Good morning ladies and gentlemen

Thank you for this given opportunities, I would like to present my paper about Underwater Source Separation Using Multistage Independent Component analysis in semi-anechoic water tank

Ok, lets begin from the background of this experiment, start from this picture

Slide 2

When we have some underwater acoustic experiment on open sea, this is what we will face, a huge noise environment as shown in this picture. Our sensors will capture all this kind of noises. It is quite difficult to separate our purposes of experiment while we can’t control this intermittent variables.

Slide 3

But it is not stop there, we will face another difficulties, Imagine when we have to repeat the experiment in different time, or different place, this is what we will get.

Slide 4

We will have different physical properties, different salinity, different depths, different temperature, and since that physical properties affected the propagation of sound in the water, we will find more difficulties to deducted our experiment

Slide 5

So to overcome this problem we capture small amount of this sea and put it in some water tank covered with absorber to reduce the sound reflection.

By using this method, we will have advantages, first we can manage all this physical properties like salinity and temperature, second the experiment will be much cost effective, since we can simulate and optimizing our experiment in lab scale before bring that to the open sea.

Slide 6

After we finish with the experiment place, then I will continue to the methods of source separation

Slide 7

When I made this experiment I want to use some methods of source separation which can minimize the dependencies to physical variables of the medium. On that time I found that this method was interesting. Since Blind source separation BSS/ ICA could reconstructed and estimated sources signal without any prior information of the sources itself.

The process of sound mixing is modelling by this equation (X=A.S). X is vector of sources, A is mixing matrix, and S is vector of our recorded signal. Our jobs here was to estimate the inverse of mixing matrix, so we can decomposed the original signal from our recorded one. Anyway this math model was called Linear mixture.

Slide 8

There are two important assumptions that have to be taken for performing ICA, first the sources signal have to be statistically independent each other. One of the approaches to determine this feature is through the covariance of each signal.

Second, distribution of signal have to be non-gaussian. We can use kurtosis value to measure non-gaussianity of the signal.

It is important to select the signal to be fulfill this 2 assumption before have the experiment.

Slide 9

In this paper we apply 3 different methods of ICA, which is TDICA, FDICA, and MSICA.

TDICA Is just like casual fast ICA, we estimate demixing matrix (W) right after we got the recorded signal

Slide 10

In FDICA we transformed the recorded signal into frequency domain and assemble the time series signal in each frequency bin. After that we estimated demixing signal and and acquired the estimated source signal in frequency domain. And last we transformed it back to time domain signal.

Slide 11

Next method is MSICA, in this method we combine two methods of TDICA and FDICA and performed it in sequence, that’s why this method was called multistage. First we estimated the signal in frequency domain through FDICA, then the output will come to TDICA, and we will acquired final estimated signal.

Slide 12

So the 2 boxes have completed, we have place of experiment in water tank, and we can evaluate three different method with controlled variable.

Slide 13

And this is the experiment setup, the tank dimension was 2m x 1m x 1 m 2 sensors and 2 sources are set to place as shown on the picture. The distance of between sensors and sources was vary from 85 to 150 (x values of this picture).

Slide 14

This is list of signals that used for the experiment. Each of the signals represent narrow band, wideband, and non-stationer signal. One of the assumption that had to be taken in ICA is non-gaussian signal and independent signal. Each of the signal have been choose to fulfill this assumption.

Slide 16

This is the result of TDICA methods. The quality of source separation represent by the MSE value, which measures the similarity of estimated signal to its original signal. This graph shown that the separation performance was increased as the distance between sources and sensors increased. The ping and ship combination shows the best source separation performances.

Slide 17

This picture shows one of the separation result, the first row is Time-freq representation of signal before mixing, second row shows the recorded signal, and third row shows the result of estimated signal. we can see in third row that the signal still not decomposed perfectly, but each of the element got amplified and the other get diminished, although it is still not perfect.

Slide 18

In FDICA we found that the result of separation shows more stable performances. It is not affected by the addition of the distances, although the lowest MSE value also occurred in the ping-ship combination signal

Slide 19

This is MSICA result. And it is show quite similar result to FDICA

Slide 20

If we combine all the result and put it in one graph, we have this picture. By this graph it is shown clearly that the performances of TDICA is decaying as the distances increased, while the other 2 methods shows more robust performances.

Slide 21

There are two explanation of this event. First, the performance of source separation decrease as the intensity of signal get attenuated by propagation. Second, as the distance increase, the ratio of indirect sound to direct sound got bigger due to the scattering. Since the FDICA and MSICA shown stable performances, first assumption got weak, and second assumption is more dependable.

Conclusion

Additional

Maximum likelihood is a method for estimating statistical model of given observation. So its just like fitting some distribution model of acquired data.

We got the data, and we need to estimate the other data that probably appear. Then we have to estimate the distribution of data that may appear, take some parameter of data that already acquired and, put it on likelihood function, after that we will have some function to estimate the data that may be acquired if we take some measurement

Goal of the maximum likelihood estimation:

1. Determine best model parameters (reality) that fit given data
2. Compare multiple models to determine the best fit to data

What it does:

1. Maximizes log-likelihood function to estimate parameters
2. Uses information theory to compare model fits

Maximum of likelihood function is difficult in practice, so the method maximizes log-likelihood instead

Procedure:

1. Collect data (x1,…,xn)
2. Assume what distribution it is, each distribution have its own parameter
3. Maximize the parameter (finding maximum likelihood estimate) through the likelihood function
4. Use log likelihood function, as log is monotonical function
5. Take derivative and set to zero, so we find max likelihood estimator