# Demo Abstract: Signal Reconstruction with SubNyquist Sampling using Wireless Sensor Networks

Andria Pazarloglou, Stephen George, Radu Stoleru, Ricardo Gutierrez-Osuna
Department of Computer Science and Engineering, Texas A&M University
{andria, mikegeorge, stoleru, rgutier}@cse.tamu.edu

## ABSTRACT

Data streaming for sensor networks is an emerging class of application that can offer significant safety and utility enhancements in a number of critical environments. In one such application, disaster response, a wireless sensor network may assist in the detection of living victims trapped in collapsed buildings. As a step towards this goal, this paper addresses the very high bandwidth and processing requirements posed by streaming applications, and demonstrates a sub-Nyquist sampling procedure for reconstructing higher-resolution speech signals.

## 1. INTRODUCTION

In the area of disaster response, wireless sensor networks (WSN) are often found in applications requiring simple, periodic sensing and reporting. However, sensing platforms can perform more sophisticated duties and acoustic sensors may help first responders search for victims in collapsed and damaged buildings. When a building is damaged in a disaster, it may become unstable or even collapse. First responders must rapidly locate victims and help them escape. A WSN can assist rescuers by extending acoustic monitoring into unsafe or nearly inaccessible locations. Simple but powerful sensors listen for signs of life and exfiltrate these sounds when detected. However, the use of WSNs present challenges. The most extreme challenge comes from the use of low-power embedded systems to collect and transmit streaming audio data. Other challenges include limited, congested bandwidth, the loss of network connectivity due to environmental factors, unreliable hardware, and limited energy resources.

Current approaches for streaming voice data focus on obtaining understandable speech from WSNs with energy consumption constraints. As examples, [1] and [2] demonstrate real-time voice streaming in mines and military surveillance operations, respectively. Both projects sample at above Nyquist rates and use audio codecs to reduce transmissions. At the other extreme, sophisticated signal processing hardware with tunable sensors is used by Greenstein et al. [3] to reduce the amount of data being transmitted.

## 2. HIGH RESOLUTION SIGNAL RECONSTRUCTION

To provide responders with the information they need, we strive for acceptable sound, rather than optimal reconstruction. An acceptable sound signal has some predefined qualitative characteristics, such as delivering audio accurate enough to be understood by a listener. Given that normal speech has a frequency range from 300-3500 Hz, with most signal energy in the 500-2500 Hz range, sampling rates in existing applications are in the 2-4 kHz [1] [2] range. Here, we propose a distributed sampling method that can be much lower depending on the number of the nodes that simultaneously hear the event. We base our reconstruction approach on the fact that streams from multiple lower quality sources provide different pieces of information about the original signal, provided their sampling times are randomly shifted [4].

Figure 1 illustrates the proposed sampling and reconstruction scheme. The original signal is denoted by $x(t)$ and the discrete-time signals $s_i(k), i = 0, \ldots, n-1$ are noisy versions of $x(t)$ produced in each node after sampling; $k$ denotes the time index at the sensor node with sampling rate $f_s$, and $k'$ denotes the time index for the higher resolution signals with sampling rate $f_{s'} = f_s \cdot r$. The $s_i(k)$ signals are then ADPCM encoded.

To reconstruct the original, high-resolution signal $x(t)$, the received streams are decoded and then time-aligned before being fused. The signals are upsampled through interpolation to compensate for differences between nodes.
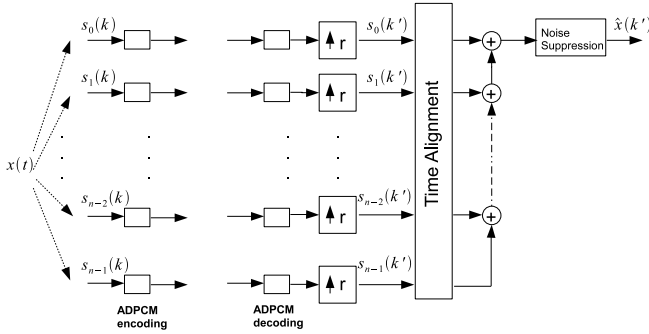
**Figure 1: Distributed Data Streaming and Signal Reconstruction**

Time delay estimation of the resulting signals $s_n(k')$ is required in order to align them properly in the time domain. To estimate the time delay between two waveforms $x_1$ and $x_2$, we use a generalized cross-correlation $(GCC)$ function $\hat{d}_{GCC}(m) = \sum_{k=0}^{N-1} W[k]G_{x_1x_2}[k]e^{j2\pi mk/N}$, where $G_{x_1x_2}[k]$ is the cross-spectrum of $x_1$ and $x_2$, $X_i[k]$ is the discrete Fourier transform (DFT) of $x_i(n)$, $N$ is the length of the DFT and $W[k]$ denotes the weighting function or prefilter. In our implementation we use a constant weighting function $W[k]$. The aligned signals are superimposed and then output after passing through a denoising filter. We developed two methods: i) interpolating the low resolution signals, followed by an average between two samples: $x(k') = \frac{1}{N}\sum_n s_n(k' - d_n)$, where $d_n$ is the delay for the signal $s_n$; and ii) upsampling (e.g., filling the spaces between samples with zeros) as:

$$s_n(k') = \begin{cases} s_n(k), & \text{if } k' = k \cdot r, k = 0, \pm1, \pm2, .. \\ 0, & \text{otherwise} \end{cases}$$

followed by $x(k') = \sum_n s_n(k' - d_n)$ and replacing zero samples with linear interpolations.

## 3. DEMONSTRATION

We demonstrate the reconstruction of a high-resolution speech signal from multiple sub-Nyquist audio streams using a network of six MicaZ motes placed randomly around an acoustic source, as depicted in Figure 2. Each node communicates to a base station on a dedicated frequency.

An audio signal of a male voice is generated and the motes begin sampling at 1 kHz. The audio data is encoded using PCM, compressed, and then transmitted wirelessly to each mote's base station. The base station passes the packets via a serial link to a computer for decoding and MATLAB processing. The six resulting audio streams are then combined into a single audio
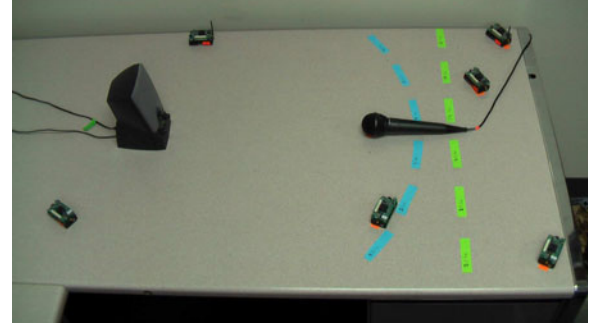


**Figure 2: Demonstration setup**

file. Figure 3 shows a low resolution signal (i.e., acquired by a single node), the reconstructed signal from all six nodes, the signal acquired by a professional microphone and the original reference signal, a good quality recording made in the mote field. We also present the quality (e.g., NRMSE and PSNR) of the reconstructed signals.
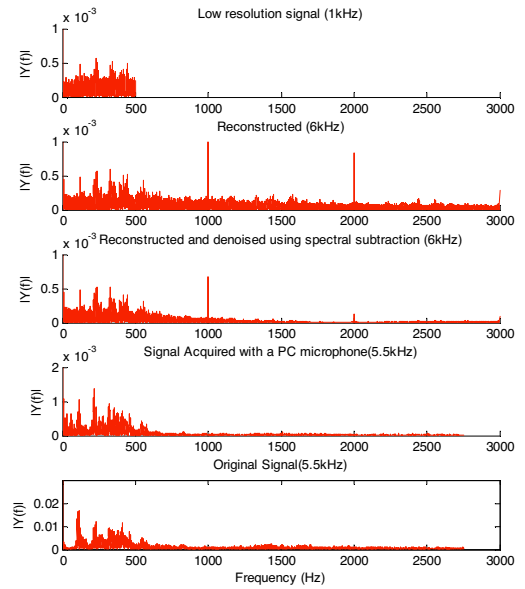


**Figure 3: The acquired (from a single node), reconstructed and original signals.**

## 4. REFERENCES

[1] R. Mangharam, A. Rowe, R. Rajkumar, and R. Suzuki, "Voice over sensor networks," in *RTSS '06*, 2006.

[2] J. Zhang, G. Zhou, S. H. Son, and J. A. Stankovic, "Ears on the ground: An acoustic streaming service in wireless sensor networks," in *IPSN '06*, 2006.

[3] B. Greenstein, C. Mar, A. Pesterev, S. Farshchi, E. Kohler, J. Judy, and D. Estrin, "Capturing high-frequency phenomena using a bandwidth-limited sensor network," in *SenSys '06*, 2006.

[4] A. Pazarloglou, R. Stoleru, and R. Gutierrez-Osuna, "High-resolution speech signal reconstruction in wireless sensor networks," in *IRSN*, 2009.