

Microsoft
WinHEC
2005

Microphone Array Support in Windows Longhorn

Ivan Tashev
Software Architect
Microsoft Research
ivantash @ microsoft.com
Microsoft Corporation

Hakon Strande
Program Manager
WMDG
hakons @ microsoft.com
Microsoft Corporation

Session Outline

- Sound capturing in Windows
 - Why we need good sound capture?
 - Problems and solutions
- Microphone array
 - How it works?
 - MicArray support in Windows codenamed “Longhorn”
 - Design considerations
 - For laptops, tablets, monitors, and standalone devices
 - Supported geometries
- Call to action

Good Sound Capture and Windows

- Windows audio system
 - Does a good job in sound playback:
 - Hi-fi quality for stereo, 5.1, etc.
 - Multiple output devices supported
 - Sound capture:
 - Still on the level of one microphone
- Why we need good sound capturing:
 - Real time communication (RTC) for home and office:
 - Windows and MSN Messengers
 - Solutions built on top of the Microsoft RTC stack
 - Enterprise level Collaboration and Groupware
 - Live Meeting, OneNote
 - Automatic Speech Recognition (ASR)
 - Voice commands – attractive for tablets and handheld devices
 - Dictation – emerging scenario

Problems and Solution

- A single microphone:
 - Picks up too much ambient noise and reverberation
 - Adds electronic noise
 - Provides bad sound quality
- The noise and reverberation:
 - Decrease the intelligibility during RTC session
 - Confuse the automatic speech recognition engine
- As a result:
 - Users are forced to use headsets
 - Nobody likes to wear them, proven in multiple user studies
 - Inconvenient: users are tethered to the computer
 - You have to put it on and off
 - Concerns about appearance during video sessions

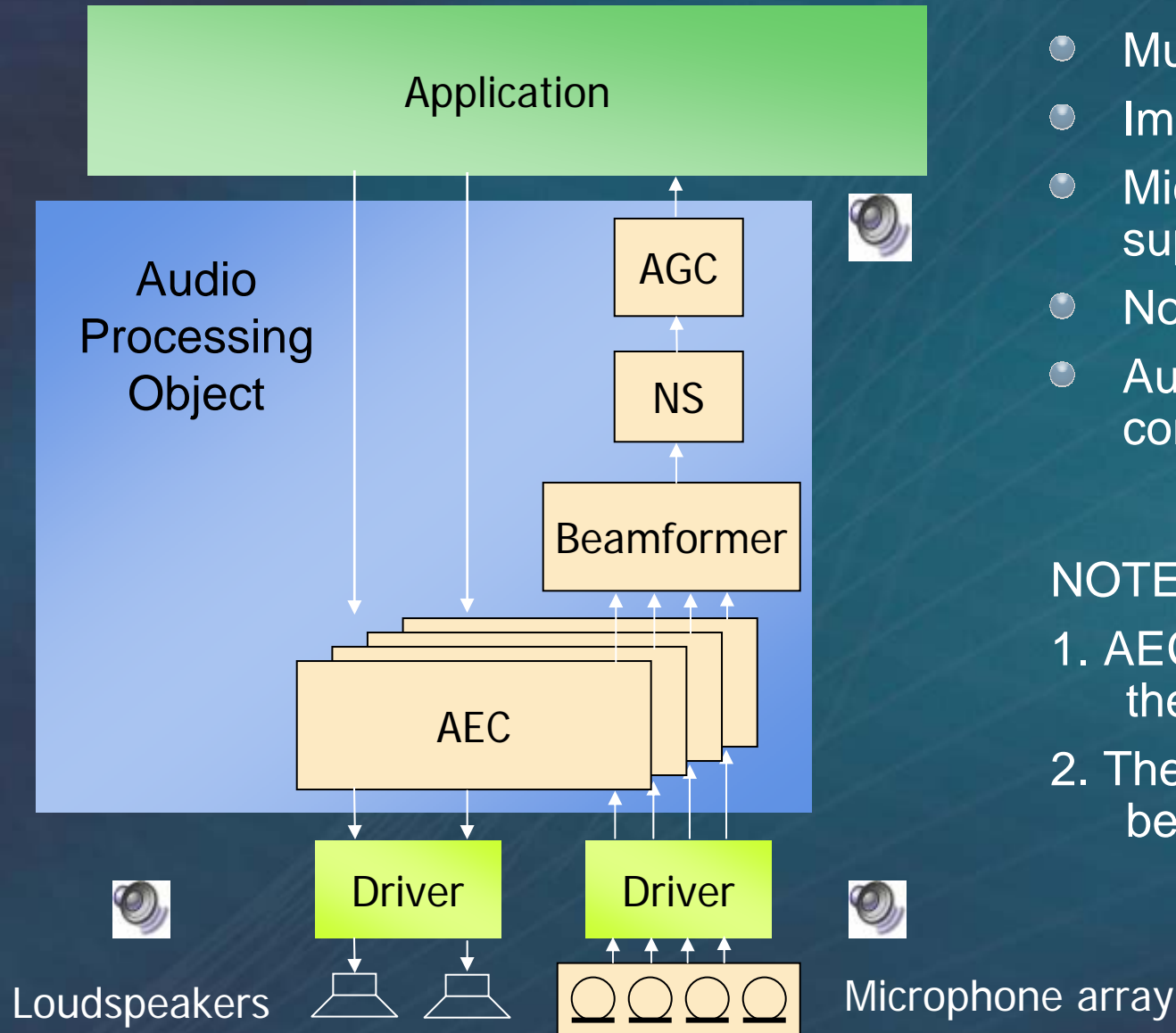
Problems and Solution (con't)

- The solution is to use multiple microphones
 - Signal processing algorithm combines the signals
 - The microphone array acts an acoustical antenna:
 - Better than a physical directional microphone
 - Electronically steerable
 - Reduces the captured ambient noise and reverberation

Just Talk to Your Computer, We Do the Rest

- “Just talk to your computer” concept:
 - Users do not wear headsets
 - Microphone array captures the sound:
 - External or integrated into the monitor for desktops
 - Integrated into the tablet/laptop
 - The operating systems does the signal processing providing good sound quality for all applications
- Integrated solution
 - Windows Longhorn provides complete audio stack:
 - Acoustic Echo Cancellation (AEC)
 - Microphone Array Support
 - Noise Suppression
 - Automatic Gain Control
 - Wideband quality of the captured sound
 - Can be used by RTC, ASR, any other voice enabled application

Sound Processing Architecture



- Multichannel Capture
- Improved AEC
- Microphone array support
- Noise suppression
- Automatic gain control

NOTE:

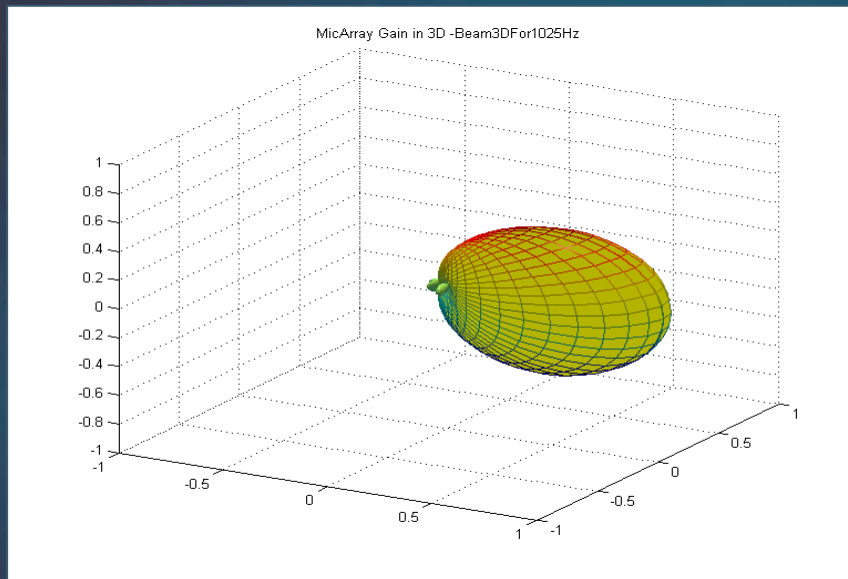
1. AEC happens before the beamformer
2. The beamformer can't be in the driver

Microphone Arrays Terminology

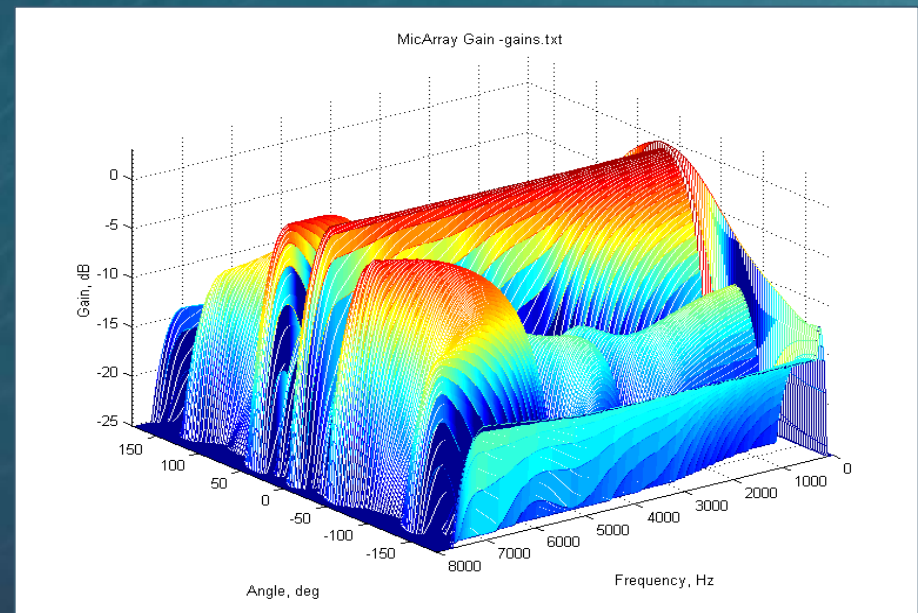
- Microphone array:
 - Set of closely positioned microphones
- Beamforming
 - Technology to make the microphone array to listen to given direction
 - Beamsteering is the ability to change the target direction
- How does it work?
 - Listen to various directions to determine where is the sound source using a scanning beam
 - Point the main capturing beam towards this direction
- Microphone array parameters
 - Directivity index: The portion of the captured white noise re: omnidirectional mike in the center of the microphone array
 - Ambient noise gain: The directivity index weighted with the typical ambient noise spectrum
 - A-weighted Ambient noise gain

Microphone Array Directivity

Beamshape at 1000 Hz for eight element array



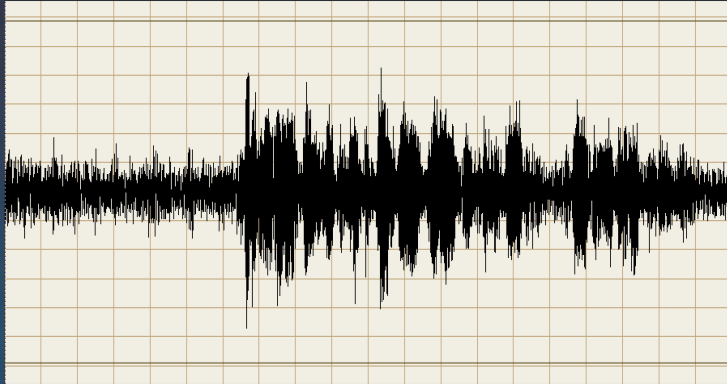
Directivity vs. frequency for four element array



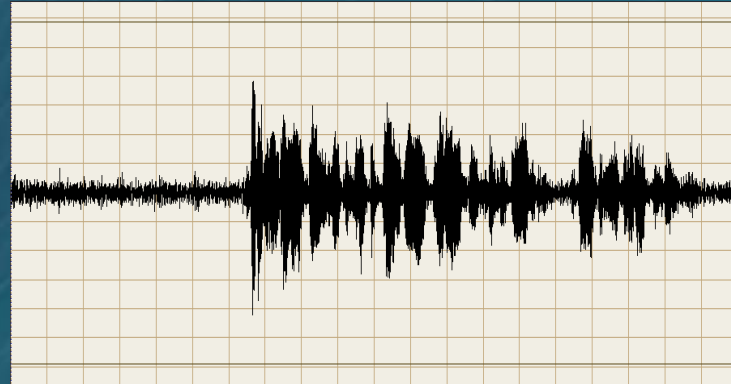
Microphone Array: Example

Person speaking at 3 feet from microphones

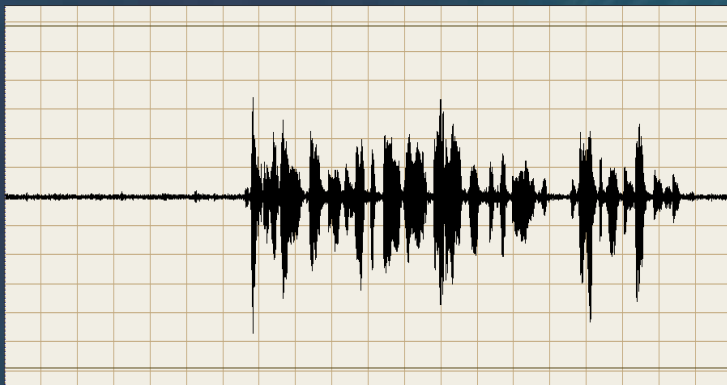
Typical \$10 PC microphone SNR=10.3 dB



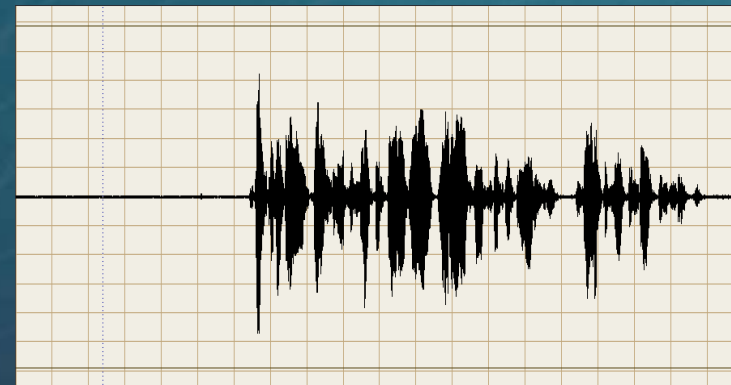
PC mic + WinXP noise reduction SNR=18.4 dB



HW DSP, SNR=34.4dB



Longhorn algorithm with USB array SNR=42.5dB



Supported MicArray Geometries

- The MicArray geometry is:
 - Number, type and position of the microphones
 - Critical for good results
 - Provided by MicArrDescriptor data structure
 - Support set of recommended geometries planned for Longhorn

Microphone array	Elements	Type	NG, dB	NGA, db	DI, dB
Linear, small	2	cardioid	-12.7	-6.0	7.4
Linear, big	2	cardioid	-12.9	-6.7	7.1
Linear, 4el	4	cardioid	-13.1	-7.6	10.1
L-shaped	4	cardioid	-12.9	-7.0	10.2
Circular, small	8	cardioid	-13.7	-8.0	7.8
Circular, big	8	cardioid	-14.2	-8.2	7.2
Circular, big, omni	8	omni	-12.6	-8.2	8.8

MicArrays for laptops, tablets, desktops

- Two element linear arrays
 - Large (200 mm) and small (100 mm)
 - Good for distances up to two feet, quiet office
- Four element arrays
 - Linear (190 mm) or L-shaped (for tablets)
 - For office/cubicle, normal noise levels – up to 6 feet
 - Under high noise levels – up to two feet
- Best place – the upper bezel



MicArrays for conference rooms

- Supported three circular 8 element arrays
 - Diameter 100 mm, cardioid microphones pointing outward
 - Diameter 170 mm, cardioid microphones pointing outward
 - Diameter 170 mm, omnidirectional microphones pointing up
- Designed for capturing meetings
 - Best place: the center of the conference room table
 - Work 360 degrees, up to 8 feet distance



Microphone Array Hardware

- Microphone Array:
 - Just a multi-channel microphone
 - No signal processing in the device
 - Allows building inexpensive devices
- Interface with the computer:
 - Digital USB interface
 - Suitable mostly for external devices
 - Provides guaranteed quality and guaranteed quality
 - The USB device provides the microphone array descriptor
 - Analog multi-channel audio input by HD Audio compatible audio chipsets
 - Suitable for integrated microphone arrays
 - Less expensive solution
 - The microphone array descriptor is in the ACPI table

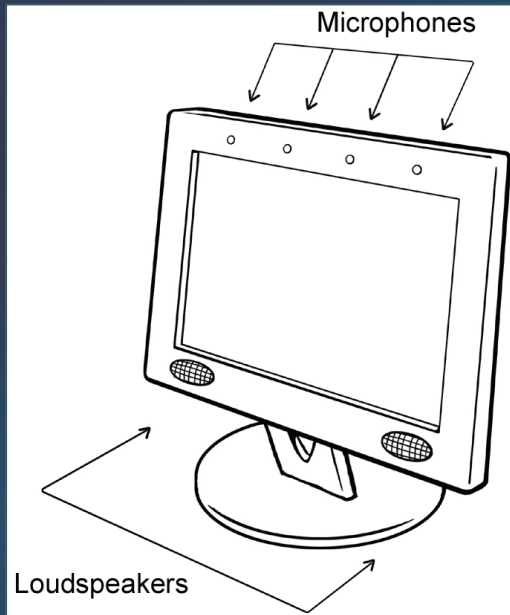
Microphones, Preamplifiers, ADCs

- Microphones:
 - Unidirectional strongly preferred
 - Low noise (better than 60 dB SNR)
 - Low manufacturing tolerances (± 4 dB in sensitivity)
- Preamplifiers:
 - Work band 200 – 7,000 Hz
 - High pass filter at 150 Hz/-3 dB, 18 dB/oct
- Analog to digital converters
 - Integrated anti-aliasing filters
 - Recommended sampling rate – 16,000 Hz
 - Synchronized sampling times – better than 1 μ s

Microphone Arrays Design: Tips and Tricks

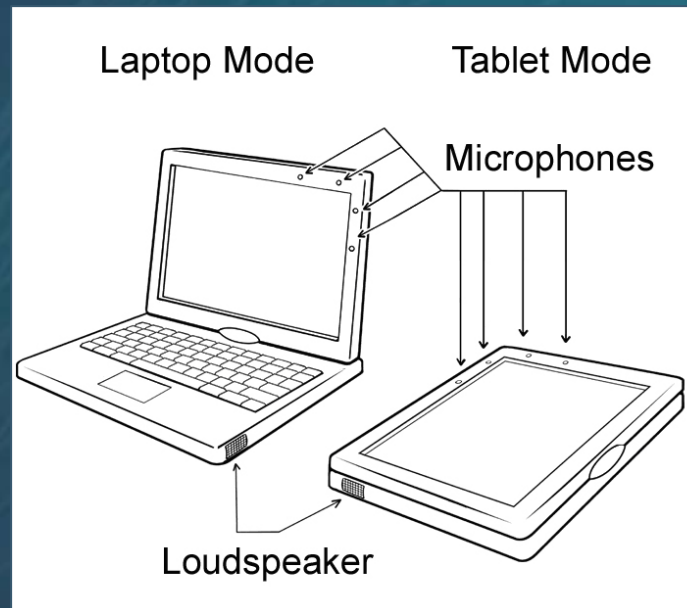
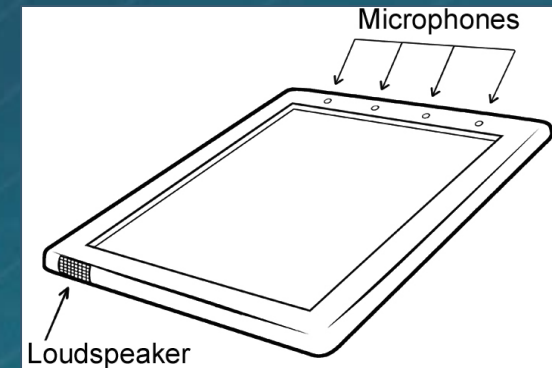
- Best place, general rules:
 - As far as possible from the loudspeaker(s)
 - Away from the keyboard
 - As close as possible to the speaker's mouth
- For laptops/convertibles
 - Consider both laptop and tablet mode
 - Right and left hand users should not cover the microphones
- In the office/cubicle
 - Best place is on the top of the monitor
- In the conference room:
 - Place it in the middle of the conference room table
- Acoustical design and construction
 - Insulate from rattles, vibrations, internal noises
 - Directional microphones need sound access from the back – vents
 - Protect from dust and humidity
 - Good quality loudspeaker goes well with the MicArray for RTC
 - AEC can't remove the nonlinear distortions

Microphone Arrays: Design Examples



Monitor, four elements array on the upper bezel, speakers in the lower part

Tablet: four element array on top, speaker on the other side



Laptop/tablet convertible: L-shaped array, works in both cases, away from the user's hands and the loudspeaker

Call To Action

- For system manufacturers:
 - Integrate microphone arrays in your laptops, tablets
 - Consider value-add, up-sell opportunity that external microphone arrays create for your PC products
 - Use Microsoft UAA compliant audio technologies
- For firmware engineers:
 - Ensure compatibility of the MicArray firmware with UAA USB requirements
- For device manufacturers:
 - Consider the business opportunities in manufacturing of external UAA compatible USB microphone arrays

Call To Action (con't)

- For driver developers:
 - Ensure that your driver supports the property set defined to pass the MicArray characteristics to OS
 - Enable multi-channel capture
 - Use WaveRT miniport for glitch resilience
- For application developers:
 - Take advantage of the high quality sound capture
 - Use Microsoft Audio Stack to benefit from integrated sound processing
 - For real-time communication applications, use Microsoft RTC APIs for better sound quality, establishing the connection, and encoding and decoding the audio

Community Resources

- **Windows Hardware & Driver Central (WHDC)**
 - www.microsoft.com/whdc/default.mspx
- **Technical Communities**
 - www.microsoft.com/communities/products/default.mspx
- **Non-Microsoft Community Sites**
 - www.microsoft.com/communities/related/default.mspx
- **Microsoft Public Newsgroups**
 - www.microsoft.com/communities/newsgroups
- **Technical Chats and Webcasts**
 - www.microsoft.com/communities/chats/default.mspx
 - www.microsoft.com/webcasts
- **Microsoft Blogs**
 - www.microsoft.com/communities/blogs

Additional Resources

- Email: micarrex @ microsoft.com
- Email: uaa @ microsoft.com
- Web Resources:
 - Whitepaper:
<http://www.microsoft.com/whdc/device/audio/default.msp>
- Related Sessions
 - Natural Input on Mobile PC Systems
 - Windows Audio/Video Excellence Requirements

questions

Microsoft[®]

Your potential. Our passion.[™]

© 2005 Microsoft Corporation. All rights reserved.

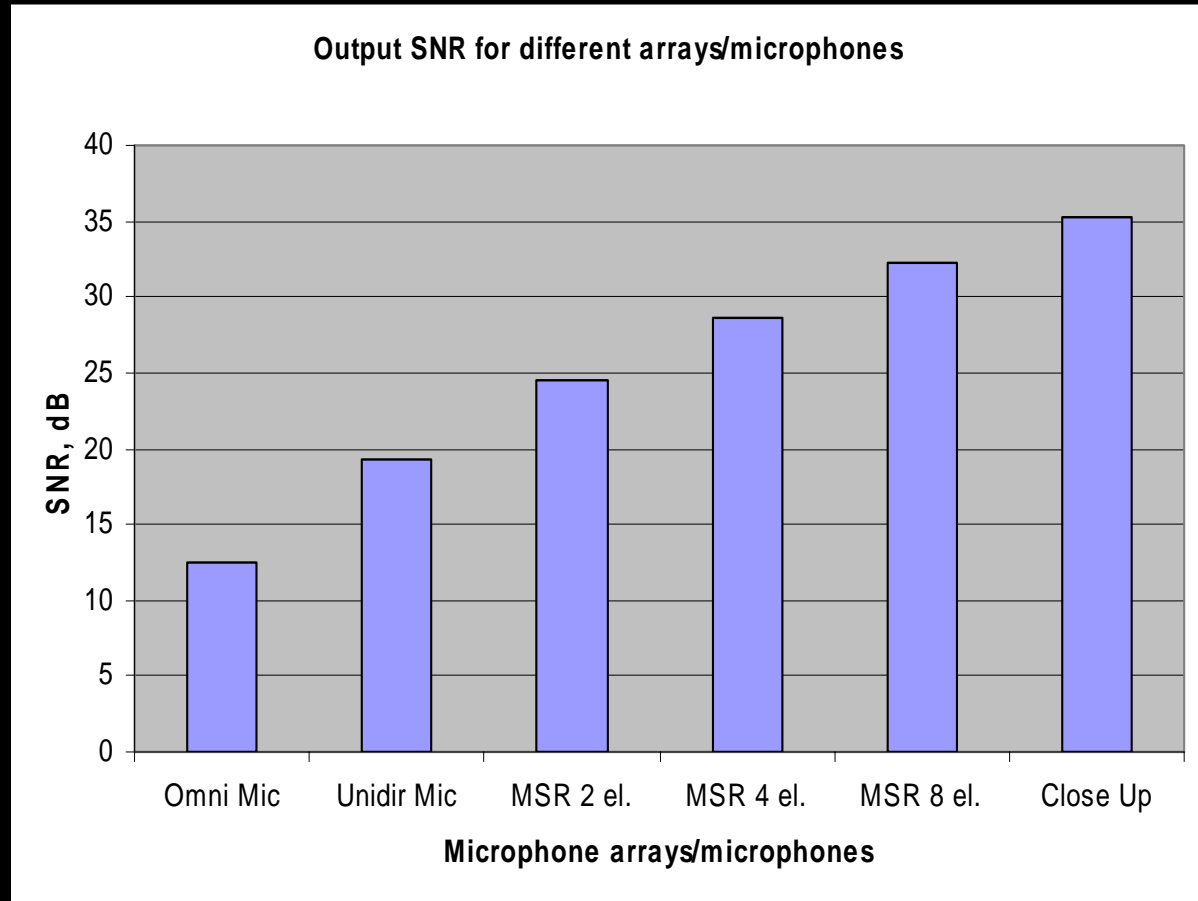
This presentation is for informational purposes only. Microsoft makes no warranties, express or implied, in this summary.

Backup slides

How many microphones?

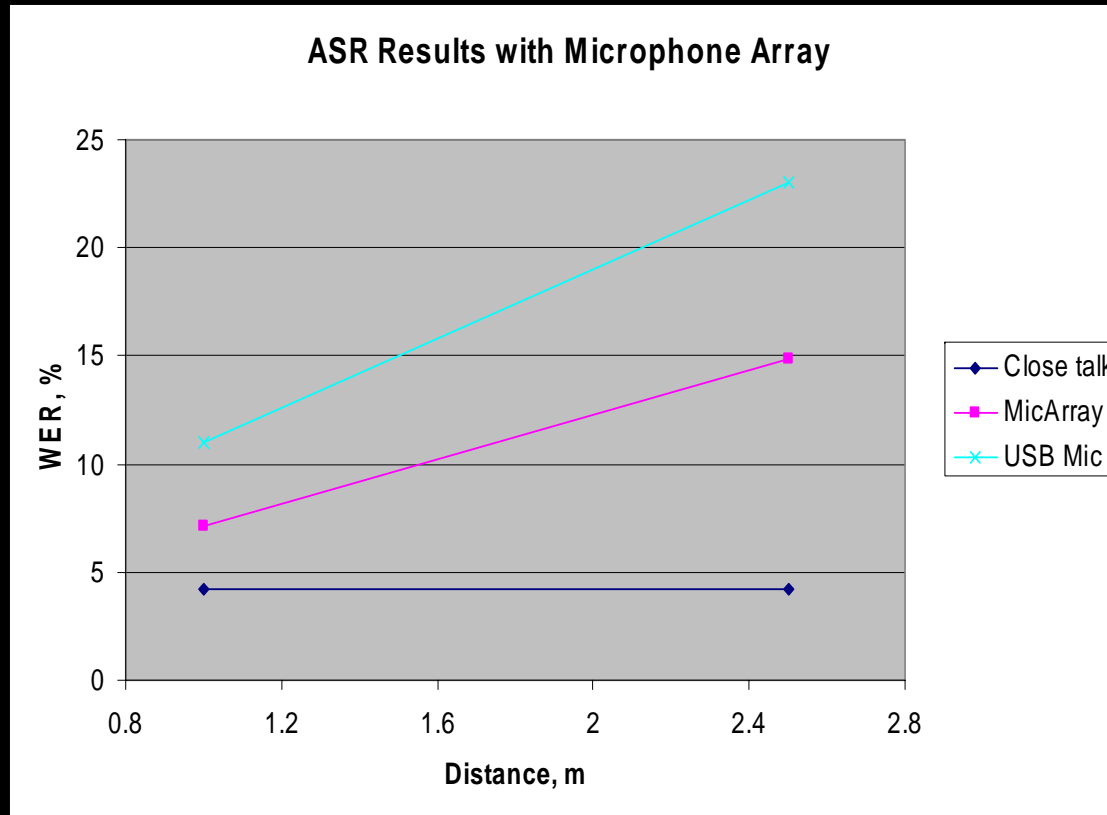
Device	SNR, dB
Omni Mic	12.51
Unidir Mic	19.3
2 elements	24.43
4 elements	28.69
8 elements	32.23
Close talk	35.24

Two for quiet office
Four for general use
Eight for heavy noise



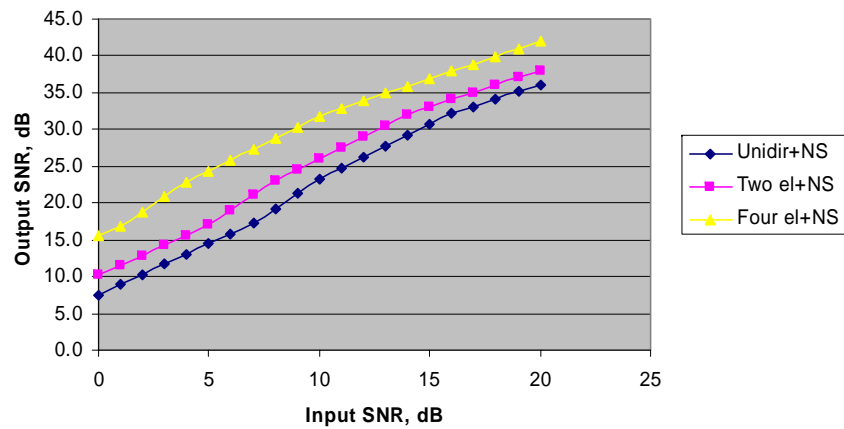
Results – speech recognition

Four element linear array, normal office conditions

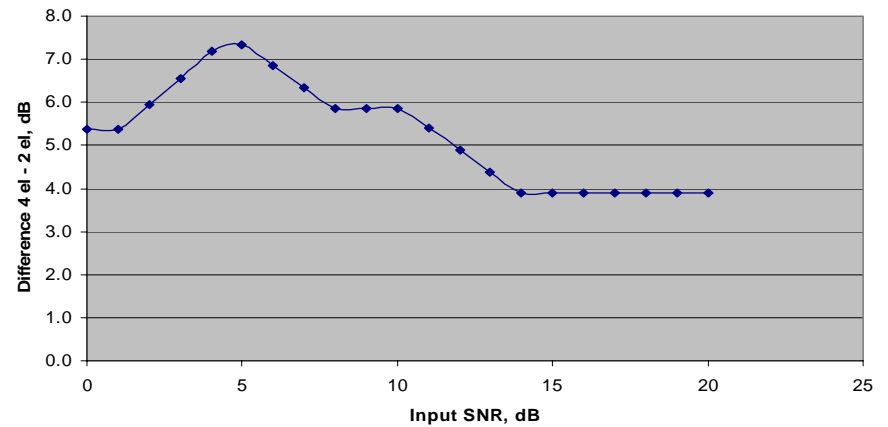


Two vs. Four microphones

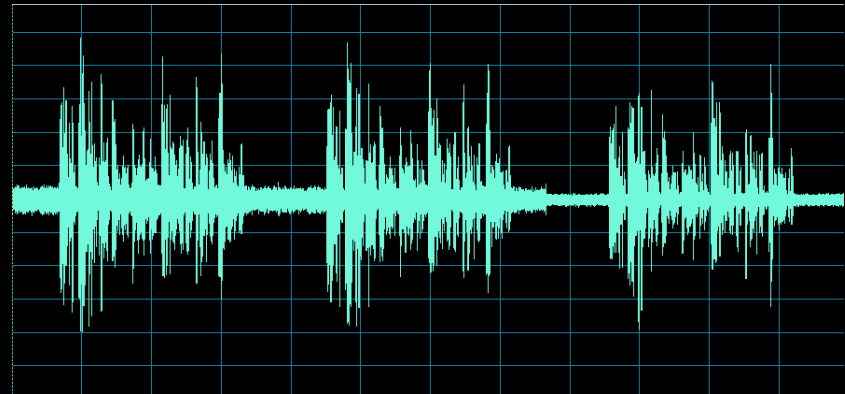
Noise suppression: one mic, 2 and 4 element micarrays



Comparison 2 and 4 element micarrays

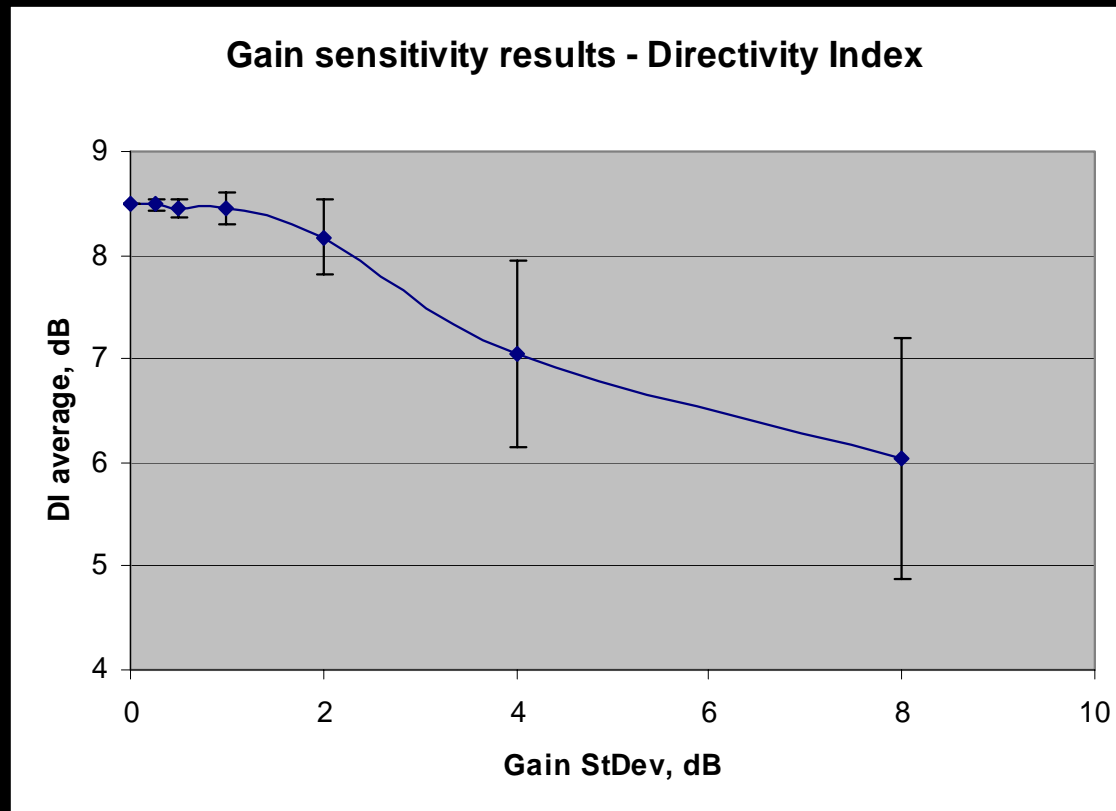


Four element is best in the typical noise conditions in a normal office: around 5 dB SNR



Sensitivity to manufacturing tolerances

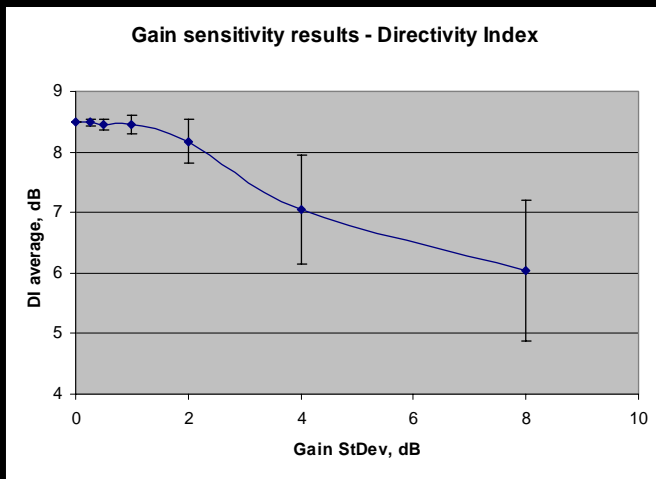
- Simulation of variances in the gain response with given deviation



Eight element circular array with WM55 microphones

Sensitivity to manufacturing tolerances

- Simulation of variances in the phase response with given deviation



Eight element circular array with WM55 microphones

