Return to Main

Objectives

Microphones:
Condenser
Electret

Directionality:
Omnidirectional
Bidirectional
Unidirectional
Other

Adaptive Filtering:
Echo Cancellation
Arrays

On-Line Resources:
Microphone Basics
Echo Cancellation Software
CAIP Arrays

LECTURE 10: ACOUSTIC TRANSDUCERS

- Objectives:
  - Introduce the basic types of microphones
  - Understand microphone impedance and other physical parameters
  - Learn how these influence the speech signal
  - Introduce the concept of adaptive filtering to improve signal quality
  - Introduce advanced technology such as microphone arrays

This material follows the course textbook closely:

> X. Huang, A. Acero, and H.W. Hon, *Spoken Language Processing - A Guide to Theory, Algorithm, and System Development*, Prentice Hall, Upper Saddle River, New Jersey, USA, ISBN: 0-13-022616-5, 2001.

This textbook contains an excellent discussion of these topics. Another good reference source for this material is:

> G.S.K. Wong and T.F.W. Embleton (Eds.), *AIP Handbook of Condensor Microphones: Theory, Calibration, and Measurements*, American Institute of Physics, New York, New York, USA, ISBN: 1-56396-284-5, 1995.

This is one of the definitive publications on condensor microphones.

ECE 8463: FUNDAMENTALS OF SPEECH RECOGNITION

Professor Joseph Picone
Department of Electrical and Computer Engineering
Mississippi State University

email: picone@isip.msstate.edu
phone/fax: 601-325-3149; office: 413 Simrall
URL: http://www.isip.msstate.edu/resources/courses/ece_8463

Modern speech understanding systems merge interdisciplinary technologies from Signal Processing, Pattern Recognition, Natural Language, and Linguistics into a unified statistical framework. These systems, which have applications in a wide range of signal processing problems, represent a revolution in Digital Signal Processing (DSP). Once a field dominated by vector-oriented processors and linear algebra-based mathematics, the current generation of DSP-based systems rely on sophisticated statistical models implemented using a complex software paradigm. Such systems are now capable of understanding continuous speech input for vocabularies of hundreds of thousands of words in operational environments.

In this course, we will explore the core components of modern statistically-based speech recognition systems. We will view speech recognition problem in terms of three tasks: signal modeling, network searching, and language understanding. We will conclude our discussion with an overview of state-of-the-art systems, and a review of available resources to support further research and technology development.

Tar files containing a compilation of all the notes are available. However, these files are large and will require a substantial amount of time to download. A tar file of the html version of the notes is available here. These were generated using wget:

```
wget -np -k -m http://www.isip.msstate.edu/publications/courses/ece_8463/lectures/current
```

A pdf file containing the entire set of lecture notes is available here. These were generated using Adobe Acrobat.

Questions or comments about the material presented here can be directed to help@isip.msstate.edu.

19: Linear Discriminant Analysis
(html, pdf)

LECTURE 10: ACOUSTIC TRANSDUCERS

- Objectives:
  - Introduce the basic types of microphones
  - Understand microphone impedance and other physical parameters
  - Learn how these influence the speech signal
  - Introduce the concept of adaptive filtering to improve signal quality
  - Introduce advanced technology such as microphone arrays

This material follows the course textbook closely:

> X. Huang, A. Acero, and H.W. Hon, *Spoken Language Processing - A Guide to Theory, Algorithm, and System Development*, Prentice Hall, Upper Saddle River, New Jersey, USA, ISBN: 0-13-022616-5, 2001.

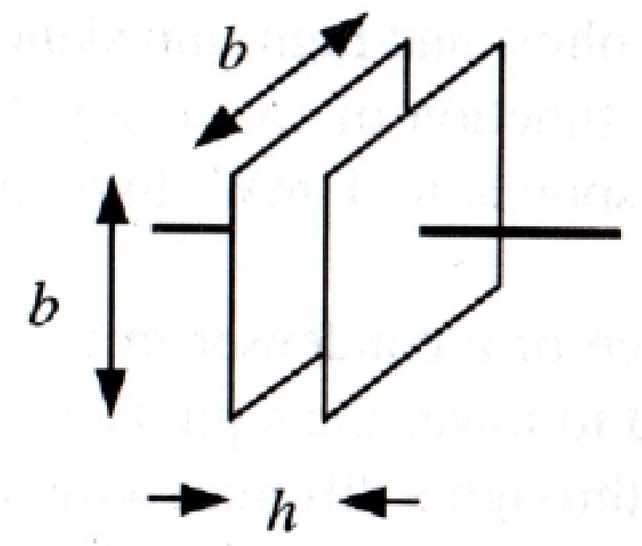
This textbook contains an excellent discussion of these topics. Another good reference source for this material is:

> G.S.K. Wong and T.F.W. Embleton (Eds.), *AIP Handbook of Condensor Microphones: Theory, Calibration, and Measurements*, American Institute of Physics, New York, New York, USA, ISBN: 1-56396-284-5, 1995.

This is one of the definitive publications on condensor microphones.

CONDENSER MICROPHONES

A *condensor* microphone has a capacitor consisting of a pair of metal plates separated by an insulating material called a dielectric:



One of its plates is free to move in response to changes in sound pressure. The sensitivity of the microphone is related to its polarizing voltage and distance of separation between these plates.

Key design equations for this type of microphone are:

$$C = \varepsilon_0 \pi b^2 / h$$

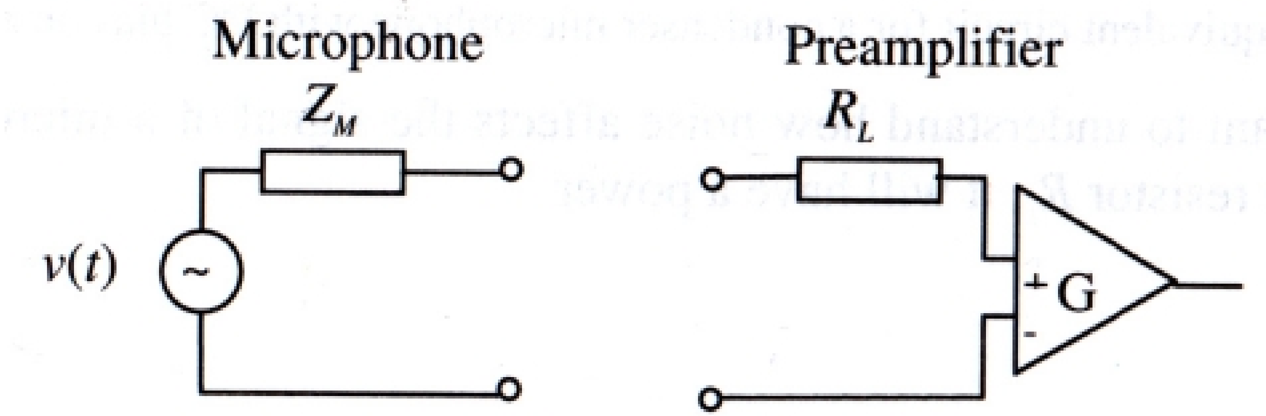
$$Q = CV_{cc}$$

$$\nabla V = \nabla h V_{cc} / h$$

Thus, the sensitivity depends on the polarizing voltage, V_{cc}, which explains why many microphones operate at large voltages (often 100V or more).

ELECTRET MICROPHONES

- An *electret* microphone is a specific type of condenser microphone that does not require a special polarizing voltage because its diaphram or back plate is permanently charged.
- Electret microphones are small, cheap, durable, and offer good performance at high frequencies. Most modern telephone handsets use electrets.
- The electrical equivalent circuit for a microphone is:



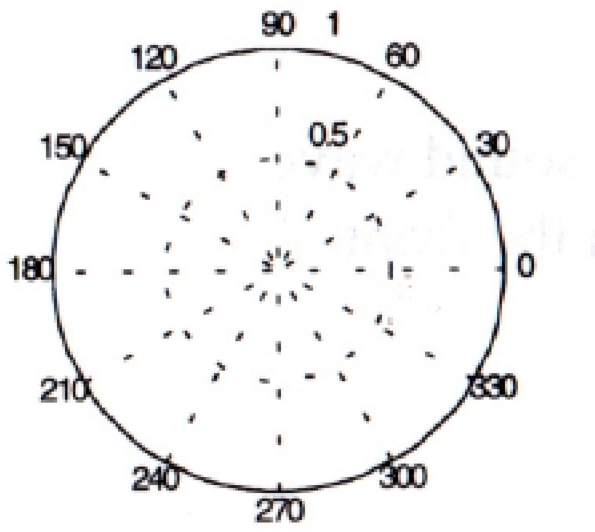
*Bridging*refers to maximizing the output voltage by increasing the load impedance and/or decreasing the microphone impedance.

- Low impedance microphones have an impedance of 100 to 300 Ohms (more expensive); high impedance microphones have an impedance of 600 to 1000 Ohms (less expensive).
- Condenser microphones require a DC bias; balanced cables (XLR) are prefered over unbalanced lines (1/8" mini plugs) because the cables are more resistant to thermal and RF noise.
- Digital and wireless microphones are popular alternatives.

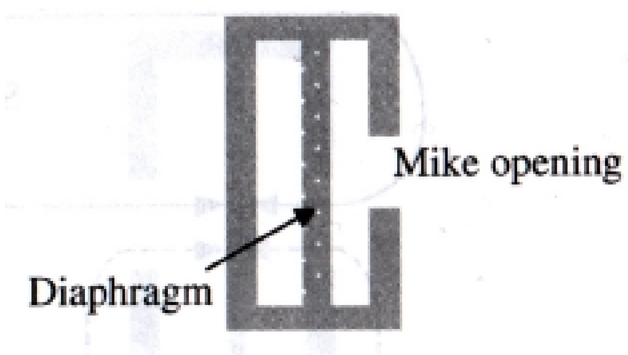
OMNIDIRECTIONAL SENSITIVITY PATTERNS

A microphone's directionality pattern can be described in terms similar to what we use for antennae. Its sensitivity can be measured as a function of direction. Microphones can be classified as *omnidirectional* (nondirectional) and *directional* (e.g., bidirectional and unidirectional).

The polar response of an ideal omnidirectional microphone is:



A cross-sectional view of the microphone:

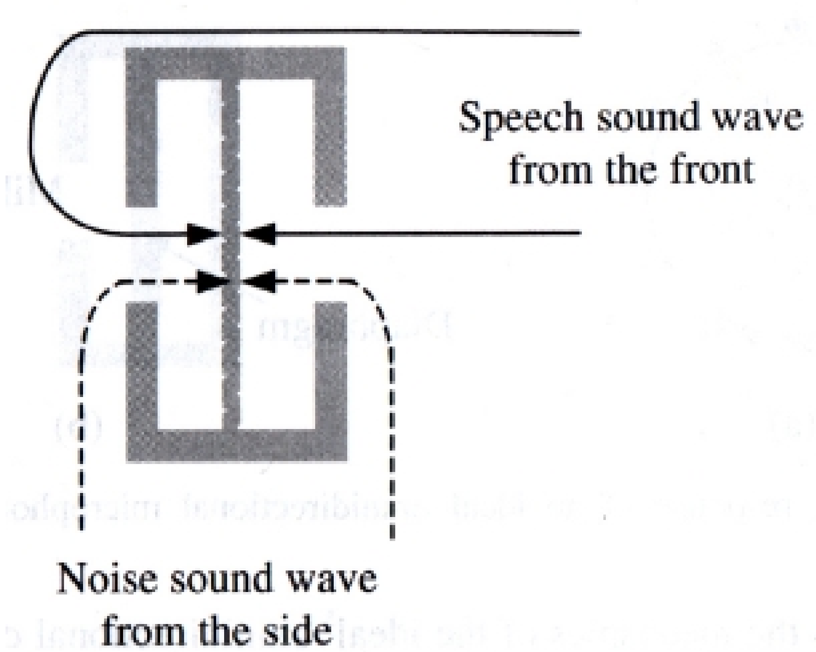


reveals that the diaphragm is designed to be sensitive to signals emanating from any direction.

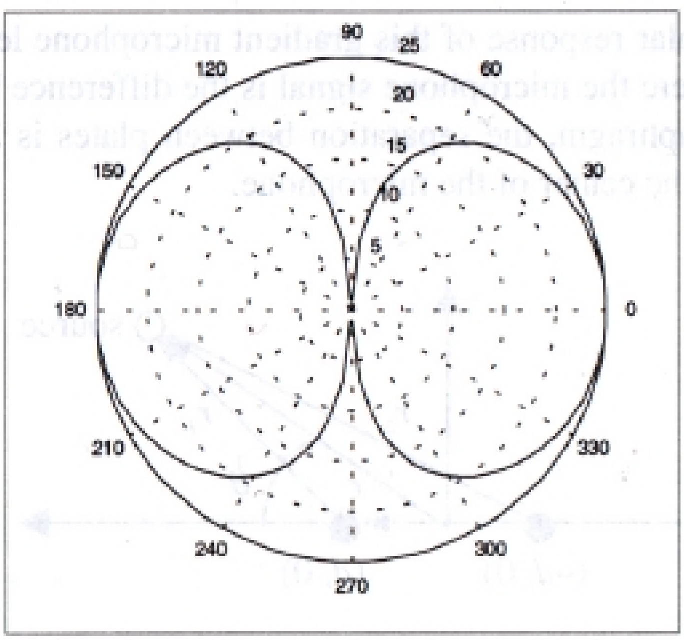
BIDIRECTIONAL SENSITIVITY PATTERNS

A bidirectional microphone is a *noise-cancelling* microphone (such as the Sennheiser HD 414 close-talking microphone that is so popular in speech research).

A bidirectional microphone uses properties of a gradient microphone to achieve noise cancellation. Sound pressure never arrives at the front and the back of the microphone at the same time. However, noise, which arrives from the side, does, and hence is cancelled.



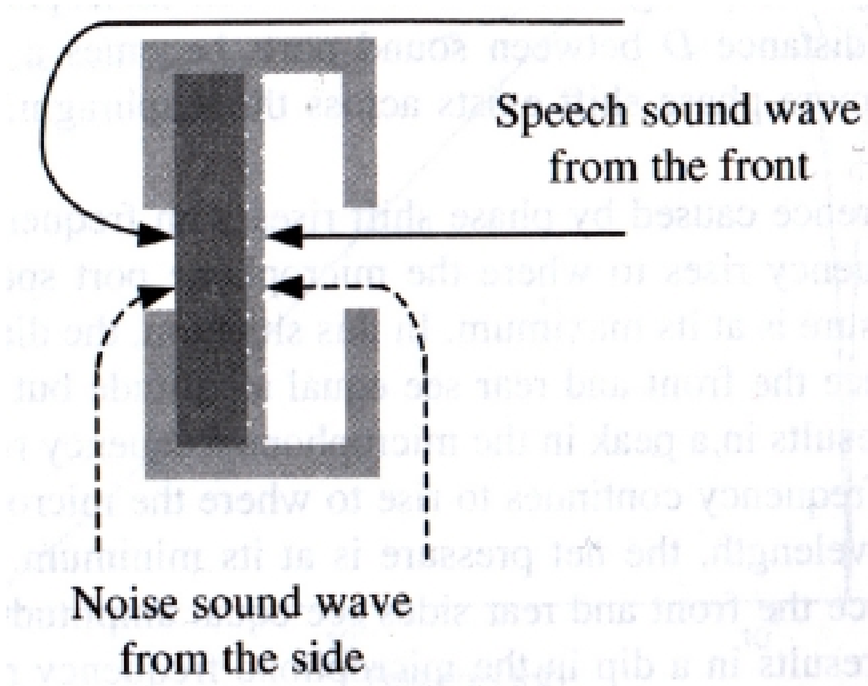
These microphones have directional sensitivity patterns:



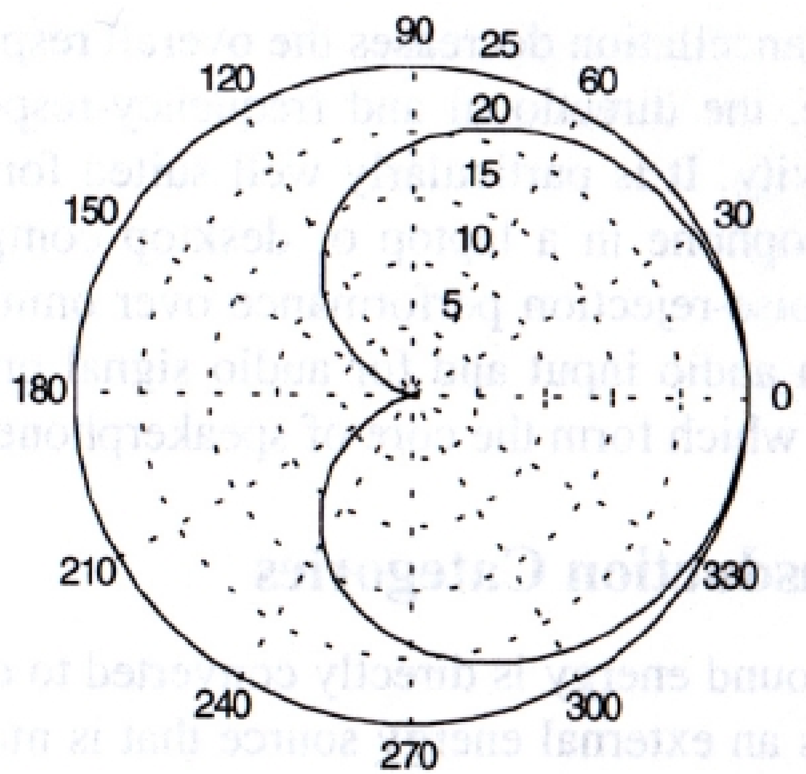
Such microphones are very sensitive to placement, and cannot be used interchangeably with recognition systems (microphone independence?).

UNIDIRECTIONAL SENSITIVITY PATTERNS

Unidirectional microphones, which are popular in computer applications involving desktop microphones, are similar to close-talking microphones:



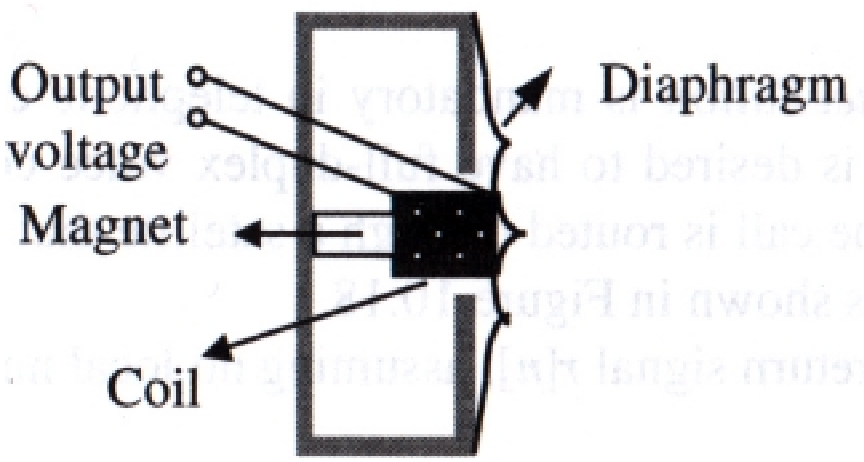
These are often referred to as *cardiod* microphones. They have the following sensitivity pattern:



PIEZOELECTRIC AND OTHER TYPES OF MICROPHONES

Microphones can be classified in terms of how they create an electrical signal:

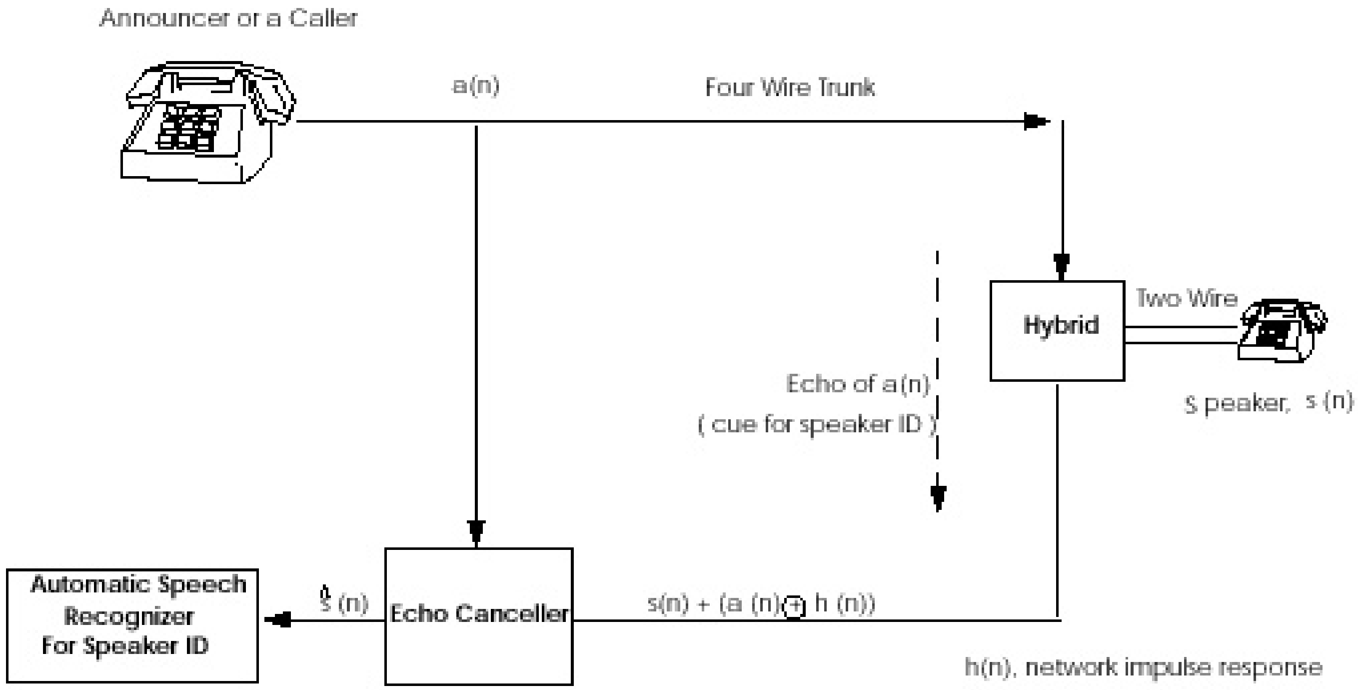
- **Electromagnetic**: Ribbons (thin metal ribbon suspended between magnets), dynamic (moving coil)



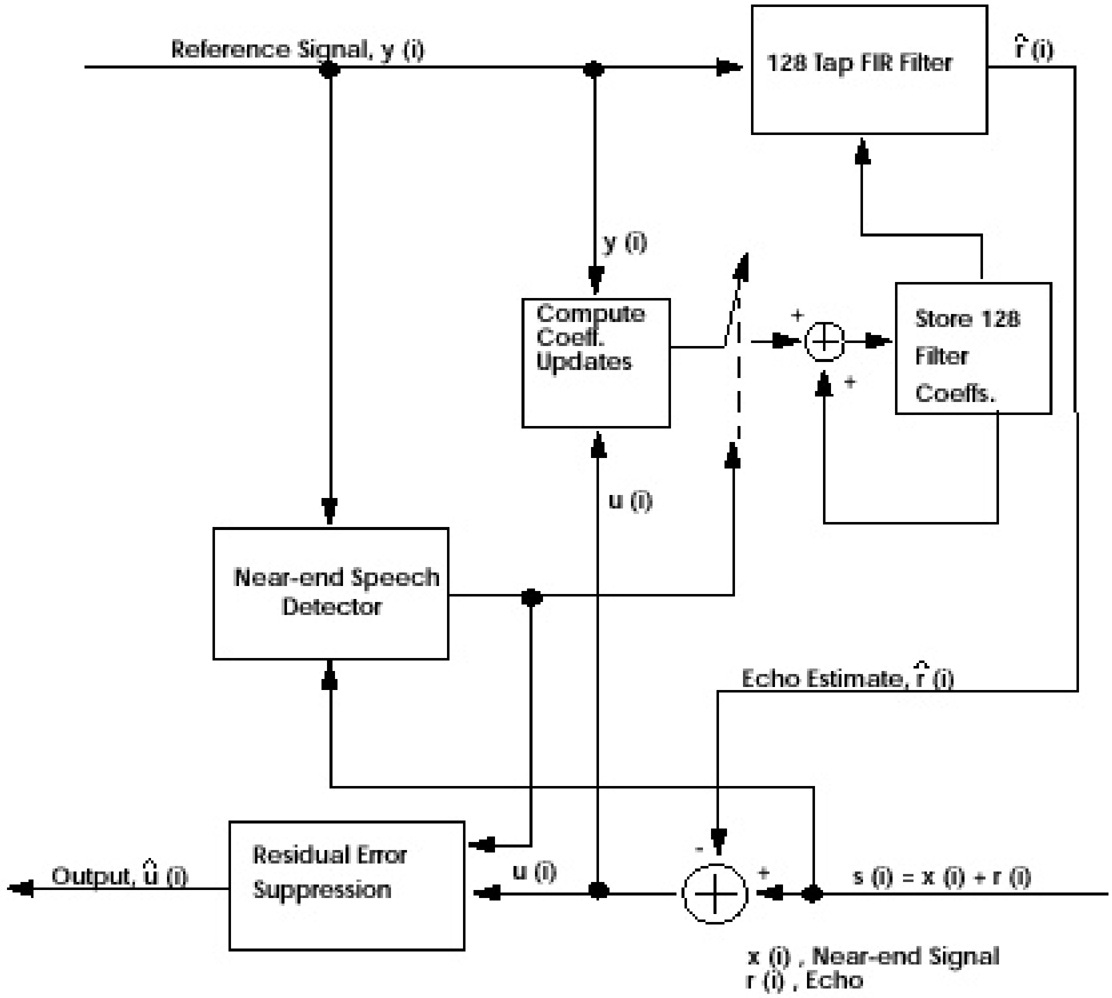
- **Electrostatic**: condensers, electrets, etc.
- **Piezoelectric**: based on variations of electric resistance of their sensor induced by changes in sound pressure. Carbon button microphones (old AT&T handsets) and desktop "far-field" microphones are popular examples of these. Lower sensitivity, more distortion, and non-flat frequency responses are characteristics of these microphones.

ECHO CANCELLATION

Echo cancellers are often required in speech recognition systems due to analog impairments present in the telephone system:



These are implemented using simple FIR adaptive filtering techniques:

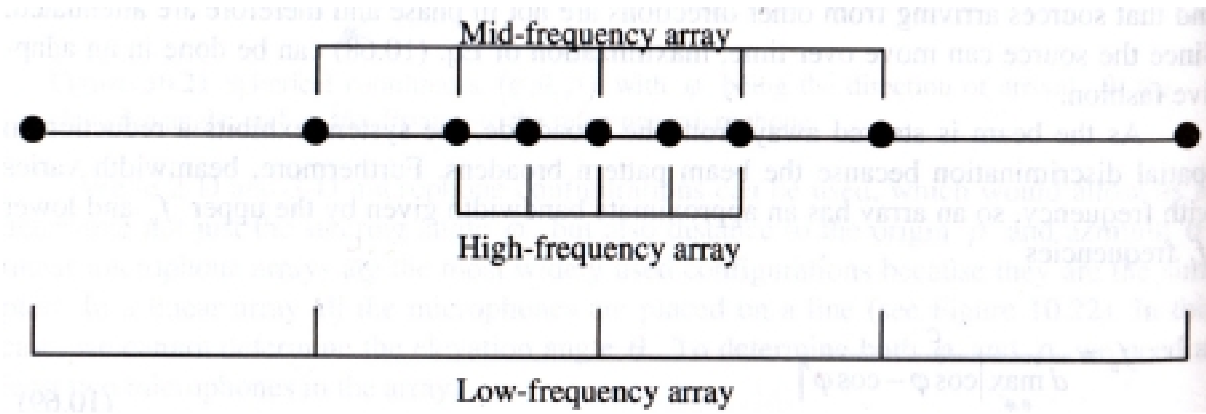


The coefficients of these filters are computed using a least mean-squared error approach (LMS). Such systems are used to allow users to speak during a prompt (barge-in), which is a very important feature of a practical recognition system. A reference implementation of a standard FIR echo canceller is available on-line, along with many educational resources and conference papers on applications to speech recognition.

MICROPHONE ARRAYS AND BEAM STEERING

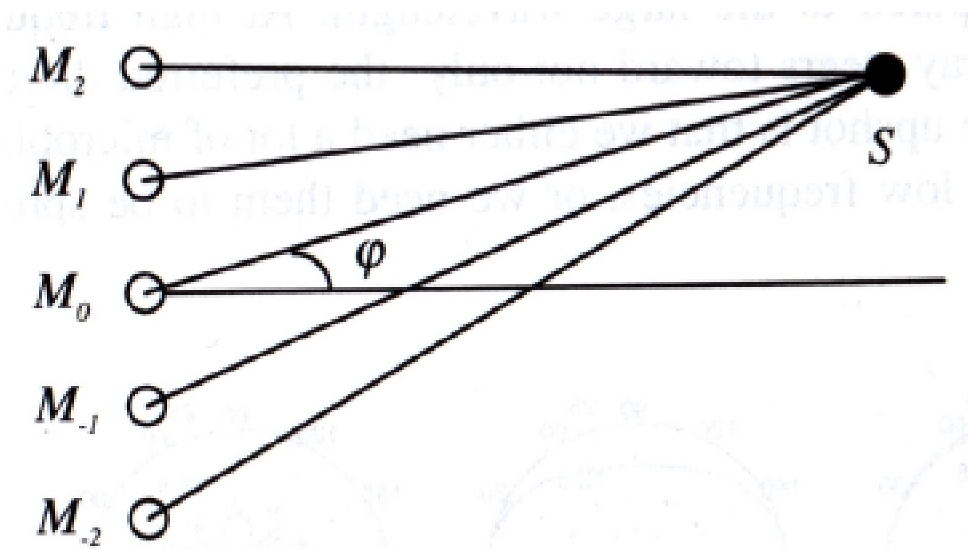
Question: Can we improve performance using multiple microphones?

The goal of a *microphone array* is to localize a sound source by directing a group of microphones to be most sensitive in a specific direction. The procedure is completely analogous to analog antenna theory.



Steering of the array has several uses: increase SNR, direct video in teleconferencing, enhance the human interface (hands-free).

Steering the array amounts to adjusting the delays in each microphone. The most common implementation is the *delay and sum beamformer*:



An example of a sensitivity pattern for this type of array is:

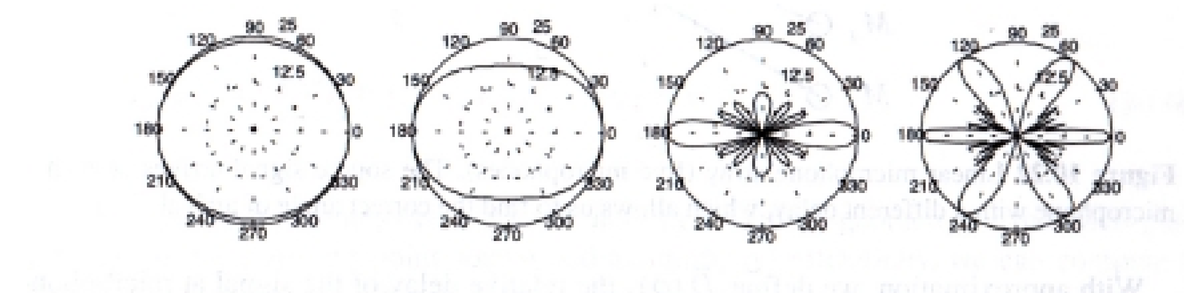


Figure 10.23 Polar pattern of a microphone array with steering angle of $\varphi' = 0$, five microphones spaced 5 cm apart for 400, 880, 4400, and 8000 Hz from left to right, respectively, for a source located at 5 m.

A 3D array can be used to localize sound to a single point in a room (direction and distance). 2D arrays are most commonly used to enhance SNR. Unfortunately, performance increases slowly as a function of the number of microphones

(1 dB rule). Hence, this technology is impractical for many consumer applications.

Two microphone versions of this idea based on adaptive filtering are popular in automotive applications for noise suppression.

Search

Search

Browse

Online Catalog

Free Catalog
Newsletter
Employment
Testimonials
Terms/Policies
Privacy Policy
Contact Us

Microphones, Part 1

Basic Microphone Theory

A Microphone is a generic term that refers to any element that transforms acoustic energy (sound) into electrical energy (electricity (audio signal)).

One type of microphone commonly used is the dynamic microphone. A dynamic microphone is much like a miniature loudspeaker. A flexibly-mounted diaphragm is coupled to a coil of fine wire. The coil is mounted in the air gap of magnet such that it is free to move back and forth within the gap. When sound strikes the diaphragm, the diaphragm surface vibrates in response. The motion of the diaphragm couples directly to the coil, which moves back and forth in the field of the magnet. As the coil cuts through lines of magnetic force in the gap, a small electrical current is induced in the wire. The magnitude and direction of that current is directly related to the motion of the coil, and the current thus is an electrical representation of the incident sound wave. Dynamic microphones are highly dependable, rugged, and reliable. They are extremely common in stage use, where physical durability is important. They are also reasonably insensitive to environmental factors, and thus find extensive use in outdoor applications.

Next to the dynamic microphone, the most common type of microphone is the condenser. Basic things you should know about condenser microphones: they need an external power supply. Dynamic microphones do not need this sort of power, but condensers do. Some condenser microphones have a battery attachment that is either part of the microphone housing or on the end of the cable, as part of the connector. Others need power delivered to them through the cable. This power system is known as phantom power. Some mixers provide phantom power, some don't. If they do, there is usually a main phantom switch somewhere. What phantom power does is send a voltage of +48VDC through pins two and three of the microphone cable to the microphone capsule. The return voltage comes back via the shield. Condenser microphones are also slightly less durable than dynamic microphones, so they aren't used as much in on-stage, touring, or reinforcement applications, although some condensers are used for micing drum kits. Mostly, condensers are used in studio applications.

A variant of the condenser microphone is the electret condenser. Electret condenser microphones don't require phantom power to charge the diaphragm (like the condenser), but they do require a power supply for their in-microphone preamplifier. "In-microphone preamplifier" does not mean that they deliver a line-level output; they don't; it's for a sort of "pre-preamplifier." Increasingly, more electret condenser mics use phantom power to power their preamplifiers.

People generally like condenser mics because the sound is usually a bit cleaner and more predictable than dynamic mics. Of course, though, they are generally more expensive.

There are several other types of microphones, but they are rarely used in sound reinforcement: Ribbon mics (very, very fragile), and Carbon mics (very bad sound quality; used in telephone handsets although more and more handsets use dynamic mics).

Home
Software
Docs
Tutorials
Demos
Databases
Dictionaries
Models
Research
Support
Mailing Lists
What's New
Search

Echo Canceller

Speech data collected over the telephone, such as the SWITCHBOARD conversational speech data corpus, contains echo caused by the process of converting a two-wire signal used in the local loop to a four-wire signal used in network transmission. See system overview for more details.

This echo could be used by the speech recognizer to gather important cues regarding the ID of the speaker or the channel conditions, thereby making the job that much easier. To eliminate this problem, we need an efficient echo cancellation technology.

We have developed an FIR echo-canceller for this purpose. A detailed description of this technology can be found in the following reference:

- Messerschmitt, David; Hedberg, David; Cole, Christopher; Haoui, Amine; Winship, Peter; "Digital Voice Echo Canceller with a TMS32020," in Digital Signal Processing Applications with the TMS320 Family, pp. 415-437, Texas Instruments, Inc., 1986.

This document can be retrieved from the Texas Instruments web site. A copy can be found on this web site as well.

We have deviated from the standard implementation of an LMS echo-canceller at places to accommodate certain problems we face. Some of the main problems we encountered during the development of the system are:

- **Double talk**: This is a condition when both the speakers talk simultaneously. If we adapt the FIR filter coefficients during double talk, the filter will diverge, causing "blips" in the output. This can be avoided by having an efficient voice activity detector (VAD). When the VAD detects near-end speech the adaptation process is suspended. This avoids the divergence problem.
- **Complex echo**: The echo-canceller performs poorly in some cases of double talk. It fails to cancel the far-end speech effectively. We attribute this to the possibility of the existence of complex echo patterns.
- **Residual Error Suppression**: We know that due to the non-linearities of the echo path of the telephone network the maximum suppression possible is limited to about 40dB. So, in cases when the return signal power falls below a threshold based on the reference signal power, it is suggested that we zero the output. This process however creates a choppyness in the background. To make the background more uniform, we decided to make the output equal to a scaled version of the reference signal when the near-end signal is not present.
- **Length of the filter**: Unfortunately the length of the FIR filter has to depend on the maximum delay in echo signal in the data set we are using. If we consider international telephone conversations, the round trip delay is typically an order of magnitude more than that for domestic calls. We would like our system to automatically choose the length of the filter depending on the maximum round-trip delay the user specifies. Also another unanswered question is the relationship between the adaptation rate constant and the filter length. From the experiments we performed, there seems to an inverse relationship between the two quantities.

You can download the following from our site:

- Tar File: download a C++ implementation in compressed gzip format.
- Source Code: view the C++ source code distribution.
- Example Data: some example data to verify your implementation.
- System Overview: a system overview in pdf format.
- TI DSP Application Note: an excellent application note describing the theory and implementation of an LMS echo canceller. The implementation included here is based on this application note and references it heavily.

CAIP Microphone Array Research

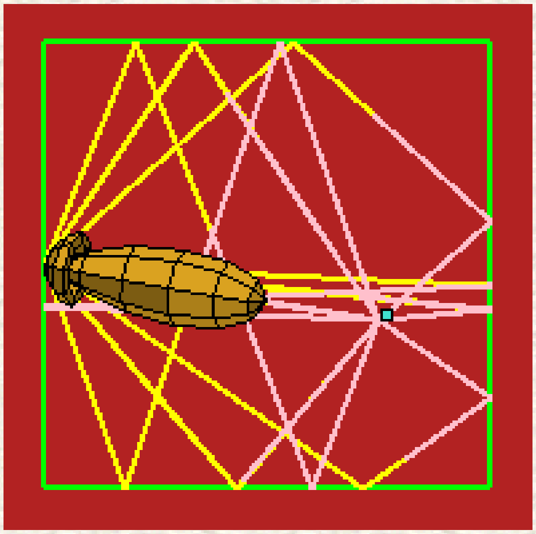
Last Updated July, 1998

Contents

Return to the Multimedia Lab Home Page or the CAIP Home Page.

Maintained by Chris Alvino <*chrsvino@caip.rutgers.edu*>

Introduction



The Center for Computer Aids for Industrial Productivity (CAIP Center) at Rutgers University-New Brunswick continues to conduct advanced research in the area of high quality sound capture as part of a larger interest in multimedia-related research. The current research has a particular focus on active and passive microphone arrays. Microphone arrays enable the capture of high quality audio waveforms from remote sound sources, under adverse acoustic conditions. In particular, arrays allow the tracking and recording of moving human talkers without requiring the use of a cumbersome tethered microphone.

Recent advances in circuit technology have made digital signal processing systems capable of real-time processing of multiple audio channels practical to implement and readily available. This has been combined with an ever-increasing demand in voice controlled applications that exhibit robust, environment independent performance. Microphone array systems have many uses for sound capture and acoustic source location in applications such as cellular telephony, video-teleconferencing, and audio interfacing with PC systems.

If you're looking for sound samples from the paper given at the San Diego ASA meeting, click here.