Microphone Array Support in Windows

April 21, 2014 Revision

Abstract
Under less than ideal conditions, even the best microphone embedded in a laptop or monitor does a poor job of capturing sound. An array of microphones can do a better job of isolating a sound source and rejecting ambient noise and reverberation. This paper provides information about the advantages that microphone arrays can offer, and about the support for microphone arrays that was introduced with the Microsoft® Windows Vista™ operating system.

If you are a laptop or computer monitor manufacturer, or a designer working to provide better quality captured-sound by integrating microphone arrays, or if you are a hardware manufacturer designing Windows-based external USB Audio microphone arrays, then this paper provides the design guidelines for building microphone arrays that will work well with Windows.

This information applies to the following operating systems:

Windows Vista and later

References and resources discussed here are listed at the end of this paper.

The current draft of this paper is available on the WHDC web site at:
http://www.microsoft.com/whdc/device/audio/default.mspx

Contents

Introduction

Under less than ideal conditions, even the best microphone embedded in a laptop or monitor does a poor job of capturing sound. An array of microphones can do a better job of isolating a sound source and rejecting ambient noise and reverberation.

Because of the advantages that microphone arrays can offer to improve sound-capture for PC computing, Microsoft has created support for microphone arrays in the Windows operating system. The support includes:

- A class driver to support USB Audio devices.
- Algorithms to support several tested microphone array geometries.
- The ability to identify microphone array geometries based on descriptors as reported by a USB device.

This paper describes the research and implementation details that provide the foundation for the Windows support for microphone arrays. It also provides specific design and implementation guidelines for good quality, cost-effective microphone array designs that will work well with Windows.

Microphone Arrays as PC Product Solutions: An Overview

PCs and other computing devices can usually play sounds well, but they do a poor job of capturing sound. With the processing power, storage capacities, broadband connections, and speech-recognition engines available today, computing devices can use better sound capture to deliver more value to customers.

With current PC-based audio technology, it is possible to provide better live communication than phones, much better record/playback or note-taking devices than tape recorders, and better command of the user interface than remote controls. For all applications that use sound, end users could benefit from better sound capturing. Consider, for example, all of these real-time communication applications:

- Microsoft Windows Messenger, MSN® Messenger, and all other applications built on top of the Microsoft Real-Time-Communication stack, such as AOL Instant Messenger, other applications for VoIP, and enhanced telephony.
- Enterprise solutions for collaboration and groupware applications, such as Live Meeting, the meeting recording capabilities in Microsoft OneNote®, and voice-messaging applications.

Robust speech-recognition technologies are still under development, but many Windows-based applications already have voice commands integration that work satisfactorily, but only when the user wears a headset with a close-talk microphone that enables decent sound-capturing quality. Such technologies are convenient for tablet PCs and handheld devices, where otherwise users have to type with a stylus.

Windows and Microphone Array Solutions

Most PCs or laptops still have just a single microphone. This is a poor solution for capturing speech, because the microphone picks up too much ambient noise and adds too much electronic noise. The captured signal also includes the room reverberation, which decreases intelligibility and confuses speech recognition algorithms. Signal processing techniques have their own limitations for removing stationary noise and reverberation from a single channel. As a result, users are typically forced into using “tethered” or wireless close-proximity microphone headsets to achieve decent sound-capturing quality.

Numerous studies show that users don't like to wear headsets or to be tethered to the computer. In many scenarios, headsets are not an option. For example, walking with a headset and a Tablet PC in your hand feels awkward. Using an array of microphones with PCs and other computing devices can alleviate the problems caused by using only one microphone. The goal—"Wear no close-proximity microphone gear; just talk to your computer"—implies mobility and freedom of movement.

Microphone array solutions should follow these basic design guidelines:

- Implement the characteristics of the tested microphone array geometries that are supported in Windows as summarized in Table 1 in this paper.
- Meet the design requirements for low noise, directionality, and low manufacturing tolerances for microphones and preamplifiers, as summarized in Table 2 in this paper.
- Follow the recommendations for analog-to-digital converters for sampling rate, sampling synchronization, and anti-aliasing filters as summarized in Table 3 in this paper.
- Choose and place the appropriate number of microphones based on the usage scenario, with recommended choices illustrated in Figure 11 of this paper.
- Observe the acoustical and construction considerations, to insulate from environmental factors that will affect performance, as summarized near the end of this paper.

Windows includes microphone array support as part of a complete audio subsystem that provides these advances:

- Improved acoustic echo cancellation
- Microphone array support
- Stationary noise suppressor
- Automatic gain control
- Wideband quality of sound capturing and processing

Microphone array processing is linear and doesn't introduce distortions to the signal, so the microphone array output is good for a human listener and friendly for the speech recognition engine. The Windows audio stack can be used both for real-time communication applications such as Windows Messenger and for speech recognition-enabled applications such as voice commands and dictation.

The rest of this paper explores the technical details supporting the design and implementation of good quality microphone arrays that will work well with Windows.

About Microphone Arrays

A microphone array is a set of closely positioned microphones. Microphone arrays achieve better directionality than a single microphone by taking advantage of the fact that an incoming acoustic wave arrives at each of the microphones at a slightly different time. The chief concepts that are used in the design of microphone arrays are beam forming, array directivity, and beam width.

Beam Forming

By combining the signals from all microphones, the microphone array can act like a highly directional microphone, forming what is called a "beam." This microphone array beam can be electronically managed to point to the originator of the sound, which is referred to in this paper as the "speaker."

In real time, the microphone array engine searches for the speaker position and acts as if it points a beam at the current speaker. The higher directivity of the microphone array reduces the amount of captured ambient noises and reverberated waves. More details about the algorithm for beamform design can be found in reference [1]. (Numbered references are listed at the end of this paper.)

Array Directivity

Because the microphone array output contains less noise and has less reverberation than a single microphone, the stationary noise suppressor does a better job than it would with a signal from a single microphone. A typical directivity pattern of a microphone array beam for 1,000 Hz is shown on Figure 1. The pattern is more directive than even that of an expensive, high-quality hypercardioid microphone.

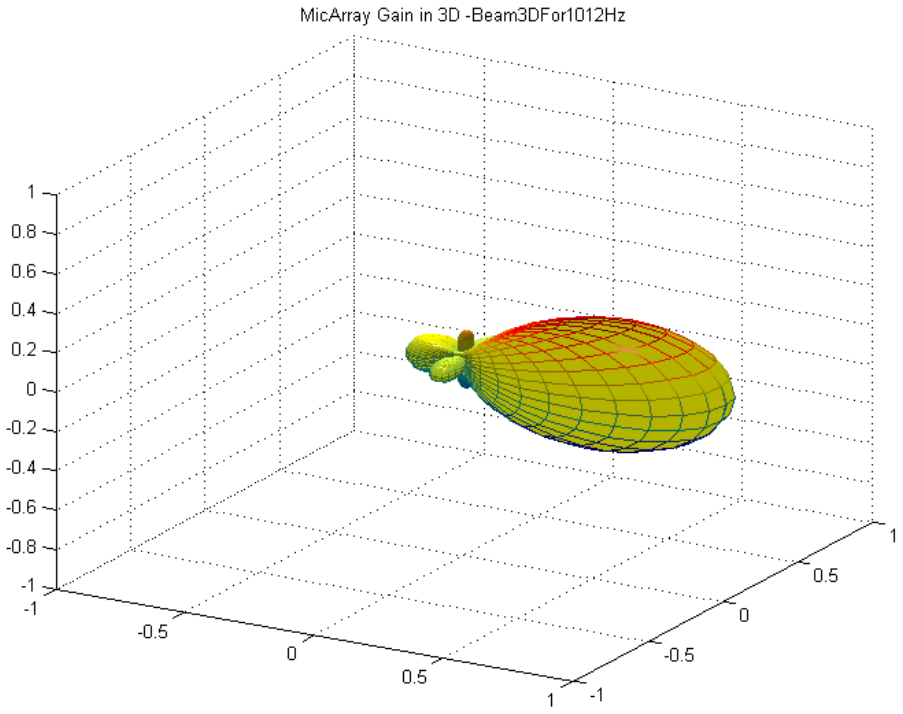

Figure 1. Microphone array directivity pattern in three dimensions

During sound capturing, the microphone array software searches for the speaker's position and aims the capturing beam in that direction. If the person speaking moves, the beam will follow the sound source. The "mechanical" equivalent is to have two highly directional microphones: one constantly scans the work space and measures the sound level, and the other—the capturing microphone—points to the direction with highest sound level; that is, to the speaker.

Constant Beam Width

The normal work band for speech capturing is from 200 Hz to 7,000 Hz, so wavelengths can vary by a factor of 35. This makes it difficult to provide a constant width of the microphone array beam across the entire work band. The problem is somewhat simpler in a typical office environment, where most of the noise energy is in the lower part of the

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.