Ear Beamer

Aaron Lucia, Niket Gupta, Matteo Puzella, Nathan Dunn

Extracting Target Position with the Kinect

Deliverable: Extract and display (x,y) components of potential targets within room

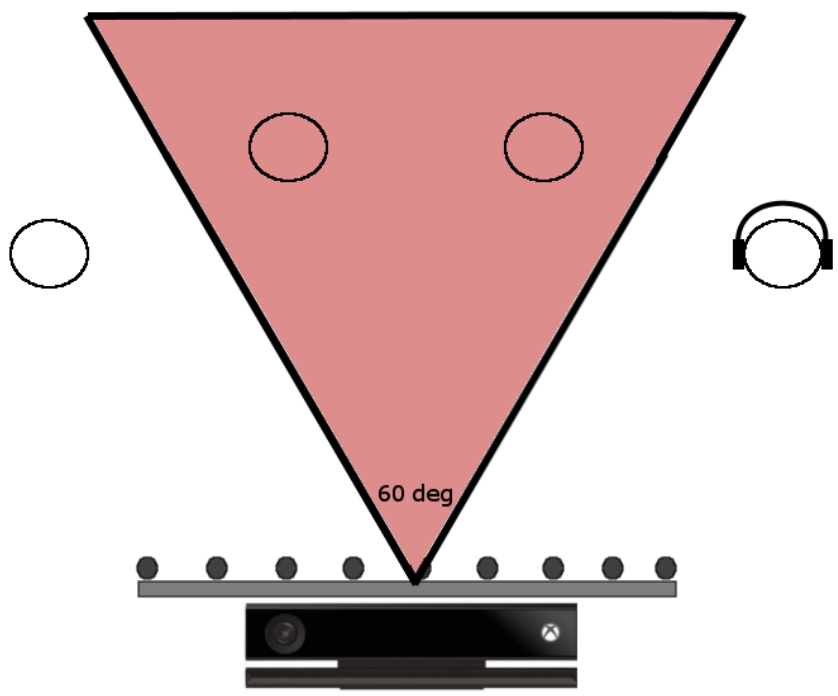
- Dynamically add/remove targets as they enter/exit Kinect detection range

Kinect Limitations:

Kinect Skeletal Tracking relies on the infrared depth sensor

- 512 x 424 depth image resolution
- Approximately 60º angle of view in testing

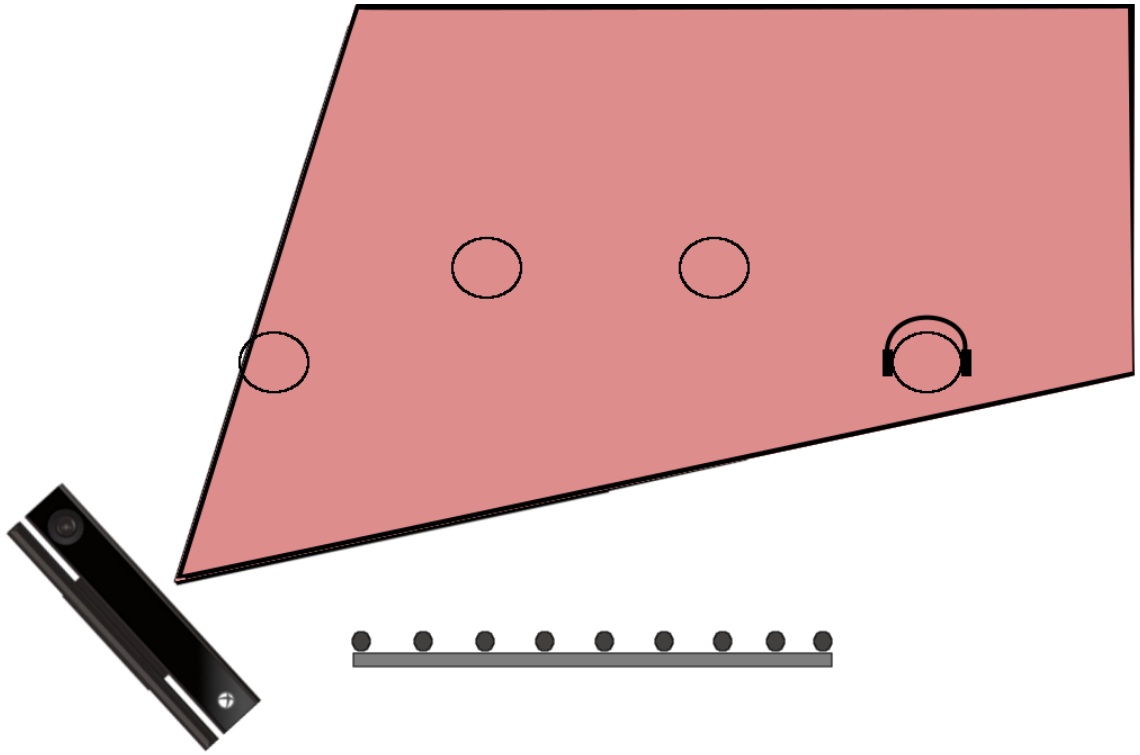
Cannot capture all possible targets when aligned with microphones



Overcoming Kinect Limitations:

Solution: Offset Kinect Camera from microphone array to optimally place targets within range

Coordinates returned from Kinect can be converted to a coordinate system based on the microphone array position

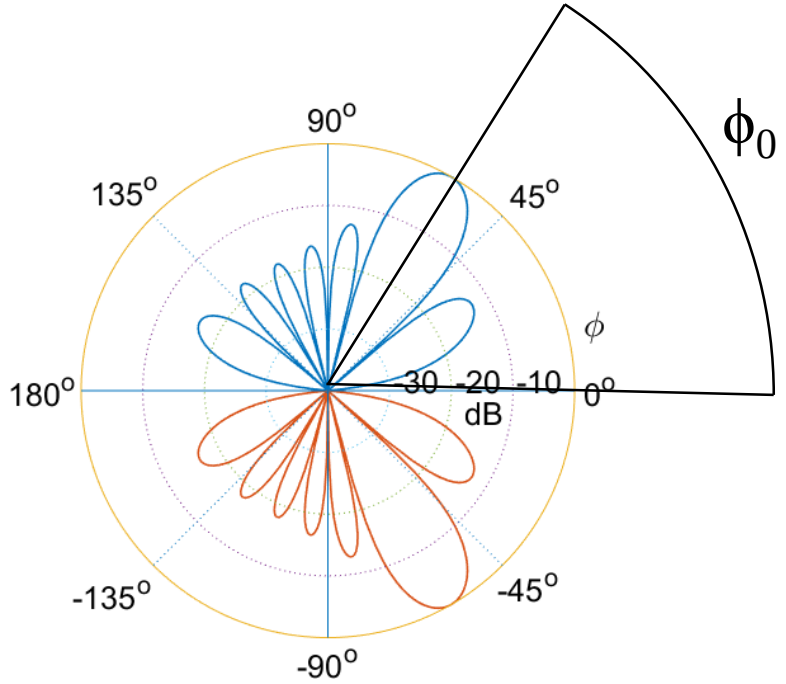


Recap: Beamwidth of Uniform Linear Array

Can calculate the beamwidth, $\Delta\phi_{3dB}$, of the main beam of the array pattern[1]:

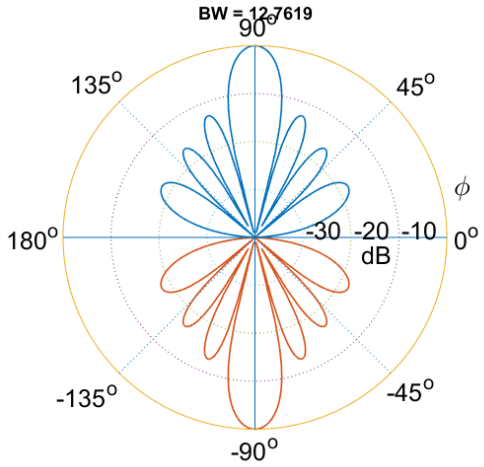
$$\Delta\phi_{3dB} \approx \frac{0.886}{\sin\phi_0}\frac{\lambda}{Nd}b$$

- ϕ_0 - Steering Angle
- N - Number of Array Elements
- d - Distance between array elements
- b - broadening factor (dependent on array weights)



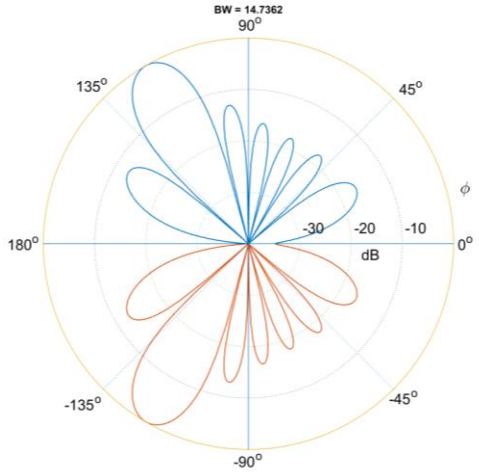
1. Orfanidis, Sophocles. *Electromagnetic Waves and Antennas*. Rutgers University, 1 Aug. 2016

Visualizing How Beamwidth Changes



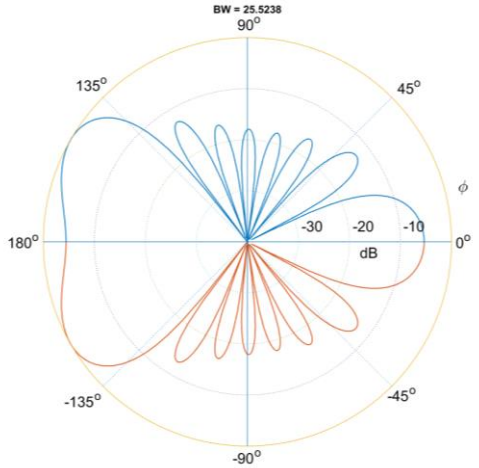
$\phi_0 = 90^o$

$\Delta\phi_{3dB} = 12.7^o$



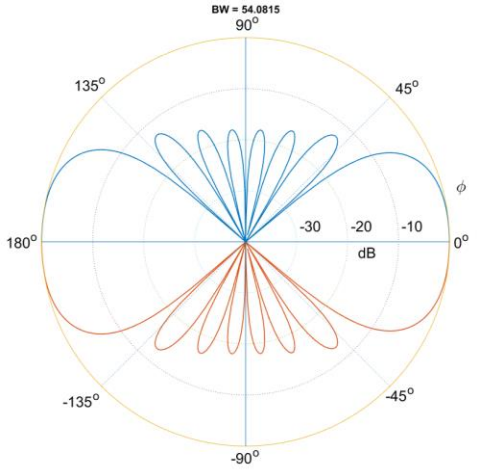
$\phi_0 = 120^o$

$\Delta\phi_{3dB} = 14.7^o$



$\phi_0 = 150^o$

$\Delta\phi_{3dB} = 25.5^o$

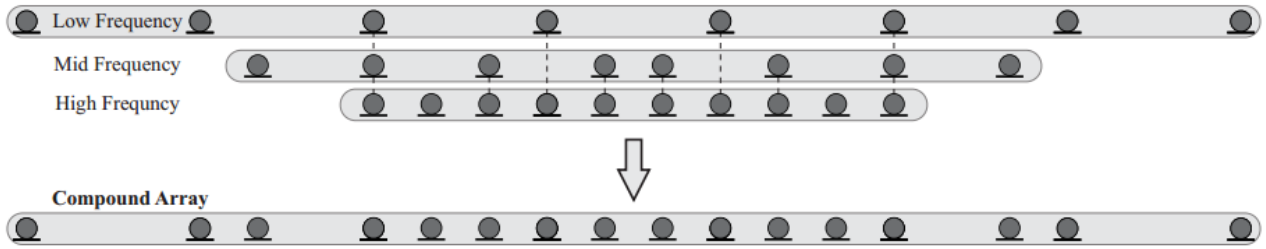


$\phi_0 = 150^o$

$\Delta\phi_{3dB} = 54.0^o$

Array responses for an 8 element array with d = 0.5λ

Analyzing SDP16's Array



Microphone Distance

3d

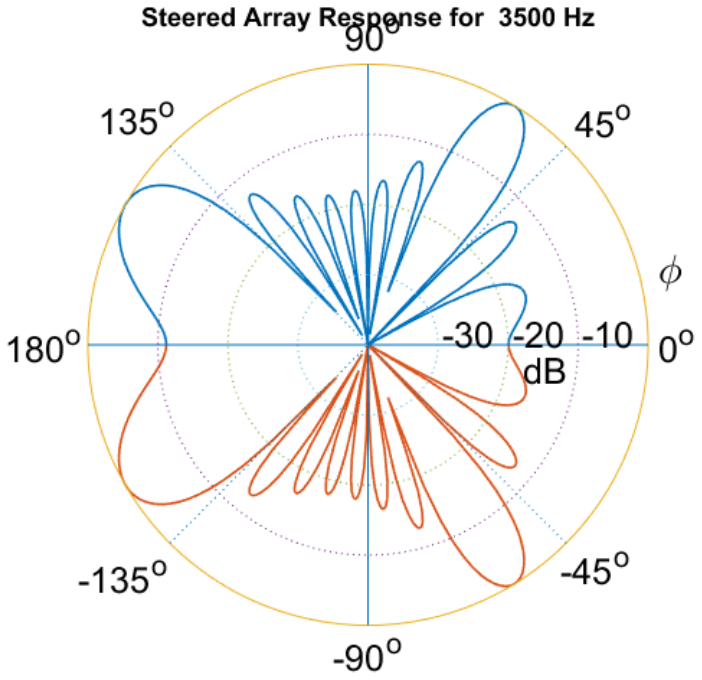
2d

1d

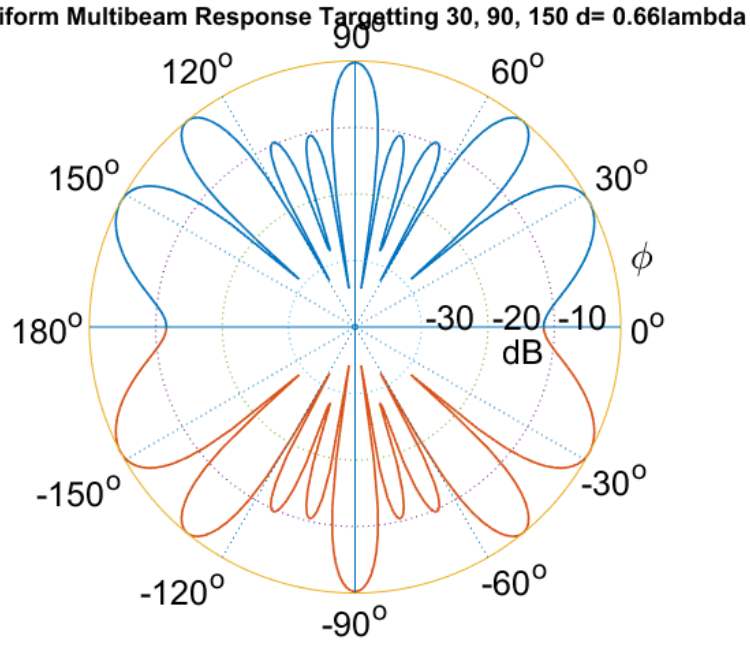
Weighting Method	Octave	High/Low Freq	d/lambda ratio	Angle	Beamwidth
Uniform	Low	1000	0.617	90	10.282
				150	20.564
		600	0.3703	90	17.137
				150	34.2748
	Mid	1700	0.699	90	9.0727
				150	15.4237
		1000	0.4114	90	18.1455
				150	30.8473
	High	3500	**0.72**	90	8.81
				150	17.62
		1700	**0.35**	90	18.14
				150	36.291

In every frequency band, half of the band exceeds the Nyquist Criteria, $d < \lambda/2$

Aliasing when Main Beam is steered:



Particularly problematic when beamforming is performed in parallel:



Creating a New Array

Constraints:

Given a frequency band [f_l, f_h], partition into N bands

- For each subband i, $[f_{i,l}, f_{i,h}]$, enforce microphone distance $d_i < \lambda_{i,h}/2$
- All microphone distances d_i must be power of two multiples of $d = \lambda_h/2$

Each frequency band becomes an octave. Need $\log_2(\frac{f_h}{f_l})$ octaves

For human speech (600 – 3500 Hz):

Octave	Frequency (Hz)	Microphone Distance (cm)
1	600 – 1200	14.1
2	1200 - 2400	7.1
3	2400 - 4800	3.5

Creating New Array, and decreasing beamwidth

Within each subband $[f_{i,l}, f_{i,h}]$

- Microphone spacing for lowest frequency is d = $\lambda_{i,l}/4$

At d = $\lambda_{i,l}/4$, beamwidth is less than ideal:

Steered Direction	Beamwidth
90	25
150	51

Solution: Limit the speech bandwidth to lower two octaves

- Allows for more microphones to be used in each band

Creating New Array, and decreasing beamwidth

An ASHA funded study found that intelligibility of speech remained relatively unaffected when higher frequency components (above 2.8 kHz) were filtered out

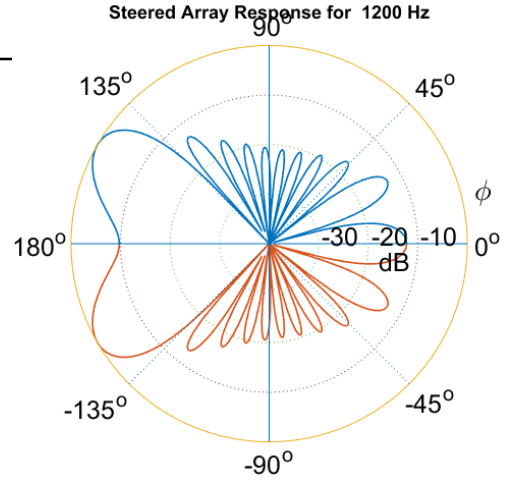
Unfiltered

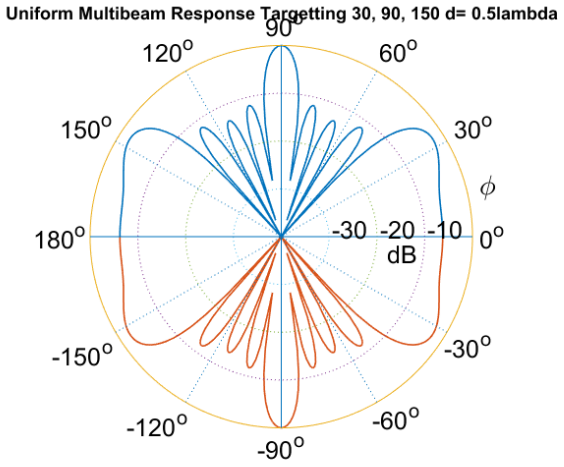
Band Pass Filtered
(600 – 2400 Hz)

Milner P, Braida LD, Durlach NI, Levitt H: Perception of filtered speech by hearing-impaired listeners. ASHA Reports Number 14, Speech Recognition by the Hearing Impaired. Rockville, MD, Speech-Language-Hearing Association, October 1984, pp. 3047

Performance

Weighting	Octave	High/Low Freq	d/lambda ratio	Angle	Beamwidth
Uniform	Low	1200	0.5	90	9.3
				150	18.6
		600	0.25	90	18.6
				150	37.1
	High	2400	0.5	90	9.3
				150	18.6
		1200	0.25	90	18.6
				150	37.1

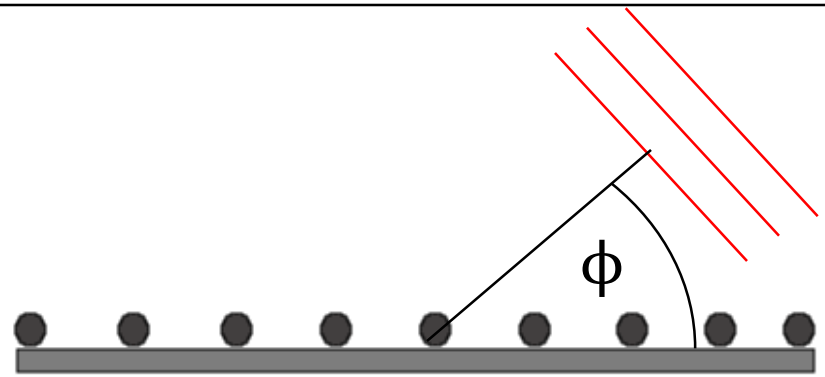


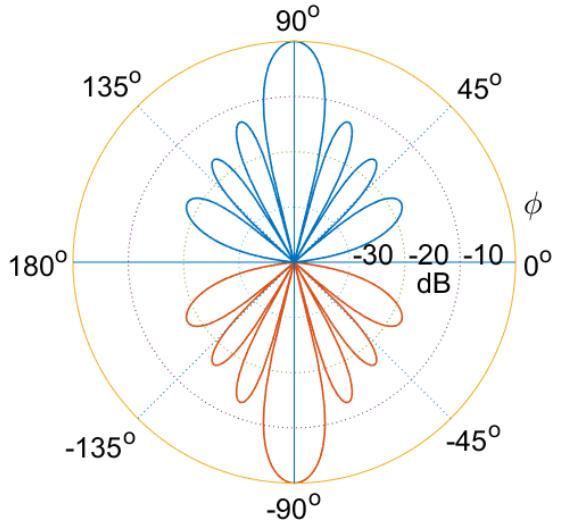


Recap: Beamforming on Uniform Linear Array

Array Factor for a 1 – Dimensional Array

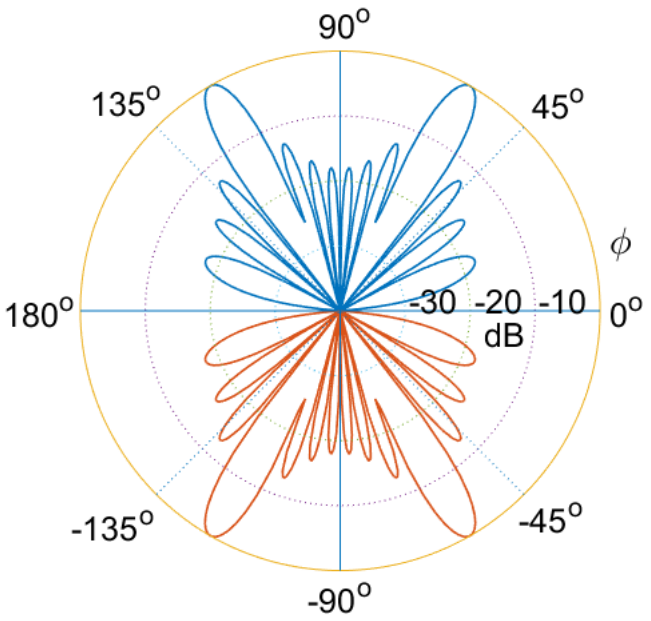
$$A(\phi) = \sum_{n} a_n e^{jnkd\cos\phi}$$





Recap: Aliasing

- When microphone distance d exceeds λ/2, aliasing can occur.
- Lobes equal in magnitude to the main



Array responses for an 8 element array with $\phi_0 = 60^o$, d = λ