras for visual estimation of sparsity locations to reduce the number of frequency samples, thus speeding up the sounding process. EV-Sounding achieves both *high-resolution* CIR measurements and *spectrum agility*. We prototype an EV-Sounding system on *COTS* devices. Our real-world experimental results and extensive simulations validate EV-Sounding's performance.**

***Index Terms*—Channel sounding, wireless measurement.**

## I. Introduction

### A. Motivation

Electronic channel sounding has long played a crucial role in the development of wireless communications systems. It sends a known signal to probe the wireless channel in order to acquire knowledge of the channel's characteristics that can be described by the channel impulse response (CIR) or channel frequency response (CFR). Without information such as the CFR, wireless transceivers cannot efficiently perform equalization or other signal filtering. For example, if communications take place on a frequency-selective band without equalizing the received signal based on the band's CFR, the resulting bit error rate will be excessive [1]. While existing electronic channel sounding techniques have been effective in traditional wireless communications, they cannot be applied well in emerging scenarios. For example:

– *Opportunistic Spectrum Access:* Due to the increasing scarcity of available spectrum, opportunistically using frequencies (e.g., as in dynamic spectrum access (DSA) networks [2]–[5]) is very promising for the future. Channel sounding needs to be performed in an *agile* manner on limited frequencies to avoid interference with licensed users.

– *Wireless Localization:* Many RSSI-based localization techniques suffer from severe RSSI fluctuation caused by multipath effects [6], [7]. CIR measurements can help extract Line-Of-Sight (LOS) RSSI in order to improve range estimation significantly. This positioning technology requires channel sounding to be performed in a *timely* manner on *commercial off-the-shelf (COTS)* wireless devices with *high-resolution* CIR.

– *Security Key Establishment:* It is hard to establish secret keys among wireless users without a fixed infrastructure, especially in a dynamic mobile wireless environment. A link's CIR can be leveraged for cryptographic services [8]–[12]. *Rapid* channel sounding with *high-resolution* CIR on *COTS* wireless devices is also necessary for such purposes.

In summary, a timely frequency agile channel sounding with high resolution on COTS wireless devices is very desirable.

| Existing Solutions | Shortcomings |
|---|---|
| Direct pulse sounding [13], [14] | Unsuitable for COTS devices |
| Sliding correlator sounding [15], [16] | Low resolution; not spectrum agile |
| Swept frequency sounding [17], [18] | Time-consuming; not spectrum agile |
| Compressed channel sounding [19], [20] | Not spectrum agile |

TABLE I: Existing channel sounding techniques and their shortcomings

There are three conventional channel sounding techniques: direct RF pulse sounding, sliding correlator sounding, and swept frequency sounding. However, they cannot provide the desirable characteristics described above. Direct RF pulse sounding suffers from vulnerability to noise and interference, which entails complex hardware for COTS wireless devices. Sliding correlator sounding transmits a pseudo-noise (PN) signal whose power is uniformly distributed over a frequency band, but this technique can induce interference with some bands, it is not frequency agile, and interference mitigation can lower CIR resolution. Swept-frequency sounding can achieve

The work was supported in part by China 973 Project 2011CB302800, the National Science Foundation of China (NSFC) under grant No. 61070221, and the US National Science Foundation (NSF) under Grant No. CNS0916584, CNS1065136 and CNS1218876. Any opinions, findings, conclusions, and recommendations in this paper are those of the authors and do not necessarily reflect the views of the funding agencies.

978-14799-3360-0/14/$31.00 ©2014 IEEE

high-resolution CIR by stepping through a large range of frequencies, but this is time-consuming and not frequency agile.

Recently, compressed sensing (CS) [21] has been used for channel sounding [19], [22]. However, there are two major problems with CS. First, there is no reliable way to determine the CIR's actual sparsity (i.e., the number of paths). Second, CS imposes a random sampling requirement. In a network with spectrum regulations, this requirement cannot be met due to licensed users' spectrum presence. These channel sounding techniques and their shortcomings are described in Table I. Technological details are given in Section II.

### B. Our Contributions

In this paper, we propose a novel approach for channel sounding that is suited to new trends in wireless technologies and free from the shortcomings discussed above. In light of sensor fusion techniques and the proliferation of COTS cameras, we propose *EV-Sounding*, which assists Electronic channel sounding by leveraging Visual signals. EV-Sounding has two parts. First, we estimate the number of paths as well as the path lengths from video cameras. Ray tracing [23] is used for visual estimation. Second, we leverage restricted linear regression to fully exploit the path information acquired by cameras. As such path information enables us to estimate the CIR with a certain degree of accuracy (e.g., the number and positions of the CIR peaks), the remaining uncertainties (e.g., the amplitudes of the peaks) can be removed using a small number of measurements in the frequency domain via restricted linear regression. We design frequency selection algorithms to minimize the number of measurements. EV-Sounding is able to greatly reduce the number of frequencies we need to scan before we acquire the high-resolution CIR.

In summary, this paper makes the following contributions:

– We propose a methodology that leverages visual cameras to achieve fast, high-resolution, and frequency agile channel sounding on COTS wireless devices;

– We propose an effective and practical sampling algorithm to select frequencies. Specifically, we adapt it to situations where the frequency availability dynamically changes. We use the visualized path information to solve the linear regression problem;

– We prototype EV-Sounding on the COTS USRP platform and conduct real-world experiments in indoor environments to evaluate its performance. Further, we perform extensive simulations to evaluate EV-Sounding's robustness when used in noisy environments.

The remainder of this paper is structured as follows. Section II reviews related work. Section III presents our EV-Sounding system design. Section IV describes our EV-Sounding system implementation. Section V presents our experimental and simulation results. Section VI concludes.

## II. Related Work

In this section, we describe three categories of related research work: channel sounding, compressed channel sounding, and reconstruction of 3D visual models from 2D images.

The first category is channel sounding. In the literature, sliding correlator sounding and frequency domain sounding have received much attention recently. Parsons et al. review various channel sounding techniques [24]. Pirkl and Durgin describe quantitative design parameters for sliding correlator sounding [15]. Kivinen describes a wideband channel sounder based on PN sequences and presents measurements at 60 GHz [25]. Chen et al. address channel sounding in DSA networks by modifying the sliding correlator technique [2]. For frequency domain channel sounding, also called swept-frequency sounding, several works use vector network analyzers to measure the channel with high resolution [26]–[28]. Unlike these works, this paper leverages visual information to assist "traditional" channel sounding.

The second category is compressed channel sounding. Compressed sensing (CS) has been proposed for various applications [21]. In communications, CS is mainly used for sparse channel estimation. Berger et al. overview such applications in [20], [29]. Paredes et al. apply CS for sparse channel estimation in ultra-wideband channels [30]. Bajwa et al. discuss CS for multipath estimation, including four specific types of multipath channels [22]. Liu et al. introduce CS for sensing idle channels [3]. In all these works, CS requires random sampling, which does not achieve spectrum access agility.

The last category is reconstruction of 3D visual models from 2D camera imagery. The 2D camera images are taken in a 3D environment such as an indoor room or an outdoor environment. Works in this category aim to determine, or "reconstruct," the 3D environment from the 2D images taken in the environment. Techniques for such reconstruction are widely used in computer vision. There are two main approaches for reconstruction: structure from motion [31] and random sample consensus (RANSAC) [32], [33]. Ma et al. [34] performed reconstruction in indoor environments. Izadi et al. [35] used Microsoft Kinect cameras for reconstruction in indoor environments. Pollefeys et al. [36] performed reconstruction in outdoor urban environments. In contrast, this paper leverages visual technologies to reconstruct 3D indoor models.

## III. EV-Sounding Design

In this section, we first discuss the methodology for our EV-Sounding system. We illustrate the workflow of our system following this methodology. Then we discuss two important components in our system: visual path extraction and electronic frequency selection.

### A. Methodology

The CIR is the intrinsic property of a channel. For the CIR $\mathbf{h}$ and the input signal $\mathbf{x}$, the output signal $\mathbf{y} = \mathbf{x} * \mathbf{h}$, where $*$ denotes the convolution operation. The CIR can be acquired by transmitting an impulse and collecting all the echoes at the receiver end. Thus, the CIR has a physically meaningful interpretation: each peak in the CIR corresponds to a reflection path of electromagnetic waves and the position of each peak is determined by the travel time along each path.

Directly measuring the CIR is very hard. This requires transmitting a signal as close to an impulse as possible, which entails prohibitively high bandwidth processing capabilities for the underlying hardware. Thus, we attack this problem from a different perspective: the frequency domain. Each CIR has an equivalent counterpart in the frequency domain, the CFR $\mathbf{H} = \mathcal{F}(\mathbf{h})$, where $\mathcal{F}(\cdot)$ is the Fourier transform. If $\mathbf{H}$ and $\mathbf{h}$ are discretized (each of them is an $N$-point vector), we can write $\mathbf{H} = \mathbf{D}\mathbf{h}$, where $\mathbf{D}$ is an $N \times N$ discrete Fourier transform matrix. If we can measure $\mathbf{H}$ accurately, we can have $\mathbf{h} = \mathcal{F}^{-1}(\mathbf{H}) = \mathbf{D}^{-1}\mathbf{H}$. In the frequency domain, we can measure the frequency response on a part of the spectrum at one time. Suppose the baseband is $[-f_b, f_b]$ and the carrier frequency is $f_c$. By dividing the reception spectrum by the transmission spectrum, we can obtain the frequency response on $[f_c - f_b, f_c + f_b]$. If we change $f_c$, we obtain the frequency response on different frequencies. Then we put these frequency responses together to get the CFR for the entire bandwidth. Here, for simplicity, we suppose $f_b$ is very small. Thus, each time we "virtually" measure the response at one frequency. In this paper, $\mathbf{D}$ is also called the *measurement matrix* because we recover $\mathbf{h}$ from measurements of $\mathbf{H}$.

In order to exactly recover $\mathbf{h}$, we need to know all the $N$-element values of $\mathbf{H}$. $N$ is the number of specific frequencies that need to be measured and can be large (e.g., $N = 1024$). We can view this problem as follows. To solve the equation $\mathbf{H} = \mathbf{D}\mathbf{h}$ for $\mathbf{h}$, we need to know the entire $\mathbf{H}$. This means that we need to perform a full frequency scan and obtain the frequency response for all $N$ frequencies. However, much previous literature shows that the CIR is sparse (i.e., with few non-zero elements), normally with under 10 paths [22]. Physically, the CIR sparsity means that only a limited number of reflection paths contribute significantly to the CIR. Mathematically, this means that many elements in $\mathbf{h}$ are zeros. If we know that there are only $K$ non-zero elements (as well as these elements' positions), we only need $K$ element values in $\mathbf{H}$ to solve the equation $\mathbf{H} = \mathbf{D}\mathbf{h}$. $K$ is normally small ($K \leq 10$) and in order to combat measurement noise, we make more measurements [37] (i.e., we acquire $4K$ element values in $\mathbf{H}$). Thus, we have an over-determined equation and we can use linear regression to find an optimal solution that minimizes the $\ell_2$ norm.

Now we consider visual assistance. We can model the 3D environment from the images captured by video cameras and calculate the possible reflection paths. Though it might be hard for us to directly "read" reflection coefficients from the images, we can estimate the number of paths and the path lengths between the sender and the receiver. The echo arrival time is the length of each path divided by the speed of light. Thus, we can estimate the positions of non-zero values in the CIR. If there are $m$ paths, then ideally there will be $m$ non-zero elements in $\mathbf{h}$. If we consider visual measurement errors and assume surfaces are not smooth, a limited number of elements neighboring the $m$ non-zero elements may also be non-zero. Yet the total number of possible non-zero elements is far less than $N$ (e.g., $(c_n + 1)m$, where $c_n$ is the number of non-zero neighboring elements). With path information from the visual side, the necessary frequency response measurements are reduced from $N$ to $(c_n + 1)m$. This greatly accelerates the sounding process.

Mathematically, any $K = (c_n + 1)m$ element values of $\mathbf{H}$ can give us an exact solution of $\mathbf{h}$. With the redundancy requirement, about $4K$ values suffice. But which $4K$ values of $\mathbf{H}$ can give us the best performance in terms of linear regression? From linear regression theory, we know that the submatrix $\mathbf{D}_S$ of $\mathbf{D}$ corresponding to the $4K$ non-zero values of $\mathbf{H}$ plays a crucial role. The smaller the condition number $\kappa(\mathbf{D}_S)$ is, the more accurate $\mathbf{h}$ is. Thus we need to select frequencies to measure (i.e., select the rows of $\mathbf{D}$) to minimize the condition number $\kappa(\mathbf{D}_S)$. We develop algorithms to find suitable $\mathbf{D}_S$s later in this paper.

*Remark:* Our method is tailored for timely channel sounding in a radio environment with limited sampling frequencies. It needs much fewer measurements in the frequency domain than a full frequency scan. Unlike CS, which requires a random selection of frequency measurements and is hence impossible in networks where available frequencies are not randomly distributed, our method does not require complete randomness. We only need to select specific frequencies in available frequency sets carefully and use linear regression to recover the best possible CIR.

### B. Workflow

Fig. 1 decribes the main components and workflow for realizing our EV-Sounding technique.

The V-Component contains two parts: the visual data collector and the path extractor. The visual data collector is installed on a camera, which takes snapshots of the sounding scenario. Based on the snapshots, the path extractor models the 3D environment and calculates multipaths according to the positions of the transmitter and the receiver in the model.

The E-Component performs two tasks: frequency selection and electronic data collection. The frequency selector determines the specific frequencies to be sampled. The electronic data collector is a conventional swept-frequency sounder that measures the frequency responses of the chosen frequencies.

The signal processor uses the visualized path information and the incomplete frequency responses to estimate the CIR. Knowing sparsity locations, it estimates the CIR via linear regression on the electronic data.

The basic workflow of our channel sounder is as follows. We use the camera to take snapshots of the measurement environment. All visual data are sent to a back-end server for path extraction. Different multipaths are identified and the number of paths and the path lengths are sent to the frequency selector and the signal processor. The frequency selector chooses sampling frequencies and tunes the swept-frequency sounder to measure the CFRs on these frequencies. The signal processor uses path information from the V-Component and the frequency response from the E-Component to estimate the CIR.

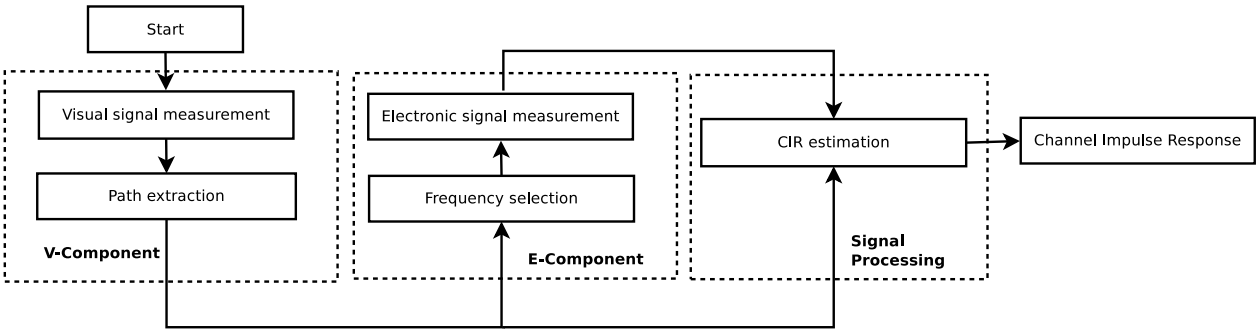



Fig. 1: Workflow

### C. Path Extraction

Path extraction is the key part of the V-Component, which needs to quickly obtain multipath information from the sounding scenario. In this paper, we only consider indoor scenarios for our prototype due to visual technology limitations. First, we estimate the 3D parameters of the sounding scenario. Then we use a ray tracing model for indoor electromagnetic wave propagation to estimate the principal path.

Overall, the V-Component estimates the indoor structure using several snapshots of the scenario from different angles that cover all critical edges determining the structure's geometric size. For example, the length, width, and height of a room determine its geometric size. We estimate these geometric parameters from snapshots with the help of reference shapes.

First, the camera is calibrated via the chessboard calibration method [38] to determine its intrinsic and distortion parameters and all snapshots are rectified. Then the reference shapes on each plane – the walls, ceiling, and floor – are used to calculate the planar homography matrix, which is a projective mapping from real-world coordinates to the image plane. Knowing each reference shape's size, the lengths of the real-world planar edges can be determined. In this way, visual technology helps estimate the physical size of the indoor scenario, which determines the multipath effect on electromagnetic wave propagation. In a more complex scenario, the indoor structure may not be regular, or there may be obstacles or reflectors that affect multipath propagation. In this case, we need more sophisticated methods to model this environment such as depth maps and RANSAC [32], [33]. Camera calibration only needs to be performed once if the camera's position is fixed.

Having reconstructed the 3D model for the sounding scenario, we use ray tracing with the mirror method [23] to estimate possible multipaths. If the scenario is relatively "simple," the mirror method can effectively obtain information about each principal path. In a "complex" scenario, we apply the ray-launching method [39], which discretizes the transmitting antenna's output signal into a large number of rays and tracks each ray to calculate its transmission, reflection, and diffraction loss. Then we can determine what paths contribute to the receiving signal.

Though we can obtain the CIR with just the V-Component, this estimation is flawed, especially regarding the accuracy of principal paths' magnitudes as we cannot precisely determine transmission and reflection loss. Thus, we only extract the approximate number and locations of the principal paths and feed this information to the accurate CIR reconstruction module in the E-Component. There, the effect of the estimation error introduced by the V-Component can be eliminated somewhat.

### D. Frequency Selection

When sampling the frequency response, our system needs not scan all available frequencies as a subset thereof suffice to obtain the CIR. We only need to select the subset. As each frequency corresponds to one row of the measurement matrix, frequency selection entails submatrix selection from a given matrix. For compressed sensing (CS), the selected submatrix should satisfy the restricted isometry property (RIP). CS randomly selects frequencies to meet this property. For linear regression, the measurement matrix needs to be well conditioned. The condition number $\kappa(\cdot)$ measures the matrix's sensitivity to perturbations. Small condition numbers are preferable to large ones. Thus, our goal is to find a set of frequencies among those available that form a measurement matrix with the smallest condition number. However, this is a very hard problem. To the best of our knowledge, no polynomial time algorithms have ever been proposed. In this paper, we use locally optimal algorithms to find submatrices with low condition numbers.

We classify frequency selection into two cases based on characteristics of underlying available frequencies. In the first case, before we start sounding, we know all the available frequencies during the entire measurement period. The availability of these frequencies is invariant during this period. In the second case, before each sampling, we only know a few frequencies that are available during this sampling period. We select one of the unused frequencies as the sampling point. This is true in certain scenarios where available channels dynamically change over time such as in DSA networks. Each time we want to access the spectrum, we need to first sense the available frequencies. For each situation, we use a different algorithm to select frequencies.

*Static Availability:* Suppose $M$ frequencies are available during the entire sounding period. Based on visually provided sparsity location information, we can build a measurement matrix $\mathbf{D}_{M\times N}$. Suppose $K$ frequencies are needed to estimate the CIR to a certain degree of accuracy. Then we need to choose a set of frequencies $F$ corresponding to a $K$-row submatrix $\mathbf{D}_S$ such that $\kappa(\mathbf{D}_S)$ is minimized. We leverage the

algorithm proposed in [40] to select $F$. This algorithm is given in Algorithm 1.

**Algorithm 1** Frequency_Select_Static($\mathbf{D}_{M\times N}$, $K$)

$\mathbf{D}_S \leftarrow \mathbf{D}_{M\times N}$; $i \leftarrow 0$; $F \leftarrow \emptyset$
**for** $i \leq (M-K)$ **do**
  $r \leftarrow$ number of rows in $\mathbf{D}_S$; $j \leftarrow 0$
  **for** $j \leq r$ **do**
    $\mathbf{D}_i \leftarrow \mathbf{D}_S$
    Delete $j$-th row from $\mathbf{D}_i$
    $C(i) \leftarrow \kappa(\mathbf{D}_i)$
    $j \leftarrow j+1$
  **end for**
  Find $i_{min}$ that minimizes $C(i)$
  Delete row $i_{min}$ from $\mathbf{D}_S$
  $i \leftarrow i+1$
**end for**
$F \leftarrow$ frequencies correspond to rows in $\mathbf{D}_S$
**return** $F$

*Dynamic Availability:* In this scenario, we assume the required frequency set is $F$ with size $K$. We divide the entire measurement period into time slots of equal duration. During each time slot, we sample only one frequency from available frequencies that have been sensed. During a time slot, there may be no available frequencies (i.e., all frequencies sensed in the current time slot have already been sampled), so we can skip this time slot. Hence the total number of time slots may exceed $K$. We propose a similar locally optimal approach to construct $\mathbf{D}_S$, which is shown in Algorithm 2. At each time slot $t$, a set of frequencies $C$ is available for access. We select a frequency $c \in C$ that has not already been sampled such that adding the corresponding row in $\mathbf{D}_S$ minimizes the condition number $\kappa(\mathbf{D}_S)$. Then we actually add the row that corresponds to frequency $c$ into $\mathbf{D}_S$. We also add $c$ to $F$.

**Algorithm 2** Frequency_Select_Dynamic($K$)

$\mathbf{D}_S \leftarrow \mathbf{D}_{0\times N}$; $i \leftarrow 0$; $t \leftarrow 0$; $F \leftarrow \emptyset$
**for** $i \leq K$ **do**
  $C \leftarrow$ available frequencies at current time slot $t$
  **if** $C - F \neq \emptyset$ **then**
    $c \leftarrow$ frequency in $C-F$ such that adding corresponding row $R$ to $\mathbf{D}_S$ minimizes $\kappa(\mathbf{D}_S)$
    Add row $R$ (corresponding to $c$) to $\mathbf{D}_S$
    $F \leftarrow F \cup \{c\}$
    $i \leftarrow i+1$
  **end if**
  $t \leftarrow t+1$
**end for**
**return** $F$

## IV. Implementation

In this section, we describe the implementation of our EV-Sounding system, including the visual and electronic parts.

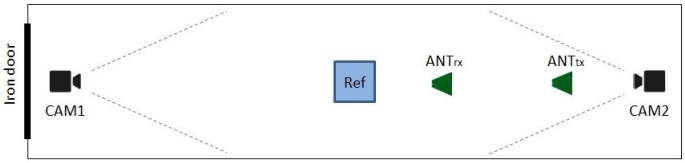



Fig. 2: Top view of the corridor scenario schematic

### A. V-Implementation

We prototype the V-Component based on the design principles described in Section III. The cameras used in our system are two programmable SJM-200C generic digital cameras with adjustable resolutions and whose transmission speeds are at most 480 Mbps. A laptop that serves a back-end server collects the images taken by the cameras and processes them.

As an example, we deploy our V-Component in a empty corridor scenario whose top view is shown in Fig. 2. $ANT_{tx}$ and $ANT_{rx}$ are the transmitting and receiving antennas, respectively. $CAM1$ and $CAM2$ are two spots where the snapshots are taken to cover the entire communication scenario. $Ref$ is a reference shape for measurement. In this scenario, we calibrate the camera via the chessboard method. Also, we use reference rectangles on the walls and the floor to calculate the planar homography matrix. In the corridor example, we estimate the size of the surrounding wall planes in order to reconstruct them.

### B. E-Implementation

We implement the E-Component on the USRP platform. The USRP is a low-cost software defined radio platform developed by Ettus Research [41]. We build a swept-frequency sounder based on the USRP N210 with an SBX daughterboard, which is a wide bandwidth transceiver that supports full duplex operation. Specifications for the N210 and SBX are given in Tables II and III, respectively.

TABLE II: USRP N210 Specifications

| | |
|---|---|
| ADC Rate | 100 MS/s |
| ADC Resolution | 14 bits |
| DAC Rate | 400 MS/s |
| DAC Resolution | 16 bits |
| FPGA | Spartan 3A-DSP 3400 |
| Interface | Gigabit Ethernet |

TABLE III: SBX Daughterboard Specifications

| | |
|---|---|
| Band Range | 400–4400 MHz |
| Output Power | Maximum 100 mW |
| IF Bandwidth | 40 MHz |

For the antenna, we use two log-periodic PCB directional antennas with frequency ranges from 850 MHz to 6.5 GHz. Both antennas have 5 dBi gains within their forward radiation cones. A very small amount of energy emanates from the sides and back of the antenna. The beamwidth is $\sim 120^\circ$ in the horizontal plane and $\sim 160^\circ$ in the vertical plane.

The setup for the E-Component is shown in Fig. 3. The USRP is connected to a laptop that serves as a back-end

server via an Ethernet cable. The two antennas are connected to the USRP via two low-loss SMA cables. The server sends commands to the USRP and stores the digitized data for post-processing.

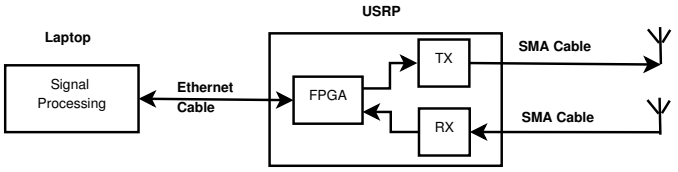



Fig. 3: Setup of the swept-frequency sounder

In practice, the measured frequency response $\mathbf{H}(f)$ is not the real frequency response of the wireless channel. $\mathbf{H}(f)$ is actually the product of the frequency responses of each part of the entire measurement system. $\mathbf{H}(f)$ can be represented as

$$\mathbf{H}(f) = \mathbf{H}_u(f) \cdot \mathbf{H}_{cab}(f) \cdot \mathbf{H}_{ant}(f) \cdot \mathbf{H}_{cha}(f),$$

where $\mathbf{H}_u(f)$, $\mathbf{H}_{cab}(f)$, $\mathbf{H}_{ant}(f)$, and $\mathbf{H}_{cha}(f)$ denote the frequency responses of the USRP, the cables, the antennas, and the wireless channel, respectively. Since we only care about $\mathbf{H}_{cha}(f)$, we need to calibrate the other terms. Let $\mathbf{H}_{cal}(f)$ represent the multiplication of these terms (i.e., $\mathbf{H}_{cal}(f) = \mathbf{H}_u(f) \cdot \mathbf{H}_{cab}(f) \cdot \mathbf{H}_{ant}(f)$).

Calibration is performed in an anechoic chamber where the two antennas are placed at a line-of-sight distance $D_{cal}$. The chamber is full of "absorbing" materials; hence, there is only one path between the TX and RX. Suppose the measured frequency response for the whole system including the wireless channel in the chamber is $\mathbf{H}_{chamber}(f)$. Then

$$\mathbf{H}_{chamber}(f) = \mathbf{H}_{cal}(f) \cdot \mathrm{e}^{-j2\pi f D_{cal}/C},$$

where $C$ is the velocity of electromagnetic propagation in free space. In order to obtain the CFR $\mathbf{H}_{cha}(f)$, we compute

$$\mathbf{H}_{cha}(f) = \frac{\mathbf{H}(f) \cdot \mathrm{e}^{-j2\pi f D_{cal}/C}}{\mathbf{H}_{chamber}(f)}.$$

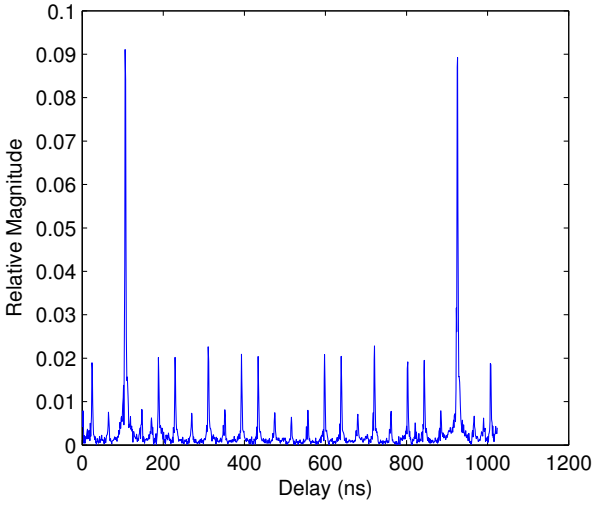
We set the TX and RX antennas at a line-of-sight distance $D_{cal} = 2.7$ m in the anechoic chamber and sweep the frequency with a range of 1024 MHz (i.e., the time resolution is about 0.98 ns). Figs. 4(a) and 4(b) show the IFFT result of the frequency response before and after calibration, respectively. Since there is only one path whose distance is 2.7 m, there should be only one peak at the position 9 of the $x$-axis, which is correctly shown in Fig. 4(b).

## V. Evaluation

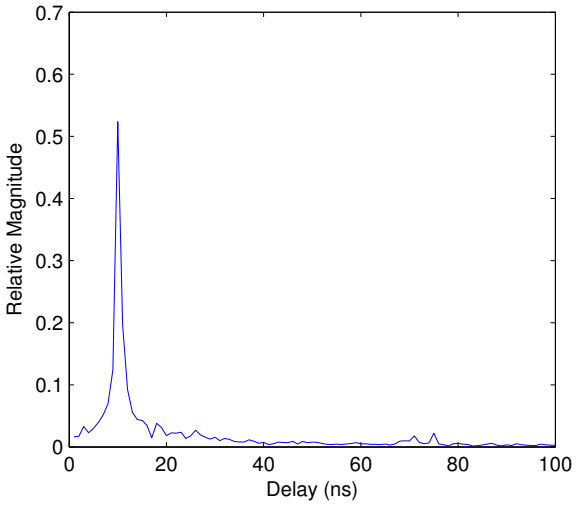
In this section, we evaluate the performance of our proposed EV-Sounding system. We conduct experiments in two indoor scenarios. We also conduct simulations to evaluate the robustness of our proposed approach under different noise conditions and sparsities.

### A. Real World Experiments

All the experiments are conducted in the 1–2.023 GHz frequency band. The band is divided into 1024 sub-channels, each of which has a bandwidth of 1 MHz.




(a) CIR without calibration




(b) CIR after calibration

Fig. 4: CIR for one 2.7m path

*1) CIR Estimation:* We conduct experiments to estimate the CIR in two indoor situations. One is a corridor, a 5.2 m $\times$ 1.2 m $\times$ 2.6 m empty space with a metal door on one side and a glass door on the other side. We choose this empty corridor to represent a simple indoor environment that facilitates visual 3D modeling. The other is a small office, a 2.7 m $\times$ 5 m $\times$ 2.5 m room with two desks and one cabinet in it. This represents a complex indoor environment.

For the empty corridor, the camera sees six paths in total including one Line-of-Sight (LoS) path and five one-time reflection Non-Line-of-Sight (NLoS) paths. For the office, five paths including one LoS path and four one-time reflection NLoS paths are obtained via ray tracing. For both scenarios, we only consider one-time reflections as the signal will be attenuated greatly after undergoing multiple reflections. Detailed path information is given in Table IV.

We first conduct conventional swept-frequency channel

sounding, which uniformly samples 1024 frequencies across the whole band and calculates the CIR $\mathbf{h}_{ref}$ via the IDFT of the CFR. We consider $\mathbf{h}_{ref}$ as an "oracle" and we use it as a reference in following experiments. The oracle CIR and our sounding result $\widehat{\mathbf{h}}$ for the corridor and the office are shown in Figs. 5a and 5b, respectively. From the figures we can see that our sounding results match the main components of the oracle CIR. Based on the arrival times, we can see there are two main paths in the corridor. The first one corresponds to the LOS path and the second one corresponds to the path reflected by the metal door. Because of our antenna's directionality, other components are not significant. Two paths with 37 ns and 69 ns arrival times, respectively, are not estimated by our sounder as these two paths cannot be seen by the camera. These paths likely correspond to the signal's passage through the walls or roof and reflection thereafter. Since these paths' relative amplitudes are very small, they can be ignored.

TABLE IV: Path for the corridor and office

| Path Type | Corridor | Office |
|---|---|---|
| LoS | 0.61m | 3 m |
| NLoS | 9.08m, 1.36m, 1.36m, 1.36m, 1.36m, 3.95m | 3.9m, 5.1m, 6.3m, 10.2m |

*2) Performance Comparison:* To evaluate the performance of our proposed approach, we compare three approaches under static availability settings:

– D-LS: This is our approach that deterministically selects frequencies among those available and uses least squares to estimate the CIR.

– R-LS: This approach only differs from ours as it randomly chooses the frequencies on the available band.

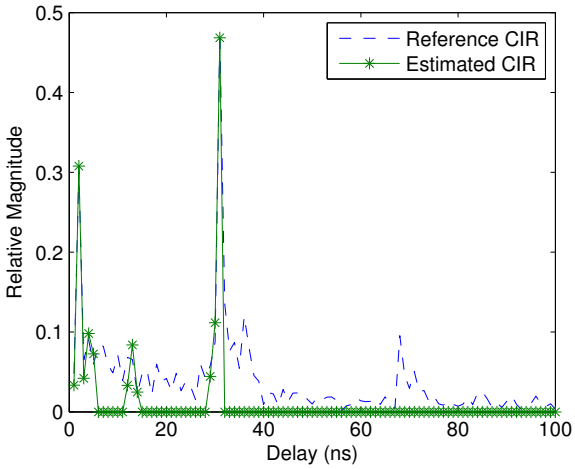
– CS: This is the compressed sensing method that randomly samples frequencies among available frequencies; it uses orthogonal matching pursuit (OMP) to reconstruct the CIR.

We use average error as a performance metric. For $N$ measurements, the average error is defined as:

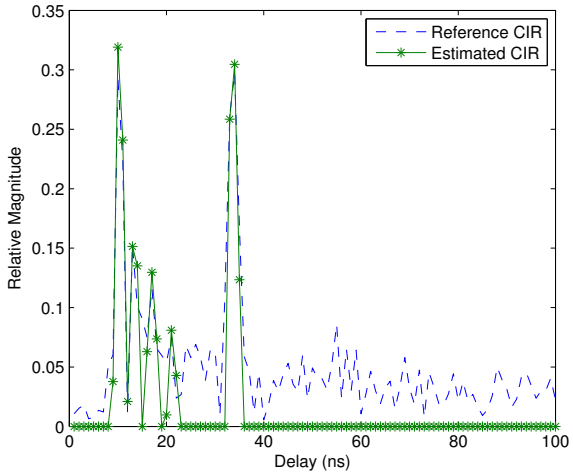
$$\frac{1}{N}\sum_{i=1}^{n}\frac{(\widehat{\mathbf{h}}_i-\mathbf{h}_{ref})^H(\widehat{\mathbf{h}}_i-\mathbf{h}_{ref})}{(\mathbf{h}_{ref})^H\mathbf{h}_{ref}}, \quad (1)$$

where $\widehat{\mathbf{h}}_i$ is the $i$-th sounding result and $^H$ denotes conjugate transpose. We use the corridor scenario as an example and run the experiment under three settings, each of which has a different frequency occupation rate. The frequency occupation rates are 0.7, 0.5, and 0.3, which means there are 300, 500, and 700 available channels, respectively. We also assume the underlying channels' availability is invariant during the sounding process. We use the three methods described above; for each method, we average its error over 100 runs. Fig. 6 shows the comparison results for different frequency occupation rates.

We can see that least squares based methods have lower average error than the CS method. By deterministically sampling among the available frequencies, we can achieve lower average error. When the available channels increase, R-LS and our D-LS are closer.




(a) CIR for the corridor



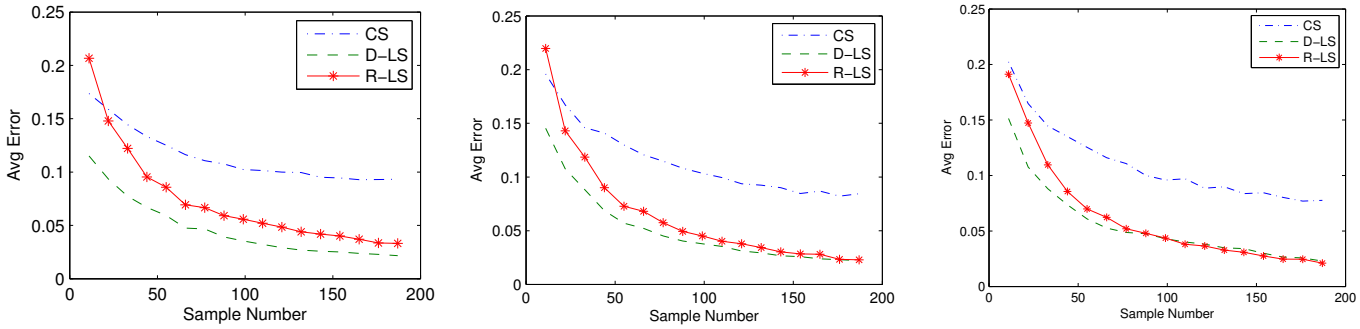

(b) CIR for the office

Fig. 5: CIR estimation

We also conduct experiments under dynamic availability settings. Practically, we model the occupation of each underlying channel as a Poisson process. For each channel $i$, we set its arrival rate $\lambda_i$ differently and vary channel availability during the sounding process. On average, there are 500 available channels over 100 runs. Fig. 7 shows the comparison of the three methods under this setting. We can see that least-squares based methods still outperform CS. Deterministic sampling (D-LS) can achieve better performance than R-LS when the number of samples is not large.

## B. Simulations

In our simulations, we further evaluate the robustness of our EV-Sounding system when impacted by electronic noise, visual noise, and sparsity. *Electronic noise* is noise associated with the electronic signal, including the channel noise and the device noise. *Visual noise* is the noise associated with the visual signal, which forms disturbances in the visually estimated path length.




(a) 70% occupation rate (b) 50% occupation rate (c) 30% occupation rate

Fig. 6: Performance for different occupation rates

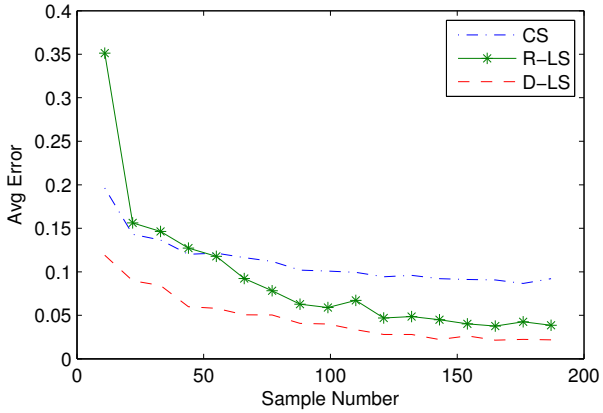



Fig. 7: Comparison of performance under dynamic availability

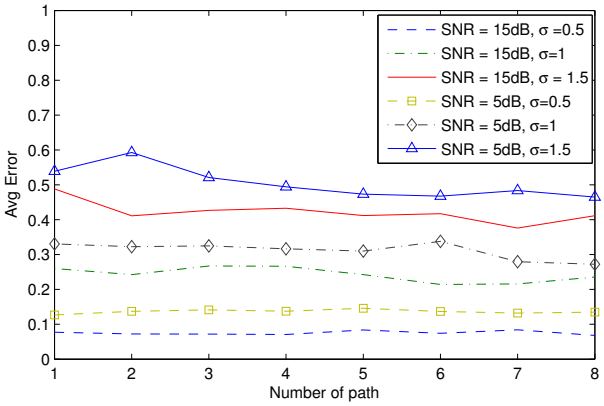



Fig. 9: Impact of multipath number

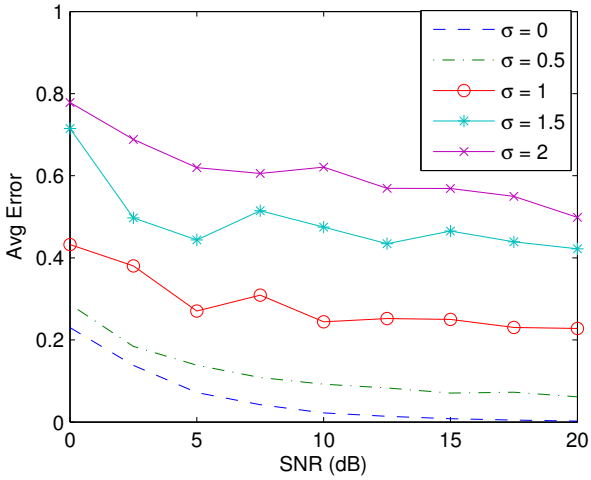



Fig. 8: Impact of electronic noise and visual noise

We create a four-path noisy wireless channel to evaluate impacts from noise in the static availability scenario. Electronic noise and visual noise are both simulated by additive white Gaussian noise (AWGN). The simulated network has 128 out of 256 channels available and 60 of the 128 channels are selected to estimate the CIR. We calculate the average error for different electronic Signal-to-Noise Ratio (SNR) under five kinds of visual noise whose standard deviation $\sigma$ ranges from 0 to 2. The error is averaged over 100 runs. The impact of both electronic noise and visual noise to our sounder is shown in Fig. 8. We can see that average error decreases as SNR increases. However, we can also see that the visual noise dominates the average error. The impact of the total number of paths is shown in Fig. 9. We simulate our EV-Sounding system where the number of paths increases from 1 to 8 in six different settings for the SNR and visual noise. We can see that for settings with the same visual noise, the average error is close. Since the variance of visual noise has three values, we can see three groups: the bottom, the middle, and the top.

We also conduct simulations for the dynamic channel availability scenario using the same settings as above. We find the results are very similar to Fig. 8.

## VI. Conclusions

In this paper, we presented EV-Sounding, a novel technique for timely, high-resolution channel sounding based on electronic and visual signals. Given the incomplete frequency responses of the channel, we aimed to accurately recover the channel impulse response (CIR). In order to cope with spectrum limitations, we designed intelligent sampling algorithms to select specific frequencies. By leveraging the visualized multipath information and our linear regression methodology, we suc-

cessfully integrated electronic and visual signals and accurately recovered the CIR. We implemented our EV-Sounding system on commercial off-the-shelf (COTS) devices and conducted real-world experiments and extensive simulations to evaluate the robustness of our EV-Sounding technique. The results showed our approach can achieve a high-resolution CIR and is efficient and robust.

It should be noted that our EV-Sounding scheme leverages computer vision technology. Although such technology has difficulty estimating multipaths accurately, especially in complicated environments such as offices full of moving people, integrating vision technology with electronic channel sounding enables us to achieve better channel sounding than state-of-the-art approaches. As computer vision technology evolves, our scheme stands to provide more accurate CIR estimations.

## ACKNOWLEDGMENT

We would like to thank Chi-Chih Chen, Jim Moncrief, Qiang Zhai, Xinfeng Li, Sihao Ding, and Siyang Cao for their help in the experiments and discussions.

## REFERENCES

[1] A. Kemp and S. Barton, "The Impact of Delay Spread on Irreducible Errors for Wideband Channels on Industrial Sites," *Wireless Personal Commun.*, vol. 34, no. 3, pp. 307–319, 2005.

[2] Q. Chen, A. M. Wyglinski, and G. J. Minden, "Frequency agile interference-aware channel sounding for dynamic spectrum access networks," in *Proc. IEEE GLOBECOM*, 2007.

[3] Q. Liu, Z. Zhao, and L. Cui, "A Wideband Compressed Spectrum Sensing Platform for Dynamic Spectrum Access Networks," in *Proc. ACM MobiCom*, 2012.

[4] Q. Zhao, L. Tong, and A. Swami, "Decentralized cognitive mac for dynamic spectrum access," in *Proc. IEEE DySPAN*, 2005.

[5] Q. Zhao and B. M. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Process. Magazine*, vol. 24, pp. 79–89, May 2007.

[6] D. Lee, K. Sowerby, and M. Neve, "Extracting Fine Multipath Detail from Measured Data at 5.8 GHz," in *Proc. IEEE Veh. Technol. Conf.*, 2004.

[7] X. Guo, D. Zhang, and N. L. M., "Localizing Multiple Objects in an RF-based Dynamic Environment," in *Proc. IEEE ICDCS*, 2012.

[8] N. Patwari, J. Croft, S. Jana, and S. K. Kasera, "High Rate Uncorrelated Bit Extraction for Shared Secret Key Generation from Channel Measurements," *IEEE Trans. on Mobile Comput.*, Jan. 2010.

[9] D. Maas, M. H. Firooz, J. Zhang, N. Patwari, and S. K. Kasera, "Channel Sounding for the Masses: Low Complexity GNU 802.11b Channel Impulse Response Estimation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 1–8, Jan. 2012.

[10] M. A. Tope and J. C. McEachen, "Unconditionally Secure Communications Over Fading Channels," in *Proc. IEEE MILCOM*, 2001.

[11] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener, "Robust Key Generation From Signal Envelopes in Wireless Networks," in *Proc. ACM CCS*, 2007.

[12] N. Patwari, J. Croft, S. Jana, and S. K. Kasera, "High Rate Uncorrelated Bit Extraction for Shared Secret Key Generation from Channel Measurements," *IEEE Trans. Mobile Comput.*, Jan. 2010.

[13] W. R. Young Jr. and L. Y. Lacy, "Echoes in Transmission at 450 Megacycles from Land-to-Car Radio Units," *Proc. IRE*, vol. 38, no. 3, pp. 255–258, 1950.

[14] T. S. Rappaport, "Characterization of UHF Multipath Radio Channels in Factory Buildings," *IEEE Trans. Antennas Propag.*, vol. 37, no. 8, pp. 1058–1069, 1989.

[15] R. J. Pirkl and G. D. Durgin, "Optimal Sliding Correlator Channel Sounder Design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3488–3497, Sep. 2008.

[16] D. C. Cox, "Delay Doppler Characteristics of Multipath Propagation at 910 MHz in a Suburban Mobile Radio Environment," *IEEE Trans. Antennas Propag.*, vol. 20, no. 5, pp. 625–635, 1972.

[17] S. J. Howard, K. Pahlavan, R. Co, and M. A. Marlboro, "Measurement and Analysis of the Indoor Radio Channel in the Frequency Domain," *IEEE Trans. Instrum. Meas.*, vol. 39, no. 5, pp. 751–755, 1990.

[18] S. S. Ghassemzadah, R. Jana, C. W. Rice, W. Turin, and V. Tarokh, "Measurement and Modeling of an Ultra-Wide Bandwidth Indoor Channel," *IEEE Trans. Commun.*, vol. 52, no. 10, pp. 1786–1796, 2004.

[19] W. U. Bajwa, J. Haupt, G. Raz, and R. Nowak, "Compressed Channel Sensing," in *Proc. CISS*, 2008.

[20] C. R. Berger, Z. Wang, J. Huang, and S. Zhou, "Application of Compressive Sensing to Sparse Channel Estimation," *IEEE Commun. Magazine*, pp. 164–174, Nov. 2010.

[21] Y. C. Eldar and G. Kutyniok, Eds., *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[22] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.

[23] W. K. Tam and V. N. Tran, "Propagation Modelling for Indoor Wireless Communication," *Electron. Commun. Eng. J.*, vol. 7, pp. 221–228, 1995.

[24] J. D. Parsons, D. A. Demery, and A. M. D. Turkmani, "Sounding techniques for wideband mobile radio channels: a review," *IEE Proc.*, vol. 138, no. 5, pp. 437–446, Oct. 1991.

[25] J. Kivinen, "60-GHz Wideband Radio Channel Sounder," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 5, pp. 1831–1838, Oct. 2007.

[26] J.-M. Conrat, P. Pajusco, and J.-Y. Thiriet, "A Multibands Wideband Propagation Channel Sounder from 2 to 60 GHz," in *Proc. IEEE IMTC*, 2006.

[27] D. N. Liu, S. Yerramalli, and U. Mitra, "On Efficient Channel Estimation for Underwater Acoustic OFDM Systems," in *Proc. ACM WUWNet*, 2009.

[28] D. Sugizaki, N. Iwakiri, , and T. Kobayashi, "Ultra-wideband Spatio-temporal Channel Sounding with Use of an OFDM Signal in an Indoor Environment," in *Proc. Progress In Electromag. Research Symp.*, 2011.

[29] C. R. Berger, S. Zhou, J. C. Preisig, and P. Willett, "Sparse Channel Estimation for Multicarrier Underwater Acoustic Communication: From Subspace Methods to Compressed Sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1708–1721, Mar. 2010.

[30] J. L. Paredes, G. R. Arce, and Z. Wang, "Ultra-Wideband Compressed Sensing: Channel Estimation," *IEEE J. Selected Topics in Signal Processing*, vol. 1, no. 3, pp. 383–395, Oct. 2007.

[31] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.

[32] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[33] M. Zuliani, C. S. Kenney, and B. S. Manjunath, "The Multiransac Algorithm and its Application to Detect Planar Homographies," in *Proc. IEEE Int'l. Conf. Image Processing (ICIP)*, Sep. 2005.

[34] W.-W. Ma, M.-H. Le, and K.-H. Jo, "3D Reconstruction and Measurement of Indoor Object Using Stereo Camera," in *Proc. Int'l. Forum of Strategic Technology*, 2011.

[35] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, "KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera," in *Proc. ACM UIST*, 2011.

[36] M. Pollefeys, D. Nistér, J.-M. Frahm, A. Akbarzadeh, P. Mordohai, B. Clipp, C. Engels, D. Gallup, S.-J. Kim, P. Merrell, C. Salmi, S. Sinha, B. Talton, L. Wang, Q. Yang, H. Stewénius, R. Yang, G. Welch, and H. Towles, "Detailed Real-Time Urban 3D Reconstruction from Video," *Int'l. J. of Computer Vision*, vol. 78, no. 2–3, pp. 143–167, 2008.

[37] E. Candes and M. Wakind., "An Introduction to Compressive Sampling," *IEEE Trans. Signal Process.*, Mar. 2008.

[38] OpenCV, http://opencv.org.

[39] G. Durgin, N. Patwari, and T. S. Rappaport, "An Advanced 3D Ray Launching Method for Wireless Propagation Prediction," in *Proc. IEEE Veh. Technol. Conf.*, 1997.

[40] C. Rakpenthai, S. Premrudeepreechacharn, S. Uatrongjit, and N. R. Watson, "An Improved PMUs Placement Method for Power System State Estimation," in *Proc. IPEC*, 2005.

[41] Ettus Research, http://www.ettus.com.