“Ghosts In the Image” – Aliasing Problems With Incoherent Synthetic Aperture Using A Sparse Array

M. Hoffmann-Kuhnt
National University of Singapore
Acoustic Research Laboratory
Tropical Marine Science Institute
12A Kent Ridge Road
Singapore, 119223
tmsmh@nus.edu.sg

M. A. Chitre
National University of Singapore
Acoustic Research Laboratory
Tropical Marine Science Institute
12A Kent Ridge Road
Singapore, 119223
tmsmac@nus.edu.sg

J. R. Potter
National University of Singapore
Acoustic Research Laboratory
Tropical Marine Science Institute
12A Kent Ridge Road
Singapore, 119223
tmsmh@nus.edu.sg

***Abstract* - Traditionally, imaging of ensonified objects is done by beamforming using a fully populated receiver array. When a sparse array is used, grating lobes appear and this is not longer possible. The grating lobe problem can be solved by using a broadband signal: here the grating lobes for different frequencies are in different locations and thus average out. When only one reflective point exists time delay beamforming can correctly identify the location. A new problem arises when multiple reflections occur, creating ambiguities in locating the correct reflective points. These ambiguities are caused by many reflections arriving within the same time window at each sensor of the array. Thus, the number of possible combinations of each arrival in the different channels increases exponentially with the number of arrivals. Most of these combinations do not have a location in physical space, but the remaining ones (expect for the original points) are spurious and place energy in 3d-space that does not align with the original source point – making the reconstruction of an object impossible. Incoherent synthetic aperture in combination with new algorithms allows distinguishing between real and false reflective points. Reflection points of an object change only little during the movement of the source creating the synthetic aperture while spurious points “flicker” on and off and change location often. These new algorithms allow thus to create an image of the object and to filter out the false aliases.**

I. INTRODUCTION

Imaging of ensonified target objects normally requires a fully populated array of sensors. When the frequency of interest is high then the spacing of the sensors can become a problem as it might be either impossible to place individual sensors closer then half the wavelength or the number of sensors required to achieve sufficient angular resolution becomes to high. Therefore, sparse array beamforming has become of great interest in situations where either only a few sensors are available or physical constraints limit the placement. In nature this case seems to appear in the context of dolphin echolocation. When a bottlenose dolphin (*Tursiops truncatus*) echolocates on a target object, the emitted signal is reflected back to the dolphin from many points on the object. The signal enters the dolphin’s auditory reception system through the lower jaw and is then transmitted to the inner ear (see [1], [2], and [3] for a review), where it is then encoded in nerve signals. Cross-modal matching to sample experiments [4], [5], [6] have shown that the dolphin is able to recognize the shape of an object through echolocation. The currently accepted theory holds that the dolphin uses its lower jaw and its inner ear to receive reflections and to process the signals to reconstruct the shape of the target object [7] [8]. Considering that an object, that has been ensonified, will return many reflections in the same time frame, the dolphin then might face the problem that with two receivers (sparse array) it might not be able to tell which of the incoming reflections come from the same reflective points on the object. Nevertheless, the animal is able to identify the shape of the object. The dolphin brain probably uses certain neural processing algorithms that allow it to correctly identify the reflective points of an object.

A more recent theory [9] [10], suggests that the dolphin might use its lower jaw as an array of receptors with the individual teeth being the receivers. In this theory each single tooth is treated as a vibration sensor and the signals of all “sensors” are transmitted via the trigeminal nerve and added up with the appropriate delay in the brain. Although individual teeth are spaced approximately 9.4 mm apart (which is more than half of the wavelength of 150 kHz), the distance between the teeth is small enough to consider it a close to fully-populated array when viewed from directly in front of the dolphin’s rostrum. But even if this theory is correct then the “array” of sensors (teeth) would be fully populated only in one direction (horizontally) but in the vertical direction it would still be only a sparse array. Given the fact that the dolphin is nevertheless able to recognize shape leaves us with the question of what neural processing might allow it to do so.

To investigate this problem we decided to simulate the properties of the sound field of a target object that had been ensonified with an acoustic signal and see whether we would be able to reconstruct the shape of the object by using a sparse four-receiver array that was setup in a 3D tetrahedral shape. A minimum of four sensors was necessary to locate a source of a single reflection in 3D space.

II. BEAMFORMING

A. Sparse Array Beamforming

Generally, beamforming can be done in either the frequency or the time domain. When data from a reflection from a specific location are beamformed using a sparse array grating lobes appear, and this leads to ambiguities as to what the location of the source of reflections is (see [11] for a detailed review of beamforming and grating lobes).

B. Frequency beamforming

One way of localizing the source of a sound with two or more receivers is to perform a Fourier transform on the signal and to beamform in the frequency domain. If the signal is narrowband containing only one particular frequency and the distance between the receivers is larger then half the wavelength, then ambiguities will arise from which direction the signal has been received. Normally the correct direction can be detected through the position of the main lobe; the grating lobes, however, have the same amplitude as the main lobe. Hence the main and grating lobes cannot be discriminated, leading to false localization or ambiguities in the beamforming process. The solution to this problem is to use a broadband signal: in this case each frequency will have its distinct pattern of grating lobes that point in different directions, but only the main lobes, that represent the correct location, will have the same angular position. By adding up all frequencies the grating lobes will average out, whereas only the main lobes will add up coherently and will provide a clear method of identifying the correct location of the signal.

C. Time Domain Beamforming

Another option to locate the origin of signals is to beamform in the time domain. In this case, the time delay of the signal for each of the receivers is calculated. Then, the identified sections of each time series with a temporal length equivalent to the length of the spatial dimensions of interest are added up and the resulting energy is assigned to this particular point in space. Another way of looking at the problem would be to say: given the speed of sound in the medium, if we would have a signal coming from a particular point in space, then at what time would this signal arrive at each of the receivers? The energy for this section in the time series would then be added up, because it would be coming from the same point in space and contribute to the total energy of this point. This process is then repeated for every point in the area of interest.

Both frequency and time domain beamforming are mathematically equivalent. Hence, the advantage of broadband signals to reject grating lobes is also applicable to time domain beamforming.

D. Time of Arrival beamforming

A third option is time-of-arrival beamforming (ToA) which can be used with discreet signals. In this method peaks within each time series are located and the time delay is measured. The source location is then computed analytically based on the delays. The only prerequisite is that a signal can be easily identified and matched unambiguously, usually associated with low distortion high signal-to-noise ratio broadband signals. ToA beamforming would not be useful if no clear signal can be detected (or the signal is unknown) or we only would want to compute the energy distribution with the particular space of interest. In this case regular time domain beamforming will provide better results then ToA beamforming, because it will combine data from the different timeline regardless of the signal, whereas ToA beamforming would require exact knowledge of the signal and its delays.

All these methods work when using a sparse array and a broadband signal with a relatively short duration and when reflections in the time series come either from one single reflective point or when the reflections from different points are separated by a delay that is larger then the time difference that separates two receivers if the reflection was coming from an "endfire" position along a line through the two receivers.

E. Multiple reflections

If a reflection comes form one specific point in space and is recorded in the time series of all receivers then beamforming allows the precise localization of the reflection in 3d space. But what happens when two or more reflections are received? If we consider that each arrival of a signal can be detected through algorithms such as matched filtering or cross-correlation, then each arrival time of a reflection can be replaced with a single peak in the timeline. For simplicity we assume peaks to be the same amplitude, a simplification that does note take into account that the matching algorithms also provide information about the quality of the fit as well as the amplitude of the reflection.

In the case where reflections from different point of a target object arrive within a time frame that is shorter than the delay between two receivers, ToA beamforming will now generate all possible combinations that have a mathematical solution and the number of spurious reflection points will increase as P^N with P being the number of peaks in each time series and N the number of receivers. Fig 1 shows the time series of two channels where a peak A and peak B arrive within a short time. All peaks are of the same amplitude as they are all a representation of the same

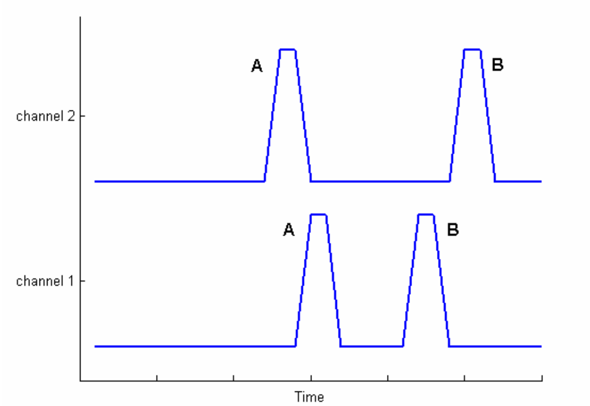

Fig 1: two peaks received in two time series

signal reflected of two different points in space. The beamforming algorithms will not be able to distinguish that A(1) in the first time series and A(2) in the second time series comes from the same point and B(1) and B(2) come from a different point. Conventional time domain beamforming will create aliases by combining A(1) with B(2) and B(1) with A(2) in addition to the two correct sources. In this simplified case four solutions are calculated. The advantage of ToA beamforming is that it allows us to exclude certain combinations by defining rules such as "if A(1) is combined with A(2), then A(1) cannot be combined with B(2)". This reduces the number of solutions from four to two in this specific case. Nevertheless, even with using ToA beamforming the number of possible combinations rises very fast with the number of reflection points and the number of receivers.

Some of these combinations do not have a solution that exists in physical space, so the actual number of points appearing after beamforming is less. Furthermore, some of the spurious points will coincide with another real reflective point and only add to that point's total energy but will not change the shape of the target object. Where these spurious points lie depends mainly on the position of the receivers in the array. The smaller the distance between the receivers is the less spurious points, but this also results in a diminished angular resolution.

A simulation of the reflective points on the surface of a diamond-shaped object is shown in Fig 2. Here, we simulated a set of 56 points in a two-dimensional plane in the shape of a diamond. Four receivers were placed in a tetrahedral shape centered in front of the object. The tetrahedral receiver array had a side length of 10 cm. We then assumed that each reflective point would be represented in the four different timelines with a single peak. The results of the beamforming algorithm are shown in Fig 3. Although the original shape of the diamond shape is still present, it is buried in the spurious points that appear as a result of the beamforming.

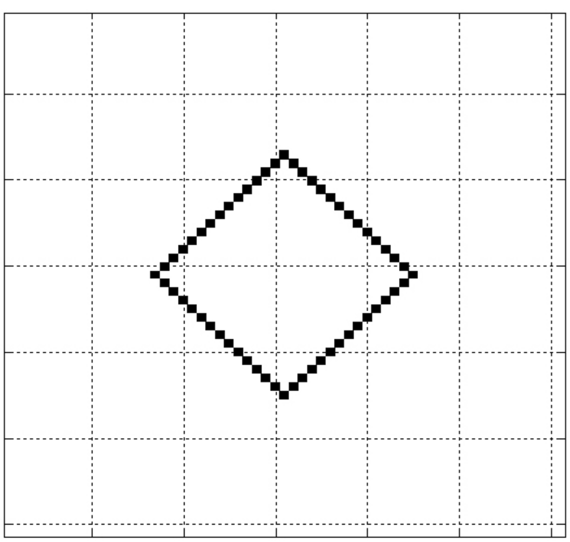

Fig 2: image of the simulated reflective points

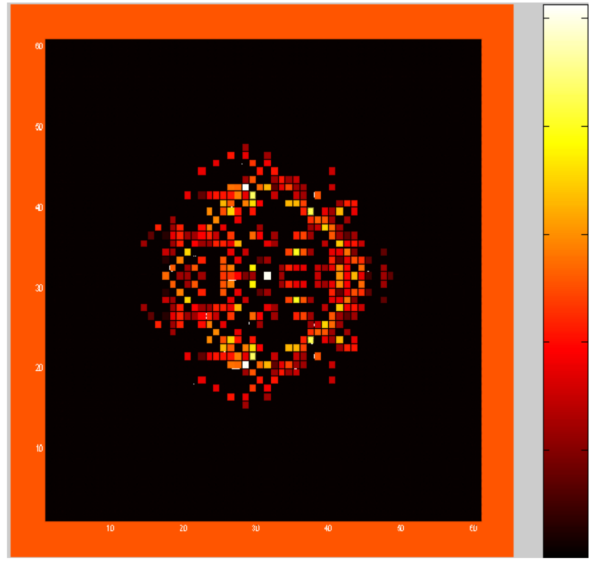

Fig 3: results of beamforming with the appearance of spurious points

III. New Algorithms

We then set out to find solutions that would be able to separate true reflective points form spurious points. Two approaches were followed. In the first one we tested whether we could apply an algorithm that had similar properties as the Marr-Poggio algorithm [12], but modified to specific constraints of acoustic domain. The second solution was based on the previous research with an echolocating dolphin: if the animal was using a synthetic aperture to investigate the target object then this would also have affect on the spurious points.

A) Iterative Enhancement

The original Marr-Poggio algorithm investigates the problem in human vision that an ambiguity exists between the correspondences of the two retinal projections when more then one point is projected. The algorithm provides a neural network based computational solution with the constraints that objects are cohesive and continuous and that each point has a unique location in space.

We followed the original algorithm by setting similar constraints that would apply to the acoustic domain: (1) Each pixel could excite adjacent pixels as shown in Fig 4 but not pixels behind or in front. In the case of the ensonified target object this would be equivalent to the fact that areas on an object that are right next to reflecting points are likely to also reflect sound in the same direction. Although this property would also apply to pixels in front and in the back, these pixels are more likely to interfere and obstruct each other. (2) Aliases that appeared on pixels located in physically distant positions were inhibited. An alias is defined as a pixel which shares at least one arrival time on a receiver with the current pixel. A schematic of the rules is shown in Fig 4. Both inhibition and excitation could be set to different strengths. Each pixel that has been assigned an energy value through the beamforming algorithm affects other pixels that also have an energy value greater then zero. Lines with a "+" depict excitation representing the higher likelihood that points adjacent to a reflection point will also reflect energy. Pixels that are not adjacent to the initial pixel are inhibited (arrows with “–”).

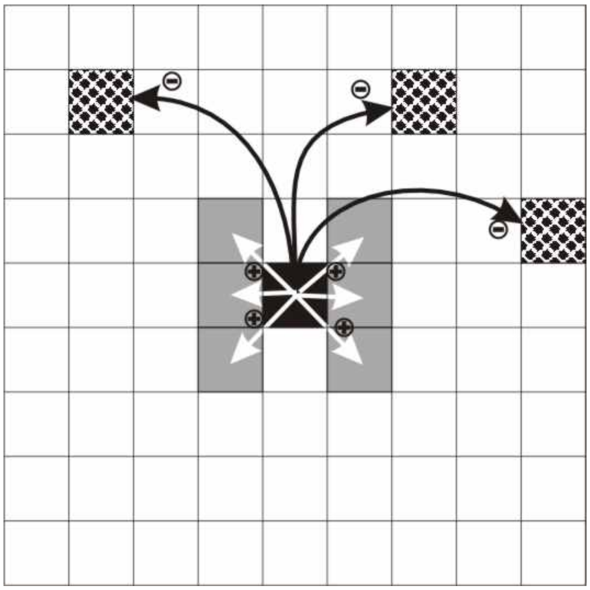

Fig 4: schematic of the rules of inhibition and excitation for each pixel.

To test the algorithm we restricted the simulation to two dimensions and two receivers. Fig 5 shows the setup of the simulated reflective points (black squares) and the position of the receivers (circles).
The iterative algorithm was then applied 200 times.

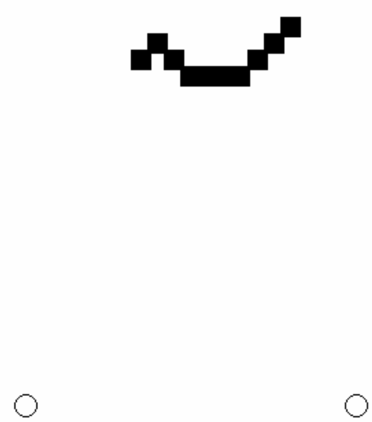

Fig 5: image of the simulated pattern before beamforming with the position of the two receivers

The results of the beamforming without applying the algorithm are show in Fig 6. Several more pixels are now assigned energy even though they don’t match the original location of points. After applying the algorithm on the matrix of pixels a closer match to the original configuration can be observed in Fig 7. The difference now is that one of the original pixels is not assigned sufficient energy. In practice this means that the algorithm has to be fine tuned to adjust to the individual settings of each image.

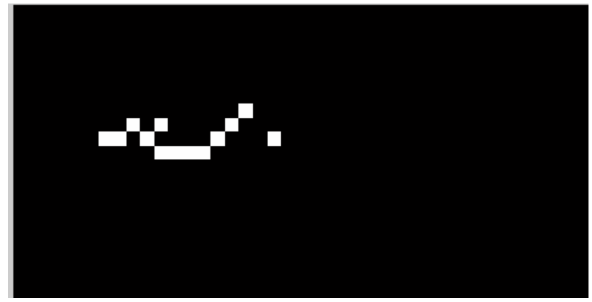

Fig 6: Image of the pattern after beamforming

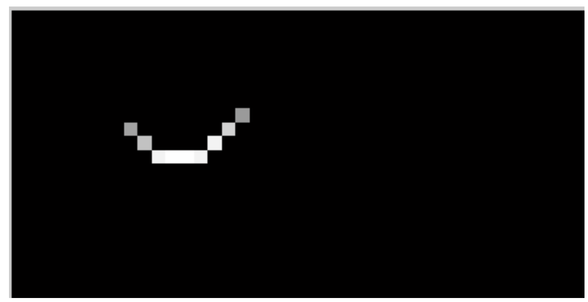

Fig. 7: Image of the pattern after applying the variation of the Marr-Paggio Algorithm

B) Synthetic Aperture

The second option mentioned previously was to simulate a synthetic aperture approach in a manner similar to the path of a dolphin approaching and echolocating on an object. Here we did not apply the Marr-Poggio variation; rather we looked at what the effect of a synthetic aperture algorithm would be. For this simulation we used the diamond shaped object again. The source position and the position of the receivers was moved along a 3D path and the virtual object was ensonified at each step. The resulting time series were the used to perform time of arrival beamforming. Each resulting distribution of reflective points (both real and spurious points) was then added up over the complete length of the approach path. Because the position of the spurious images depends mainly on the position of the receivers and their distance to the real reflective points, the spurious points change position with each new frame, while the real reflective points stay constant. This “flickering” of the aliases causes the energy to average out while the real points add up consistently. Fig 8 shows the result of a synthetic beamforming approach after 100 individual frames have been added up. Although there is still some energy in positions other than the real reflection points the overall difference in amplitude between real and spurious points allows us to set a threshold which would leave the original points intact while eliminating the aliases.

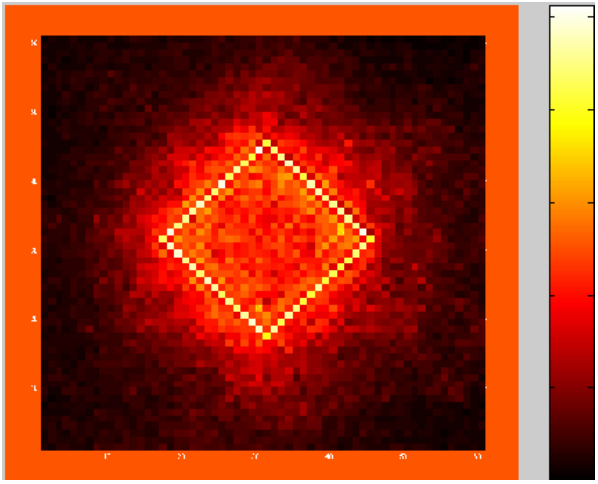

Fig 8: result of the synthetic aperture approach

IV. Conclusion

Both methods – the iterative enhancement and the synthetic aperture, have proven to enhance the processing of the beamformed data and do allow us to separate real reflection points from spurious points with the set limitations. Whether both approaches will work with data collected from real reflections remains to be seen. One of the problems with reflections of real objects is that not all points reflect continuously and the amplitude of each reflection can vary considerably. Thus either algorithm might not recreate the object in its complete shape. Furthermore, the Marr-Poggio variation was only simulated in 2D space and needs to be implemented in 3D space to provide realistic results. But a possible combination of both algorithms might provide a good representation of the original shape and might eventually provide a possible explanation of how dolphins might use similar algorithms in their neural processing to resolve the ambiguity of spurious reflections when recognizing target objects through echolocation.

REFERENCES

[1] W. W. L. Au, The sonar of dolphins. New York: Springer Verlag, 1993

[2] J. L. Aroyan, Three-dimensional modeling of hearing in Delphinus delphis. J. Acoust. Soc. Am. 110(6), 3305-3318. 2001

[3] D. R. Ketten,. Cetacean Ears. In W. W. L. Au, A. N. Popper, & R. R. Fay (Eds.) Hearing by whales and dolphins. (pp. 43-108). NY: Springer-Verlag. 2000

[4] L. M. Herman and A. A. Pack, Echoic-visual cross-modal recognition by a dolphin. In Thomas, J. R., Kastelein, R. A., Supin A. Y. (Eds.). Marine mammal sensory systems. Plenum Press, New York, pp.709-726, 1992

[5] A. Pack and L. M. Herman, Sensory integration in the bottlenosed dolphin: Immediate recognition of complex shapes across the senses of echolocation and vision. J. Acoust. Soc. Am. 98, 722-733, 1995

[6] L. M. Herman, A. A. Pack, and M. Hoffmann-Kuhnt, Seeing through sound: Dolphins (Tursiops truncatus) perceive the spatial structure of objects through echolocation. J. Comp. Psychol., 112, 292-305, 1998

[7] A. Pack, L. M. Herman, and M. Hoffmann-Kuhnt, Dolphin echolocation shape perception: From sound to object. In J. Thomas, C. Moss, & M. Vater (Eds.). Advances in the study of echolocation in bats and dolphins. Chicago: University of Chicago Press, 2004.

[8] A. Pack, L. M. Herman, M. Hoffmann-Kuhnt, and B. Branstetter, The object behind the echo: Dolphins (Tursiops truncatus) perceive object shape globally through echolocation. Behavioral Processes, 58, 1-26. 2002

[9] R. Potter and E. Taylor On novel reception models for Bottlenose dolphin echolocation. Proceedings of the Institute of Acoustics, vol.24, Biosonar, Loughborough, UK, 2001

[10] D. Goodson and M. Klinowska, "A Proposed Echolocation Receptor for the Bottlenose Dolphin (Tursiops truncatus): Modelling the Receive Directivity from Tooth and Lower Jaw Geometry", In: Sensory Abilities of Cetaceans: Laboratory and Field Evidence, J.A.Thomas and R.A.Kastelein (eds). NATO ASI Series A, v196, Plenum, NY, pp 255-267. 1990

[11] D. A. Johnson, Array signal processing : concepts and techniques / Don H. Johnson, Dan E. Dudgeon. Imprint Englewood Cliffs, NJ : P T R Prentice Hall , 1993.

[12] D Marr and T. Poggio,. A computational theory of human stereo vision. Proceedings of the Royal Society of London B, 204:301-328. 1979