General Blind Source Signal Separation

Problem Statement

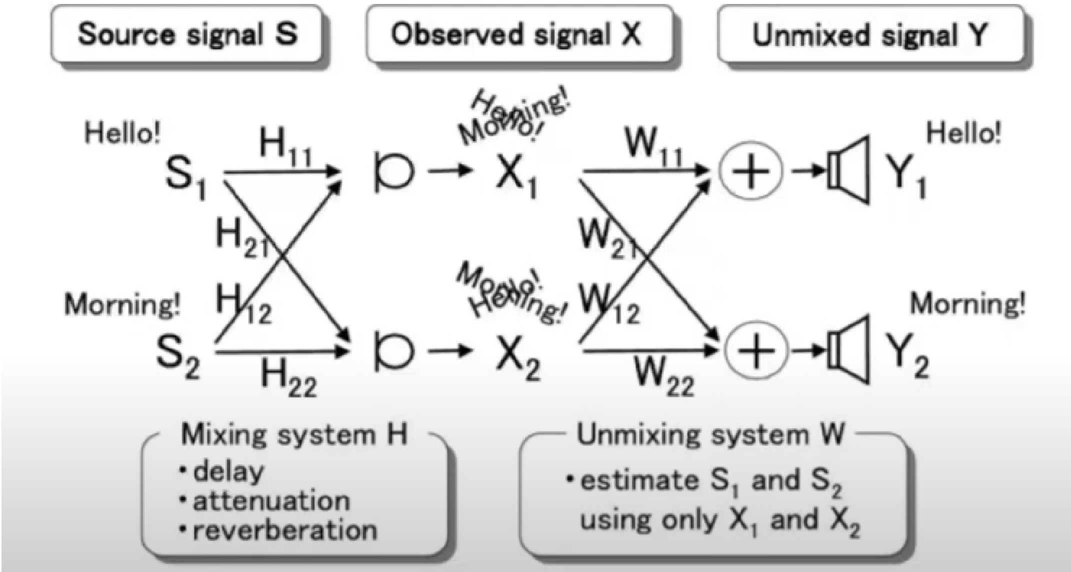
Find the separation of a set of source signals from a given set of mixed signals.

The mixed signals is represented as a column vector x ; $x \in R^n$, $n \in N$.

You have no information about how the mix happened between the sources, nor do you know any information about the source.

You can assume the following:

- The number of independent sources is equal to the number of mixtures (the number of independent sources is also n).
- Independent sources are non-gaussian signals and that they are mutually independent.


Source signal S
Observed signal X
Unmixed signal Y
Hello!
Morning!
S_1
S_2
H_{11}
H_{21}
H_{12}
H_{22}
X_1
X_2
W_{11}
W_{21}
W_{12}
W_{22}
Y_1
Y_2
Hello!
Morning!
Mixing system H
• delay
• attenuation
• reverberation
Unmixing system W
• estimate S_1 and S_2 using only X_1 and X_2

Encoding function

we have n independent sources $s_1(t)\ s_2(t)...\ s_n(t)$ represented as a vector $s(t)\ =\ [s_1(t)\ s_2(t)...\ s_n(t)]^T$.

$A;\ A\ \in\ R^{n\times n}$, is called the mixing matrix, represents the mixture coefficients where $-\ 5.0\ \le a_{ij} \le 5.0$ for $1\ \le i, j\ \le n$ and $i,\ j\ \in N$.

$$x(t)\ =\ As(t)\ ,$$

To find the source signals, we want to optimize the following equation,

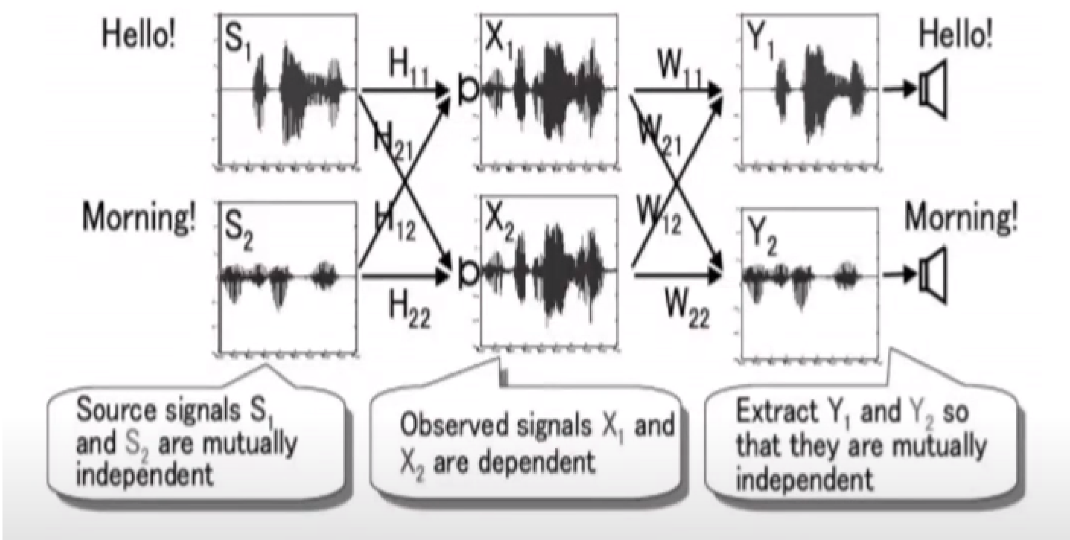
$$y(t)\ =\ Bx(t)\ ,$$

$y(t)\ =\ [y_1(t)\ y_2(t)...\ y_n(t)]^T$

$B\ \in\ R^{nxn}, -\ 5.0\ \le b_{ij} \le 5.0$

B as a chromosome, we can just flatten the matrix to a one dimensional vector, where each gene correspond to an entry in the matrix B, if $n\ =\ 2$ then a chromosome for B is

$b_{1,1}$	$b_{1,2}$	$b_{2,1}$	$b_{2,2}$



Selection method

Initial population

We choose the size of the population, let it be k, then we generate k random candidate solutions (chromosomes).

Selection

"using the helpful parts of biological evolution, and replacing the parts that might not work for us"

Tournament selection.

We choose the size of the tournament that is 20% of k, which is the size of the population, and round it to the nearest power of two.

We choose the probability p *such that* $0 < p < 1$ of each selection, and distribute the probabilities for each position in the tournament to be chosen as follows,

$p * ((1 - p)^{\alpha_i})$; α_i *is the fitness for chromosome* i

This keeps running until we have our population complete (number of selection = k), then the selected parents are passed to crossover and mutate.

Crossover

- Multiline crossover
- Multiline with random crossover

Mutation

We chose a mutation rate of 0.01, where we change a certain gene to a random value m; $-5 \leq m \leq 5$

Fitness function

"The nearer the individual to the solution of the optimization problem, the more fit is the individual"

maximization of the kurt sum, minimize the mutual information

- Kurt is used to show the independency of the components of the source signals.
- Mutual information (MI) is used to show how much common information between two random variables X and Y.

"In optimization, it is more convenient to solve minimization problems. However, the opposite perspective would be valid, too."

$let\ M\ \in\ R^{n\times n},\ n\ \in\ N$, and it's entries are the $MI(y_i,\ y_j)$

$$f(y)\ =\ \sum_{i=1}^{n}\left|kurt(y_i)\right|\ +\ \sum_{i=1,\ j=1,\ i!=j}^{n} 100e^{-10m_{ij}},\ for\ all\ m_{ij}\ <\ 0.2$$