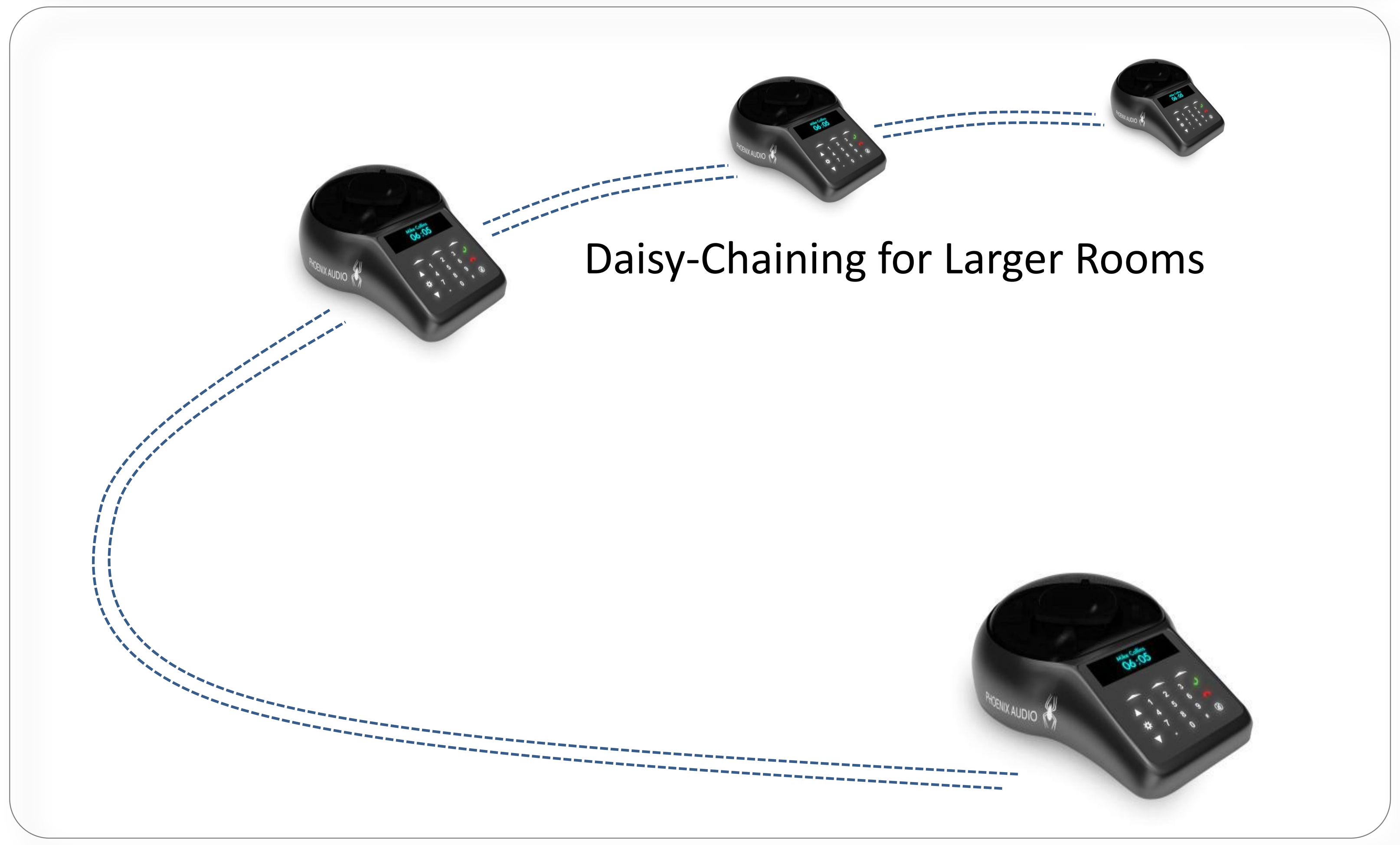
REAL-TIME MICROPHONE SELECTION IN NOISY REVERBERANT ENVIRONMENTS FOR TELECONFERENCING SYSTEMS

Israel Cohen[1], Baruch Berdugo[2] and Joseph Marash[3]

[1]*Department of Electrical Engineering, Technion – Israel Institute of Technology, Israel*
[2]*MRD Technologies Ltd., Kibbutz Usha 30031, Israel*
[3]*Phoenix Audio Technologies, 2934 N. Naomi Street, Burbank, CA 91504, USA*

Abstract

- **Teleconferencing in large rooms:** Use more than one microphone for audio pickup**.**
- **A major challenge:** Monitor the perceived quality of each microphone signal and select, at any given point in time, the microphone with the best reception.
- We present a **real-time system** for dynamically selecting the best microphone amongst randomly placed microphones in a conference room.
- We demonstrate state-of-the-art results by an **economical, practical and innovative algorithm**.



Introduction

- Microphones that are used in industrial applications are generally **not calibrated**.
- The **sensitivities** of different microphones may be quite different.
- Therefore, the **power** is not reliable for a comparison between signals measured with different microphones.
- The **signal-to-noise ratio** is also not a reliable measure to quantify the level of reverberation, since in real applications, the noise cannot be assumed uniform, nor the late reverberation is uniform.

Problem formulation

- A source signal measured at point $p_i = (x_i, y_i, z_i)$ $(i = 1,2,\dots,N)$ is given by

$$r_i(t) = s(t) * h_i(t) + n_i(t)$$

- **Perception of the amount of reverberation** in a given signal is closely related to the **direct-to-reverberation ratio.**
- For evaluating the direct-to-reverberation ratio, the impulse response $h_i(t)$ is split into **early and late parts**:

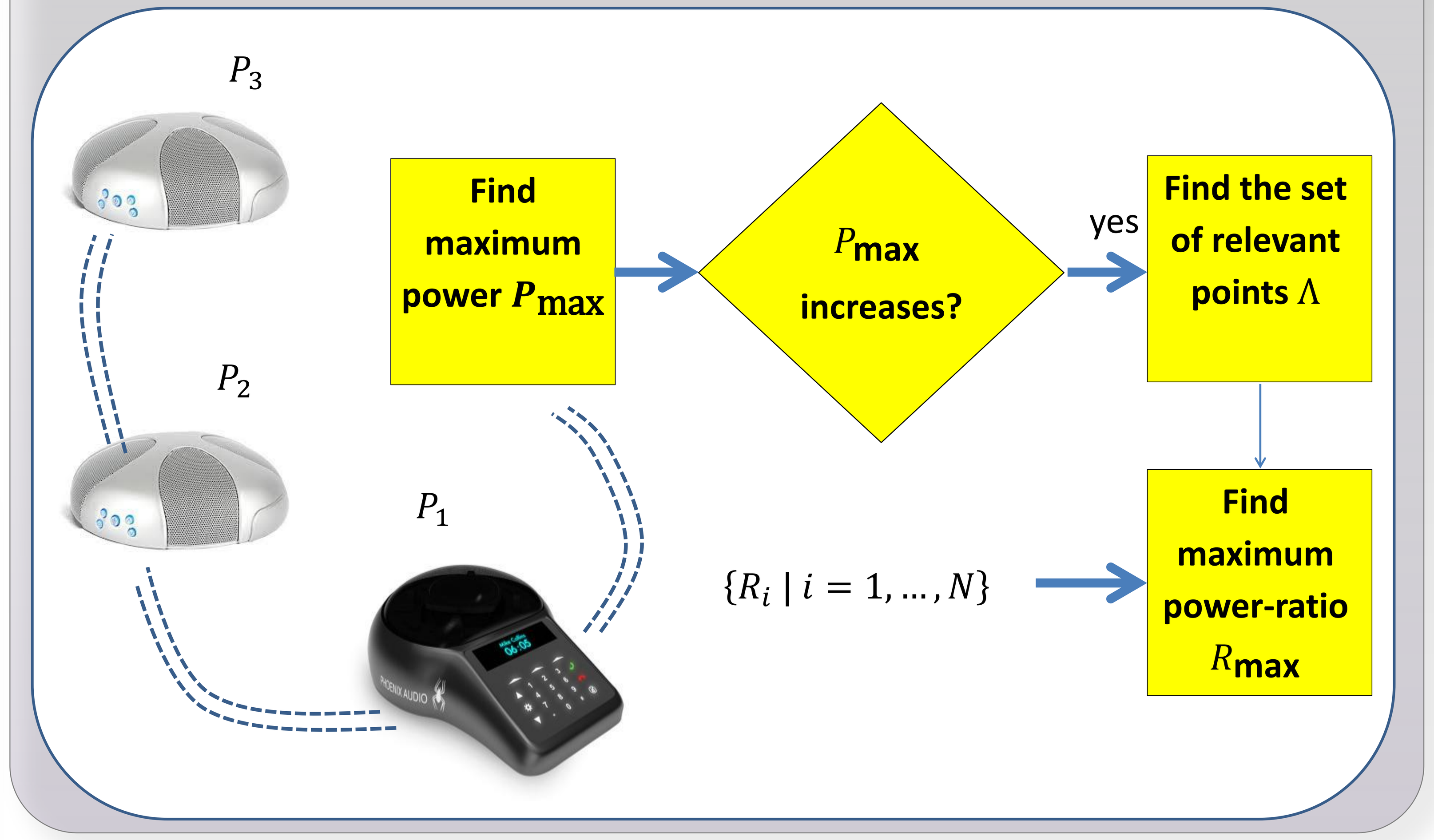
$$h_i(t) = h_i^e(t) + h_i^l(t)$$

- The direct-to-reverberation ratio is defined as the ratio between the power of the "**direct signal**" $s(t) * h_i^e(t)$ (including the early reflections) and the power of the "**reverberations**" $s(t) * h_i^l(t)$ (containing only the late reflections).
- Our objective is to determine which signal out of the given set of measured signals $\{r_i(t) \mid i = 1,2,\dots,N\}$ has the **greatest direct-to-reverberation ratio**.

Implementation

- Our system is based on **clusters of uni-directional** microphones, each looking at a different direction (for demonstration, we use four uni-directional microphones looking at direction 90 degrees apart).
- We compare the signal received by each of the microphones in a cluster (referred to as local) and compare it with the other **local microphones**.
- We derive a **reverberation quality figure** based on the assumption that direct signals are received with different levels by the local microphones, while indirect signals (reverberations) are received with a much closer level on all the local microphones.
- We then compare the reverberation quality figure between all the clusters and **select the audio source with the least amount of reverberation**.

- The proposed procedure contains two stages.
  1. The first stage is **local**: for each point we compute some features of the local signals.
  2. The second stage is **global**: we select the least reverberant signal based on the features of the local signals.
- The features include **local power** and **local power-ratio**.
- The **local power** is associated with the directional microphone that measures the strongest signal at a given point, compared to the signals that are measured by the other microphones at that point.
- The **local power-ratio** is defined as the ratio between the local maximum power and the local minimum power.



Future development

- Integration of the proposed system with **source direction indicators**, to demonstrate a combination of real-time source localization and channel selection.