U.S. PTO
61/403952
09/24/2010

PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 1 of 2

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. ______________

INVENTOR(S)

Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Manli	Zhu	Pearl River, NY
Qi	Li	New Providence, NJ

Additional inventors are being named on the ______________ *separately numbered sheets attached hereto.*

TITLE OF THE INVENTION (500 characters max):

Microphone Array Design and Implementation for Telecommunications and Handheld Devices

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☐ The address corresponding to Customer Number: ______________

OR

☒ Firm or Individual Name: Qi (Peter) Li

Address: Li Creative Technologies, Inc., 25B Hanover Road, Suite 140		
City: Florham Park	State: NJ	Zip: 07932
Country: USA	Telephone: 973-822-0048	Email: li@licreativetech.com

ENCLOSED APPLICATION PARTS *(check all that apply)*

☐ Application Data Sheet. See 37 CFR 1.76

☐ CD(s), Number of CDs ______________

☐ Drawing(s) *Number of Sheets* ______________

☒ Other (specify) descrip w/ drawings 23 pp

☐ Specification (e.g. description of the invention) *Number of Pages* ______________

Fees Due: Filing Fee of $220 ($110 for small entity). If the specification and drawings exceed 100 sheets of paper, an application size fee is also due, which is $270 ($135 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

METHOD OF PAYMENT OF THE FILING FEE AND APPLICATION SIZE FEE FOR THIS PROVISIONAL APPLICATION FOR PATENT

☒ Applicant claims small entity status. See 37 CFR 1.27.

☐ A check or money order made payable to the *Director of the United States Patent and Trademark Office* is enclosed to cover the filing fee and application size fee (if applicable).

☒ Payment by credit card. Form PTO-2038 is attached.

☐ The Director is hereby authorized to charge the filing fee and application size fee (if applicable) or credit any overpayment to Deposit Account Number: ______________ .

TOTAL FEE AMOUNT ($): 110

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

- [x] No.
- [ ] Yes, the name of the U.S. Government agency and the Government contract number are: ______

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

SIGNATURE ______ Date 09-22-2010

TYPED or PRINTED NAME Qi (Peter) Li REGISTRATION NO. ______
(if appropriate)

TELEPHONE 973-822-0048 Docket Number: ______

Effective on 12/08/2004.
Fees pursuant to the Consolidated Appropriations Act, 2005 (H.R. 4818).

FEE TRANSMITTAL
For FY 2009

☑ Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT ($) 110

Complete if Known	
Application Number	
Filing Date	09-22-2010
First Named Inventor	Manli Zhu
Examiner Name	
Art Unit	
Attorney Docket No.	

METHOD OF PAYMENT (check all that apply)

☐ Check ☑ Credit Card ☐ Money Order ☐ None ☐ Other (please identify): ____________

☐ Deposit Account Deposit Account Number: ____________ Deposit Account Name: ____________

For the above-identified deposit account, the Director is hereby authorized to: (check all that apply)

☐ Charge fee(s) indicated below ☐ Charge fee(s) indicated below, **except for the filing fee**

☐ Charge any additional fee(s) or underpayments of fee(s) under 37 CFR 1.16 and 1.17 ☐ Credit any overpayments

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

FEE CALCULATION

1. BASIC FILING, SEARCH, AND EXAMINATION FEES

	FILING FEES		SEARCH FEES		EXAMINATION FEES		
Application Type	Fee ($)	Small Entity Fee ($)	Fee ($)	Small Entity Fee ($)	Fee ($)	Small Entity Fee ($)	Fees Paid ($)
Utility	330	165	540	270	220	110	
Design	220	110	100	50	140	70	
Plant	220	110	330	165	170	85	
Reissue	330	165	540	270	650	325	
Provisional	220	110	0	0	0	0	110

2. EXCESS CLAIM FEES

Fee Description	Fee ($)	Small Entity Fee ($)
Each claim over 20 (including Reissues)	52	26
Each independent claim over 3 (including Reissues)	220	110
Multiple dependent claims	390	195

Total Claims ______ - 20 or HP = **Extra Claims** ______ x **Fee ($)** ______ = **Fee Paid ($)** ______

HP = highest number of total claims paid for, if greater than 20.

Indep. Claims ______ - 3 or HP = **Extra Claims** ______ x **Fee ($)** ______ = **Fee Paid ($)** ______

HP = highest number of independent claims paid for, if greater than 3.

Multiple Dependent Claims

Fee ($) ______ **Fee Paid ($)** ______

3. APPLICATION SIZE FEE

If the specification and drawings exceed 100 sheets of paper (excluding electronically filed sequence or computer listings under 37 CFR 1.52(e)), the application size fee due is $270 ($135 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

Total Sheets ______ - 100 = **Extra Sheets** ______ / 50 = **Number of each additional 50 or fraction thereof** ______ (round **up** to a whole number) x **Fee ($)** ______ = **Fee Paid ($)** ______

4. OTHER FEE(S) **Fees Paid ($)**

Non-English Specification, $130 fee (no small entity discount) ______

Other (e.g., late filing surcharge): ____________ ______

SUBMITTED BY

Signature	[signature]	Registration No. (Attorney/Agent)	Telephone 973-822-0048
Name (Print/Type)	Qi (Peter) Li		Date 09-22-2010

Certificate of Mailing under 37 CFR 1.8

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to:

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

on 09-22-2010 .
Date

Signature

Qi (Peter) Li
Typed or printed name of person signing Certificate

Registration Number, if applicable

973-822-0048
Telephone Number

Note: Each paper must have its own certificate of mailing, or this certificate must identify each submitted paper.

Li Creative Technologies, Inc.
25 B Hanover Road, Suite 140, Florham Park, NJ 07932, USA
Tel: (973) 822-0048; Fax: (973) 822-0399; Website: www.licreativetech.com

Written Assertion of Small Entity Status

This document states that Li Creative Technologies, Inc. (LcT) is entitled to small entity status with regards to our filing a patent application with the USPTO and therefore should be permitted to pay reduced fees.

____
Signature

Qi (Peter) Li, President 09-22-2010
Printed Name Date

Microphone Array Design and Implementation for Telecommunications and Handheld Devices

Manli Zhu and Qi Li

1. Background of the Invention

A microphone array consists of a set of microphone sensors located at different positions. The array can achieve directional gain in any preferred spatial direction and frequency band while suppressing signals from other directions and bands. The array can be implemented by filtering and summing multiple microphone outputs. Conventional array processing techniques, typically developed for applications such as radar and sonar, are generally not appropriate for hands-free or handheld speech acquisition devices. The main reason is that the desired speech signal has an extremely wide bandwidth relative to its center frequency, meaning that conventional narrowband techniques are not suitable. In the approaches to keep the constant response in the wide range of frequency, the array size is usually large; thus most of prototypes or products of microphone arrays on the market are quite large, which prevents the array products from having broader applications, such as for use in mobile and handheld communication devices.

2. Summary of the Invention

Our invention is Microphone Array Design and Implementation for Telecommunications and Handheld Devices. Our invention can be used for arbitrary directivity pattern for arbitrarily distributed microphones. Our invention can be used to design a microphone array for small, portable communication devices, such as conference phones, mobile phones, or tablet computers. To illustrate our invention, we present three applications of our invention: (1) a microphone array for a conference phone conference phone device with of eight microphones non-uniformly distributed on around a circle with diameter of 4 inches, (2) a microphone array of four microphones located at the four corners of a rectangle for a wireless phone or handheld device; and (3) a microphone array of four microphones located on the frame of a tablet computer.

3. Description of the Invention

In this section, we provide a complete description of our invention together with all the drawings necessary to understand the invention. Our invention can be used for arbitrary numbers of microphone components and arbitrary locations of the microphone components. Our can be implemented in either software or hardware or a combination

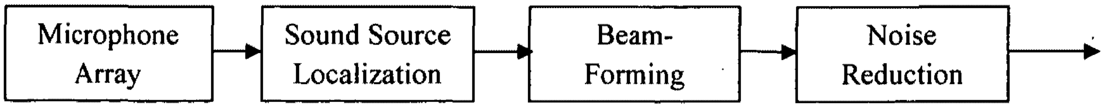


Figure 1. For the best performance, the invented microphone array system may consist of the following modules: microphone array sensors, sound source localization, beamforming, and noise reduction.

The microphone array module consists of multiple microphone components functioning as a single unit to pick up sound signals. The source localization module serves to find the spatial location of the principal sound source such that an acoustic beam can point to the sound source. The beamforming module serves to form acoustic beams in the direction of the principal sound source enhancing sound from this range and suppressing sound from all other directions. The noise reduction module serves to further reduce background noise and enhance speech. Depending on applications, a real product may use some or all of the modules.

3.1. Two-Dimensional Microphone Array Configuration

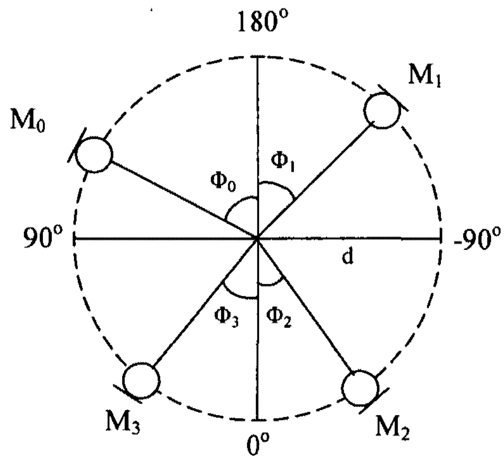


Figure. 2 Illustration of a microphone array configuration wherein N microphones sensors are arbitrarily distributed on a circle with diameter of d (N=4).

Assuming N microphones are arbitrarily distributed on a circular with diameter of d as shown in Figure 2, where only four microphones are displayed. Microphone locations are specified a acute angles from the y-axis, shown as Φ_n ($\Phi_n \geq 0$, n=1...N). The output y of the array is the filter-and-sum of the N microphone outputs, i.e., $y = \sum_{n=0}^{N-1} w_n^T x_n$, where x_n is the output of the (n+1)th microphone and w_n is the length-L filter applied to it as shown in Figure 3.

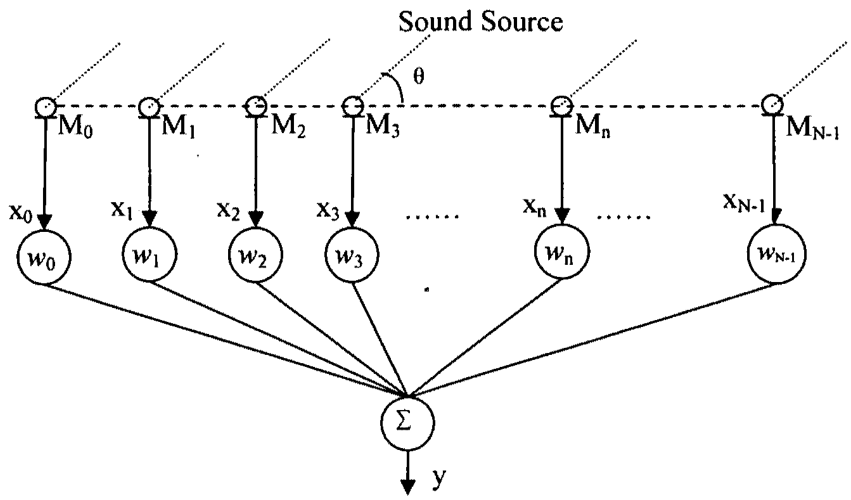


Figure 3. Illustration of filter-and-sum beamforming.

The spatial directivity pattern $H(\omega,\theta)$ for the sound source from angle θ with normalized frequency ω is defined as:

$$H(\omega,\theta)=\frac{Y(\omega,\theta)}{\overline{X}(\omega,\theta)}=\frac{\sum_{n=0}^{N-1}W_n(\omega)X_n(\omega,\theta)}{\overline{X}(\omega,\theta)} \tag{1}$$

where $\overline{X}$ is the signal received at the center of the circular array and W is the frequency response of the real-valued FIR filter w. If the sound source is far enough away from the array, the difference between the signal received by the $(n+1)^{th}$ microphone x_n and the center of the array is a pure delay τ_n. i.e., $X_n(\omega,\tau)=\overline{X}(\omega,\theta)e^{-j\omega\tau_n}$. Figure 4 illustrates the distance between origin and microphone M_1 and microphone M_3 when the incoming sound is from angle of θ.

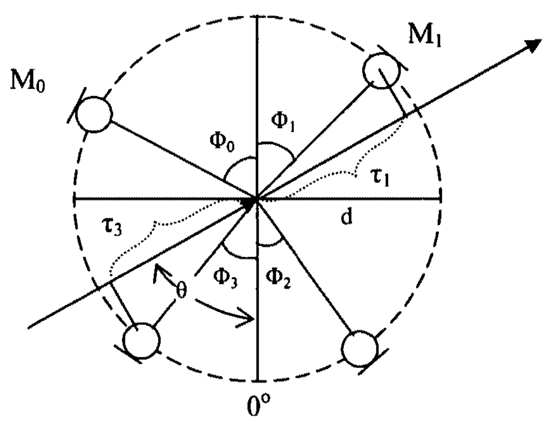


Figure 4. Illustration of τ_1 and τ_3, the distance between origin and microphone M_1 and microphone M_3 when the incoming sound is from angle of θ.

We derived the distance for each microphone, measured both in meters and in the number of samples, and summarize them in Table 1.

Table 1. Distance between each microphone and origin.
Note: d is the radius of the circle, f_s is the sampling frequency, and C is the sound speed.

Microphone	Distance (m)	Distance (number of samples)
M0	$d*\cos(\theta+\Phi_0)$	$d*\cos(\theta+\Phi_0)*f_s/C$
M1	$d*\cos(\theta-\Phi_1)$	$d*\cos(\theta-\Phi_1)*f_s/C$
M2	$-d*\cos(\theta+\Phi_2)$	$-d*\cos(\theta+\Phi_2)*f_s/C$
M3	$-d*\cos(\theta-\Phi_3)$	$-d*\cos(\theta-\Phi_3)*f_s/C$

In general, the distance and the location have the following relationship:

Table 2. Relationship of microphone position and its distance to the origin.

Microphone position	Distance (m)	Distance (number of samples)
0°	$-d*\cos(\theta)$	$-d*\cos(\theta)*f_s/C$
180°	$d*\cos(\theta)$	$d*\cos(\theta)*f_s/C$
90°	$-d*\sin(\theta)$	$-d*\sin(\theta)*f_s/C$
-90°	$d*\sin(\theta)$	$d*\sin(\theta)*f_s/C$
Φ clockwise away from 0° $(0\leq\Phi\leq 90^\circ)$	$-d*\cos(\theta-\Phi)$	$-d*\cos(\theta-\Phi)*f_s/C$
Φ anticlockwise away from 0° $(0\leq\Phi\leq 90^\circ)$	$-d*\cos(\theta+\Phi)$	$-d*\cos(\theta+\Phi)*f_s/C$
Φ clockwise away from 180° $(0\leq\Phi\leq 90^\circ)$	$d*\cos(\theta-\Phi)$	$d*\cos(\theta-\Phi)*f_s/C$
Φ anticlockwise away from 180° $(0\leq\Phi\leq 90^\circ)$	$d*\cos(\theta+\Phi)$	$d*\cos(\theta+\Phi)*f_s/C$

Now, the spatial directivity pattern H can be re-written as:

$$H(\omega,\theta)=\sum_{n=0}^{N-1}W_n(\omega)e^{-j\omega\tau_n(\theta)}=\mathbf{w}^T\mathbf{g}(\omega,\theta) \qquad (2)$$

where $\mathbf{w}^T=[w_0{}^T,w_1{}^T,w_2{}^T,w_3{}^T,\ldots,w_{N-1}{}^T]$ and $\mathbf{g}(\omega,\theta)=\{g^i(\omega,\theta)\}_{i=1\ldots NL}=\{e^{-j\omega(k+\tau_n(\theta))}\}_{i=1\ldots NL}$ is the steering vector, i=1...NL, k=mod(i-1,L) and n=floor((i-1)/L).

3.2. Extension to 3-Dimensional Sound Source

The calculation in section 3.1 is for the sound source in the same plane with the array. In real applications, the sound can come from any direction in the 3-D space. We generalize the problem as shown in Figure 5. The sound is from the 3-dimentional (3-D) space, where Ψ is the elevate angle and θ is the azimuth. We have proved that when the sound is coming from the angle of (Ψ,θ), the delay between each microphone and the center of the array is similar to Table 2 but with an extra factor sin(Ψ) as shown in Table 3. When Ψ moves from 90^0 to 0^0, sin(Ψ) changes from 1 to 0, and as the result, the difference between each microphone gets smaller and smaller. When $\Psi=0^0$, there is no difference between microphones, which means the sound reaches each microphone at the same time. Taking into account that the sample delay between microphones can only be an integer, we determine the range where all microphones are identical. As shown in Figure 6, when $\Psi<\Phi$, four microphones receive identical signals for $0^0<\theta<360^0$. Our beamforming technique enhances sound from this range and suppresses sound from all other directions, treating it as background noise.

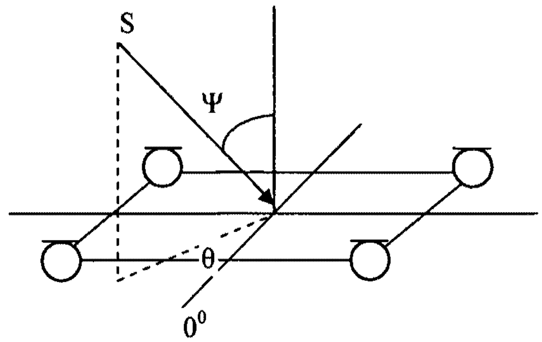


Figure 5. Illustration of 3-D sound source: The sound is from the direction (Ψ,θ), where Ψ is the elevation angle and θ is the azimuth.

Table 3 The delay between each microphone and the array center for sound from (Ψ,θ).

τ_1=-r•*fs*•sin(θ+Φ)sin(Ψ)/C	τ_2= -r•*fs*•sin(θ-Φ)sin(Ψ)/C	τ_3= r•*fs*•sin(θ+Φ)sin(Ψ)/C	τ_4= r•*fs*•sin(θ-Φ)sin(Ψ)/C

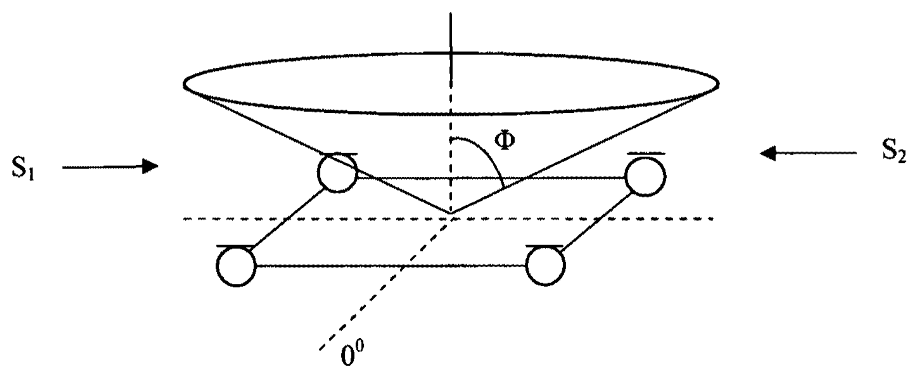


Figure 6. Illustration of the array working space: When sound comes from Ψ< Φ, four microphones receive same signals. Our beam-forming technique will enhance sound from this range and treating sound from other directions (for example S_1 and S_2) as background noise to suppress.

3.3. Least Mean Square Solution

Let the desired spatial directivity pattern be 1 in pass band and 0 in stop band. The least square cost function can be defined as

$$
\begin{aligned}
J(w) &= \int_{\Omega_p}\int_{\Theta_p} |H(\omega,\theta)-1|^2\, d\omega\, d\theta + \alpha\int_{\Omega_s}\int_{\Theta_s} |H(\omega,\theta)|^2\, d\omega\, d\theta \\
&= \int_{\Omega_p}\int_{\Theta_p} |H(\omega,\theta)|^2\, d\omega d\theta + \alpha\int_{\Omega_S}\int_{\Theta_X} |H(\omega,\theta)|^2\, d\omega d\theta - 2\int_{\Omega_p}\int_{\Theta_p} \mathrm{Re}(H(\omega,\theta))d\omega d\theta + \int_{\Omega_p}\int_{\Theta_p} 1\, d\omega d\theta
\end{aligned}
\tag{3}
$$

Replacing $|H(\omega,\theta)|^2 = w^T g(\omega,\theta) g^H(\omega,\theta) w = w^T G(\omega,\theta) w = w^T (G_R(\omega,\theta) + jG_I(\omega,\theta)) w = w^T G_R(\omega,\theta) w$ and $\mathrm{Re}(H(\omega,\theta)) = w^T g_R(\omega,\theta)$, we then have

$$J(\omega) = w^T Q w - 2w^T a + d \quad \text{, where}$$

$$Q = \int_{\Omega_P} \int_{\Theta_P} G_R(\omega,\theta) d\omega d\theta + \alpha \int_{\Omega_S} \int_{\Theta_S} G_R(\omega,\theta) d\omega d\theta$$
$$a = \int_{\Omega_P} \int_{\Theta_P} g_R(\omega,\theta) d\omega d\theta$$
$$d = \int_{\Omega_P} \int_{\Theta_P} 1 \; d\omega d\theta \qquad (4)$$

where $g_R(\omega,\theta) = \cos[\omega(k+\tau_n)]$ and $G_R(\omega,\theta) = \cos[\omega(k-l+\tau_n-\tau_m)]$.

When $\partial J / \partial w = 0$, the cost function J is minimized. The least-square estimate of w is obtained by

$$w = Q^{-1} a \qquad (5)$$

3.4. Linear Constrain

Applying linear constrains $Cw = b$, we can further constrain the spatial response to a predefined value b at angle θ_f using following equation:

$$\begin{bmatrix} \mathbf{g}_R^T(\omega_{start},\theta_f) \\ \dots \\ \mathbf{g}_R^T(\omega_{end},\theta_f) \end{bmatrix} w = \begin{bmatrix} b_{start} \\ \dots \\ b_{end} \end{bmatrix} \qquad (6)$$

Now, the design problem becomes

$$\min_w w^T Q w - 2w^T a + d \qquad \text{subject to} \quad Cw = b \qquad (7)$$

and the solution of the constrained minimization problem is equal to:

$$w = Q^{-1} C^T (C Q^{-1} C^T)^{-1} (b - C Q^{-1} a) + Q^{-1} a \qquad (8)$$

where w is the filter parameters for the designed beamformer.

3.5 Sound Source Localization

There are two categories of techniques to estimate the sound localization: one employs time difference of arrival (TDOA) and another is based on steered response power (SRP).

For an array with N microphones, a delayed, filtered and noise corrupted version of sound signals is presented in each of the microphone signals. The delay-and-sum beam former time aligns and sums all the microphone signal as $y(t,q) = \sum_{n=0}^{N} \mathrm{x_n}(\mathrm{t} + \Delta_\mathrm{n})$, where Δ is the steering delay appropriate for focusing the array to the direction of q. When the focus corresponds to the location of the sound source, the steered response power (SRP) should reach a global maximum.

Time difference of arrival can be used to estimate the sound source location. According to the sound propagation theory, the sound direction is uniquely determined by the time difference for a wave to propagate through non-linearly distributed distant microphones. Estimating the sound direction is essentially identical to estimate the TDOA, which is achieved by estimating the cross correlation.

Our preliminary research showed that TDOA-based localization is effective under low to moderate reverberation condition. The SRP approach requires shorter analysis intervals and exhibits an elevated insensitivity to environmental condition while not allowing for use under excessive multi-path. We implemented a new method called SRP-PHAT which combines the advantages of two approaches, and has a decreased sensitivity to noise and reverberations and more precise location estimates than the existing localization methods.

Figure 7 shows our experimental results. The upper plot is the value of SPR-PHAT at each angle. The minimum value corresponds to the sound location.

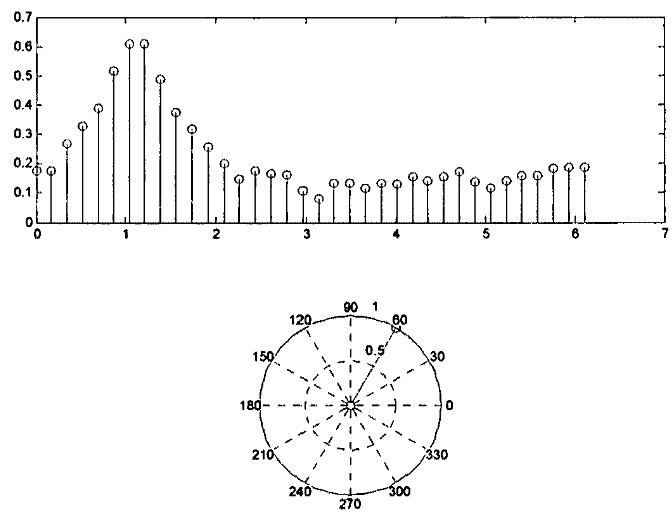


Figure 7. The upper image shows the value of SRP-PHAT for every 10°; the lower image represents the estimation and ground truth.

3.6 Adaptive Beamforming

Section 3.2 and 3.3 introduce the algorithm to derive the fixed beamforming to form the directivity pattern. We further extend it to adaptive beamforming. Adaptive beamforming can achieve better interference suppression than fixed beamforming. This is because the target direction of arrival, which is assumed to be stable in fixed beamforming, does change with the movement of the speaker. Also, the sensor gains, which are assumed uniform in fixed beamforming, exhibit significant distribution. All these factors will reduce speech quality. On the other hand, adaptive beamforming adaptively performs beam steering and null steering; therefore, the adaptive beamforming method is more robust against steering error caused by the array imperfection mentioned above.

The structure of our adaptive beamforming method is shown in Figure 8. It comprises of a fixed beamforming, a blocking matrix (BM) and a set of adaptive filters. The purpose of the blocking matrix is to block the target signal and let interfering noises through. The interfering noises are fed into an adaptive filter to minimize their influence in the output. One of the key steps in adaptive beamforming is to determine when the adaptation should be applied. Because of signal leakage, the output z of the blocking matrix may contain some weak speech signals. If the adaptation is active when speech is present, the speech will be cancelled out together with the noise; therefore, our invention uses a control module on the adaptation. This module enables adaptation according to the spectrum and energy of both noise signal and speech signal.

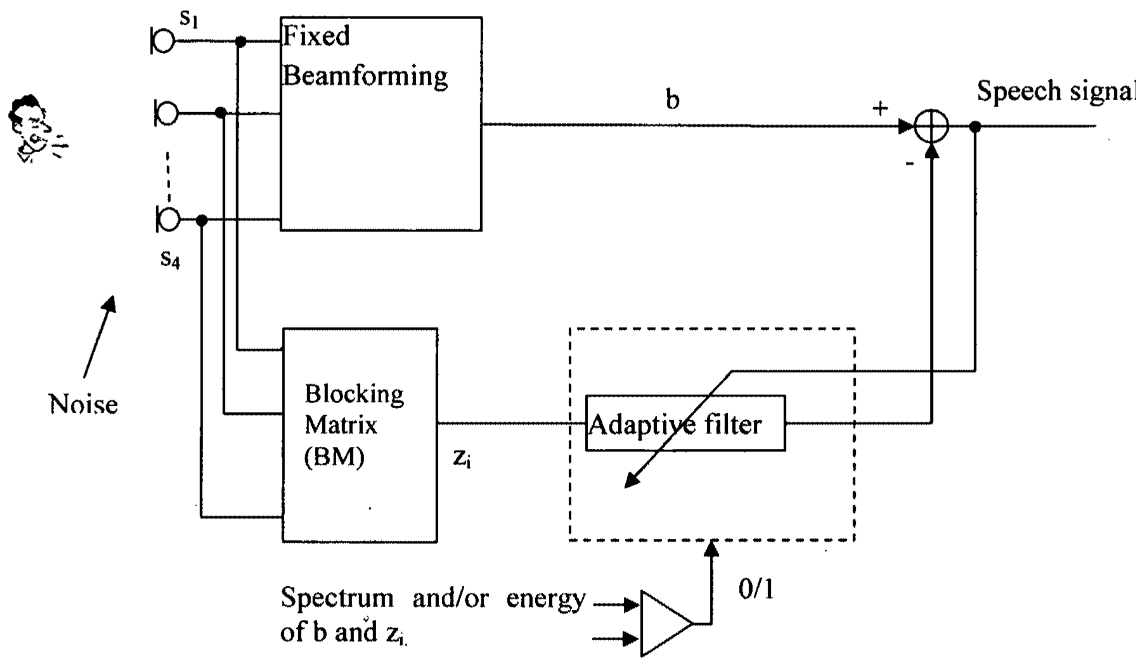


Figure 8. Diagram of adaptive beamforming: It consists of a fixed beamformer, the blocking matrix, and the adaptive filter. A control module is applied to enable/disable the adaptation process.

In Figure 8, the dotted block represents our adaptive filtering process. We developed a sub-band adaptive filtering for this invention for two reasons: firstly, it leads to a higher convergence speed than when using a full band adaptive filter. Secondly, our noise reduction algorithm is developed in sub-band, so applying sub-band adaptive filtering here provides the same framework for both beamforming and noise reduction, and saves on computational cost. Figure 9 shows the structure of our sub-band adaptive filtering. Both input signals are split into frequency sub-bands via an analysis filter bank. Each sub-band adaptive filter usually has a shorter impulse response than its full band counterpart. The step size can be adjusted individually for each sub-band, which leads to a higher convergence speed than when using a full band filter.

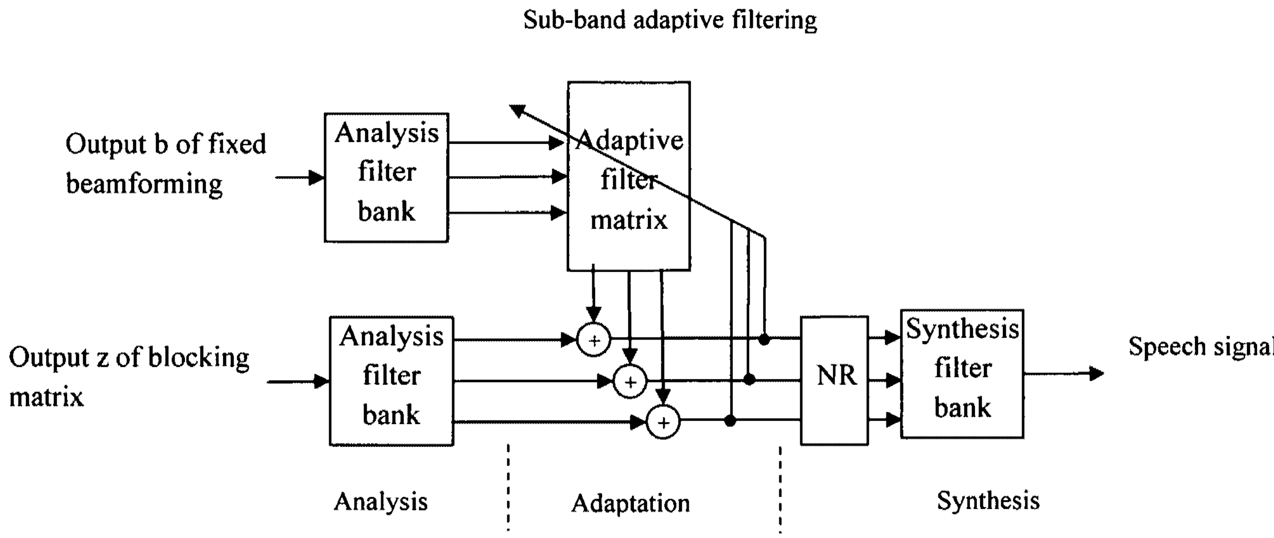


Figure 9. The structure of the sub-band adaptive filter: In the analysis step, both outputs of fixed beamforming and blocking matrix are split into sub-band through the analysis filter bank. In the adaptation step, the filter is adapted such that the output only contains speech signal. Finally, in the synthesis step, the sub-band speech signal is synthesized to full-band speech through the synthesis filter bank. Because noise reduction and beamforming are in the same sub-band framework, we applied noise reduction (NR) before synthesis to save computation. The NR module will be introduced in the next section.

To ensure the speech quality, the filter bank should not distort the sound signal by itself. We already implemented an efficient perfect-reconstruction filter bank, which can fully meet this requirement. In this implementation, all sub-band filters are factorized to operation on the prototype filter coefficients and a modulation matrix is used to take advantage of FFT. This modification ensures a minimum amount of multiply-accumulate operations. Figure 10 shows the performance of our filter bank. The blue line represents the input signal to the filter bank, and the red circle is the output of the filter band after analysis and synthesis. The output perfectly matches the input, called perfect-reconstruction filter bank.

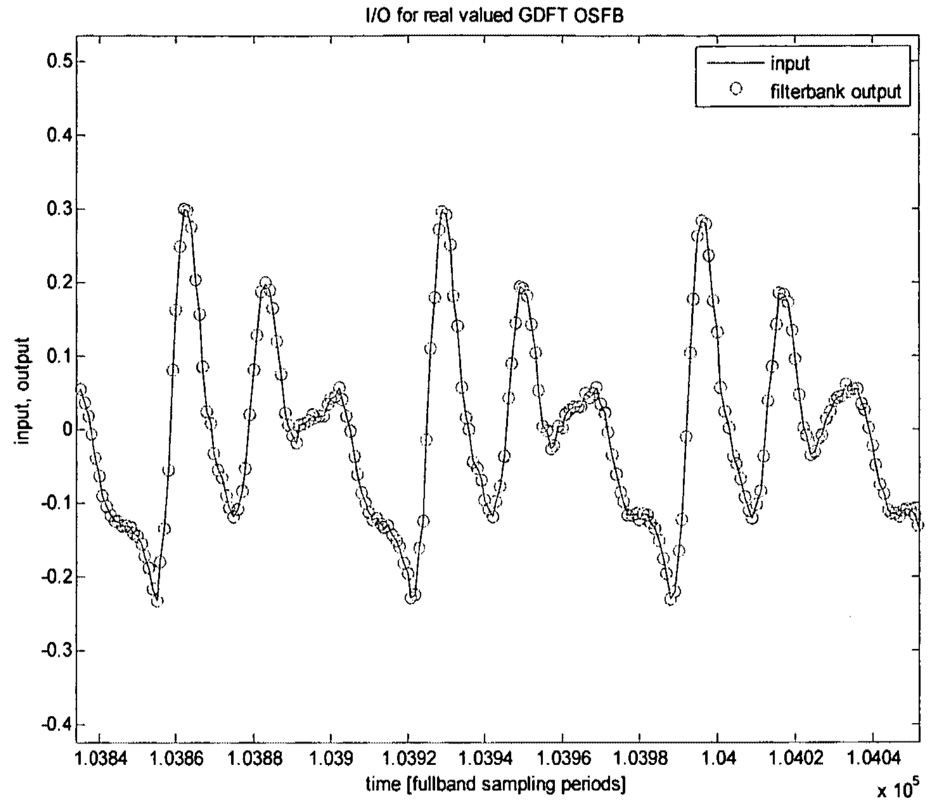


Figure 10. Perfect reconstruction filter bank input and output: The blue line represent input signal to filter bank and red circle is the output of the filter band after analysis and synthesis. The output perfectly matches the input.

The noise reduction (NR) module as shown in Figure 9 is used to further reduce background noise after adaptive beamforming. It explores the short-term and long-term statistics of speech and noise, and the wide-band and narrow-band signal-to-noise ratio (SNR) to support a Wiener gain filtering. After the spectrum of noisy-speech passes through the Wiener filter, an estimation of the clean-speech spectrum is generated. The filter bank synthesis module, as an inverse process of filter bank analysis module, reconstructs the signals of the clean speech given the estimated spectrum of the clean speech.

3.7 Noise reduction

The noise reduction module can include any kind of noise reduction algorithm, such as Wiener filter-based noise reduction, spectral subtraction noise reduction, auditory (or cochlear) transform-based noise reduction, or model-based noise reduction algorithm.

3.8 Hardware Implementation:

The structure of circuit design is shown in Figure 11. The acoustic signal is picked up by four or eight microphone components/elements arranged as a linear or circular array. First, the microphone amplifiers provide 20dB gain to boost the signal level to enhance the microphone sensitivity, and then the audio Codec provides an adjustable gain level from -74dB to 6dB before it converts the four channels of analog signals into digital signals. The pre-amplifier may not need for some applications. The Codec then transmits the digital audio signals to DSP (digital signal processing) chip for audio signal processing and computation. The DSP chip also transmits output signal to the Codec, and then the Codec converts it into analog signal, which is then amplified by speaker amplifier to drive the internal loudspeaker if it is needed.

The flash memory stores the code for the DSP chip and compressed audio signals. Once the system boots up, the DSP chip reads the code from flash memory into internal memory and starts to execute the code. During the start up stage, we can also configure the Codec by writing to the registers of the DSP chip. There are switch power regulators and linear power regulators to provide appropriate voltage and current supply for all the components on the board.

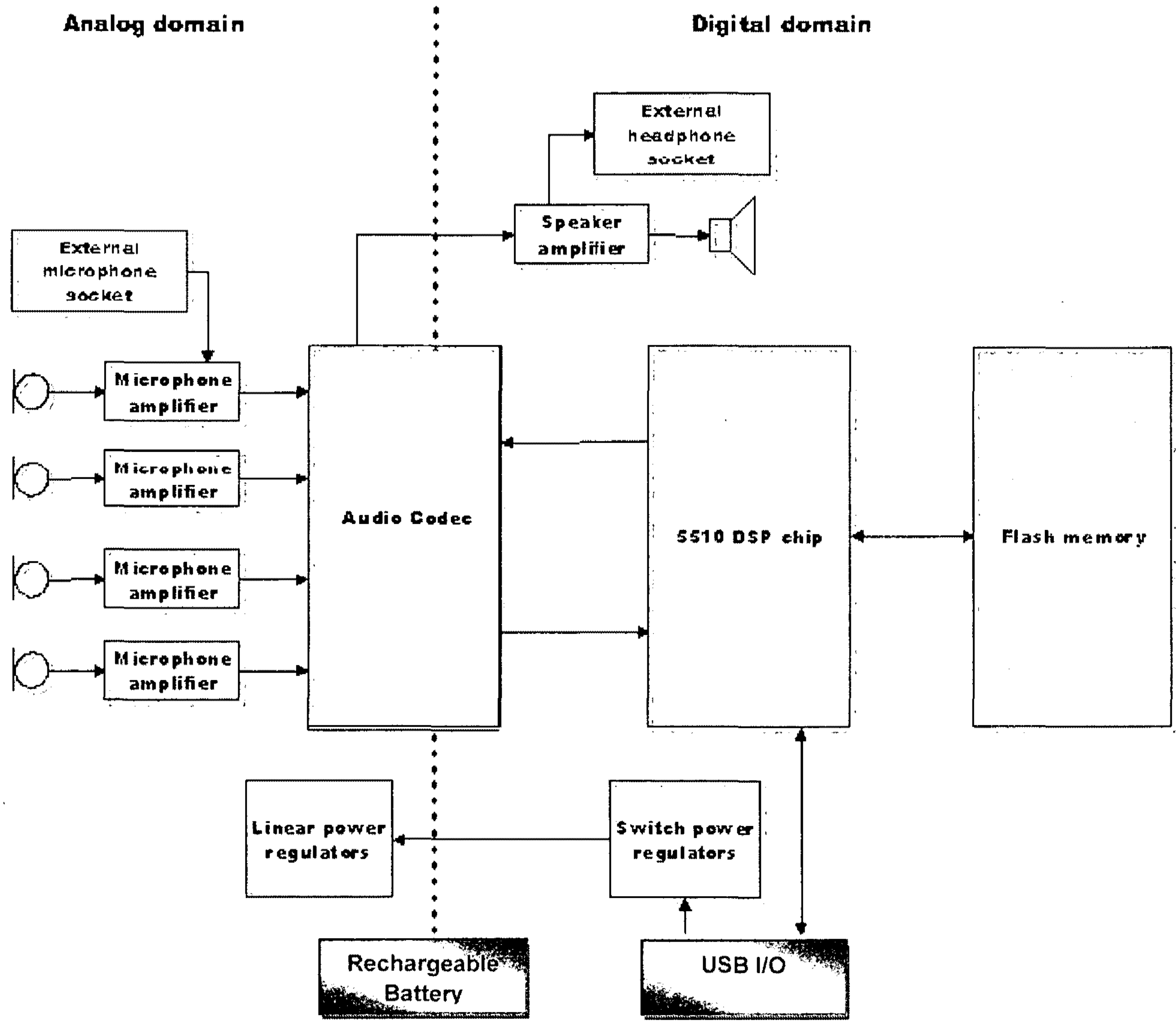


Figure 11. Hardware implementation of the invention: It consists of 3 major chips, codec, DSP, and flash memory. The USB control is built in the DSP chip. For 8-sensor microphone array, we can use two four-channel codec chips.

We will use a mixed signal circuit board (6-layer PCB). The board layout will be carefully partitioned to isolate the analog circuits from the digital circuits, because the noisy digital signal can easily contaminate the low voltage analog signal from the microphones. Although the speaker amplifier's input and output is an analog signal, it will be placed in the digital region because of its high power consumption and its switch amplifier nature. Only linear power regulators are deployed in the analog region due to their low noise property.

To ensure the quality, five power regulators are designed in the microphone array circuits. The switch power regulators can achieve efficiency to 95% of input power and have high output current capacity but their outputs are too noisy for analog circuits. The linear regulators' efficiency is determined by the ratio of the output voltage over the input voltage, which is lower than that of switch regulators in most of the cases. Our experiments showed that the regulator outputs are very stable, quiet, and suitable for the low power analog circuits.

Figure 12(A) is our 4-sensor microphone array product named CrispMic™. Figure 12(B) is the PCB design, which is very similar to the PCB of the proposed medical recorder. In the PCB, we selected two new chips from the state-of-the-art semiconductor technology. The DSP chip from TI is a low power consumption design devised especially for portable devices. The DSP chip is powerful enough to handle all the DSP computation in the proposed system. With these two chips, we can reduce both the cost and power consumption of the proposed system.

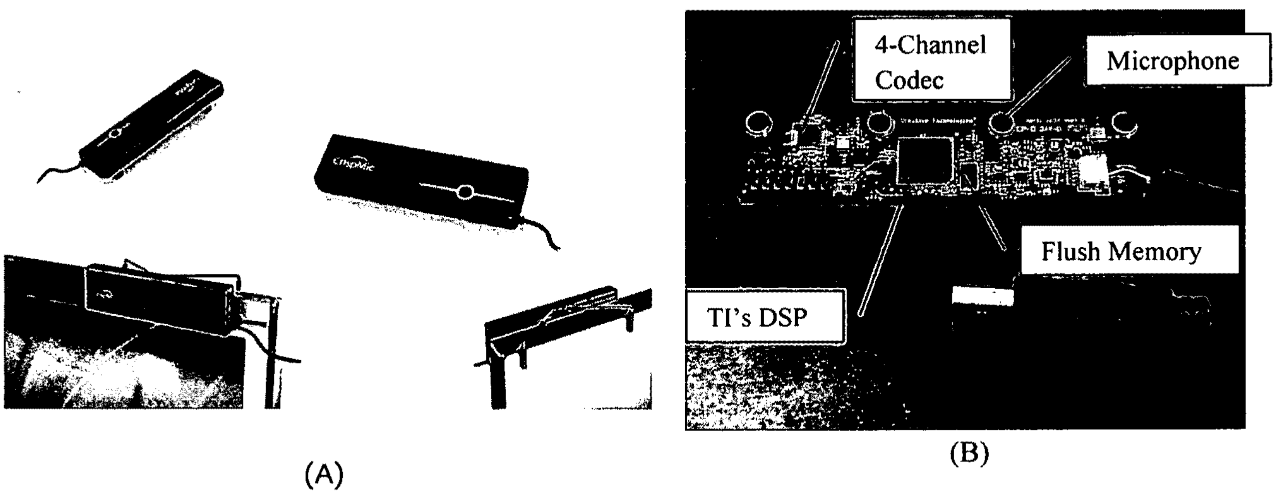


Figure 12. Microphone array and DSP hardware developed at LcT: (A) 4-sensor microphone array, 10 cm x 2.5 cm x 1.5 cm with USB interface. (B) Assembled PCB for the microphone array with TI's DSP, Codec, and memory chips. The same hardware can be used for the proposed system. Only software needs to be changed.

Figure 20 is a conference phone with 8-sensor microphone array. The microphone array on the top of the conference phone was designed based on the invented design method. The location of the microphone phones are shown in figure 4.

3.9 Software Implementation:

The beamformer design method can be implemented as software in the DSP chip using our invention. Furthermore, the software for noise reduction, echo cancellation, and USB interface can be developed using our invention.

4. Applications of Invention

In this section, we use our invention to design microphone array devices with broadband beamforming for a conference phone, a mobile phone, and a tablet computer.

4.1. Eight-Sensor Microphone Array for Conference Phone

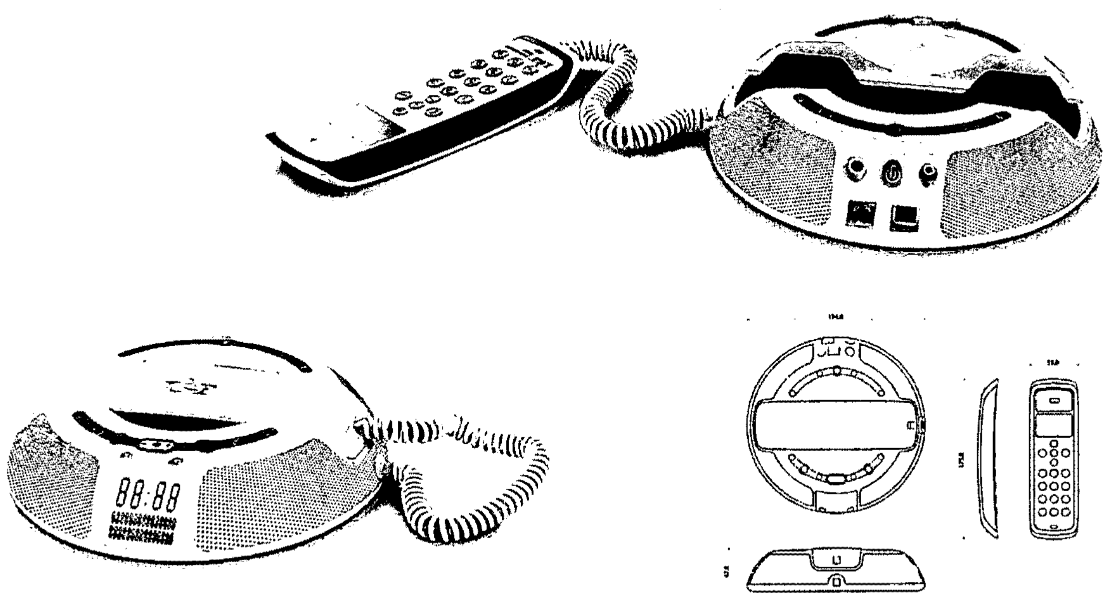


Figure 13. A conference phone with 8-snesor microphone array developed based on our invention.

Using our invention, we designed an 8-sensor circular microphone array for a conference phone (figure 13). The circular microphone array has a diameter of four inches. Eight microphones are distributed on the circle as shown in Figure 14. Microphones 4-7 are separated 90 degrees from each other, and microphones 0-3 are rotated counterclockwise 60 degrees from microphone 4-7 respectively.

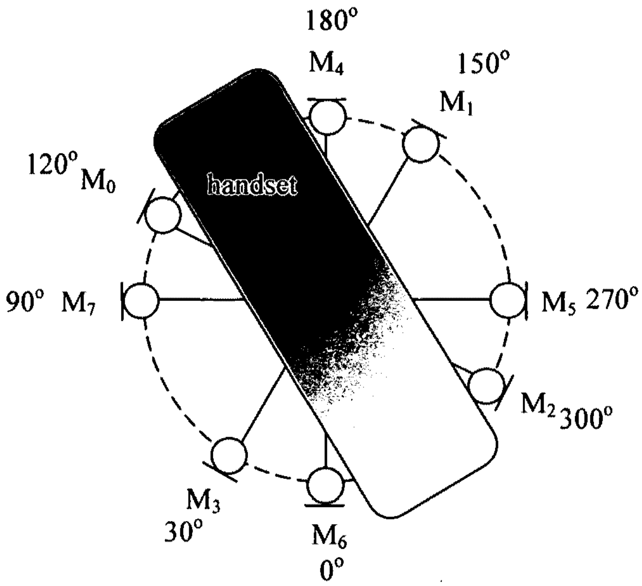


Figure 14. The layout of 8-sensor microphone array for a conference phone: Eight microphones are mounted on the surface of the conference phone. The conference phone has a removable handset on top, so our microphone array design has space for the handset.

Beamforming Parameters

The space is divided into eight regions with equal space centered at 15°, 60°, 105°, 150°, 195°, 240°, 285°, and 330° respectively as shown in Figure 15. The designed beamforming of our invention will automatically point to one of those regions according to where the sound comes from.

Computer simulation was conducted to verify the performance of our designed beamformer with the following parameters: the sampling frequency f_s = 16k, and FIR filter taper length L=20.

For the region centered at 15°

- Passband (Θ_p, Ω_p) = {300-5000Hz, -5°- 35°}, designed spatial directivity pattern is 1.
- Stopband (Θ_s, Ω_s) = {300~5000Hz, -180°~-15° + 45°~180°}, the designed spatial directivity pattern is 0.

For the region centered at 60°

- Passband (Θ_p, Ω_p) = {300-5000Hz, 40°- 80°}, designed spatial directivity pattern is 1.
- Stopband (Θ_s, Ω_s) = {300~5000Hz, -180°~30° + 90°~180°}, the designed spatial directivity pattern is 0.

The other six regions have similar parameters.

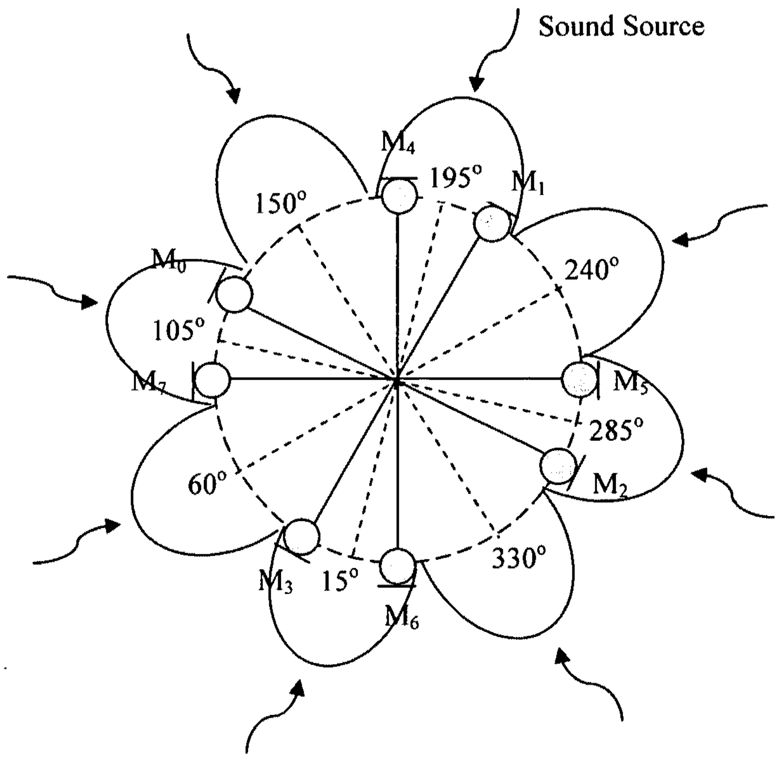
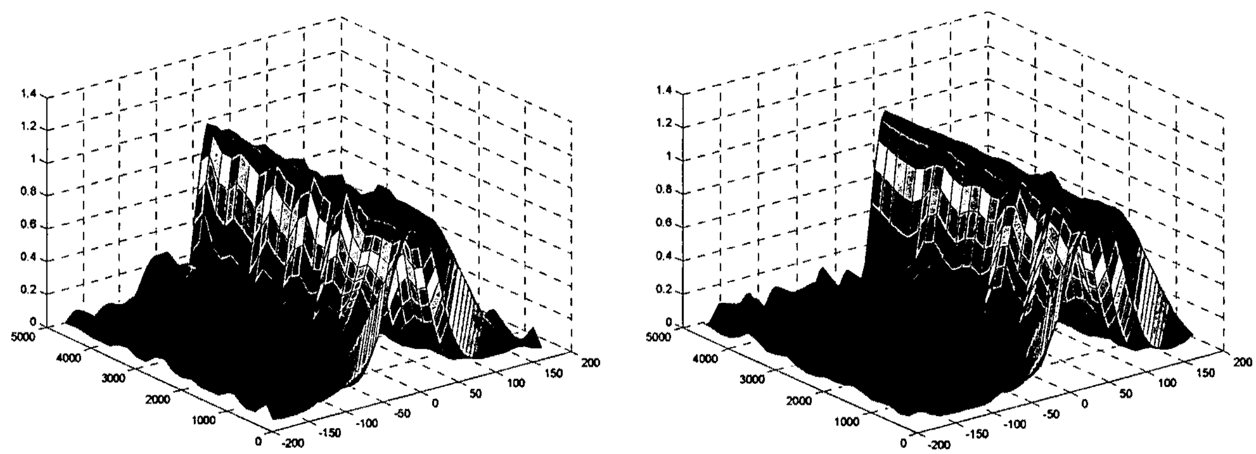


Figure 15. Illustration of the eight spatial regions that microphone array responds to.

Simulation Result



(a) (b)

Figure 16. Directivity pattern of designed microphone array for conference phone: the frequency bands of 300 Hz to 5 KHz of the sound from (a) 15° and (b) 60° are enhanced and the sound from other directions are reduced.

Our new design method calculates the filter coefficients for speech signals from each microphone and combines the filtered signals to enhance the speech from any specific direction. Because speech covers a large range of frequencies, our designed beamforming method covers broadband signals from 300Hz to 5000Hz. Figure 16 shows the directivity pattern of the designed microphone array conference phone pointing to 15° and 60°. In all frequencies, the main lobe has the same level, which means the speech signal has little distortion in frequency. Figure 17 is the averaged directivity pattern from 300Hz to

5000Hz. The main lobe is about 10dB higher than the side lobe; therefore the background sound from other directions will be highly suppressed compared to the sound in the desired pass direction.

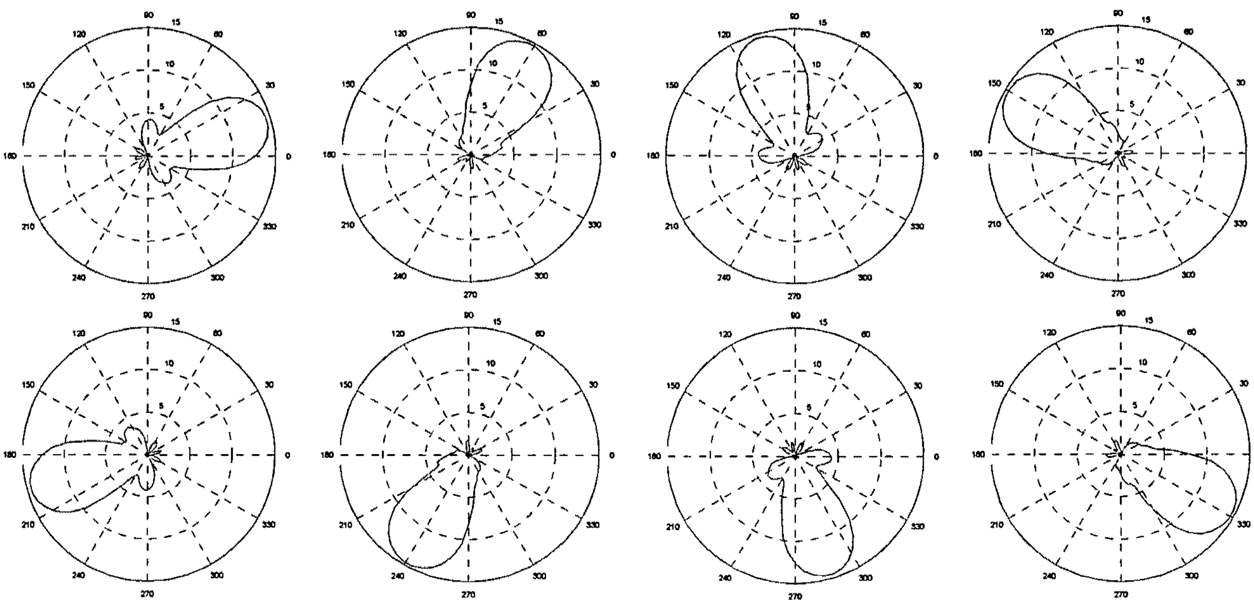


Figure 17. The directivity pattern for 8 directions, where each pattern is an average response from 300Hz to 5000Hz. Our invention can detect sound source direction and enhance the sound from that direction while suppressing the background noise from other directions

4.2. Four-Sensor Microphone Array for Wireless Phones and Handheld Devices

Using our invention, we designed a microphone array for a mobile phone (figure 18). In this design, four microphones are uniformly distributed around a circle with diameter equal to 2 inches. It is identical to putting four microphones on four corners of a square

.

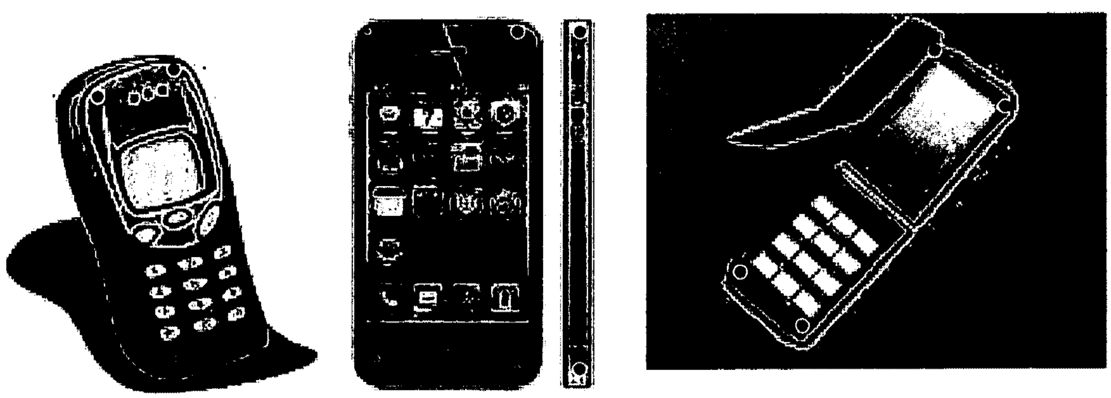


Figure 18. Invented 4-sensor microphone array for a mobile phone: the microphones can be mounted on the surface or on the edges.

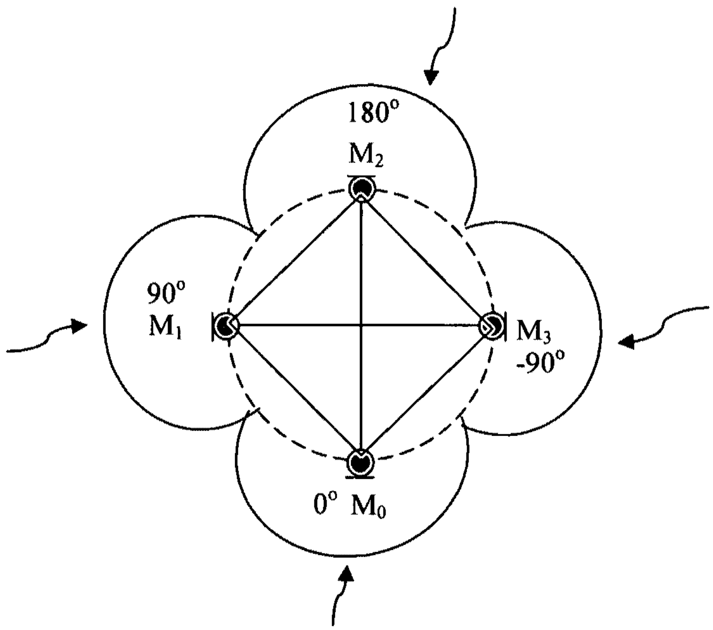


Figure 19. Illustration of the four regions that microphone array responds to.

Beamforming Parameters

The space is divided into four regions with equal space centered at -90°, 0°, 90° and 180°, respectively, as shown in Figure 19. The designed beamforming will automatically point to one of those regions according to where the sound comes from.

Computer simulation was conducted to verify the performance of our designed beamformer with the following parameters: the sampling frequency f_s = 16k, and FIR filter taper length L=20.

For region centered at 0°

- Passband (Θ_p, Ω_p) = {300-4000Hz, -20°- 20°}, designed spatial directivity pattern is 1.
- Stopband (Θ_s, Ω_s) = {300~4000Hz, -180°~-30° + 30°~180°}, the designed spatial directivity pattern is 0.

For region centered at 90°

- Passband (Θ_p, Ω_p) = {300-4000Hz, 70°- 110°}, designed spatial directivity pattern is 1.
- Stopband (Θ_s, Ω_s) = {300~4000Hz, -180°~60° + 120°~180°}, the designed spatial directivity pattern is 0.

Other two directions are similar

Simulation Result

Figure 20 shows the directivity pattern of this microphone array with respect to azimuth and frequency. The left column represents the 3-D display. It shows that the pass bands have the same height. The right column represents the 2-D display. It shows that pass bands have the same width along the frequency and demonstrates the broadband properties of the microphone array.

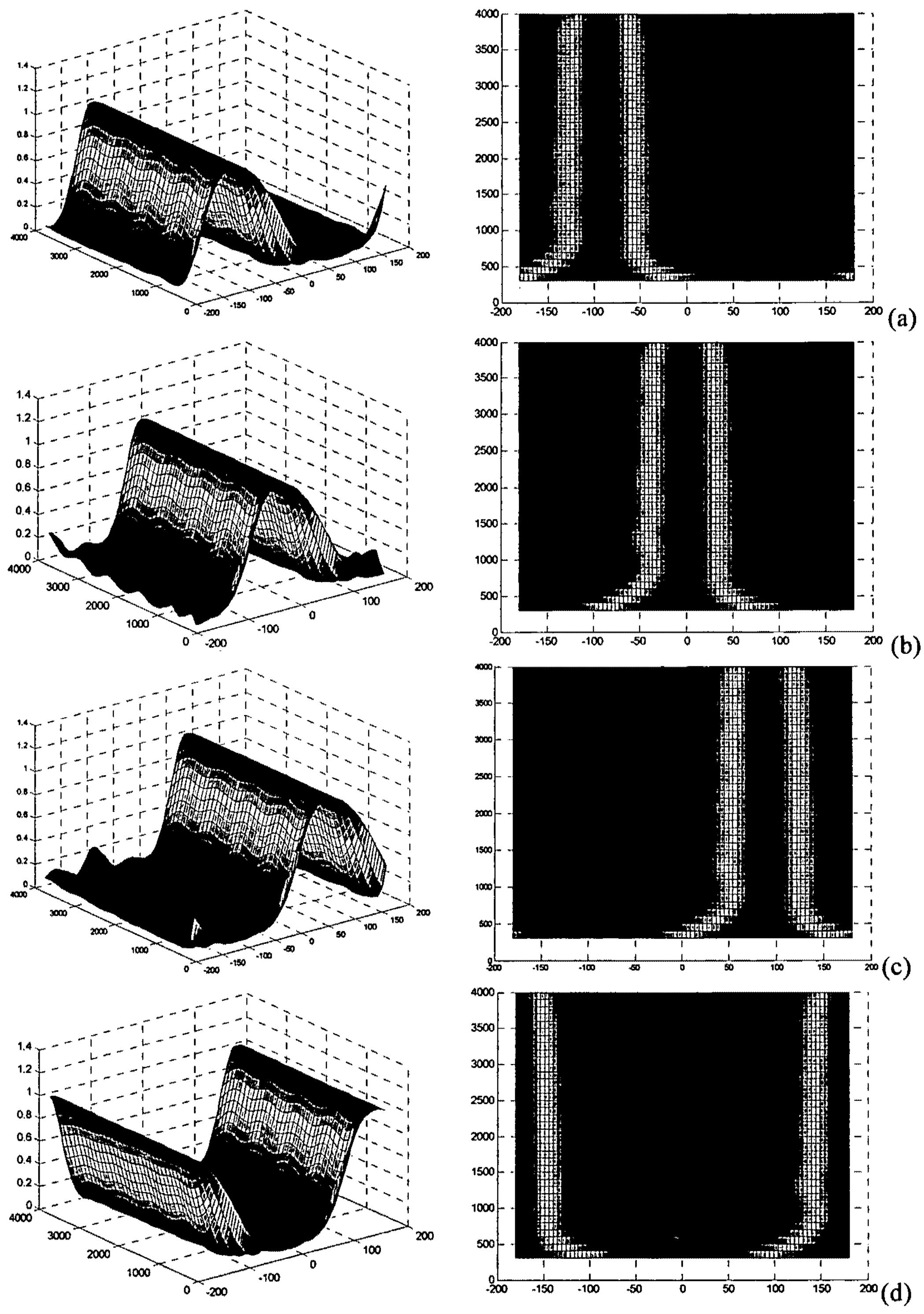


Figure 20. Show the directivity pattern of 4-sensor microphone array with respect to azimuth and frequency. The left column represents the 3-D display. It shows that the pass bands have the same height. The right column represents the 2-D display. It shows that pass bands have the same width along the frequency and demonstrates the broadband properties of the microphone array.

We note that our design method can also take care of the case where 4 microphones are in the corners of a rectangular shape; however, in theory, the 4 microphone in a square shape can give a better performance that the rectangular one.

4.3 Four-Sensor Microphone Array for Tablet Computer

The four microphones of the array are located on the frame of a tablet computer. Geometrically, they are distributed on the circle as shown in Figure 21. The radius of the circle is equal to the width of the iPad or similar product. The angle θ between Mic 2 and Mic3 is determined to avoid spatial aliasing up to 4000Hz.

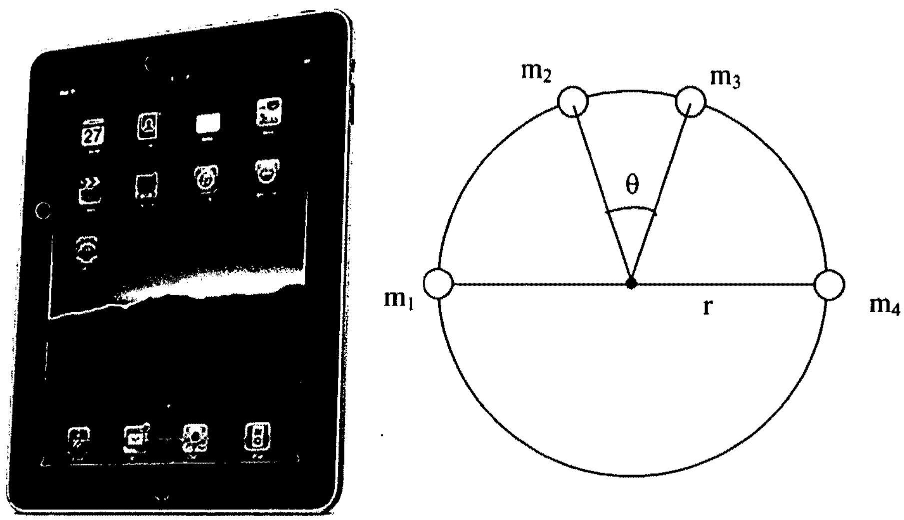


Figure 21. The microphone layout for a tablet computer. (Photo adapted from apple.com.)

The objective of this array design is to enhance the front talker's voice and suppress background noise. We form an acoustic beam pointing front. The sound within the range of $\Phi=30°$ will be enhanced compared to the sound from other directions.

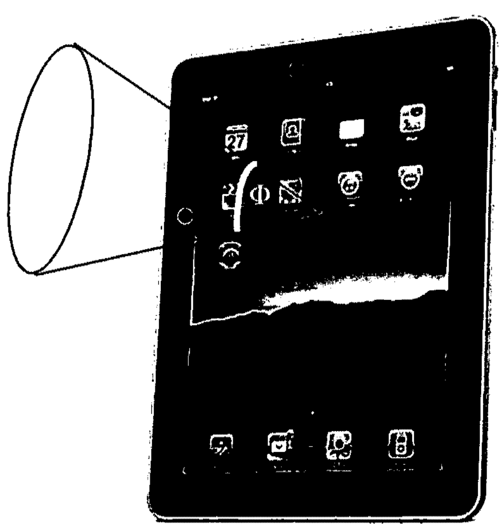
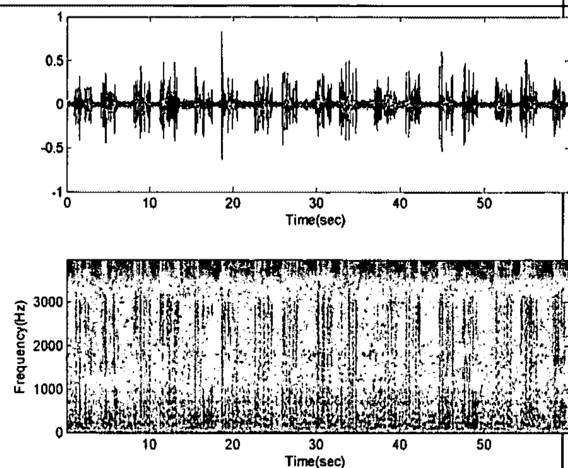
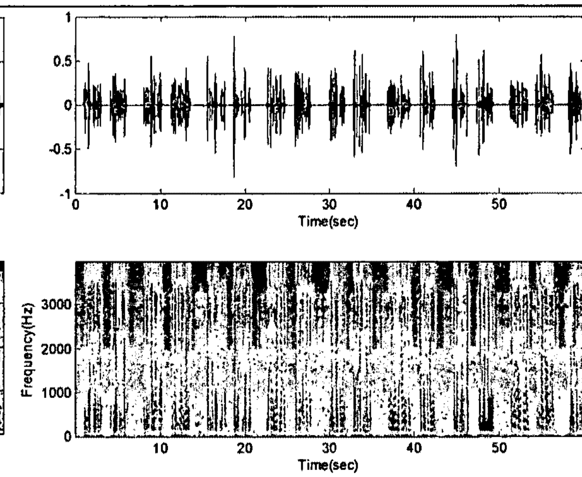
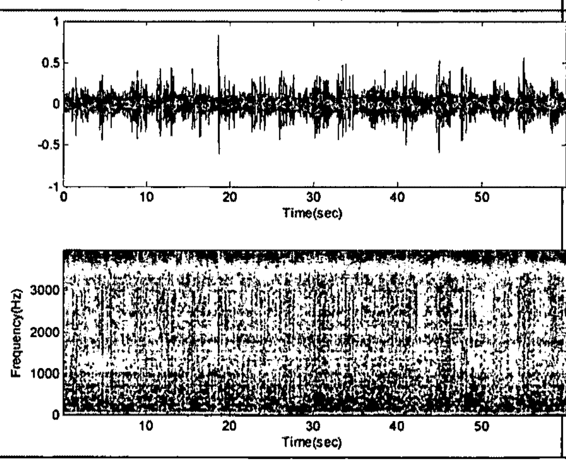
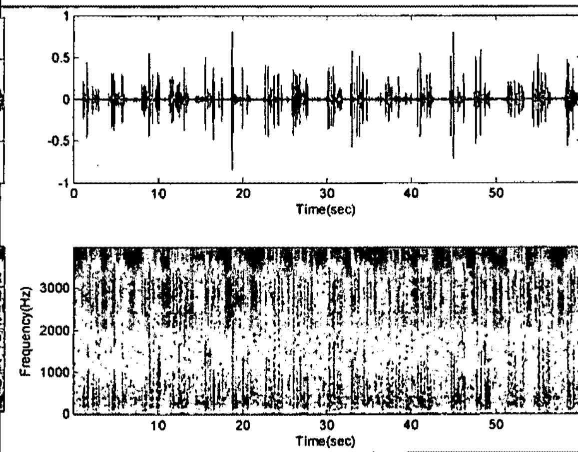


Figure 22. The acoustic beam formed on the tablet computer's microphone array. (Photo adapted from apple.com.)

One speaker is talking in front of the table computer with background noise on the side. We apply our adaptive beamforming on the array. Table 4 shows the processing result. The left column is the noisy signal (shown both in time domain and in spectral domain), i.e. the single channel input from the microphone array. The right column is the signal after beamforming and noise reduction. It is clear that LcT's design method suppresses background noise while keeping the speech clarity.

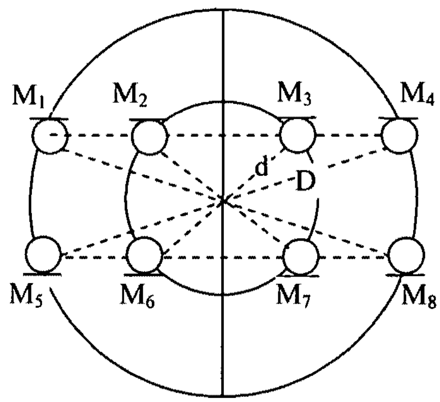
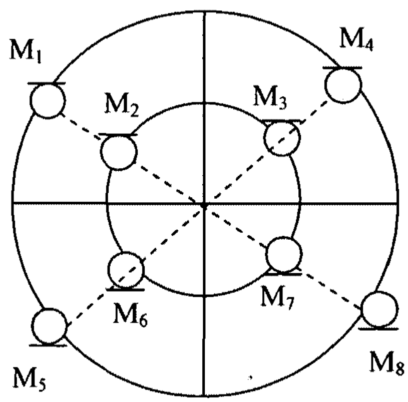
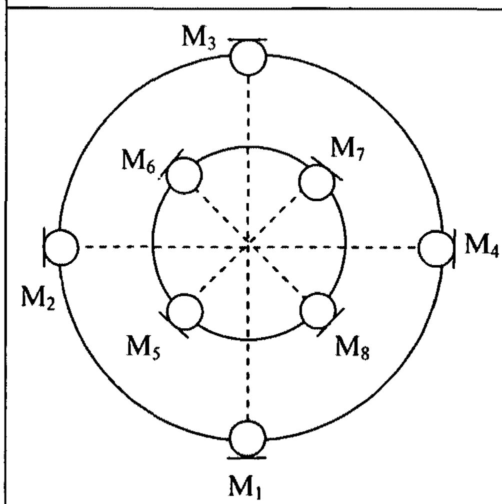
Table 4. Adaptive beamforming and noise reduction performance on microphone array for tablet computer.

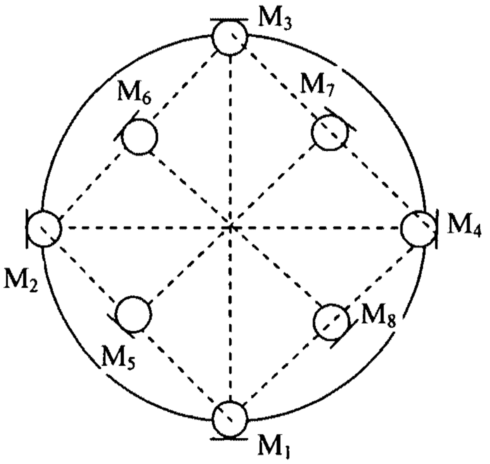
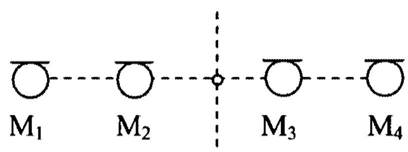
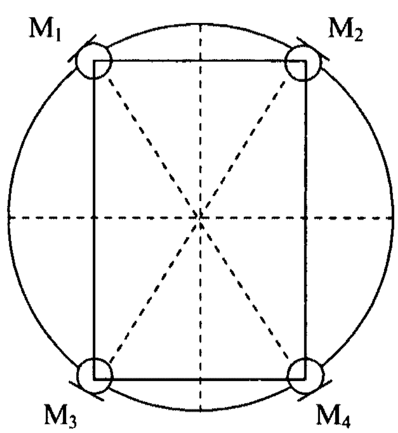
	Before Beamforming and NR	After Beamforming and NR
(A) SNR = 15dB		
(B) SNR = 0dB		

4.4. Other Microphone Configurations

Our design method is for the beamformer with an arbitrary directivity pattern for arbitrarily distributed microphones. For any specific microphone array layout, the only thing we need to define to achieve the beamformer coefficients is the value of τ_n for each microphone. To do that, we first define the center or reference point of the array, then measure the distance d_n between each microphone and the reference point, and the angle Φ_n of each microphone biased from vertical axis. In Table 5, we list several microphone layout and their corresponding value of τ_n.

Table 5. Example of microphone array layout and corresponding value of τ_n.

 (a)	$\tau_1=D*\cos(\theta+\Phi_1)$
	$\tau_2=d*\cos(\theta+\Phi_2)$
	$\tau_3=d*\cos(\theta-\Phi_3)$
	$\tau_4=D*\cos(\theta-\Phi_4)$
	$\tau_5=-D*\cos(\theta-\Phi_5)$
	$\tau_6=-d*\cos(\theta-\Phi_6)$
	$\tau_7=-d*\cos(\theta+\Phi_7)$
	$\tau_8=-D*\cos(\theta+\Phi_8)$
 (b)	$\tau_1=D*\cos(\theta+\Phi_{12})$
	$\tau_2=d*\cos(\theta+\Phi_{12})$
	$\tau_3=d*\cos(\theta-\Phi_{34})$
	$\tau_4=D*\cos(\theta-\Phi_{34})$
	$\tau_5=-D*\cos(\theta-\Phi_{56})$
	$\tau_6=-d*\cos(\theta-\Phi_{56})$
	$\tau_7=-d*\cos(\theta+\Phi_{78})$
	$\tau_8=-D*\cos(\theta+\Phi_{78})$
	$\tau_1=-D*\cos(\theta)$
	$\tau_2=-D*\sin(\theta)$
	$\tau_3=D*\cos(\theta)$
	$\tau_4=D*\sin(\theta)$
	$\tau_5=-d*\cos(\theta-\Phi_5)$
	$\tau_6=d*\cos(\theta+\Phi_6)$
	$\tau_7=d*\cos(\theta-\Phi_7)$
	$\tau_8=-d*\cos(\theta+\Phi_8)$

(c)	
 (d)	$\tau_1=-D*\cos(\theta)$
	$\tau_2=-D*\sin(\theta)$
	$\tau_3=D*\cos(\theta)$
	$\tau_4=D*\sin(\theta)$
	$\tau_5=-d*\cos(\theta-\Phi_5)$
	$\tau_6=d*\cos(\theta+\Phi_6)$
	$\tau_7=d*\cos(\theta-\Phi_7)$
	$\tau_8=-d*\cos(\theta+\Phi_8)$
 (e)	$\tau_1=-d1*\sin(\theta)$
	$\tau_2=-d2*\sin(\theta)$
	$\tau_3=d3*\cos(\theta)$
	$\tau_4=d4*\cos(\theta)$
 (f)	$\tau_1=d*\cos(\theta+ \Phi_1)$
	$\tau_2=d*\cos(\theta- \Phi_2)$
	$\tau_3=-d*\cos(\theta- \Phi_3)$
	$\tau_4=-d*\cos(\theta+ \Phi_4)$

We have proposed to apply the layout (f) on the handheld device for hands-free speech acquisition. Some application are shown in Figure 18 and Figure 23. In a hands-free and non-close talking scenario, people prefer to talk from a distance rather than having to speak close to the microphone, and people may want to talk while watching the video screen on their handheld device. In this case, if one uses a single omni-directional microphone, the speech and interfering noise from other directions are picked up equally; consequently, the signal-to-noise ratio is much lower than the close talking scenario. By utilizing our beamforming technique, the device can pick up sounds from the direction of the speaker's mouth and suppress noise from other directions. As a result, the speech quality is enhanced. The invented design method can be applied to design microphone arrays for any kind of handheld devices and any kind of microphone configurations.

Figure 23. Invented 4-sensor microphone array for a voice recorder: the microphones can be mounted on the surface or on the edges.

Reference

[1] J. Benesty, J. Chen and Y. Huang, Microphone Array Signal Processing, Springer, 2008.
[2] M. Brandstein and D. Ward, Microphone Arrays, Springer, 2001.
[3] S. Doclo and M. Moonen, Design of far-field and near-field broadband beamformers using eigenfilters, *Signal Processing*, Vol. 83, pp.2641-2673, 2003.

PATENT APPLICATION SERIAL NO.________________________

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FEE RECORD SHEET

09/27/2010 SZEWDIE1 00000028 61403952

01 FC:2005 110.00 OP

PTO-1556
(5/87)

*U.S. Government Printing Office: 2002- 489-267/69033

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
61/403,952	09/24/2010		110			

CONFIRMATION NO. 5940

FILING RECEIPT

OC000000044043346

Qi (Peter) Li
Li Creative Technologies, Inc.
Suite 140
25B Hanover Road
Florham Park, NJ 07932

Date Mailed: 10/20/2010

Receipt is acknowledged of this provisional patent application. It will not be examined for patentability and will become abandoned not later than twelve months after its filing date. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Applicant(s)

Manli Zhu, Pearl River, NY;
Qi Li, New Providence, NJ;

Power of Attorney: None

If Required, Foreign Filing License Granted: 10/18/2010

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 61/403,952**

Projected Publication Date: None, application is not eligible for pre-grant publication

Non-Publication Request: No

Early Publication Request: No

**** SMALL ENTITY ****

Title

Microphone array design and implementation for telecommunications and handheld devices

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4158).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and

Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).