24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

SPATIAL DELTA-SIGMA MODULATION FOR DIRECTIVITY CONTROL OF AN ACOUSTIC PIXEL ARRAY USING CNT

Dong-Hwan Kim and Jung-Woo Choi

School of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea
email: dh.kim92@kaist.ac.kr and jwoo@kaist.ac.kr

Abstract

Conventional sound field reproduction technique requires extensive multi-channel signal processing to control the beampattern of a speaker array. In this paper, we introduce a simplified directivity control strategy, in which a designated beampattern is synthesized through the control of individual polarity of very small speaker units. Each small speaker operates like a pixel of a display panel, and the combination of multiple pixels using a spatial binary pattern produces the designated beampattern. A Spatial Delta-Sigma Modulation (DSM) technique is used to convert the continuous spatial excitation pattern into a binary pattern. For the speaker unit, thermoacoustic speakers made up of vertically aligned carbon nanotubes (VACNT) is selected for its sound radiation characteristic less influenced by radiating area. The experiment with a prototype pixel display shows that the directivity control is possible with negligible errors.

Keywords: thermo-acoustic, carbon nanotube, beampattern, single channel array, small speaker

1. Introduction

Application of binary excitation patterns for reducing the computational complexity in array signal processing has been attempted in various ways. Takeoka[1] proposed a 1-bit array signal processing scheme based on temporal delta-sigma modulation (DSM). As well known, DSM can convert a continuous signal into an equivalent 1-bit binary signal by exploiting the oversampling and noise shaping properties[2,3]. In the scheme proposed in [1], multichannel signals are produced by applying different delays to a single channel DSM signal, so it is advantageous for steering a beampattern. Beyond the temporal DSM structure, more general space-time coupled DSM structures have been proposed by Tamura[4] and Scholnik[5], for simultaneous conversion in both space and time domain. The space-time coupled DSM also generates 1-bit multichannel signals, but the combination of spatal and temporal DSM enables efficient conversion and generation of multichannel signals at a reduced sampling rate.

In this paper, we attempt to realize an array structure that controls directivity simply by turning on/off the array elements. The on/off switching of array elements is equivalent to applying a spatial binary excitation pattern to the array. Spatial DSM is employed for synthesizing a spatial binary pattern, and all array elements receiving the same channel time signal are turned on/off according to the applied spatial binary pattern. However, the spatial DSM requires the oversampled transducer distribution in space, which implies that the size of each array element and inter-element spacing should be small enough. This is impractical for conventional dynamic loudspeakers whose sound radiation efficiency is proportional to the area of the transducer. In this work, we tackle this problem using thermo-acoustic (TA) transducers that are less influenced by the radiating area. Since the proposed array is constructed from a group of small TA transducers, we denote the array as an acoustic pixel display.

2. Solution methods

2.1 Spatial DSM

DSM[2,3] can convert a band-limited analogue signal into a binary signal through the 1-bit quantization. The quantization noise inevitably produced during the 1-bit quantization process is suppressed by utilizing oversampling and noise shaping. Usually, the DSM handles time-domain signals, but in this work, a spatial DSM structure is considered to produce a spatial binary pattern that can synthesize a desired beampattern in space. In the following sections, the basic process of the temporal DSM is reviewed and differences between the temporal and spatial DSM are explained.

2.1.1 Basic operation of temporal DSM

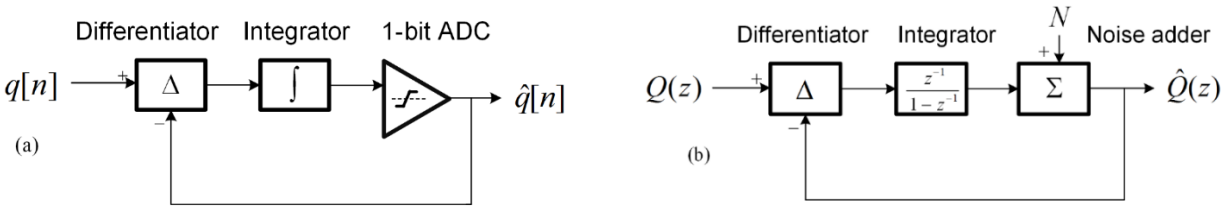

Figure 1: (a) Block diagram of DSM in the time domain. (b) Block diagram of DSM in z-domain.

The temporal DSM structure is characterized by a feedback loop combined with a low order quantizer, as shown in Figure 1(a). A discrete-time signal $q[n]$ represents an input signal $q(t)$ sampled with the sampling frequency F_s and bounded between -1 and $+1$. The output signal $\hat{q}[n]$ is a bipolar signal obtained by quantizing the input $q[n]$ using the 1-bit quantizer. Because the 1-bit quantizer is a nonlinear block, a quantization noise exists in $\hat{q}[n]$. In general, the quantization noise can have various spectral distributions, but for simple analysis it is often assumed as a white noise[6].

The relation between $q[n]$ and $\hat{q}[n]$ can be clearly seen in the z-domain, where the 1-bit quantizer is expressed as a white noise adder. From the block diagram of Figure 1(b), the quantized output $\hat{Q}(z)$ can be related to the input $Q(z)$ and quantization noise $N(z)$ as

$$\hat{Q}(z) = z^{-1}Q(z) + (1 - z^{-1})N(z), \tag{1}$$

where only the quantization noise is high-pass filtered. This noise shaping structure suppresses the quantization noise in the low frequency region of interest without altering the input signal. When an analogue low-pass filter is applied to the binary signal, the high-pass filtered noise is removed, and hence, the original signal can be recovered. However, the noise shaping alone is insufficient to suppress the quantization noise. For this reason, an oversampling is also applied to lower the power spectral density of the noise signal. The resultant SNR in the frequency band of interest is given by

$$SNR = -3.4 + 6.02B + 30\log_{10} OSR, \tag{2}$$

where B is the number of bits used for quantizer, and the OSR denotes an oversampling ratio[7].

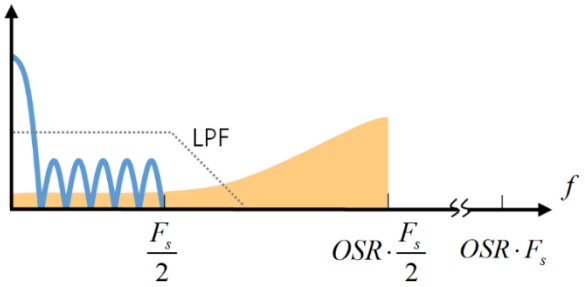

Figure 2: Signal spectrum after DSM (F_s : sampling frequency in Hz).

2.1.2 Differences between temporal and spatial DSM

The basic principle of temporal DSM described above can be adopted to a spatial DSM as well, simply by changing $q(t)$ to a space function $q(x)$. Suppose that a continuous line array is aligned along the x axis, and that the array is driven with the continuous excitation function $q(x)$. The wavenumber spectrum of the excitation function $q(x)$, given by

$$Q(k_x) = \int_{-\infty}^{\infty} q(x) e^{ik_x x} dx, \tag{3}$$

is related to a beampattern $b(\theta) = Q(k \sin\theta)$ at the wave number $k_x = k \sin\theta = 2\pi f \sin\theta / \nu$ (ν : speed of sound, θ: angle from the broadside direction). Since the beampattern is produced by the portion of wavenumber spectrum in the limited range $|k_x| \leq k$, we assume that $q(x)$ is bandlimited in the wavenumber domain. The continuous function is spatially sampled to achieve $q[n] = q(n\Delta x)$, which means that the discrete arrangement of array elements are driven by $q[n]$.

The primary objective of the spatial DSM is to convert an excitation function $q[n]$ into a binary signal, such that the binary signal can produce a desired beampattern. Unlike temporal DSM, however, there are two differences in the spatial DSM. First, the analogue low-pass filter applied for suppressing the high frequency quantization noise cannot be implemented as a circuit in the spatial DSM. Instead, the radiation circle(e.g. [8]), which suppress the propagation of high wavenumber components in the region $k_x > k$, plays a role of the natural low-pass filter.

In addition, our objective is to control the beampattern just by turning on and off each loudspeaker element. For this purpose, the final output should be in form of a unipolar signal, so the bipolar output from the DSM should be biased. The bias results in the increase of a zero wavenumber component, which can distort the generated beampattern. In this regard, we apply a negative spatial bias to the input signal $q[n]$, and then, apply a positive spatial bias to the output signal $\hat{q}[n]$ to maintain the original DC component in the unipolar spatial excitation pattern.

In implementing the spatial DSM, the spatial oversampling requires small inter-element spacing, so the size of loudspeakers should be small. Usually, small radiation area of a loudspeaker leads to the reduction in sound pressure level (SPL), so it is not feasible to realize spatial DSM with conventional loudspeakers. To circumvent this problem, we employ TA transducers, of which SPLs are less influenced by the radiating area.

2.2 TA transducers

TA loudspeakers generate sound via heat exchange with surrounding medium. In the 1910s, Arnold and Crandall[9] solved the physical mechanism of the thermo-acoustic transducers. When an AC signal is applied to a conductor, a thermal wave is produced and propagates through the thin air layer surrounding the conductor. The net effect of the thermal wave yields expansion and contraction of the medium, which leads to the change of volume velocity converted into the sound.

Among many types of conducting material, we chose CNT as thermo-acoustic material because of its extremely low heat capacity per unit area (c_h) and ease of handling[10-13]. In the recent models[10,12], the RMS sound pressure p_{rms} under the frequency of 100kHz is related to the input power W_{in}, distance r and frequency f as

$$p_{rms} = W_{in} \frac{f}{2\sqrt{2} T_0 \cdot c_p} \frac{1}{r}, \tag{4}$$

where T_0 is ambient temperature and c_p is heat capacity per unit mass of ambient gas. This model implies that the volume velocity and resultant pressure are overly dependent on the input power, not by the radiating area. This is a strong advantage for the implementation of spatial DSM, because the small transducers for realizing the spatial oversampling can produce the sound, as long as the heat concentration on the transducer is not too high.

Albeit of this advantage, TA transducers have several downsides, such as nonlinearity ($p_{rms} \propto W_{in}$) and low SPL at low frequencies ($p_{rms} \propto f$). Accordingly, the sound pressure generated from a TA loudspeaker includes high harmonic distortion, which is suppressed by the biased temporal DSM(e.g., [14]) in this work. In addition, due to its low efficiency in a low frequency region, TA transducers implemented in this work are only suitable for tweeters.

3. Acoustic pixel design

3.1 Array design

The arrays are designed to cover a high frequency range around 5 ~ 8 kHz, considering the high-pass characteristic of a TA transducer. With wavelength λ, the half-power beamwidth θ_{hpbw} of a uniformly-driven sound beam is roughly given by

$$\theta_{hpbw} \approx 0.886 \cdot \lambda / L \ \text{(degrees)} \tag{5}$$

where L is the aperture length. The maximum θ_{hpbw} was set to $20°$ such that $L = 174$ mm. To avoid the spatial aliasing, the inter-element spacing before applying DSM should satisfy the following condition at the highest operating frequency [15]:

$$\Delta x < \frac{\lambda}{|1+\sin\theta_{beam}|}, \tag{6}$$

where θ_{beam} is a beam steering angle, measured from the broadside direction. In the case of an end-fire beampattern, Δx should be smaller than the half of the wavelength at 8 kHz, so the inter-element spacing was determined as $\Delta x = 20.3$ mm. With this spacing, 8 array elements are required to sample the array aperture.

3.2 Beampattern simulation

As an exemplary target beampattern, we employed a band-limited Dolph-Chebyshev window of -20dB sidelobe level. The target excitation function is shown in Figure 3(a) as a grey line, which is a continuous function. The corresponding wavenumber spectrum is shown in Figure 3(b), where $|k_x / k|$ is the normalized wavenumber and k is the wavenumber corresponding to 8.44 kHz.

The orange-coloured unipolar pattern of Figure 3(a) represents the output of the spatial DSM. From simulations, OSR of 8.5 times showed the quantization noise less than the sidelobe level. It can be seen that the quantization noise only influences the evanescent region (i.e., outside the radiation circle), in which the sound wave does not propagate to the far-field.

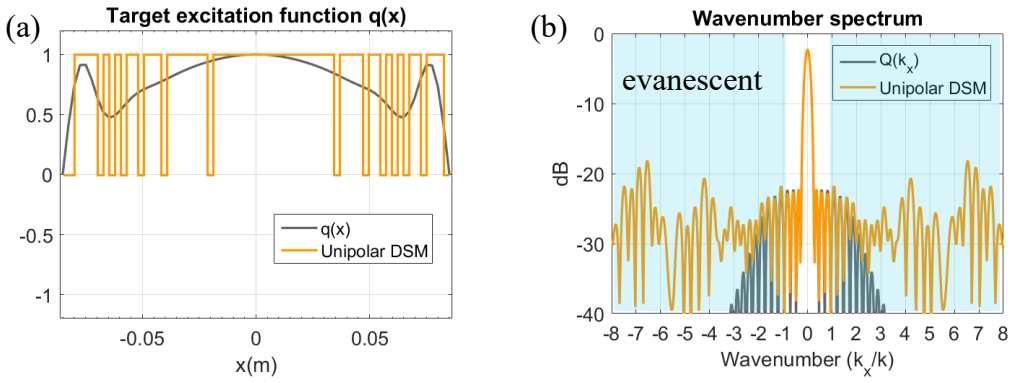

Figure 3: (a) Excitation functions before and after the spatial DSM.
(b) Wavenumber spectra of the target and DSM excitation functions.

4. Experiments

4.1 Fabrication of an acoustic pixel array

With *OSR* = 8.5, the oversampled inter-element spacing $\Delta x_{os} = \Delta x / OSR$ becomes 2.54 mm, and the total number of acoustic pixels is given by 68. A single CNT pixel was fabricated by drawing Vertically Aligned CNTs (VACNTs) over a 2-pin jumper header of 2.54 mm pitch (Figure 4(a); small sample on the left). The size of the fabricated sample is 2.54 mm × 3.00 mm. The pixel array was constructed by arranging 68 pixels on a multi-pin socket (Figure 5(a)). Each pixel was wired to a mechanical switch to turn on/off the pixel according to the unipolar pattern obtained the spatial DSM.

Due to manufacturing variances, resistances of pixels were not perfectly identical. The mean resistance and standard deviation of 68 pixels are $261 \pm 6\Omega$. All pixels were electrically connected in parallel, so the total resistance of the array becomes $3.8\,\Omega$ when all pixels are turned on.

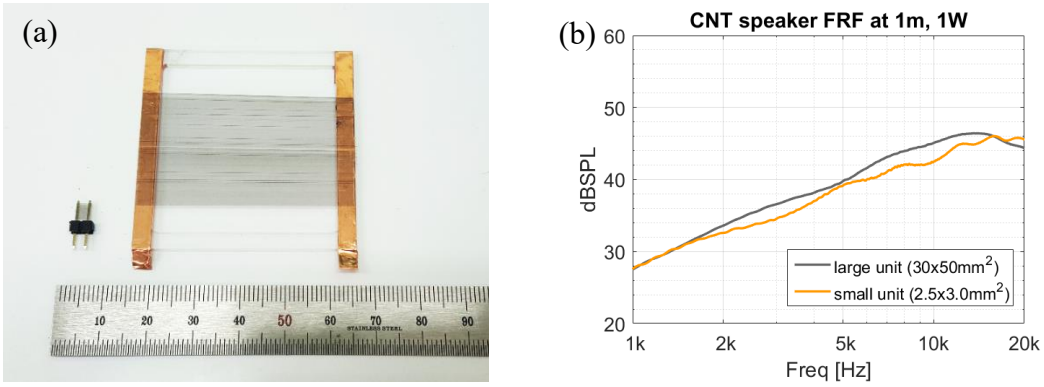

Figure 4: (a) Fabricated free-standing CNT speaker samples. Small sample on the left side (2.54 mm × 3.00 mm) and large sample on the right side (30 mm × 50 mm). (b) Measured frequency response of both samples

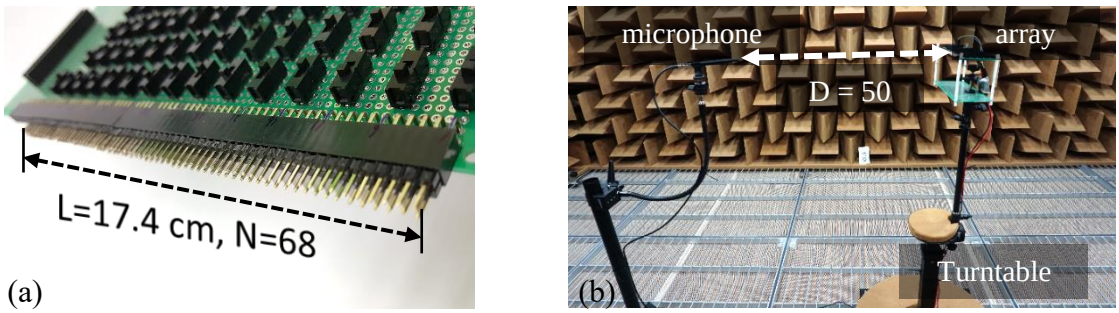

Figure 5: (a) Prototype acoustic pixel array. (b) Experimental setup.

4.2 Experimental setup and measurement

The acoustic pixel array was tested in an anechoic chamber. Figure 5(b) shows the experimental setup for the acoustic measurements. To measure the directivity at all frequencies within the frequency range of interest, we measured the frequency response for every 2 degrees using an automatic precision turntable (LT360EX, LinearX) synchronized with the acoustic analyser (CLIO fw11, Audiomatica). A logarithmic chirp signal was used for driving pixels, which was amplified by an amplifier (CLIO QCBox Model V, Audiomatica). The distance between the acoustic pixel array and the

microphone (MIC-02, Audiomatica) was configured as 0.5 m, considering the Rayleigh distance of the given array size.

When the nonlinear behaviour of TA transducers is considered, only the 2nd harmonic signal in the log-chirp measurement is relevant to the input signal. For this reason, we extracted and examined the second order harmonic component by applying a temporal window to the measured impulse response. Regarding the maximum allowable input power per unit area of a TA transducer [11, 16], the input power to each pixel is controlled. Through a series of experiments, we found that each pixel can withstand maximum input power up to 0.3W on average. The array hence can handle 20.4W in total.

As discussed in section 2.2, the sound pressure of a TA transducer is directly proportional to the input power (equation (4)). Using this relation, we converted magnitudes of all measured responses to equivalent outputs corresponding to 1W input power.

4.3 Size dependency of a single pixel

Prior to measuring the array's response, we tested the size dependency of SPL produced by a single CNT pixel. To this end, we fabricated two pixels of different sizes (Figure 4(a)). The sizes of small and large samples are 2.54 mm×3.00 mm and 30 mm×50 mm, respectively. The measured frequency responses of both samples, converted into responses for the 1W input power, are shown in Figure 4(b). It can be seen that SPL difference between two samples is less than 2dB, which ensures that the SPL of a CNT pixel is not significantly influenced by its size.

4.4 Comparison of measured beampatterns

The simulated and measured beampatterns at different frequencies are compared in Figures 6-8. Each beampattern is normalized by the sound pressure on the beam axis. We compared the performance by two measures: maximum sidelobe level (MSL) and half-power beamwidth θ_{hpbw}. The maximum difference in MSLs between the simulation and experiment is 4.5 dB, and the maximum error in θ_{hpbw} is 4 degrees. This result shows that the proposed method can reproduce a designated beampattern with an acceptable amount of errors. The major source of these errors is suspected to be the unidentical resistances across individual pixels.

One possible problem of the proposed technique is a loss of efficiency, because much of its input power is consumed for the evanescent components that do not propagate to the far-field. To check this possibility, we compared two on-axis SPLs generated by the spatial DSM and uniform excitation (Figure 9). It is turned out that the Dolph-Chebyshev array shows higher on-axis pressure (Figure 9(a)). This is suspected due to the higher side lobes of the uniform excitation, which is not small compared to the evanescent waves of spatial DSM. In the simulation, on-axis beam powers of both arrays are almost the same (Figure 9(b)), so the on-axis pressure with respect to the same input power is not much degraded compared to the uniformly excited arrays.

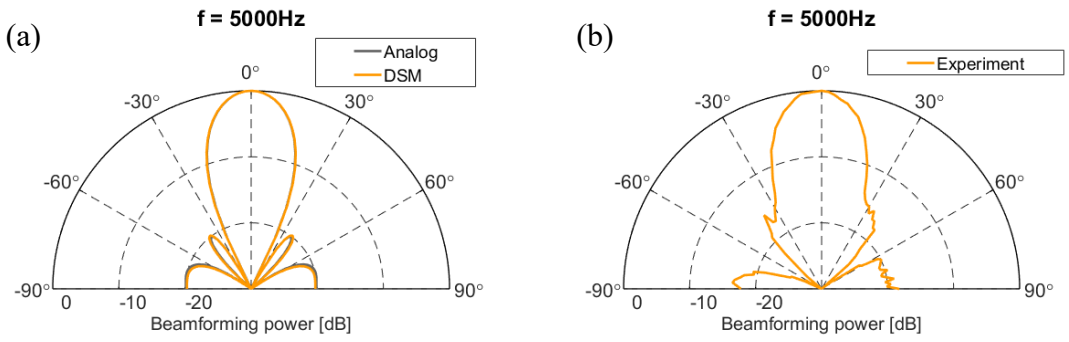

Figure 6: Simulated and measured beampatterns at 5kHz. (a) Simulated beampattern: MSL = –20dB, $\theta_{hpbw} = 20°$. (b) Measured beampattern: MSL = –15.8dB, $\theta_{hpbw} = 24°$.

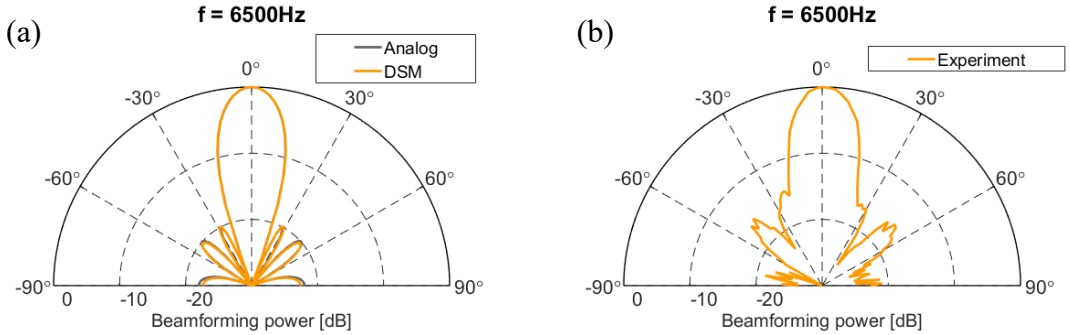

Figure 7: Simulated and measured beampatterns at 6.5kHz. (a) Simulated beampattern: MSL = –20dB, $\theta_{hpbw} = 15.4°$. (b) Measured beampattern: MSL = –15.5dB, $\theta_{hpbw} = 18°$

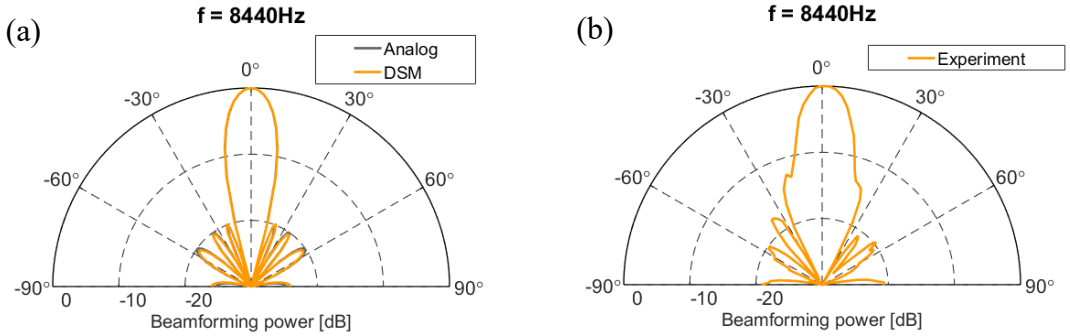

Figure 8: Simulated and measured beampatterns at 8.44kHz. (a) Simulated beampattern: MSL = –20dB, $\theta_{hpbw} = 11.9°$. (b) Measured beampattern: MSL = –17.6dB, $\theta_{hpbw} = 14°$.

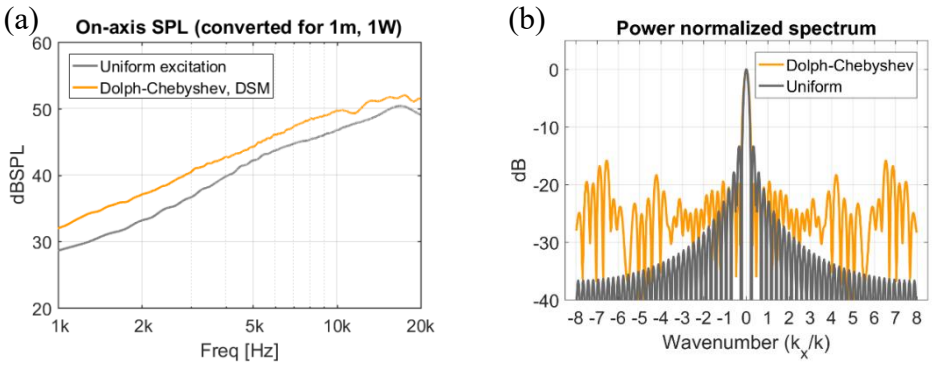

Figure 9: Efficiency comparison test. (a) On-axis pressure of uniform excitation array and Dolph-Chebyshev array after DSM. (b) Wavenumber spectrum of both arrays.

5. Conclusion

We proposed and validated a method to control a beampattern through on/off switching of acoustic pixels. The spatial DSM was introduced to obtain a binary spatial excitation function, and thermo-acoustic CNT loudspeakers were fabricated to implement small acoustic pixels. From simulations and experiments, it has been shown that the proposed method can produce the designated beampattern with an acceptable error. The proposed method can reduce the complexity and system construction cost by removing the multi-channel system. The proposed structure by itself is not adequate for steering a beampattern, which requires extra delays in the signal processing chain. Nevertheless, the simple structure enables shaping of beampatterns into various shapes just by on/off switching. Due to its simplicity, the proposed technique can be extended and combined with various space-time structures to enable more comprehensive control of beampatterns.

Acknowledgement

This research was supported by Samsung Research Funding Centre of Samsung Electronics under project No. SRFC-IT1301-04 and BK21 (Brain Korea 21) project initiated by the Ministry of Education. A part of this work was presented at the 172nd Meeting of the Acoustical Society of America, held in Hawaii, USA, 2016. This article includes detailed procedures to apply the spatial DSM and introduces how the acoustic pixel array is designed to implement the oversampled structure.

REFERENCES

1 S. Takeoka, Y. Yamasaki, "Acoustic projector using directivity controllable parametric loudspeaker array," *Proceedings of 20th international congress on acoustics*, ICA 2010. p. 921–5, (2010)

2 J. Candy and O. Benjamin, "The Structure of Quantization Noise from Sigma-Delta Modulation," *IEEE Transactions on Communications*, vol. 29, no. 9, pp. 1316-1323, Sep, (1981).

3 S. R. Norsworthy, D. A. Rich and T. R. Viswanathan, "A minimal multibit digital noise shaping architecture," *IEEE International Symposium on Circuits and Systems*, Atlanta, GA, pp. 5-8 vol.1, (1996).

4 Y. Tamura, N. Kawakami, O. Akasaka, M. Okada and K. Koyama, "Beam-forming using multidimensional sigmadelta modulation, *Proc. IEEE Ultrasonics Symp.*, 1998, pp. 1077–1080, (1998).

5 D. P. Scholnik and J. O. Coleman, "Space-time vector deltasigma modulation," *IEEE Int'l Conf. on Circuits and Systems*, Phoenix, AZ, May 26–29, (2002).

6 S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data Converters: Theory, Design, and Simulation*, Wiley-IEEE Press, New York, NY, USA, (1996).

7 V. Mladenov, P. Karampelas, G. Tsenov, and V. Vita, "Approximation Formula for Easy Calculation of Signal-to-Noise Ratio of Sigma-Delta Modulators," *ISRN Signal Processing*, vol. 2011, Article ID 731989, 7 pages, 2011. doi:10.5402/2011/731989, (2011).

8 E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*, Academic press, (1999).

9 H. D. Arnold and I. B. Crandall, "The thermophone as a precision source of sound," *Phys. Rev*. 10, 22–38, (1917).

10 A.R. Barnard, D.M. Jenkins, T.A. Brungart, T.M. McDevitt, and B.L. Kline, " Feasibility of a high-powered carbon nanotube thin-film loudspeaker," *The Journal of the Acoustical Society of America*, 134(3), EL276-EL281, (2013).

11 L. Xiao, Z. Chen, C. Feng, L. Liu, Z. Q. Bai, Y. Wang, L. Qian, Y. Zhang, Q. Li, K. Jiang, and S. Fan, "Flexible, Stretchable, Transparent Carbon Nanotube Thin Film Loudspeakers," *Nano Lett.* 8, 4539, (2008).

12 A.E. Aliev, N.K. Mayo, M. Jung de Andrade, R. O. Robles, S. Fang, R. H. Baughman, M. Zhang, Y. Chen, J.A. Lee and S.J. Kim, "Alternative nanostructures for thermophones," *ACS nano*, *9*(5), 4743-4756, (2015).

13 K. Suzuki, S. Sakakibara, M. Okada, Y. Neo, H. Mimura, Y. Inoue and T. Murata, "Study of carbonnanotube web thermoacoustic loud speakers," *Jpn. J. Appl. Phys. 50*, 01BJ10, (2011)

14 C.S. Kim, S.K. Hong, J.M. Lee, D.S. Kang, B.J. Cho and J.W. Choi, "Free‐Standing Graphene Thermophone on a Polymer-Mesh Substrate," *Small*, 12(2), 185-189, (2016)

15 L.E. Kinsler, A.R. Frey, *Fundamentals of acoustics*, Wiley, New York, NY, USA, (2009).

16 A.E. Aliev, N.K. Mayo, R.H. Baughman, D. Avirovik, S. Priya, M.R. Zarnetske and J.B. Blottman, "Thermal management of thermoacoustic sound projectors using a free-standing carbon nanotube aerogel sheet as a heat source", *Nanotechnology* 2014, 25, 293–295, (2014).