Hypothetical Neurons for the Imaging Sonar of Dolphins and Related Geometric Methods of Computing Source-Points of Echoes

by Douglas Moreman

A variety of geometric methods, introduced in this paper, for computing images from echoes might help explain the sonar of dolphins. The method begins by organizing an array into triangulating subsets, each of an appropriate number of sensors. Each subset both detects and checks possible sources of sound.

How are dolphins able to "see" by means of echoes of their sonic-clicks? Dolphins have two ears for listening to a broad frequency-range of sounds of the sea. Perhaps, rather than using those two auditory channels simultaneously for another purpose, their "sonic vision" is informed by touch-like sensors in the face.

Assuming that there are more than two points of entry of sound has allowed creation of software that computes images from simulated echoes of dolphin-clicks. The same numerical methods that produce images from simulated echoes ought to make images from files made from real echoes that are recorded via a suitable array of hydrophones. But, so far, all is theory and simulation.

The bottlenose dolphin and some related species of odontocetes produce clicks that have a prominent feature which one might call, due to its shape in a graph, a "fang." A fang is a rapid change in amplitude that 1) occurs in about 1/100,000 second and 2) is far and away a greater change than exists in any other of the few cycles of the click.

Let us suppose that the face of a dolphin has "echotriggers" – sensors that are exquisitely sensitive to the passing-by of a fang. Perhaps they evolved from touch or vibration sensors. A typical human finger contains more than 1000 touch/vibration sensors so, perhaps it is plausible that the face of a dolphin contains 1000 echotriggers. Surely such a multitude could provide more information for triangulation than can just two ears.

A practical man-made, dolphin-inspired, system can function with fewer than 1000 hydrophones. Methods introduced here have produced images from simulated echoes arriving at an array of just nine hydrophones.

We assume, for now, that the face of a dolphin has an array of echotriggers and each of these is connected by its own axon directly into the brain. Experiments done in simulation suggest that subsets of six echotriggers might ordinarily be adequate for both computing directions to possible sources of sounds and for checking those directions.

Nature might do some imaging operations, not to say it does do, using just two echotriggers. Suppose that the two axons of a pair of echotriggers connect to a hypothetical "ditoa-neuron."

A ditoa-neuron has exactly two input dendrites and one output axon and is such that the output fires only after the nearly-enough simultaneous firing of the two inputs. "Ditoa" comes from Difference-In-Time-Of-Arrival. A time-of-arrival is called a "toa." A ditoa is a difference, one toa minus another toa, and can be negative.

The "fang" is not the only feature of a click that can be used in computing images from echoes of waves but it is simple and memorable and no other feature will be used here.

978-0-933957-40-4 ©2013 MTS

Suppose there are two echotriggers M1 and M2 on the head, spaced equally to the right and to the left of the fore-aft axis of the head, respectively. From each of M1 and M2 its axon leads into the brain wherein it splits into dendrites that connect to ditoa-neurons. There might be more than one ditoa-neuron that receives input from both M1 and M2, but we first consider just one of them, call it "N."

Imagine there is just one sound-reflecting point P and it is located to the right front of the head of our imagined dolphin. Let t1 and t2 denote the time-of-travel of a fang from P to M1 and to M2, respectively. Since P is nearer to M1, t1 < t2.

Suppose that the ditoa-neuron N is on the left side of the brain and has an input that is a branch of an axon from M1, and that s1 denotes the time-of-travel, of an axonal pulse from M1 to N. Likewise, suppose N can receive an input from M2 after a time-delay of s2. Since the path from M2 to N is shorter than the path from M1 to N, s2 < s1. Note the possibility that t1+s1 = t2+s2.

If a fang leaves P at time T0, it arrives at M1 at time T0+t1 and generates a pulse at M1 that arrives at ditoa-neuron N after an extra duration s1. N receives that pulse at time T0 + (t1+s1). Likewise, N receives via M2 a pulse at time T0 + (t2+s2).

Let T1 and T2 denote, respectively, the toas T0 + t1 and T0 + t2, the times-of-arrival of the fang from P at sensors M1 and M2 respectively.

If a pulse from echotrigger M1 and a pulse from M2 arrive at ditoa-neuron N at nearly enough the same instant, then N is stimulated to fire, sending a pulse to another level in the brain. When, nearly enough, the difference (T2+s2) - (T1+s1) equals zero, N fires. So, when the ditoa T2 - T1 equals, nearly enough, s1 - s2, the ditoa-neuron N fires. The difference s1 - s2 is internal to the dolphin and T2 - T1 is imposed on the dolphin by the environment.

Let the distance from a point A of 3-Space to a point B of 3-Space be denoted by d(A,B).

The set of all possible reflecting points P that could produce exactly the same t1,s1,T1 and t2,s2, and T2 constitute a subset of a surface H12 determined by distances d(M1,P) and d(M2,P) and a number u computable from the times and the speed of sound in the medium. H12 is the set of all points Q such that d(M2,Q) - d(M1,Q) = u.

Let Speed denote the speed of waves in the medium.

It can be shown that u = Speed * (s1 - s2). H12 can be called a "hard-wired hyperbola" of the system. It depends upon the locations of M1 and M2 and upon the axonal lengths that determine the time-delays s1 and s2. When externally impinging T1 and T2 are such that the ditoa T2 - T1 is near enough to s1 - s2, the ditoa-neuron N fires.

N would be just one of a possibly linear array of ditoa-neurons that receive input from M1 and M2. Call that array DitoaLine(M1,M2). Which of the ditoa-neurons in DitoaLine(M1,M2) fires as a result of a fang from a point P could signal where P is located in a left-to-right sense. If u=0 then a ditoa-neuron fires that indicates P is straight ahead. As u is more negative a neuron fires that indicate further P is to the left of straight ahead. As u is more positive, the further P is to the right.

It was to simplify discussion that we imagined echotriggers M1 and M2 to be equally spaced about the fore-aft axis of the head. Similar reasoning would apply to any two sensors. Our instinct might be to separate M1 and M2 as much as possible, but, let us imagine that M1 and M2 are together, right and left, on the dolphin's chin and that M3 is a third echotrigger on the chin, positioned so that Angle (M2,M1,M3) is right. The DitoaLine(M1,M3) determines direction in the up-down sense.

Together, DitoaLine(M1,M2) and DitoaLine(M1,M3) could determine, at least roughly, a direction from the head to a reflecting point P. Another, planar, set of neurons G could each receive input an input from DitoaLine(M1,M2) and an input from a ditoa-neuron in DitoaLine(M1,M3). G could represent, to some deeper part of the brain, the direction to P, the source of the fang.

We see here indications of methods available to, but not necessarily really functioning in, Nature. Such indications have suggested simulate-able methods for a device, let's call it the "Device," that will compute images from echoes. The sensors of Device are hydrophones, not echotriggers. Device might have just nine hydrophones and be seemingly easy to build but, so far, all results are from simulations.

Two pairs of hydrophones can determine a direction towards an isolated source of waves but, it seems, cannot do imaging. Two is not enough sensors to eliminate phantoms - indications of a reflecting point where none exists. In the Device, sensors work together in small subsets, called "scopions." In many simulations, scopions have consisted of exactly 6 sensors. The extra sensors, above three, allow reduction in the number of phantoms.

In a "click-event" there is one click and its echoes at the sensors are digitized and sent to a computer that computes points to put into an image.

A software routine called FeatureDetector does this: for each sensor M, it detects fangs in the output of M, and records the toa of each detected fang. The output of FeatureDetector goes to a software routine called HitDetection which does this: for each scopion S it computes upon the toas of fangs in the outputs of the sensors in S, to detect "hitpoints," which are point-representations of possible origins of those fangs. It avoids or eliminates many of those phantom-producing wrong associations of a fang on one sensor in a scopion S with a fang on another sensor in S.

Typically, each scopion produces just a few hitpoints and these might be spread sparsely about in space in front of Device. But, the hitpoints of some large number of scopions might combine into a useful image.

In most simulations so far, a target object, often a simple fish, is modeled with enough points to simulate smooth surfaces. Each sensor, by contrast, is modeled as a point.

To simulate an echo at a sensor M: for each target-point P, the time of travel from the clicker-point to a reflector point P, and thence to M is computed. That time is used to appropriately time-delay a copy of the click. The hundreds of resulting click-copies are summed to represent the echo at that sensor M. This is done for every sensor M in the array.

All simulations programs have been written from the ground up, using only the tools that came with the compiler.

There is a variety of alternative sub-methods for computing representations of sources of sound. One sub-variety, call it Surfaces, detects prospective hitpoints via surfaces of geometry and checks those points via a separate operation. Another sub-variety, call it Grids, searches a pre-computed grid and, in effect, combines detecting and checking into the search.

In the methods of Surfaces, representations of sources of fangs are computed by means of the geometry of conic surfaces that are themselves each determined by toas at known locations and, in case of active sonar-imaging, by the position and time of emission of a click.

Hyperboloids and ellipsoids, and also spheres, are herein called “conicoids.” Each conicoid can be generated by rotating a planar conic section about an axis of that conic section, just as a sphere can be generated by spinning a circle about one of its diameters.

From times-of-arrival of echoes at a few sensors, a point P might be computed as the intersection of some conicoids. Suppose M1 and M2 are sensors and T1 and T2 are the toas of a fang on M1 and a fang on M2 respectively. The set H12 of all points Q such that d(M2,Q) - d(M1,Q) = (T2-T1)*Speed is an “hyperboloid”. M1 and M2 are the foci of H12. If M3 is a third sensor and the hyperboloid H13 is defined similarly to H12 then the intersection of H12 and H13 might indicate, at least roughly, a direction towards the source of that fang.

A method or act of computing, via an intersection of conicoids, a point P to represent a source of a wave is called triconication. Tri-conic-ation can be done using three conicoids. It is analogous to, but also quite different from, “tri-angulation.” Triconication can be executed by any one of several numerical algorithms.

Triconication upon the outputs of sensors in a small (e.g. 3 or 4 member) set can be used, and probably is used in conscious computations by some people, in what might be called “echo-location,” to locate sources of waves. Perhaps, for example, the submarine Scorpion was located, at least roughly, by triconication upon the sharp sounds sent across the Atlantic Ocean by its imploding hull.

While echolocation using triconication can locate isolated source-points of waves, there are two reasons to expect echolocation to fail as a basis of sonic imaging:

1) it produces too few points to properly depict an object, and
2) the points it produces are obscured by phantoms.

1) Points produced by echolocation are few for two reasons:

A) There is Destructive Overlap: when two or more wave-reflectors are at nearly enough the same distance from a sensor, there is overlap of waves that destroys time-of-arrival information on that sensor.

B) There is the Features-per-Duration Limitation. For example, in a computer-simulation of a “fish-finder” a test fish, though it might be 24 inches long, “seen” sideways can easily span less than 6 inches from its nearest point to the Device to its farthest point from the Device. The differences in time-of-arrival from the farthest point and the time-of-arrival from the nearest point of the fish would be, in seconds, less than d = 6/Speed where Speed is about 60,000 inches per second. In that simulation, the duration of a fang is p = 1/100,000 second. It seems unlikely that, on any one sensor, the number of fangs from that fish could exceed d/p=10. So, in the given duration d, a sensor’s output would be limited to no more than 10 fangs resulting from some dolphin-like sonic illumination of that fish. And, in the detection of representations of possible sources of waves, a triconicating set of sensors would detect no more than 10 points. Ten is not enough for a proper image of a fish. And, due to Destructive Overlap on each sensor, rather than 10, a more realistically anticipated number of image-points on that fish per triconicating set, per click, would be less than 10. In simulations, zero has been a frequent result.

The problem of Too Few Points can be reduced by increasing the number of scopions. It might be reasonable to guess, for this reason and others, that the face of a dolphin has hundreds of echotriggers.

2) The Echoplasm of Phantoms, a second reason for failure of echolocation in imaging objects rather than just locating objects, can be briefly illustrated. Suppose that M1, M2, M3 are three detecting sensors of a scopion S. Suppose there are 100 discrete, reflecting points. Suppose that representations of source-points of waves are to be detected via toas of fangs on M1, M2 and

M3. Suppose T1 is a toa on M1 and there are 100 toas T2 on M2, and 100 toas T3 on M3. There might be only one source of the fang that produced T1 but there might be 1*100*100 or 10,000 triples (T1,T2,T3) to consider for that T1. Of these, at least 9,900 would be "phantoms," Given a ratio of 99 phantom points per good image-point, the true object would be obscured in a cloud of echoplasm.

It seems remarkable that such difficulties have, largely, fallen to fairly simple techniques derived from geometry. The difficulities were not daunting, because no other approach seemed likely, and because dolphins had shown that imaging, via their clicks, is possible. In fact, dolphins show us that one dolphin can "see" via clicks of another dolphin - pretty much eliminating active sonar from our list of possible mechanisms for dolphin-sonar, though not from our practical applications.

Simulations confirm, as well as simulations can, that the problem of Too Few Points is largely reduced by increasing the number of scopions. The problem can be further reduced by clicking more than once per picture, changing slightly, between clicks, the position of the clicker, by internal means or by moving the head.

Simulations show that the other reason for doubt, the Echoplasm of Phantoms, is reduced by using a sufficiently large number of sensors per scopion. And, that number has been, in simulated applications so far, usually no greater than six. How they are used is, of course, also important.

Suppose that, in Device, 6 is the number of sensors in every scopion and for a scopion S, the sensors of S are denoted by S1, S2,...,S6.

A Grids method was generating computer-animations in 2005. It uses active sonar. It works like this:

First, suppose there is a suitable notion of distance defined on 6-Space and Er() is a function such that for points X and Y in 6-Space, Er(X,Y) is the distance from X to Y.

A Grid is superimposed on the space in front of Device. Before being used for imaging, for each sensor M and for each gridpoint P is pre-computed GridTime(P,M), the sum of the times-of-travel of a fang from Clicker to P and from P to M. So, for every grid-point P and every scopion S there is GridKey(P,S) which is a six-term sequence t = (t1, t2, ...,t6) such that tK = GridTime(P,SK) for K= 1 to 6. The sequence t is a point in the above-mentioned 6-Space.

In a click event, for every scopion S, if T1 is a toa on S1, T2 is a toa on S2,..., and T6 is a toa on S6 then the six-term sequence (T1,T2,...,T6) is an "EchoKey" of S for that click-event.

The distance-function Er( ) can be used like this: in case, for some scopion S, each of X and Y is an EchoKey of S then Er(X,Y) can be regarded as the error in taking Y to represent X.

Suppose a click-event has produced recorded echoes on all the sensors of Device and FeatureDetector has sent, for every sensor M, a list of toas of fangs on M to HitDetector.

HitDetector does this: For every Scopion S, for every EchoKey T of S (computed by a subroutine EchoKeyDetector), Grid is searched for a "grid-hitpoint" P that minimizes Er(T,GridKey(P,S)).

The set of all the grid-hitpoints P could used to compute an image, but, there can be this problem: it may be, in terms of computational resources, impossible to make the Grid fine enough for acceptable quality of such images.

So, a second searching procedure moves P off of the Grid and into uncharted space so that Er(T, P) is reduced further.

If the error fails to fall below some maxError2, then T is classified as a phantom. Earlier, a similar check against some maxError1, used in the search on the Grid could have, in some cases, reached the same conclusion and saved the time of the second search.

In simulations, a Grids method seemed, by 2005, adequate for low-resolution applications such as “fish-finding.”

The methods of EchoKeyDetection and many other details omitted here exist on the Internet. The search expression "mystery of the sonar of dolphins" finds them efficiently. The original website is DolphinInspiredSonar.com.

More recently, a Surfaces method, “feature-based passive” (FPB) sonar, uses hyperboloids, to detect prospective hitpoints and checks those points using a separate procedure. Feature-based passive (FBP) sonar uses the known shape of a feature of the click but does not use the time or the place of emission of a click.

Simulations suggest that FBP, applied to the output of a square array of just nine hydrophones, each just 0.5 inch from its nearest neighbor, can compute images, of low resolution, of an object that a dolphin is clicking on.