

# ANÁLISE DE COMPONENTES INDEPENDENTES (ICA) E SUA APLICAÇÃO EM UMA SITUAÇÃO PRÁTICA



Capitão de Corveta Leonardo Cesar Cassiles de Souza

## 1 INTRODUÇÃO

Em diversos problemas de processamento de sinais é desejável encontrar uma transformação dos dados de modo que sua estrutura seja mais acessível. Na grande maioria dos casos, não existem muitas informações a respeito dos dados e, portanto, o aprendizado deve ser efetuado de forma cega.

Quando o meio de propagação e os sinais originais são desconhecidos e apenas os sinais recebidos são medidos, assumimos que o conjunto de sinais medidos é formado por uma combinação linear de fontes desses sinais. O objetivo do método é recuperar essas fontes utilizando, para isso, somente os sinais medidos.

Em [2] define-se que a combinação linear das fontes é provocada por uma matriz de mistura. Para melhor entendimento, segue um exemplo na forma matricial, sendo  $A$  a matriz de mistura,  $x$  os sinais medidos e  $s$  as fontes de sinais:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mathbf{X} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \quad \mathbf{x} = \mathbf{A}\mathbf{s}$$

Logo, conseguimos um modelo inverso para chegarmos às fontes por meio dos sinais medidos. Ou seja, a combinação linear dos sinais medidos é provocada por uma matriz de separação e, desta forma, chegamos às fontes. Chamamos a matriz de separação de  $W$ , sendo  $W = A^{-1}$  e  $N$  o número de fontes.

$$\mathbf{s} = \mathbf{W}\mathbf{x} \quad s_i = \sum_{j=1}^N (w_j x_j)$$

O desafio proposto pela Análise de Componentes Independentes, ou ICA, é encontrar, por meio de algoritmos matemáticos que envolvem Estatística, a matriz de separação, por meio da qual é possível descobrir as fontes dos sinais originais [1] e [2].

## 2 APLICAÇÃO DA TÉCNICA

Consideramos, para este trabalho, um conjunto de dados, fornecido pelo IPqM. Os dados foram coletados utilizando um arranjo cilíndrico de hidrofones (CHA) com 32 staves, sendo cada stave com três hidrofones em série, conforme mostrado na Figura 1.



Figura 1: CHA com 32 staves e 96 hidrofones, utilizado pelo IPqM.

Esses dados se referem ao monitoramento do cenário do tráfego marítimo próximo à Escola Naval.

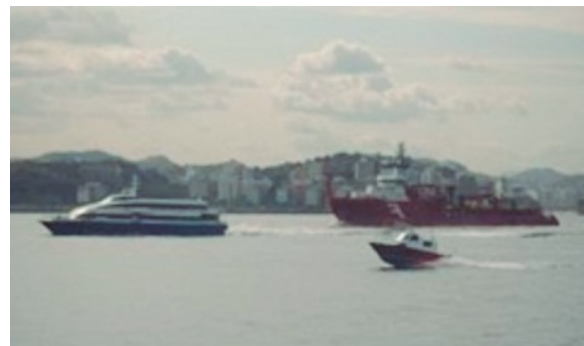


Figura 2: Fotografia no momento da gravação dos dados com 3 contatos a serem analisados.

Para comprovar a eficiência da técnica, comparamos a análise espectral do sinal após ser processado pelo algoritmo FastICA [1] e [2] com a análise espectral do sinal após a aplicação da técnica de estimação da direção de chegada, conhecida como “Atraso e Soma” [4]. Esta é uma técnica de conformação de sinais amplamente conhecida e difundida no meio científico [1]. Por meio dela, identificamos a região de interesse através de um gráfico Marcação x Tempo, apresentado na Figura 3. Nesta Figura, o eixo horizontal representa a marcação, em graus, e o eixo vertical representa o tempo, em segundos. A energia em cada marcação tem uma representação em escala de cores. Quanto maior for a tonalidade amarela, maior será a probabilidade de um alvo estar presente naquela marcação. Com isto, é possível fazer uma observação visual da variação da direção do contato.

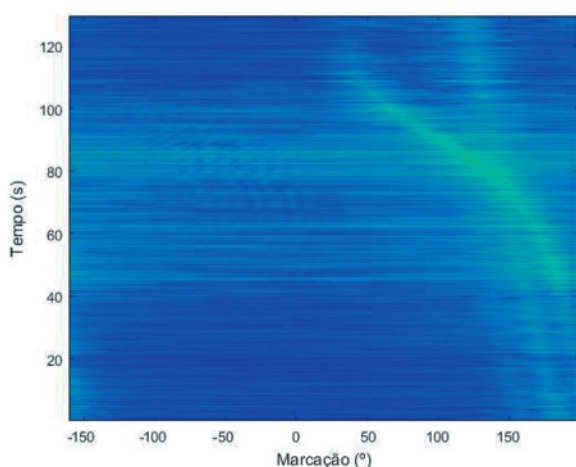


Figura 3: Gráfico Marcação x Tempo do conjunto de dados.

Na Figura 3, identificamos duas marcações de interesse para as análises. A marcação 150°, que representa a região onde ocorre maior permanência das embarcações e um provável cruzamento entre as mesmas, e a marcação 065°, que representa a região onde um dos contatos se isola. Sua análise é importante para tentar estimar frequências e rotações desse contato.

Realizamos, então, as análises espectral, LOFAR (Low Frequency Analysis Recording) e DEMON (Detection Envelope Modulation On Noise) [5] do sinal obtido no conjunto de dados, nas marcações 150° e 065°.

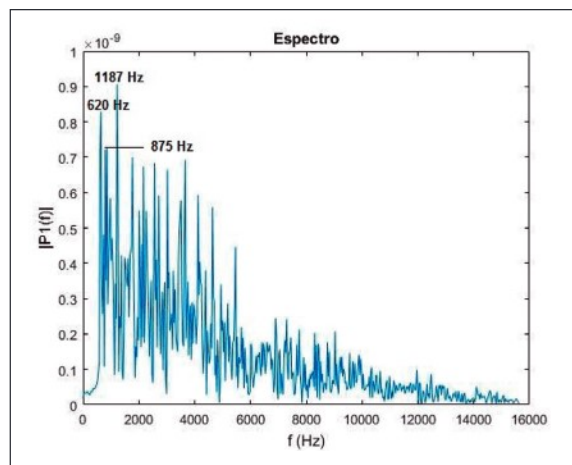


Figura 4: Análise espectral do sinal na marcação 150°.

Na Figura 4, observamos três tons característicos de eixo constante em 620 Hz, 875 Hz e 1187 Hz. As frequências mais elevadas, 1187 e 620 Hz, podem estar relacionadas aos contatos que permanecem próximos à marcação 150° durante todo o período de aquisição dos dados.

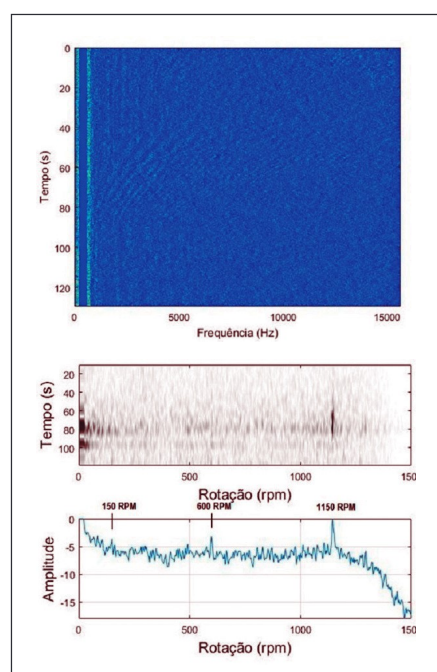


Figura 5: Análises LOFAR e DEMON do sinal na marcação 150°.

Por meio da análise DEMON, na Figura 5, é possível identificar três rotações que se destacam: 150, 600 e 1150 RPM, caracterizando os três possíveis contatos apresentados na Figura 2.

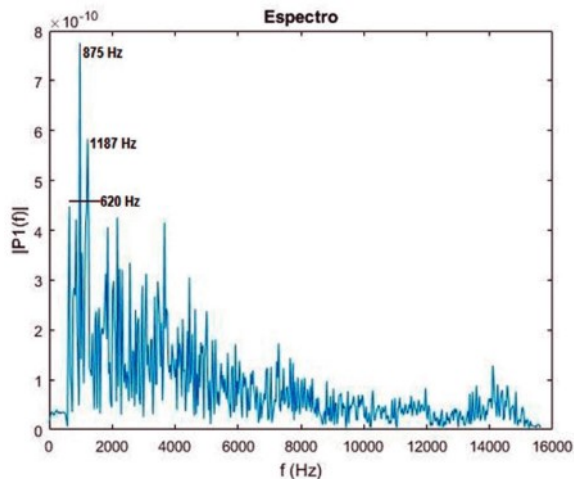


Figura 6: Análise espectral do sinal na marcação 065°.

Na Figura 6, observamos os mesmos três tons característicos. Porém, há uma variação de intensidade da densidade espectral. Na marcação 065°, a frequência de 875 Hz passa a ter maior densidade espectral. Esse fato indica que essa frequência pode estar relacionada com o contato que se isolou nesta marcação.

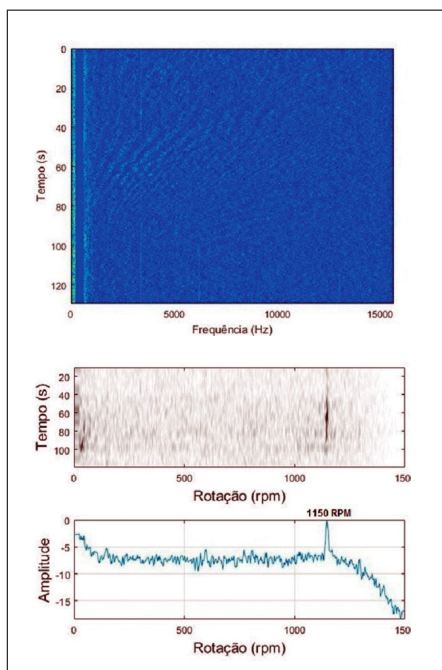


Figura 7: Análises LOFAR e DEMON do sinal na marcação 065°.

Na Figura 7, apenas a rotação de 1150 RPM permanece com elevada amplitude do sinal. As outras duas rotações ainda permanecem características, contudo um pouco mais atenuadas. Esse fato indica que a rotação de 1150 RPM provavelmente se refere ao contato que se isolou na marcação 065°.

Em seguida, aplicamos o algoritmo FastICA nos sinais medidos no intuito de separar as fontes geradoras dos sinais. Consideramos os sinais medidos para compor a matriz de mistura como sinais conformados nas marcações 140°, 150° e 160°, a partir do conjunto de dados fornecido pelo IpqM.

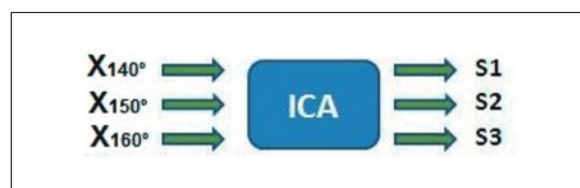


Figura 8: Esquema didático da aplicação do algoritmo FastICA.

Após separarmos as fontes, realizamos a análise espectral juntamente com as análises LOFAR e DEMON. O resultado é, então, comparado com a análise anterior, comprovando a eficiência da técnica de separação cega.

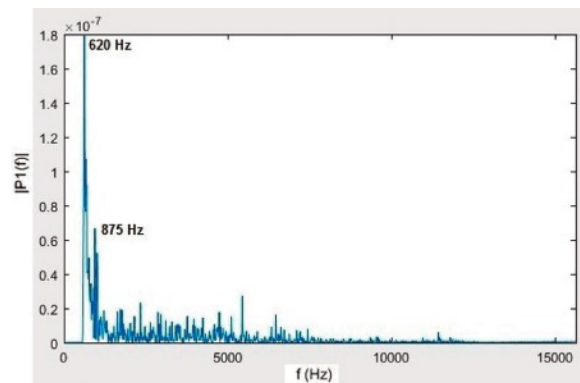


Figura 9: Análise espectral da fonte s1.

Por meio da Figura 9, percebemos que a frequência de 620 Hz se destacou sobre as demais, demonstrando que o algoritmo FastICA foi eficiente na separação cega.

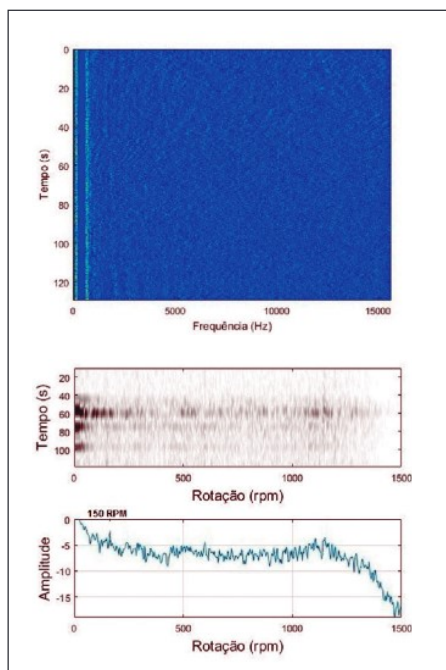


Figura 10: Análises LOFAR e DEMON da fonte s1.

Na análise DEMON, apresentada na Figura 10, também ficou evidente a melhora na identificação da fonte de sinal. A rotação de 150 RPM foi destacada e as demais foram atenuadas.

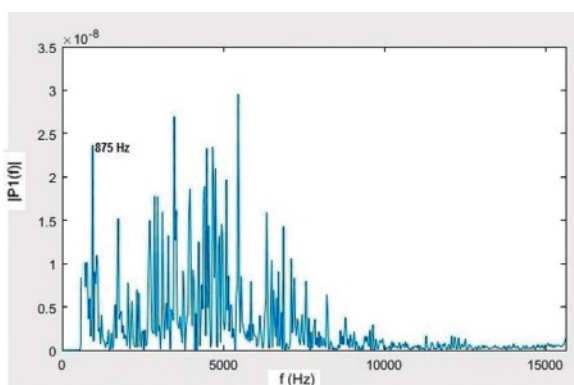


Figura 11: Análise espectral da fonte s2.

Após a separação cega, a análise espectral da segunda componente independente, apresentada na Figura 11, evidenciou a frequência de 875 Hz, além de outras frequências mais altas, proporcionando uma melhor relação sinal/ruído.

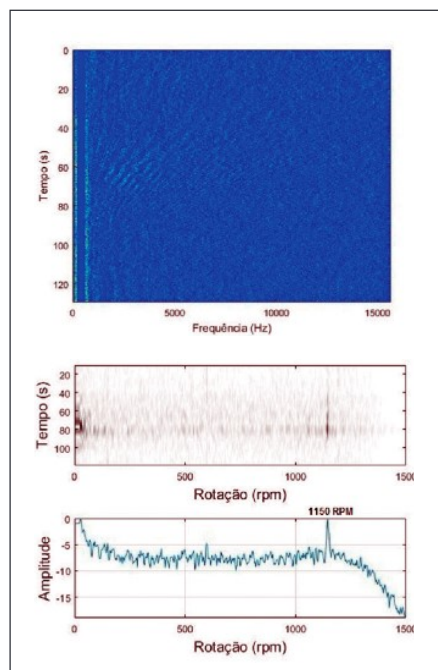


Figura 12: Análises LOFAR e DEMON da fonte s2.

Na análise DEMON da segunda componente independente, apresentada na Figura 12, percebemos claramente o destaque da rotação de 1150 RPM. Isso demonstra que o algoritmo FastICA atenuou as interferências e separou a fonte de maneira eficiente.

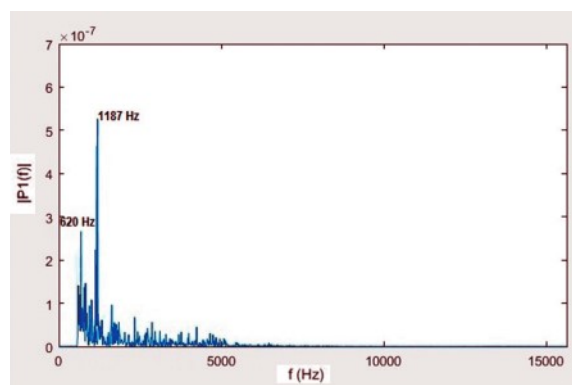


Figura 13: Análise espectral da fonte s3.

A Figura 13 apresentou a análise espectral da terceira componente independente, destacando a frequência de 1187 Hz. Neste caso, a aplicação do algoritmo FastICA também foi bem-sucedida, melhorando a relação sinal/ruído.

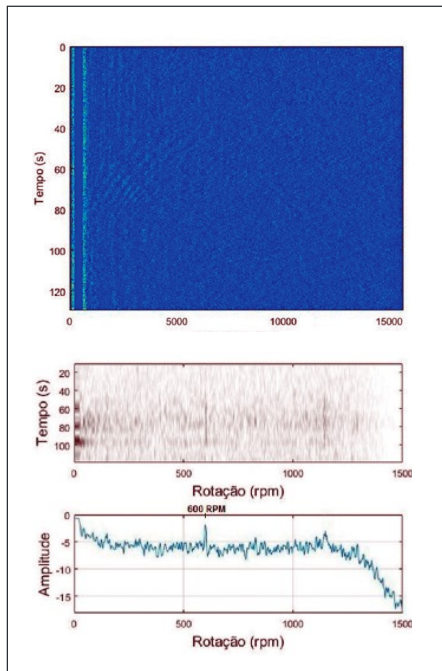


Figura 14: Análises LOFAR e DEMON da fonte s3.

Na análise DEMON da terceira componente independente, apresentada na Figura 14, ficou evidente que a rotação de 1150 RPM ficou mais destacada que as demais. Esse fato demonstra que o algoritmo FastICA também separou essa fonte de maneira eficiente.

### 3 CONCLUSÃO

A Análise de Componentes Independentes (ICA) tem se mostrado uma técnica muito eficiente para a separação cega de sinais. Sua utilização se reveste de importância pelo fato de que todos os sinais recebidos pelos diversos sistemas de sonares estão suscetíveis a interferências das mais variadas fontes acústicas e do ruído ambiente, dificultando a classificação e identificação dos alvos. Com isso, torna-se imperiosa a busca da recuperação das fontes de sinais de interesse.

Neste trabalho, foi utilizada a ICA, especificamente o algoritmo FastICA, para a realização da separação das fontes, a partir do sinal misturado fornecido pelo IPqM.

Os resultados obtidos após a aplicação da técnica tiveram um bom desempenho para os dados do conjunto. Como foi fornecida uma fotografia do momento da aquisição dos dados, verificou-se a existência de um tráfego de, pelo menos, três contatos. Portanto,

para a aplicação do algoritmo FastICA, definimos uma matriz de mistura com três sinais conformados na região de interesse. Como resultado da separação cega, obtivemos três componentes independentes do sinal, que correspondem às três fontes especificadas no ambiente de gravação dos sinais. A partir da separação dos contatos, todas as análises foram melhoradas, comprovando a eficiência do algoritmo.

Futuros trabalhos poderão ser realizados usando outros algoritmos de separação cega de fontes com o objetivo de comparar com o algoritmo FastICA que foi implementado neste trabalho.

### REFERÊNCIAS

DE MOURA, Natanael Nunes. Detecção e classificação de sinais de sonar passivo usando métodos de separação cega de fontes. 2013. Tese de Doutorado. Universidade Federal do Rio de Janeiro.

HYVÄRINEN, Aapo; KARHUNEN, Juha; OJA, Erkki. Independent component analysis. John Wiley & Sons, 2004.

JOHNSON, D. H., DUDGEON, D. E. Array signal processing: Concepts and Techniques. Prentice Hall Signal Processing Series, 1993.

LIU, Wei; WEISS, Stephan. Wideband beamforming: concepts and techniques. John Wiley & Sons, 2010.

NIELSEN, R. O. Sonar signal processing. Artech House Inc, Nortwood, 1991.