



República Federativa do Brasil  
Ministério da Economia  
Instituto Nacional da Propriedade Industrial

**(11) BR 112014003663-2 B1**



**(22) Data do Depósito: 14/08/2012**

**(45) Data de Concessão: 21/12/2021**

**(54) Título:** MATRIZES DE MIXAGEM IDEAL E USO DE DESCORRELACIONADORES NO PROCESSAMENTO DE ÁUDIO ESPACIAL

**(51) Int.Cl.:** G10L 19/00.

**(30) Prioridade Unionista:** 17/08/2011 US 61/524,647; 21/02/2012 EP 12156351.4.

**(73) Titular(es):** FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E. V..

**(72) Inventor(es):** JUHA VILKAMO; TOM BAECKSTROEM; FABIAN KUCH; ACHIM KUNTZ.

**(86) Pedido PCT:** PCT EP2012065861 de 14/08/2012

**(87) Publicação PCT:** WO 2013/024085 de 21/02/2013

**(85) Data do Início da Fase Nacional:** 17/02/2014

**(57) Resumo:** MATRIZES DE MIXAGEM IDEAL E USO DE DESCORRELACIONADORES NO PROCESSAMENTO DE ÁUDIO ESPACIAL Um aparelho para gerar um sinal de saída de áudio com dois ou mais canais de saída de áudio de um sinal de entrada de áudio com dois canais de entrada de áudio ou mais é fornecido. O aparelho compreende um provedor (110) e um processador de sinal (120) . O provedor (110) é adaptado para fornecer as primeiras propriedades de covariância do sinal de entrada de áudio. O processador de sinal (120) é adaptado para gerar o sinal de saída de áudio aplicando uma regra de mixagem em, pelo menos, dois de dois canais de entrada de áudio ou mais. O processador de sinal (120) é configurado para determinar a regra de mixagem com base nas primeiras propriedades de covariância do sinal de entrada de áudio e com base nas segundas propriedades de covariância do sinal de saída de áudio, as segundas propriedades de covariância diferentes das primeiras propriedades de covariância

"MATRIZES DE MIXAGEM IDEAL E USO DE  
DESCORRELACIONADORES NO PROCESSAMENTO DE ÁUDIO ESPACIAL".

DESCRIÇÃO

A presente invenção refere-se ao processamento do  
5 sinal de áudio e, em particular, a um aparelho e um método que  
emprega matrizes de mixagem ideal e, além disso, ao uso de  
descorrelacionadores no processamento de áudio espacial.

O processamento de áudio se torna mais e mais  
importante. No processamento perceptual de áudio espacial, uma  
10 suposição típica é a de que o aspecto espacial de um som  
reproduzido por um alto-falante é determinado especialmente pelas  
energias e as dependências alinhadas pelo tempo entre os canais de  
áudio nas faixas de frequência perceptuais. Isso é encontrado na  
observação que estas características, quando reproduzidas nos  
15 alto-falantes, transferem em diferenças de nível interaural,  
diferenças de tempo interaural e coerências interaurais, que são  
os sinais binaurais da percepção espacial. A partir deste  
conceito, vários métodos de processamento espaciais surgiram,  
incluindo *upmixing*, vide

20 [1] C. Faller, "Multiple-Loudspeaker Playback of  
Stereo Signals", Journal of the Audio Engineering Society, Vol.  
54, No. 11, pp. 1051-1064, June 2006,

e microfonia espacial, vide, por exemplo,

25 [2] V. Pulkki, "Spatial Sound Reproduction with  
Directional Audio Coding", Journal of the Audio Engineering  
Society, Vol. 55, No. 6, pp. 503-516, June 2007; e

[3] C. Tournery, C. Faller, F. Küch, J. Herre,  
"Converting Stereo Microphone Signals Directly to MPEG Surround",

tipicamente sinais do microfone de primeira ordem, que são por meios da mixagem, posição da amplitude e decorrelação processada para perceptualmente aproximar um campo de som medido. No *upmixing* (vide [1]), os canais de entrada estéreo são, novamente, como 5 função de tempo e frequência distribuídos de forma adaptativa a uma configuração surround.

É um objetivo da presente invenção fornecer conceitos melhorados para gerar a partir de um conjunto de canais de entrada um conjunto de canais de saída com propriedades 10 definidas. O objetivo da presente invenção é solucionado por um aparelho de acordo com a reivindicação 1, por um método, de acordo com a reivindicação 25 e por um programa de computador, de acordo com a reivindicação 26.

Um aparelho para gerar um sinal de saída de áudio 15 com dois ou mais canais de saída de áudio a partir um sinal de entrada de áudio com dois ou mais canais de entrada de áudio é fornecido. O aparelho compreende um provedor e um processador de sinal. O provedor é adaptado para fornecer primeiras propriedades de covariância do sinal de entrada de áudio. O processador de 20 sinal é adaptado para gerar o sinal de saída de áudio aplicando uma regra de mixagem em, pelo menos, dois dos dois canais de entrada de áudio ou mais. O processador de sinal é configurado para determinar a regra de mixagem com base nas primeiras propriedades de covariância do sinal de entrada de áudio e com 25 base nas segundas propriedades de covariância do sinal de saída de áudio, as segundas propriedades de covariância diferentes das primeiras propriedades de covariância.

Por exemplo, as energias do canal e as

dependências alinhadas pelo tempo podem ser expressas pela parte real da matriz de covariância do sinal, por exemplo, nas faixas de frequência perceptuais. A seguir, um conceito geralmente aplicável para processar o som espacial neste domínio é apresentado. O 5 conceito compreende uma solução de mixagem adaptativa para atingir dadas propriedades de covariância alvo (as segundas propriedades de covariância), por exemplo, uma dada matriz de covariância alvo, pelo melhor uso dos componentes independentes nos canais de entrada. Em uma aplicação, meios podem ser fornecidos para injetar 10 a quantidade necessária da energia de som descorrelacionada, quando o alvo não é obtido de outra forma. Este conceito é robusto em sua função e pode ser aplicado em vários casos de uso. As propriedades de covariância alvo podem, por exemplo, ser fornecidas por um usuário. Por exemplo, um aparelho de acordo com 15 uma utilização pode ter meios de modo que um usuário possa inserir as propriedades de covariância.

De acordo com uma aplicação, o provedor pode ser adaptado para fornecer as primeiras propriedades de covariância, caracterizado pelas primeiras propriedades de covariância terem um 20 primeiro estado para uma primeira posição de tempo-frequência e, em que as primeiras propriedades de covariância têm um segundo estado, diferente do primeiro estado, para uma segunda posição de tempo-frequência, diferente da primeira posição de tempo-frequência. O provedor não necessariamente precisa realizar a 25 análise para obter as propriedades de covariância, mas pode fornecer estes dados de um armazenamento, uma entrada do usuário ou de fontes semelhantes.

Em outra aplicação, o processador de sinal pode

ser adaptado para determinar a regra de mixagem com base nas segundas propriedades de covariância, em que as segundas propriedades de covariância têm um terceiro estado para uma terceira posição de tempo-frequência e, em que as segundas 5 propriedades de covariância têm um quarto estado, diferente do terceiro estado para uma quarta posição de tempo-frequência, diferente da terceira posição de tempo-frequência.

De acordo com outra aplicação, o processador de sinal é adaptado para gerar o sinal de saída de áudio aplicando a 10 regra de mixagem de modo que cada um de dois ou mais canais de saída de áudio depende de cada um de dois ou mais canais de entrada de áudio.

Em outra aplicação, o processador de sinal pode ser adaptado para determinar a regra de mixagem de modo que uma 15 medição de erro seja reduzida. Uma medição de erro pode, por exemplo, ser um sinal de diferença absoluta entre um sinal de saída de referência e um sinal de saída real.

Em uma aplicação, uma medição de erro pode, por exemplo, ser uma medição dependendo de

$$20 \quad ||\mathbf{y}_{\text{ref}} - \mathbf{y}||^2$$

caracterizado por  $\mathbf{y}$  ser o sinal de saída de áudio, em que

$$\mathbf{y}_{\text{ref}} = \mathbf{Q}\mathbf{x},$$

em que  $\mathbf{x}$  especifica o sinal de entrada de áudio e 25 em que  $\mathbf{Q}$  é uma matriz de mapeamento, que pode ser específica da aplicação, de modo que  $\mathbf{y}_{\text{ref}}$  especifica um sinal de saída de áudio alvo de referência.

De acordo com outra aplicação, o processador de

sinal pode ser adaptado para determinar a regra de mixagem de modo que

$$e = E [\|\mathbf{y}_{\text{ref}} - \mathbf{y}\|^2]$$

seja reduzido, caracterizado por  $E$  ser um

5 operador de expectativa, em que  $\mathbf{y}_{\text{ref}}$  é um ponto de referência definido e em que  $\mathbf{y}$  é o sinal de saída de áudio.

De acordo com outra aplicação, o processador de sinal pode ser configurado para determinar a regra de mixagem determinando as segundas propriedades de covariância, 10 caracterizado pelo processador de sinal poder ser configurado para determinar as segundas propriedades de covariância com base nas primeiras propriedades de covariância.

De acordo com outra aplicação, o processador de sinal pode ser adaptado para determinar uma matriz de mixagem como 15 a regra de mixagem, caracterizado pelo processador de sinal poder ser adaptado para determinar a matriz de mixagem com base nas primeiras propriedades de covariância e com base nas segundas propriedades de covariância.

Em outra aplicação, o provedor pode ser adaptado 20 para analisar as primeiras propriedades de covariância determinando uma primeira matriz de covariância do sinal de entrada de áudio e caracterizado pelo processador de sinal poder ser configurado para determinar a regra de mixagem com base em uma segunda matriz de covariância do sinal de saída de áudio como as 25 segundas propriedades de covariância.

De acordo com outra aplicação, o provedor pode ser adaptado para determinar a primeira matriz de covariância de modo que cada valor da diagonal da primeira matriz de covariância

possa indicar uma energia de um dos canais de entrada de áudio e de modo que cada valor da primeira matriz de covariância que não é um valor da diagonal possa indicar uma correlação intercanal entre um primeiro canal de entrada de áudio e um segundo canal de 5 entrada de áudio diferente.

De acordo com outra aplicação, o processador de sinal pode ser configurado para determinar a regra de mixagem com base na segunda matriz de covariância, caracterizado por cada valor da diagonal da segunda matriz de covariância poder indicar 10 uma energia de um dos canais de saída de áudio e em que cada valor da segunda matriz de covariância que não é um valor da diagonal possa indicar uma correlação intercanal entre um primeiro canal de saída de áudio e um segundo canal de saída de áudio.

De acordo com outra aplicação, o processador de 15 sinal pode ser adaptado para determinar a matriz de mixagem de modo que:

$$\mathbf{M} = \mathbf{K}_y \mathbf{P} \mathbf{K}_x^{-1}$$

de modo que

$$\mathbf{K}_x \mathbf{K}_x^T = \mathbf{C}_x$$

$$\mathbf{K}_y \mathbf{K}_y^T = \mathbf{C}_y$$

20 caracterizado por  $\mathbf{M}$  ser a matriz de mixagem, em que  $\mathbf{C}_x$  é a primeira matriz de covariância, em que  $\mathbf{C}_y$  é a segunda matriz de covariância, em que  $\mathbf{K}_x^T$  é uma primeira matriz transposta de uma primeira matriz decomposta  $\mathbf{K}_x$ , em que  $\mathbf{K}_y^T$  é uma segunda matriz transposta de uma segunda matriz decomposta  $\mathbf{K}_y$ , em que  $\mathbf{K}_x^{-1}$  25 é uma matriz inversa da primeira matriz decomposta  $\mathbf{K}_x$  e em que  $\mathbf{P}$  é uma primeira matriz unitária.

Em outra aplicação, o processador de sinal pode

ser adaptado para determinar a matriz de mixagem de modo que

$$\mathbf{M} = \mathbf{K}_y \mathbf{P} \mathbf{K}_x^{-1}$$

caracterizado por

$$\mathbf{P} = \mathbf{V} \mathbf{U}^T$$

5 em que  $\mathbf{U}^T$  é uma terceira matriz transposta de uma segunda matriz unitária  $\mathbf{U}$ , em que  $\mathbf{V}$  é uma terceira matriz unitária, em que

$$\mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y$$

em que  $\mathbf{Q}^T$  é uma quarta matriz transposta da 10 matriz de *downmix*  $\mathbf{Q}$ , em que  $\mathbf{V}^T$  é uma quinta matriz transposta da terceira matriz unitária  $\mathbf{V}$  e, em que  $\mathbf{S}$  é uma matriz diagonal.

De acordo com outra aplicação, o processador de sinal é adaptado para determinar uma matriz de mixagem como a regra de mixagem, caracterizado pelo processador de sinal ser 15 adaptado para determinar a matriz de mixagem com base nas primeiras propriedades de covariância e com base nas segundas propriedades de covariância, em que o provedor é para fornecer ou analisar as primeiras propriedades de covariância determinando uma primeira matriz de covariância do sinal de entrada de áudio e, em 20 que o processador de sinal é configurado para determinar a regra de mixagem com base em uma segunda matriz de covariância do sinal de saída de áudio como as segundas propriedades de covariância, em que processador de sinal é configurado para modificar, pelo menos, alguns valores da diagonal de uma matriz diagonal  $\mathbf{S}_x$  quando os 25 valores da matriz diagonal  $\mathbf{S}_x$  são zero ou menores do que uma valor limite predeterminado, de modo que os valores sejam maiores do que ou iguais ao valor limite, em que o processador de sinal é adaptado para determinar a matriz de mixagem com base na matriz

diagonal. Entretanto, o valor limite não precisa ser necessariamente predeterminado, mas pode ainda depender de uma função.

Em outra aplicação, o processador de sinal é configurado para modificar, pelo menos, alguns valores da diagonal da matriz diagonal  $\mathbf{S}_x$ , caracterizado por  $\mathbf{K}_x = \mathbf{U}_x \mathbf{S}_x \mathbf{V}_x^T$  e, em que  $\mathbf{C}_x = \mathbf{K}_x \mathbf{K}_x^T$ , em que  $\mathbf{C}_x$  é a primeira matriz de covariância, em que  $\mathbf{S}_x$  é a matriz diagonal, em que  $\mathbf{U}_x$  é uma segunda matriz,  $\mathbf{V}_x^T$  é uma terceira matriz transposta e em que  $\mathbf{K}_x^T$  é uma quarta matriz transposta da quinta matriz  $\mathbf{K}_x$ . As matrizes  $\mathbf{V}_x$  e  $\mathbf{U}_x$  podem ser matrizes unitárias.

De acordo com outra aplicação, o processador de sinal é adaptado para gerar o sinal de saída de áudio aplicando a regra de mixagem em, pelo menos, dois dos dois canais de entrada de áudio ou mais para obter um sinal intermediário  $\mathbf{y}' = \hat{\mathbf{M}}\mathbf{x}$  e adicionando um sinal residual  $r$  ao sinal intermediário para obter o sinal de saída de áudio.

Em outra aplicação, o processador de sinal é adaptado para determinar a matriz de mixagem com base em uma matriz de ganho diagonal  $\mathbf{G}$  e uma matriz intermediária  $\hat{\mathbf{M}}$ , de modo que  $\mathbf{M}' = \mathbf{G}\hat{\mathbf{M}}$ , caracterizada pela matriz de ganho diagonal ter o valor

$$\mathbf{G}(i,i) = \sqrt{\frac{\mathbf{C}_y(i,i)}{\hat{\mathbf{C}}_y(i,i)}}$$

$$\text{onde } \hat{\mathbf{C}}_y = \hat{\mathbf{M}}\mathbf{C}_x\hat{\mathbf{M}}^T,$$

em que  $\mathbf{M}'$  é a matriz de mixagem, em que  $\mathbf{G}$  é a matriz de ganho diagonal e em que  $\hat{\mathbf{M}}$  é a matriz intermediária, em que  $\mathbf{C}_y$  é a segunda matriz de covariância e em que  $\hat{\mathbf{M}}^T$  é uma quinta

matriz transposta da matriz  $\hat{\mathbf{M}}$ .

As aplicações preferenciais da presente invenção serão explicadas com referência às figuras em que:

A Figura 1 ilustra um aparelho para gerar um sinal de saída de áudio com dois ou mais canais de saída de áudio de um sinal de entrada de áudio com dois ou mais canais de entrada de áudio de acordo com uma aplicação,

A Figura 2 descreve um processador de sinal de acordo com uma aplicação,

A Figura 3 mostra um exemplo para aplicar uma combinação linear de vetores  $\mathbf{L}$  e  $\mathbf{R}$  para atingir um novo conjunto de vetor  $\mathbf{R}'$  e  $\mathbf{L}'$ ,

A Figura 4 ilustra um diagrama em blocos de um aparelho de acordo com outra aplicação,

A Figura 5 mostra um diagrama que descreve um sinal de microfone de coincidência estéreo no codificador Surround MPEG de acordo com uma aplicação,

A Figura 6 descreve um aparelho de acordo com outra utilização referente ao nível de correção/ICC downmix para um codificador SAM-para-MPS,

A Figura 7 descreve um aparelho de acordo com uma utilização para uma melhoria para pequenas matrizes de microfone espaçadas,

A Figura 8 ilustra um aparelho de acordo com outra utilização para melhoria cega da qualidade do som espacial na reprodução estéreo ou multicanal,

A Figura 9 ilustra a melhoria das configurações estreitas do alto-falante,

A Figura 10 descreve um a utilização que fornece a interpretação da Codificação de Áudio Direcional com base em um sinal de microfone em formato B,

A Figura 11 ilustra a tabela 1 que mostra 5 exemplos numéricos de uma aplicação, e

A Figura 12 descreve a lista 1 que mostra uma implementação Matlab de um método de acordo com uma aplicação.

A Figura 1 ilustra um aparelho para gerar um sinal de saída de áudio com dois ou mais canais de saída de áudio 10 de um sinal de entrada de áudio com dois ou mais canais de entrada de áudio de acordo com uma aplicação. O aparelho compreende um provedor 110 e um processador de sinal 120. O provedor 110 é adaptado para receber o sinal de entrada de áudio com dois ou mais canais de entrada de áudio. Além disso, o provedor 110 é adaptado 15 para analisar as primeiras propriedades de covariância do sinal de entrada de áudio. O provedor 110 é ainda adaptado para fornecer as primeiras propriedades de covariância ao processador de sinal 120. O processador de sinal 120 é ainda adaptado para receber o sinal de entrada de áudio. O processador de sinal 120 é ainda adaptado 20 para gerar o sinal de saída de áudio aplicando uma regra de mixagem em, pelo menos, dois dos dois canais de entrada ou mais do sinal de entrada de áudio. O processador de sinal 120 é configurado para determinar a regra de mixagem com base nas primeiras propriedades de covariância do sinal de entrada de áudio 25 e com base nas segundas propriedades de covariância do sinal de saída de áudio, as segundas propriedades de covariância diferente das primeiras propriedades de covariância.

A Figura 2 ilustra um processador de sinal de

acordo com uma aplicação. O processador de sinal compreende uma unidade de formulação da matriz de mixagem ideal 210 e uma unidade de mixagem 220. A unidade de formulação da matriz de mixagem ideal 210 formula uma matriz de mixagem ideal. Para isso, a unidade de 5 formulação da matriz de mixagem ideal 210 utiliza as primeiras propriedades de covariância 230 (por exemplo, propriedades de covariância de entrada) de um sinal de entrada de áudio da faixa de frequência estéreo ou multicanal conforme recebido, por exemplo, por um provedor 110 da utilização da Figura 1. Além 10 disso, a unidade de formulação da matriz de mixagem ideal 210 determina a matriz de mixagem com base nas segundas propriedades de covariância 240, por exemplo, uma matriz de covariância alvo, que pode ser dependente do pedido. A matriz de mixagem ideal que é formulada pela unidade de formulação da matriz de mixagem ideal 15 210 pode ser utilizada como uma matriz de mapeamento de canal. A matriz de mixagem ideal pode então ser fornecida à unidade de mistura 220. A unidade de mistura 220 se aplica à matriz de mixagem ideal na entrada da faixa de frequência estéreo ou multicanal para obter uma saída da faixa de frequência estéreo ou 20 multicanal do sinal de saída de áudio. O sinal de saída de áudio tem as segundas propriedades de covariância desejadas (propriedades de covariância alvo).

Para explicar as aplicações da presente invenção em mais detalhes, as definições são introduzidas. Agora, os sinais 25 de entrada e saída complexos com média zero  $\mathbf{x}_i(t, f)$  e  $\mathbf{y}_j(t, f)$  são definidos, caracterizado por  $t$  ser o índice de tempo, em que  $f$  é o índice de frequência, em que  $i$  é o índice do canal de entrada e em que  $j$  é o índice do canal de saída. Além disso, os vetores do

sinal do sinal de entrada de áudio  $\mathbf{x}$  e do sinal de saída de áudio  $\mathbf{y}$  são definidos:

$$\mathbf{x}_{N_x}(t, f) = \begin{bmatrix} x_1(t, f) \\ x_2(t, f) \\ \vdots \\ x_{N_x}(t, f) \end{bmatrix} \quad \mathbf{y}_{N_y}(t, f) = \begin{bmatrix} y_1(t, f) \\ y_2(t, f) \\ \vdots \\ y_{N_y}(t, f) \end{bmatrix} \quad (1)$$

onde  $N_x$  e  $N_y$  são o número total de canais de entrada e saída. Além disso,  $N = \max(N_y, N_x)$  e sinais preenchidos por 0 de dimensão igual são definidos:

$$\begin{aligned} \mathbf{x}(t, f) &= \begin{bmatrix} \mathbf{x}_{N_x}(t, f) \\ \mathbf{0}_{(N-N_x) \times 1} \end{bmatrix} \\ \mathbf{y}(t, f) &= \begin{bmatrix} \mathbf{y}_{N_y}(t, f) \\ \mathbf{0}_{(N-N_y) \times 1} \end{bmatrix}. \end{aligned} \quad (2)$$

Os sinais preenchidos por zero podem ser utilizados na formulação até quando a solução derivada for estendida a diferentes comprimentos do vetor.

Conforme foi explicado acima, a medição amplamente utilizada para descrever o aspecto espacial de um som multicanal é a combinação das energias do canal e as dependências alinhadas pelo tempo. Estas propriedades são compreendidas na parte real das matrizes de covariância, definido como:

$$\begin{aligned} \mathbf{C}_x &= E[\operatorname{Re}\{\mathbf{x}\mathbf{x}^H\}] \\ \mathbf{C}_y &= E[\operatorname{Re}\{\mathbf{y}\mathbf{y}^H\}] \end{aligned} \quad (3)$$

Na equação (3) e a seguir,  $E[\cdot]$  é o operador de expectativa,  $\operatorname{Re}\{\cdot\}$  é o operador da parte real e  $\mathbf{x}^H$  e  $\mathbf{y}^H$  são as transposições conjugadas de  $\mathbf{x}$  e  $\mathbf{y}$ . O operador de expectativa  $E[\cdot]$  é um operador matemático. Nas aplicações práticas é substituído por uma estimativa como uma média sobre um determinado intervalo de tempo. Seguindo as seções, o uso do termo matriz de covariância se

refere a esta definição de valor real.  $\mathbf{C}_x$  e  $\mathbf{C}_y$  são simétricas e semidefinida positiva e, assim, as matrizes reais  $\mathbf{K}_x$  e  $\mathbf{K}_y$  podem ser definidas, de modo que:

$$\begin{aligned} \mathbf{C}_x &= \mathbf{K}_x \mathbf{K}_x^T \\ \mathbf{C}_y &= \mathbf{K}_y \mathbf{K}_y^T. \end{aligned} \quad (4)$$

[7] Golub, G.H. and Van Loan, C.F., "Matrix computations", Johns Hopkins Univ Press, 1996.

10 Deve ser observado que há um número infinito de decomposições que realizam a equação (4). Para qualquer matriz ortogonal  $\mathbf{P}_x$  e  $\mathbf{P}_y$ , as matrizes  $\mathbf{K}_x\mathbf{P}_x$  e  $\mathbf{K}_y\mathbf{P}_y$  também realizam a condição desde que

$$\begin{aligned} \mathbf{K}_x \mathbf{P}_x \mathbf{P}_x^T \mathbf{K}_x^T &= \mathbf{K}_x \mathbf{K}_x^T = \mathbf{C}_x \\ \mathbf{K}_y \mathbf{P}_y \mathbf{P}_y^T \mathbf{K}_y^T &= \mathbf{K}_y \mathbf{K}_y^T = \mathbf{C}_y. \end{aligned} \quad (5)$$

15

nos casos de estéreo utilizados, a matriz de covariância é geralmente dada na forma das energias do canal e a correlação intercanal (ICC | *inter-channel correlation*), por exemplo, em [1, 3, 4]. Os valores da diagonal de  $\mathbf{C}_x$  são as energias do canal e o ICC entre os dois canais é

$$\text{ICC}_x = \frac{\mathbf{C}_x(1,2)}{\sqrt{\mathbf{C}_x(1,1)\mathbf{C}_x(2,2)}} \quad (6)$$

e correspondentemente para  $C_y$ . Os índices nos parênteses denotam a linha e a coluna da matriz.

A definição restante é a matriz de mapeamento determinada pela utilização  $Q$ , que compreende a informação, cujos canais de entrada devem ser utilizados na composição de cada canal

de saída. Com  $\mathbf{Q}$  pode-se definir um sinal de referência

$$\mathbf{y}_{\text{ref}} = \mathbf{Q}\mathbf{x}. \quad (7)$$

A matriz de mapeamento  $\mathbf{Q}$  pode compreender mudanças na dimensionalidade e escala, combinação e reorganização dos canais. Devido à definição de preenchimento zero dos sinais,  $\mathbf{Q}$  é aqui uma matriz quadrada  $N \times N$  que pode compreender zero fileiras e colunas. Alguns exemplos de  $\mathbf{Q}$  são:

- Melhoria espacial:  $\mathbf{Q} = \mathbf{I}$ , nas aplicações, onde a saída deve ser o mais parecida com a entrada.
- *Downmixing*:  $\mathbf{Q}$  é uma matriz de *downmix*.
- Síntese espacial dos sinais do microfone de primeira ordem:  $\mathbf{Q}$  pode ser, por exemplo, uma matriz de mixagem do microfone Ambisonic, que significa que  $\mathbf{y}_{\text{ref}}$  é um conjunto de sinais de microfone virtual.

A seguir, é formulado como gerar um sinal  $\mathbf{y}$  a partir de um sinal  $\mathbf{x}$ , com uma restrição que  $\mathbf{y}$  tem a matriz de covariância definida pela utilização  $\mathbf{C}_y$ . A utilização também define uma matriz de mapeamento  $\mathbf{Q}$  que dá um ponto de referência para a otimização. O sinal de entrada  $\mathbf{x}$  tem a matriz de covariância medida  $\mathbf{C}_x$ . Conforme definido, os conceitos propostos para realizar esta transformação estão utilizando principalmente um conceito de apenas mixagem ideal dos canais, pois utilizar os descorrelacionadores tipicamente comprehende a qualidade do sinal e depois, pela injeção de energia descorrelacionada quando o objetivo não é de outra forma obtido.

A relação entrada-saída de acordo com estes

conceitos pode ser escrita como

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{r} \quad (8)$$

onde  $\mathbf{M}$  é uma matriz de mixagem real de acordo com o conceito primário e  $\mathbf{r}$  é um sinal residual de acordo com o conceito secundário.

A seguir, conceitos são propostos para a modificação da matriz de covariância.

Primeiro, a tarefa de acordo com o conceito primário é solucionada apenas pela mixagem cruzada dos canais de entrada. A equação (8) então simplifica para

$$\mathbf{y} = \mathbf{M}\mathbf{x}. \quad (9)$$

Das equações (3) e (9), uma tem

$$\begin{aligned} \mathbf{C}_y &= E[\text{Re}\{\mathbf{y}\mathbf{y}^H\}] \\ &= E[\text{Re}\{\mathbf{M}\mathbf{x}\mathbf{x}^H\mathbf{M}^T\}] = \mathbf{M}\mathbf{C}_x\mathbf{M}^T. \end{aligned} \quad (10)$$

Das equações (5) e (10) segue que

$$\mathbf{K}_y\mathbf{P}_y\mathbf{P}_y^T\mathbf{K}_y^T = \mathbf{M}\mathbf{K}_x\mathbf{P}_x\mathbf{P}_x^T\mathbf{K}_x^T\mathbf{M}^T \quad (11)$$

15

do qual um conjunto de soluções para  $\mathbf{M}$  que realiza a equação (10) segue

$$\mathbf{M} = \mathbf{K}_y\mathbf{P}_y\mathbf{P}_x^T\mathbf{K}_x^{-1} = \mathbf{K}_y\mathbf{P}\mathbf{K}_x^{-1} \quad (12)$$

A condição para estas soluções é que  $\mathbf{K}_x^{-1}$  existe. A 20 matriz ortogonal  $\mathbf{P} = \mathbf{P}_y\mathbf{P}_x^T$  é o parâmetro livre restante.

A seguir, é descrito como uma matriz  $\mathbf{P}$  fornece uma matriz ideal  $\mathbf{M}$ . De todos  $\mathbf{M}$  na equação (12), busca-se um que produza uma saída mais próxima ao ponto de referência definido  $\mathbf{y}_{\text{ref}}$ , isto é, que reduza

$$25 \quad e = E[\|\mathbf{y}_{\text{ref}} - \mathbf{y}\|^2] \quad (13a)$$

isto é, que reduza

$$e = E[\|\mathbf{y}_{\text{ref}} - \mathbf{y}\|^2] = E[\|\mathbf{Q}\mathbf{x} - \mathbf{M}\mathbf{x}\|^2]. \quad (13)$$

Agora, um sinal  $\mathbf{w}$  é definido, de modo que  $E[\text{Re}\{\mathbf{w}\mathbf{w}^H\}] = \mathbf{I}$ .  $\mathbf{w}$  possa ser escolhido de modo que  $\mathbf{x} = \mathbf{K}_x \mathbf{w}$ , desde que

$$\begin{aligned} E[\text{Re}\{\mathbf{x}\mathbf{x}^H\}] &= E[\text{Re}\{\mathbf{K}_x \mathbf{w} \mathbf{w}^H \mathbf{K}_x^T\}] \\ &= \mathbf{K}_x E[\text{Re}\{\mathbf{w}\mathbf{w}^H\}] \mathbf{K}_x^T \quad (14) \\ &= \mathbf{K}_x \mathbf{K}_x^T = \mathbf{C}_x. \end{aligned}$$

5

Então, segue-se que

$$\mathbf{M}\mathbf{x} = \mathbf{M}\mathbf{K}_x \mathbf{w} = \mathbf{K}_y \mathbf{P}\mathbf{w}. \quad (15)$$

A equação (13) pode ser escrita como

$$\begin{aligned} e &= E[\|\mathbf{Q}\mathbf{x} - \mathbf{M}\mathbf{x}\|^2] \\ &= E[\|\mathbf{Q}\mathbf{K}_x \mathbf{w} - \mathbf{K}_y \mathbf{P}\mathbf{w}\|^2] \quad (16) \\ &= E[\|(\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y \mathbf{P})\mathbf{w}\|^2] \\ &= E[\mathbf{w}^H (\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y \mathbf{P})^T (\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y \mathbf{P}) \mathbf{w}]. \end{aligned}$$

De  $E[\text{Re}\{\mathbf{w}\mathbf{w}^H\}] = \mathbf{I}$ , pode ser prontamente mostrado

10 para uma matriz simétrica real  $\mathbf{A}$  que

$E[\mathbf{w}^H \mathbf{A}\mathbf{w}] = \text{tr}(\mathbf{A})$ , que é o traço da matriz. Segue

que a equação (16) toma a forma

$$e = \text{tr}[(\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y \mathbf{P})^T (\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y \mathbf{P})]. \quad (17)$$

Para traços da matriz, pode ser prontamente

15 confirmado que

$$\begin{aligned} \text{tr}(\mathbf{A} + \mathbf{B}) &= \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B}) \\ \text{tr}(\mathbf{A}) &= \text{tr}(\mathbf{A}^T) \\ \text{tr}(\mathbf{P}^T \mathbf{A} \mathbf{P}) &= \text{tr}(\mathbf{A}). \quad (18) \end{aligned}$$

Utilizando estas propriedades, a equação (17) toma a forma

$$\begin{aligned} e &= \text{tr}(\mathbf{K}_x^T \mathbf{Q}^T \mathbf{Q} \mathbf{K}_x) + \text{tr}(\mathbf{K}_y^T \mathbf{K}_y) \\ &\quad - 2\text{tr}(\mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y \mathbf{P}). \quad (19) \end{aligned}$$

20 Apenas o último termo depende de  $\mathbf{P}$ . O problema de otimização é assim

$$\mathbf{P} = \arg \min_{\mathbf{P}} e = \arg \max_{\mathbf{P}} [\text{tr}(\mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y \mathbf{P})]. \quad (20)$$

Pode ser prontamente mostrado para uma matriz diagonal não negativa  $\mathbf{S}$  e qualquer matriz ortogonal  $\mathbf{P}_s$  que

5  $\text{tr}(\mathbf{S}) \geq \text{tr}(\mathbf{S}\mathbf{P}_s). \quad (21)$

Assim, definindo a decomposição do valor singular  $\mathbf{USV}^T = \mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y$ , onde  $\mathbf{S}$  é não negativo e diagonal e  $\mathbf{U}$  e  $\mathbf{V}$  são ortogonais, segue que

$$\begin{aligned} \text{tr}(\mathbf{S}) &\geq \text{tr}(\mathbf{SV}^T \mathbf{PU}) = \text{tr}(\mathbf{USV}^T \mathbf{PUU}^T) \\ &= \text{tr}(\mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y \mathbf{P}) \end{aligned} \quad (22)$$

10 para qualquer ortogonal  $\mathbf{P}$ . A igualdade é válida para

$$\boxed{\mathbf{P} = \mathbf{VU}^T} \quad (23)$$

pelo qual este  $\mathbf{P}$  produz o máximo de  $\text{tr}(\mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y \mathbf{P})$  e o mínimo da medição de erro na equação (13).

15 Um aparelho de acordo com uma utilização determina uma matriz de mixagem ideal  $\mathbf{M}$ , de modo que um erro seja reduzido. Deve-se observar que as propriedades de covariância do sinal de entrada de áudio e do sinal de saída de áudio podem variar para diferentes posições de tempo-frequência. Para isso, um provedor de um aparelho de acordo com uma utilização é adaptada para analisar as propriedades de covariância do canal de entrada de áudio que pode ser diferente para diferentes posições de tempo-frequência. Além disso, o processador de sinal de um aparelho de acordo com uma utilização é adaptada para determinar uma regra de mixagem, por exemplo, uma matriz de mixagem  $\mathbf{M}$  com base nas segundas propriedades de covariância do sinal de saída de áudio, caracterizado pelas segundas propriedades de covariância poderem

ter diferentes valores para diferentes posições de tempo-frequência.

Como a matriz de mixagem determinada  $\mathbf{M}$  é aplicada em cada um dos canais de entrada de áudio do sinal de 5 entrada de áudio e, como cada um dos canais de saída de áudio resultantes do sinal de saída de áudio pode assim depender de cada um dos canais de entrada de áudio, um processador de sinal de um aparelho de acordo com uma utilização é portanto adaptada para gerar o sinal de saída de áudio aplicando a regra de mixagem de 10 modo que cada um de dois ou mais canais de saída de áudio depende de cada um de dois ou mais canais de entrada de áudio do sinal de entrada de áudio.

De acordo com outra aplicação, é proposto utilizar a decorrelação quando  $\mathbf{K}_x^{-1}$  não existe ou é instável. Nas 15 aplicações descritas acima, uma solução foi fornecida para determinar uma matriz de mixagem ideal onde foi suposto que  $\mathbf{K}_x^{-1}$  existe. Entretanto,  $\mathbf{K}_x^{-1}$  pode nem sempre existir ou seu inverso pode implicar multiplicadores muito grandes se alguns dos componentes principais em  $\mathbf{x}$  forem muito pequenos. Uma forma 20 efetiva para regularizar o inverso é empregar a decomposição do valor singular  $\mathbf{K}_x = \mathbf{U}_x \mathbf{S}_x \mathbf{V}_x^T$ . Consequentemente, o inverso é

$$\mathbf{K}_x^{-1} = \mathbf{V}_x \mathbf{S}_x^{-1} \mathbf{U}_x^T. \quad (24)$$

Problemas surgem quando alguns dos valores da diagonal da matriz diagonal não negativa  $\mathbf{S}_x$  são zero ou muito 25 pequenos. Um conceito que robustamente regulariza o inverso é então substituir estes valores por valores maiores. O resultado deste procedimento é  $\hat{\mathbf{S}}_x$  e o inverso correspondente  $\hat{\mathbf{K}}_x^{-1} = \mathbf{V}_x \hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T$  e a matriz de mixagem correspondente  $\hat{\mathbf{M}} = \mathbf{K}_y \mathbf{P} \hat{\mathbf{K}}_x^{-1}$ .

Esta regularização efetivamente significa que dentro do processo de mixagem, a amplificação de alguns dos pequenos componentes principais em  $\mathbf{x}$  é reduzida e, consequentemente, sua integridade para o sinal de saída  $\mathbf{y}$  também é 5 reduzida e a covariância alvo  $\mathbf{C}_y$  não é atingida no geral.

Dessa forma, de acordo com uma aplicação, o processador de sinal pode ser configurado para modificar, pelo menos, alguns valores da diagonal de uma matriz diagonal  $\mathbf{S}_x$ , caracterizado pelos valores da matriz diagonal  $\mathbf{S}_x$  serem zero ou 10 menores do que um valor limite (o valor limite pode ser predeterminado ou pode depender de uma função), de modo que os valores sejam maiores ou iguais ao valor limite, em que o processador de sinal pode ser adaptado para determinar a matriz de mixagem com base na matriz diagonal.

15 De acordo com uma aplicação, o processador de sinal pode ser configurado para modificar, pelo menos, alguns valores da diagonal da matriz diagonal  $\mathbf{S}_x$ , caracterizado por  $\mathbf{K}_x = \mathbf{U}_x \mathbf{S}_x \mathbf{V}_x^T$ , e em que  $\mathbf{C}_x = \mathbf{K}_x \mathbf{K}_x^T$  em que  $\mathbf{C}_x$  é a primeira matriz de covariância, em que  $\mathbf{S}_x$  é a matriz diagonal, em que  $\mathbf{U}_x$  é uma segunda 20 matriz,  $\mathbf{V}_x^T$  é uma terceira matriz transposta e em que  $\mathbf{K}_x^T$  é uma quarta matriz transposta da quinta matriz  $\mathbf{K}_x$ .

A perda acima de um componente do sinal pode ser completamente compensada com um sinal residual  $\mathbf{r}$ . A relação original entrada-saída será elaborada com o inverso regularizado.

$$25 \quad \begin{aligned} \mathbf{y} &= \hat{\mathbf{M}}\mathbf{x} + \mathbf{r} = \mathbf{K}_y \hat{\mathbf{K}}_x^{-1} \mathbf{x} + \mathbf{r} \\ &= \mathbf{K}_y \mathbf{P} \mathbf{V}_x \hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T \mathbf{x} + \mathbf{r} \end{aligned} \quad (25)$$

Agora, um componente aditivo  $\mathbf{c}$  é definido de modo que ao invés de  $\hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T \mathbf{x}$ , um tenha  $\hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T \mathbf{x} + \mathbf{c}$ . Além disso, um sinal

independente  $\mathbf{w}'$  é definido, de modo que  $E[\text{Re}\{\mathbf{w}'\mathbf{w}'^H\}] = \mathbf{I}$  e

$$\mathbf{c} = \sqrt{\mathbf{I} - (\hat{\mathbf{S}}_x^{-1}\mathbf{S}_x)^2}\mathbf{w}'. \quad (26)$$

Pode ser prontamente mostrado que um sinal

$$\begin{aligned} \mathbf{y}' &= \mathbf{K}_y \mathbf{P} \mathbf{V}_x (\hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T \mathbf{x} + \mathbf{c}) \\ &= \hat{\mathbf{M}} \mathbf{x} + \mathbf{K}_y \mathbf{P} \mathbf{V}_x \mathbf{c} \end{aligned} \quad (27)$$

5 tenha covariância  $\mathbf{c}_y$ . O sinal residual para compensar a regularização é então

$$\mathbf{r} = \mathbf{K}_y \mathbf{P} \mathbf{V}_x \mathbf{c}. \quad (28)$$

Das equações (27) e (28), segue que

$$\mathbf{C}_r = E[\text{Re}\{\mathbf{r}\mathbf{r}^H\}] = \mathbf{C}_y - \hat{\mathbf{M}} \mathbf{C}_x \hat{\mathbf{M}}^T. \quad (29)$$

10 Como  $\mathbf{c}$  foi definido como um sinal estocástico, segue que a propriedade relevante de  $\mathbf{r}$  é sua matriz de covariância. Assim, qualquer sinal que for independente com relação a  $\mathbf{x}$  que é processado para ter a covariância  $\mathbf{C}_r$  serve como um sinal residual que idealmente reconstrói a matriz de 15 covariância alvo  $\mathbf{C}_y$  em situações quando a regularização conforme descrito foi utilizada. Tal sinal residual pode ser prontamente gerado utilizando os descorrelacionadores e o método de mixagem do canal proposto.

Encontrar analiticamente o equilíbrio ótimo entre 20 a quantidade de energia descorrelacionada e a amplificação de pequenos componentes de sinal não é simples. Isso é porque depende dos fatores específicos da utilização como a estabilidade das propriedades estatísticas do sinal de entrada, janela de análise aplicada e o SNR do sinal de entrada. Entretanto, é bastante 25 simples ajustar uma função heurística para realizar este equilíbrio sem desvantagens óbvias, como foi feito no exemplo de código fornecido abaixo.

De acordo com isso, o processador de sinal de um aparelho de acordo com uma utilização pode ser adaptado para gerar o sinal de saída de áudio aplicando a regra de mixagem em, pelo menos, dois dos dois sinais de entrada de áudio ou mais, para 5 obter um sinal intermediário  $\mathbf{y}' = \hat{\mathbf{M}}\mathbf{x}$  e adicionando um sinal residual  $\mathbf{r}$  ao sinal intermediário para obter o sinal de saída de áudio.

Foi mostrado que quando a regularização do inverso de  $\mathbf{K}_x$  é aplicada, os componentes do sinal ausentes na 10 saída geral podem ser completamente complementados com um sinal residual  $\mathbf{r}$  com covariância  $\mathbf{C}_r$ . Por estes meios, pode-se garantir que a covariância alvo  $\mathbf{C}_y$  seja sempre obtida. A seguir, uma forma de gerar um sinal residual correspondente  $\mathbf{r}$  é apresentada. Este compreende as etapas a seguir:

15 1. Gerar um conjunto de sinais tanto quanto canais de saída. O sinal  $\mathbf{y}_{ref} = \mathbf{Q}\mathbf{x}$  pode ser empregado, pois tem tantos canais quanto sinal de saída e cada sinal de saída contém um sinal apropriado para este canal particular.

20 2. Descorrelacionar este sinal. Há várias formas de descorrelacionar, incluindo filtros passa tudo, circunvoluções com rajadas de ruído e atraso pseudoaleatório nas faixas de frequência.

25 3. Medir (ou supor) a matriz de covariância do sinal descorrelacionado. A medição é mais simples e mais robusta, mas desde que os sinais são de descorrelacionadores, eles poderiam ser supostos como incoerentes. Então, apenas a medição de energia seria suficiente.

4. Aplicar o método proposto para gerar uma

matriz de mixagem que, quando aplicada ao sinal descorrelacionado, gera um sinal de saída com a matriz de covariância  $\mathbf{C}_r$ . Utilize aqui uma matriz de mapeamento  $\mathbf{Q} = \mathbf{I}$ , pois se deseja afetar minimamente o conteúdo do sinal.

5. Processar o sinal dos descorrelacionadores com esta matriz de mixagem e inseri-lo ao sinal de saída para complementar a falta de componentes do sinal. Desse modo, o alvo  $\mathbf{C}_y$  é obtido.

Em uma utilização alternativa os canais descorrelacionados são anexos ao sinal de entrada (pelo menos um) antes de formular a matriz de mixagem ideal. Neste caso, a entrada e a saída é da mesma dimensão e desde que o sinal de entrada tenha tantos componentes de sinal independentes quanto canais de entrada, não há necessidade de utilizar um sinal residual  $r$ .

Quando os descorrelacionadores são utilizados desta forma, o uso de descorrelacionadores é “invisível” ao conceito proposto, pois os canais descorrelacionados são canais de entrada como qualquer outro.

Se o uso de descorrelacionadores for indesejável,  
 20 pelo menos as energias alvo do canal podem ser obtidas  
 multiplicando-se as fileiras do  $\hat{\mathbf{M}}$  de modo que  

$$\mathbf{M}' = \mathbf{G}\hat{\mathbf{M}} \quad (30)$$

$$\mathbf{M}' = \mathbf{G}\mathbf{M} \quad (30)$$

onde  $\mathbf{G}$  é uma matriz de ganho diagonal com valores

$$\mathbf{G}(i,i) = \sqrt{\frac{\mathbf{C}_y(i,i)}{\hat{\mathbf{C}}_y(i,i)}} \quad (31)$$

Em muitas aplicações, o número de canais de entrada e saída é diferente. Conforme descrito na Equação (2), o

preenchimento zero do sinal com uma dimensão menor é aplicado para ter a mesma dimensão que o mais alto. O preenchimento zero implica na sobrecarga computacional, pois algumas fileiras ou colunas no **M** resultante correspondem aos canais com energia zero definida.

5 Matematicamente, equivalente a utilizar o primeiro preenchimento zero e finalmente cortando **M** na dimensão relevante  $N_y \times N_x$ , a sobrecarga pode ser reduzida introduzindo a matriz **A** que é uma matriz de identidade anexa ao zero na dimensão  $N_y \times N_x$ , por exemplo,

$$\Lambda_{3 \times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (32)$$

10

Quando **P** é redefinido de modo que

$$\mathbf{P} = \mathbf{V} \Lambda \mathbf{U}^T \quad (33)$$

o **M** resultante seja uma matriz de mixagem  $N_y \times N_x$  15 que é a mesma que a parte relevante do **M** do caso de preenchimento zero. Consequentemente, **C<sub>x</sub>**, **C<sub>y</sub>**, **K<sub>x</sub>** e **K<sub>y</sub>** pode ser de dimensão natural e a matriz de mapeamento **Q** é de dimensão  $N_y \times N_x$ .

A entrada matriz de covariância é sempre decomponível em  $\mathbf{C}_x = \mathbf{K}_x \mathbf{K}_x^T$  pois é uma medição semidefinida positiva 20 de um sinal real. É, entretanto, possível definir tais matrizes de covariância alvo que não são decomponíveis pela razão que representam dependências do canal impossíveis. Há conceitos para garantir a decomposição, como ajustar os valores próprios negativos a zeros e normalizar a energia, vide, por exemplo,

25 [8] R. Rebonato, P. Jäckel, "The most general methodology to create a valid correlation matrix for risk management and option pricing purposes", Journal of Risk, Vol. 2, No. 2, pp. 17-28, 2000.

Entretanto, o uso mais significativo do conceito proposto é solicitar apenas as possíveis matrizes de covariância.

Para resumir o que foi mencionado acima, a tarefa comum pode ser reescrita conforme segue. Primeiramente, tem-se um sinal de entrada com uma determinada matriz de covariância. Depois, a utilização define dois parâmetros: a matriz de covariância alvo e uma regra, cujos canais de entrada devem ser utilizados na composição de cada canal de saída. Para realizar esta transformação, é proposto utilizar os seguintes conceitos: O conceito primário, conforme ilustrado pela Figura 2, é que a covariância alvo é obtida utilizando uma solução de mixagem ideal dos canais de entrada. Este conceito é considerado primário, pois evita o uso do descorrelacionador, que geralmente compromete a qualidade do sinal. O conceito secundário ocorre quando não há componentes suficientes independentes da energia razoável disponível. A energia descorrelacionada é injetada para compensar a falta destes componentes. Juntos, estes dois conceitos fornecem meios para realizar o ajuste robusto da matriz de covariância em qualquer cenário determinado.

A principal utilização esperada do conceito proposto está no campo da microfonia espacial [2,3], que é o campo onde os problemas relacionados ao sinal covariância são particularmente evidentes devido às limitações físicas dos microfones direcionais. Ainda, os casos do uso esperado incluem melhoria estéreo e multicanal, extração de ambiente, *upmixing* e *downmixing*.

Na descrição acima, as definições foram determinadas, seguidas pela derivação do conceito proposto.

Primeiramente, a solução da mixagem cruzada foi fornecida, então o conceito de injeção da energia do som correlacionado foi determinado. Depois disso, uma descrição do conceito com um diferente número de canais de entrada e saída foi fornecida e 5 também considerações na decomposição da matriz de covariância. A seguir, casos de uso prático são fornecidos e um conjunto de exemplos numéricos e a conclusão são apresentados. Além disso, um código de exemplo Matlab com funcionalidade completa de acordo com esta função é fornecido.

10 A característica espacial percebida de um som estéreo ou som multicanal é amplamente definida pela matriz de covariância do sinal nas faixas de frequência. Um conceito foi fornecido para, de forma ideal e adaptável, mixar um conjunto de canais de entrada com determinadas propriedades de covariância em 15 um conjunto de canais de saída com propriedades de covariância que podem ser definidas de forma arbitrária. Outro conceito foi fornecido para injetar a energia descorrelacionada apenas onde necessário quando componentes de som independentes de energia razoável não estão disponíveis. O conceito tem uma ampla variedade 20 de aplicações no campo de processamento do sinal de áudio espacial.

As energias do canal e as dependências entre os canais (ou a matriz de covariância) de um sinal multicanal podem ser controladas por apenas mixando linearmente e de forma variável 25 por tempo os canais dependendo das características de entrada e as características alvo desejadas. Este conceito pode ser ilustrado com uma representação do fator do sinal onde o ângulo entre os vetores corresponde à dependência do canal e a amplitude do vetor

é igual ao nível do sinal.

A Figura 3 ilustra um exemplo para aplicar uma combinação linear de vetores **L** e **R** para atingir um novo conjunto de vetor **R'** e **L'**. Semelhantemente, os níveis de canal de áudio e 5 sua dependência podem ser modificados com a combinação linear. A solução geral não inclui vetores, mas uma formulação da matriz que é ideal para qualquer número de canal.

A matriz de mixagem para sinais estéreos pode ser prontamente formulada também de forma trigonométrica, como pode 10 ser visto na Figura 3. Os resultados são os mesmos que com a matemática da matriz, mas a formulação é diferente.

Se os canais de entrada são altamente dependentes, atingir a matriz de covariância alvo é possível apenas utilizando descorrelacionadores. Um procedimento para 15 injetar os descorrelacionadores apenas onde necessário, por exemplo, de forma ideal, também foi fornecido.

A Figura 4 ilustra um diagrama em blocos de um aparelho de uma utilização aplicando a técnica de mixagem. O aparelho compreende um módulo de análise da matriz de covariância 20 410 e um processador de sinal (não mostrado), caracterizado pelo processador de sinal compreender um módulo de formulação da matriz de mixagem 420 e um módulo de utilização da matriz de mixagem 430. Propriedades de covariância de entrada de uma entrada da faixa de frequência estéreo ou multicanal são analisadas por um módulo de 25 análise da matriz de covariância 410. O resultado da análise da matriz de covariância é inserido em um módulo de formulação da matriz de mixagem 420.

O módulo de formulação da matriz de mixagem 420

formula uma matriz de mixagem com base na análise do resultado da matriz de covariância, com base em uma matriz de covariância alvo e possivelmente também com base em um critério de erro.

O módulo de formulação da matriz de mixagem 420 5 insere a matriz de mixagem em um módulo de utilização da matriz de mixagem 430. O módulo de utilização da matriz de mixagem 430 aplica a matriz de mixagem na entrada da faixa de frequência estéreo ou multicanal para obter uma saída da faixa de frequência estéreo ou multicanal com, por exemplo, predefinido, propriedades 10 de covariância alvo dependendo da matriz de covariância alvo.

Resumindo o que foi mencionado acima, a finalidade geral do conceito é melhorar, ajustar e/ou sintetizar o som espacial com um grau extremo de otimização em termos de qualidade do som. O alvo, por exemplo, as segundas propriedades de 15 covariância, é definido pela aplicação.

Também aplicável em faixa completa, o conceito é perceptualmente significativo especialmente no processamento da faixa de frequência.

Descorrelacionadores são utilizados a fim de 20 melhorar (reduzir) a correlação intercanal. Eles fazem isso, mas estão propensos a comprometer a qualidade geral do som, especialmente com um componente de som transiente.

O conceito proposto evita, ou em alguma utilização reduz, o uso de descorrelacionadores. O resultado é a 25 mesma característica espacial, mas sem perda da qualidade do som.

Entre outros usos, a tecnologia pode ser empregada em um Codificador SAM-para-MPS.

O conceito proposto foi implementado para

melhorar uma técnica de microfone que gera o fluxo de dados MPEG Surround (MPEG | *Moving Picture Experts Group*) fora de um sinal de microfones coincidentes estéreos de primeira ordem, vide, por exemplo, [3]. O processo inclui estimar do sinal estéreo a direção e a difusão do campo do som nas faixas de frequência e criar tal fluxo de dados MPEG Surround que, quando decodificado no receptor final, produza um campo de som que perceptualmente aproxima o campo de som original.

Na Figura 5, um diagrama é ilustrado descrevendo um sinal de microfone de coincidência estéreo ao codificador MPEG Surround de acordo com uma aplicação, que emprega o conceito proposto para criar o sinal de *downmix* MPEG Surround no determinado sinal do microfone. Todo o processamento é realizado nas faixas de frequência.

Um módulo de determinação de dados espaciais 520 é adaptado para formular os dados de informação da configuração compreendendo dados surround espaciais e ICC *downmix* e/ou níveis com base na informação de direção e difusão dependendo de um modelo do campo de som 510. O modelo do campo de som tem como base uma análise do microfone ICC's e níveis de um sinal de microfone estéreo. O módulo de determinação de dados espaciais 520 então fornece os ICC's de *downmix* alvo e níveis a um módulo de formulação da matriz de mixagem 530. Além disso, o módulo de determinação de dados espaciais 520 pode ser adaptado para formular os dados surround espaciais e ICC's *downmix* e níveis como a informação lateral espacial MPEG Surround. O módulo de formulação da matriz de mixagem 530 então formula uma matriz de mixagem com base nos dados de informação da configuração

fornecidos, por exemplo, ICC's *downmix* alvo e níveis, e insere a matriz em um módulo de mixagem 540. O módulo de mixagem 540 aplica a matriz de mixagem no sinal estéreo do microfone. Desse modo, um sinal é gerado com ICC's e níveis alvo. O sinal com os ICC's e 5 níveis alvo é então fornecido a um codificador de núcleo 550. Em uma aplicação, os módulos 520, 530 e 540 são submódulos de um processador de sinal.

Dentro do processo conduzido por um aparelho de acordo com a Figura 5, um *downmix* estéreo MPEG Surround deve ser 10 gerado. Isso inclui uma necessidade de ajustar os níveis e os ICC's do sinal estéreo determinado com impacto mínimo na qualidade do som. O conceito proposto de mixagem cruzada foi aplicado para esta finalidade e o benefício perceptual da técnica anterior em [3] foi observado.

15 A Figura 6 ilustra um aparelho de acordo com outra utilização referente à correção de ICC *downmix/nível* para um codificador SAM-para-MPS. Uma análise de ICC e nível é conduzida no módulo 602 e o modelo do campo de som 610 depende da Análise do ICC e nível pelo módulo 602. O módulo 620 corresponde ao módulo 520, o módulo 630 corresponde ao módulo 530 e o módulo 640 corresponde ao módulo 540 da Figura 5, respectivamente. O mesmo se 20 aplica para o codificador de núcleo 650 que corresponde ao codificador de núcleo 550 da Figura 5. O conceito descrito acima 25 pode ser integrado em um codificador SAM-para-MPS para criar a partir dos sinais do microfone o MPS *downmix* com ICC e níveis exatamente corretos. O conceito descrito acima também é aplicável à interpretação direta SAM-para-multicanal sem MPS para fornecer síntese espacial ideal enquanto reduz a quantidade do uso do

descorrelacionador.

Melhorias são esperadas com relação à distância fonte, localização da fonte, estabilidade, conforto e envolvimento da audição.

5 A Figura 7 descreve um aparelho de acordo com uma utilização para uma melhoria para pequenas matrizes de microfone espaçadas. Um módulo 705 é adaptado para conduzir uma análise da matriz de covariância de um sinal do microfone de entrada para obter uma matriz de covariância do microfone. A matriz de 10 covariância do microfone é inserida em um módulo de formulação da matriz de mixagem 730. Além disso, a matriz de covariância do microfone é utilizada para derivar um modelo do campo de som 710. O modelo do campo de som 710 pode se basear em outras fontes além da matriz de covariância.

15 Informação de direção e difusão com base no modelo do campo de som é então inserida ao módulo de formulação da matriz de covariância alvo 720 para gerar uma matriz de covariância alvo. O módulo de formulação da matriz de covariância alvo 720 então insere a matriz de covariância alvo ao módulo de 20 formulação da matriz de mixagem 730.

O módulo de formulação da matriz de mixagem 730 é adaptado para gerar a matriz de e insere a matriz de mixagem gerada ao módulo de utilização da matriz de mixagem 740. O módulo de utilização da matriz de mixagem 740 é adaptado para aplicar a 25 matriz mixagem no sinal do microfone de entrada para obter um sinal do microfone de saída com as propriedades de covariância alvo. Em uma aplicação, os módulos 720, 730 e 740 são submódulos de um processador de sinal.

Tal aparelho segue o conceito em DirAC e SAM, que é estimar a direção e difusão do campo de som original e para criar esta saída que melhor reproduz a direção e difusão estimada. Este procedimento de processamento de sinal exige grandes ajustes 5 da matriz de covariância para fornecer a imagem espacial correta. O conceito processado é a solução. Pelo conceito proposto, a distância da fonte, a localização da fonte e/ou a separação da fonte, o conforto e envolvimento da audição.

A Figura 8 ilustra um exemplo que mostra uma 10 utilização para melhoria cega da qualidade do som espacial na reprodução estéreo ou multicanal. No módulo 805, uma análise da matriz de covariância, por exemplo, uma Análise de ICC ou nível de conteúdo estéreo ou multicanal é conduzida. Então, uma regra de melhoria é aplicada no módulo de melhoria 815, por exemplo, para 15 obter as saídas ICC das entradas ICC. Um módulo de formulação da matriz de mixagem 830 gera uma matriz de mixagem com base na análise da matriz de covariância conduzida pelo módulo 805 e com base na informação derivada para aplicar a regra de melhoria que foi conduzida no módulo de melhoria 815. A matriz de mixagem é 20 então aplicada no conteúdo estéreo ou multicanal no módulo 840 para obter o conteúdo estéreo ou multicanal ajustado com as propriedades de covariância alvo.

Referente ao som multicanal, por exemplo, mixagens e gravações, é muito comum encontrar a subotimização 25 perceptual no som espacial, especialmente em termos de ICC muito alto. Uma consequência típica tem a qualidade reduzida com relação à largura, envolvimento, distância, separação da fonte, localização da fonte e/ou estabilidade da fonte e conforto da

audição. Foi testado informalmente que o conceito pode melhorar estas propriedades com itens que têm ICC's desnecessariamente altos. Melhorias observadas são largura, distância da fonte, localização/separação da fonte, envolvimento e conforto da

5 audição.

A Figura 9 ilustra outra utilização para melhoria das configurações estreitas do alto-falante (por exemplo, tablets, TV). O conceito proposto é provavelmente útil como uma fermenta para melhorar a qualidade estéreo nas configurações de reprodução

10 onde um angulo do alto-falante é muito estreito (por exemplo, tablets). O conceito proposto fornecerá:

- replanejar as fontes dentro do arco determinado para corresponder a uma configuração mais ampla do alto-falante
- aumentar o ICC para melhor corresponder ao de

15 uma configuração do alto-falante mais ampla

- fornecer um melhor ponto inicial para realizar o cancelamento da diafonia, por exemplo, utilizando o cancelamento da diafonia apenas quando não há forma direta para criar os sinais binaurais desejados.

20 Melhorias são esperadas com relação à largura e com relação ao cancelamento da diafonia regular, qualidade do som e robustez.

Em outra utilização exemplo ilustrada pela Figura 10, uma utilização é descrita fornecendo renderização da

25 Codificação do Áudio Direcionado (DirAC) ideal com base em um sinal de microfone em formato B.

A utilização da Figura 10 é com base na constatação de que a renderização das unidades do estado da

A utilização da Figura 10 é com base na constatação de que a renderização das unidades do estado da técnica DirAC com base nos sinais do microfone coincidentes aplicam a descorrelação na extensão desnecessária, assim, 5 comprometendo a qualidade do áudio. Por exemplo, se o campo de som é analisado como difuso, a correlação total é aplicada em todos os canais, embora um formato B forneça já três componentes de som incoerentes no caso de um campo de som horizontal (W, X, Y). Este efeito está presente nos graus variantes exceto quando a difusão 10 for zero.

Além disso, os sistemas descritos acima utilizando os microfones virtuais não garantem a matriz de covariância de saída correta (níveis e correlações do canal), pois os microfones virtuais realizam o som de forma diferente 15 dependendo do ângulo da fonte, posicionamento do alto-falante e difusão do campo de som.

O conceito proposto soluciona ambas as questões. Há duas alternativas: fornecer os canais descorrelacionados como canais de entrada extra (como na figura abaixo); ou utilizar um 20 conceito de mixagem do descorrelacionador.

Na Figura 10, um módulo 1005 conduz uma análise da matriz de covariância. Um módulo de formulação da matriz de covariância alvo 1018 não considera apenas um modelo do campo de som, mas também leva em conta uma configuração do alto-falante ao 25 formular uma matriz de covariância alvo. Além disso, um módulo de formulação da matriz de mixagem 1030 gera uma matriz de mixagem não apenas com base em uma análise da matriz de covariância e a matriz de covariância alvo, mas também com base nos critérios de

otimização, por exemplo, uma matriz de mixagem em formato B para microfone virtual fornecida por um módulo 1032. O modelo do campo de som 1010 pode corresponder ao modelo do campo de som 710 da Figura 7. O módulo de utilização da matriz de mixagem 1040 pode 5 corresponder ao módulo de utilização da matriz de mixagem 740 da Figura 7.

Em outra utilização exemplo, uma utilização é fornecida para o ajuste espacial nos métodos de conversão do canal, por exemplo, *downmix*. A conversão do canal por exemplo, 10 fazendo 5.1 *downmix* de 22.2 faixas de áudio automática inclui quebrar os canais. Isso pode incluir uma perda ou mudança da imagem espacial que pode ser direcionada com o conceito proposto. Novamente, há duas alternativas: a primeira utiliza o conceito no domínio do número de canais mais altos, mas definindo os canais de 15 energia zero para os canais ausentes do número inferior; a outra formula a solução da matriz diretamente para diferentes números de canal.

A Figura 11 ilustra a tabela 1, que fornece exemplos numéricos dos conceitos descritos acima. Quando um sinal 20 com covariância  $\mathbf{C}_x$  é processado com uma matriz de mixagem  $\mathbf{M}$  e complementado com um possível sinal residual com  $\mathbf{C}_r$ , o sinal de saída tem covariância  $\mathbf{C}_y$ . Embora estes exemplos numéricos sejam estáticos, o caso do uso típico do método proposto é dinâmico. A ordem do canal é suposta L, R, C, Ls, Rs, (Lr, Rr).

25 A Tabela 1 mostra um conjunto de exemplos numericamente para ilustrar o comportamento do conceito proposto em alguns casos esperados. As matrizes foram formuladas com o código Matlab fornecido na lista 1. A lista 1 é ilustrada na

Figura 12.

A lista 1 de Figura 12 ilustra uma implementação Matlab do conceito proposto. O código Matlab foi utilizado nos exemplos numéricos e fornece a funcionalidade geral do conceito 5 proposto.

Embora as matrizes sejam ilustradas estáticas, em aplicações típicas elas variam em tempo e frequência. O critério de desenho é por definição cumprido se um sinal com covariância  $\mathbf{C}_x$  for processado com uma matriz de mixagem  $\mathbf{M}$  e completada com um 10 possível sinal residual com  $\mathbf{C}_r$  o sinal de saída tem a covariância definida  $\mathbf{C}_y$ .

A primeira e a segunda fileira da tabela ilustram um caso de uso da melhoria estéreo por meios da correlação dos 15 sinais. Na primeira fileira há um pequeno, mas razoável componente incoerente entre os dois canais e assim a saída completamente incoerente é obtida com apenas mixagem do canal. Na segunda fileira, a correlação da entrada é muito alta, por exemplo, o menor componente principal é muito pequeno. A amplificação deste nos graus extremos não é desejável e assim o limitador embutido 20 começa a exigir a injeção da energia correlacionada ao invés, por exemplo,  $\mathbf{C}_r$  é agora não -zero.

A terceira fileira mostra uma caixa estéreo a 5,0 *upmixing*. Neste exemplo, a matriz de covariância alvo é definida de modo que o componente incoerente da mixagem estéreo é 25 igualmente e incoerentemente distribuída nos alto-falantes lateral e traseiro e o componente coerente é colocado no alto-falante central. O sinal residual é novamente não-zero desde que a dimensão do sinal seja aumentada.

A quarta fileira mostra uma caixa simples de 5,0 a 7,0 *upmixing* onde os dois canais traseiros originais fazem *upmixing* nos quatro novos canais traseiros, incoerentemente. Este exemplo ilustra que o processamento foca nestes canais onde 5 ajustes são solicitados.

A quinta fileira descreve um a caixa de *downmixing* de sinal 5,0 ao estéreo. *Downmixing* passivo, como aplicar uma matriz de mixagem de *downmix* estática **Q**, amplificaria os componentes coerentes sobre os componentes incoerentes. Aqui a 10 matriz de covariância alvo foi definida para preservar a energia, que é realizada pelo **M** resultante.

A sexta e a sétima fileiras ilustram a caixa de uso da microfonia espacial coincidente. As matrizes de covariância de entrada **C<sub>x</sub>** são o resultado de colocação dos microfones 15 coincidentes de primeira ordem ideais em um campo difuso ideal. Na sexta fileira os ângulos entre os microfones são iguais e, na sétima fileira os microfones estão voltados aos ângulos de uma configuração de 5,0. Em ambos os casos, os maiores valores da diagonal de **C<sub>x</sub>** ilustram a desvantagem inerente da técnica de 20 microfones coincidente da primeira ordem passiva na caixa ideal, a matriz de covariância que melhor representa um campo difuso é diagonal e este foi então definido como o alvo. Em ambos os casos, a razão do resultado da energia correlacionada sobre toda a energia é exatamente 2/5. Isso acontece porque há três componentes 25 do sinal independentes disponíveis nos sinais do microfone coincidente horizontais de primeira ordem e dois devem ser adicionados para atingir a matriz de covariância alvo diagonal de cinco canais.

A percepção espacial na reprodução estéreo e multicanal foi identificada para depender especialmente da matriz de covariância do sinal nas faixas de frequência perceptualmente relevantes.

5 Um conceito para controlar a matriz de covariância de um sinal pela mixagem cruzada ideal dos canais foi apresentado. Meios para injetar a energia descorrelacionada onde necessário nos casos quando componentes do sinal independentes suficientes da energia razoável não estão disponíveis foram  
10 apresentados.

O conceito foi observado como robusto em sua finalidade e uma ampla variedade das utilizações semelhantes foi identificada.

A seguir, aplicações são apresentadas, como gerar  
15  $\mathbf{C}_y$  com base em  $\mathbf{C}_x$ . Como um primeiro exemplo, estéreo a 5.0 *upmixing* é considerado. Referente ao estéreo-para-5,0 *upmixing*, no *upmixing*,  $\mathbf{C}_x$  é uma matriz  $2 \times 2$  e  $\mathbf{C}_y$  é uma matriz  $5 \times 5$  (neste exemplo, o canal subwoofer não é considerado). As etapas para gerar  $\mathbf{C}_y$  com base em  $\mathbf{C}_x$ , em cada parte de tempo-frequência, no contexto de  
20 *upmixing*, pode, por exemplo, ser conforme segue:

1. Estimar a energia ambiente e direta no canal esquerdo e direito. O ambiente é caracterizado por um componente incoerente entre os canais que tem energia igual em ambos os canais. A energia direta é a restante quando a parte da energia  
25 ambiente é removida da energia total, por exemplo, o componente de energia coerente, possivelmente com diferentes energias nos canais esquerdo e direito.

2. Estimar um ângulo do componente direito. Isso

é feito utilizando uma lei da posição da amplitude inversamente. Há uma razão de posição da amplitude no componente direito e há apenas um ângulo entre alto-falantes dianteiros que correspondem a ele.

5

3. Gerar uma matriz  $5 \times 5$  de zeros como  $\mathbf{C}_y$ .

4. Colocar a quantidade de energia direta na diagonal de  $\mathbf{C}_y$  correspondente a dois alto-falantes mais próximos da direção analisada. A distribuição da energia entre estes pode ser adquirida pelas leis de posição da amplitude. A posição da amplitude é coerente, então adicione à não diagonal correspondente a raiz quadrada do produto das energias dos dois canais.

10

5. Adicione à diagonal de  $\mathbf{C}_y$ , correspondente aos canais L, R, Ls e Rs, a quantidade de energia que corresponde à energia do componente do ambiente. A distribuição igual é uma boa escolha. Agora se tem o alvo  $\mathbf{C}_y$ .

15

Como outro exemplo, a melhoria é considerada. O objetivo é aumentar as qualidades perceptuais como largura ou envolvimento ajustando a coerência intercanal a zero. Aqui, dois diferentes exemplos são determinados, em duas formas de realizar a melhoria. Para a primeira forma, seleciona-se uma melhoria da caixa de uso estéreo, assim  $\mathbf{C}_x$  e  $\mathbf{C}_y$  são matrizes  $2 \times 2$ . As etapas são conforme segue:

20

1. Formular ICC (o valor de covariância normalizado entre -1 e 1, por exemplo com a fórmula fornecida).

25

2. Ajustar ICC por uma função. Por exemplo  $ICC_{new} = \text{sinal}(ICC) * ICC^2$ . Esse é um ajuste leve. Ou  $ICC_{new} = \text{sinal}(ICC) * \max(0, \text{abs}(ICC) * 10 - 9)$ . Este é um ajuste maior.

3. Formular  $\mathbf{C}_y$  de modo que os valores da diagonal

sejam os mesmos como em  $\mathbf{C}_x$ , mas o valor não-diagonal é formulado utilizando  $ICC_{new}$ , com a mesma fórmula como na etapa 1, mas inversamente.

5 No cenário acima, o sinal residual não é necessário, pois o ajuste de ICC é sinalizado de modo que o sistema não solicite grande amplificação de pequenos componentes do sinal.

O segundo tipo de implementação do método nesta caixa de uso, é conforme segue. Tem-se um sinal de entrada do 10 canal N, assim  $\mathbf{C}_x$  e  $\mathbf{C}_y$  são matrizes NxN.

1. Formular  $\mathbf{C}_y$  de  $\mathbf{C}_x$  ajustando simplesmente os valores da diagonal em  $\mathbf{C}_y$  o mesmo como em  $\mathbf{C}_x$ , e os valores não-diagonais a zero.

2. Permitir o método de compensação de ganho no 15 método proposto, ao invés de utilizar os residuais. A regularização no inverso de  $\mathbf{K}_x$  cuida para que o sistema seja estável. A compensação de ganho cuida para que as energias sejam preservadas.

As duas formas descritas para fazer a melhoria 20 fornecem resultados semelhantes. O último é mais fácil de implementar na caixa de uso multicanal.

Finalmente, como um terceiro exemplo, o modelo direto/difusão, por exemplo, a Codificação de Áudio Direcional (DirAC), é considerado

25 DirAC e também os Microfones de Áudio Espaciais (SAM | *Spatial Audio Microphones*), fornecem uma interpretação de um campo de som com direção e difusão dos parâmetros. A direção é o ângulo de chegada do componente de som direto. A difusão é um

valor entre 0 e 1, que fornece a informação do tamanho da quantidade da energia de som total difusa, por exemplo, suposto para chegar incoerentemente de todas as direções. Esta é uma aproximação do campo de som, mas quando aplicado nas faixas de 5 frequência perceptuais, uma representação perceptualmente boa do campo de som é fornecida. A direção, difusão e toda a energia do campo de som conhecidos são supostos em um domínio de tempo-frequência. Estes são formulados utilizando a informação na matriz de covariância do microfone  $\mathbf{C}_x$ . Tem-se uma configuração do alto-falante do canal N. As etapas para gerar  $\mathbf{C}_y$  são semelhantes ao 10 *upmixing*, conforme segue:

1. Gerar uma matriz NxN de zeros como  $\mathbf{C}_y$ .
2. Colocar a quantidade de energia direta, que é  $(1 - \text{difusão}) * \text{energia total}$ , ao diagonal de  $\mathbf{C}_y$  correspondente a 15 dois alto-falantes mais próximos da direção analisada. A distribuição da energia entre estas pode ser adquirida pelas leis da posição da amplitude. A posição da amplitude é coerente, então adicione ao não-diagonal correspondente uma raiz quadrada dos produtos das energias dos dois canais.
- 20 3. Distribuir ao diagonal de  $\mathbf{C}_y$  a quantidade de energia difusa, que é difusão \* energia total. A distribuição pode ser feita, por exemplo, de modo que mais energia é colocada nestas direções onde os alto-falantes são insuficientes. Agora se tem o alvo  $\mathbf{C}_y$ .

25 Embora alguns aspectos tenham sido descritos no contexto de um aparelho, fica claro que estes aspectos também representam uma descrição do método correspondente, onde um bloco ou dispositivo corresponde a uma etapa do método ou a uma função

de uma etapa do método. De modo análogo, os aspectos descritos no contexto de uma etapa do método também representam uma descrição de um bloco correspondente ou item ou função de um aparelho correspondente.

5 Dependendo de certas exigências de implantação, as aplicações da invenção podem ser implementadas em hardware ou em software. A implementação pode ser realizada utilizando um meio de armazenamento digital, por exemplo, um disquete, um DVD, um CD, uma memória ROM, uma PROM, uma EPROM, uma EEPROM ou uma memória 10 FLASH, tendo sinais de controle legíveis eletronicamente armazenados nele, que cooperam (ou são capazes de cooperar) com um sistema de computador programável de modo que o respectivo método seja realizado.

15 Algumas aplicações de acordo com a invenção compreendem um suporte de dados tendo sinais de controle eletronicamente legíveis que são capazes de cooperar com um sistema de computador programável, de modo que um dos métodos descritos aqui seja realizado.

20 Geralmente, as aplicações da presente invenção podem ser implementadas como um produto de programa de computador com um código do produto, o código do produto sendo operativo para realizar um dos métodos quando o produto do programa de computador for executado em um computador. O código do produto pode, por exemplo, ser armazenado em um suporte legível por máquina.

25 Outras aplicações compreendem o programa de computador para realizar um dos métodos descritos aqui, armazenado em um suporte legível por máquina ou um meio de armazenamento não transitório.

Em outras palavras, uma aplicação do método inventivo é, portanto, um programa de computador tendo um código do produto para realizar um dos métodos descritos aqui, quando o programa de computador é executado em um computador.

5 Outra aplicação do método inventivo é, portanto, um suporte de dados (ou um meio de armazenamento digital, ou um meio legível por computador) compreendendo, gravado nele, o programa de computador para realizar um dos métodos descritos aqui.

10 Outra aplicação do método inventivo é, portanto, um fluxo de dados ou uma sequência de sinais representando o programa de computador para realizar um dos métodos descritos aqui. O fluxo de dados ou a sequência de sinais pode, por exemplo, ser configurado para ser transferido através de uma conexão de 15 comunicação de dados, por exemplo, através da Internet.

Uma aplicação adicional compreende um meio de processamento, por exemplo, um computador, ou um dispositivo lógico programável, configurado ou adaptado para realizar um dos métodos descritos aqui.

20 Uma aplicação adicional compreende um computador tendo instalado nele o programa de computador para realizar um dos métodos descritos aqui.

Em algumas aplicações, um dispositivo lógico programável (por exemplo, um arranjo de portas lógicas 25 programáveis) pode ser utilizado para realizar algumas ou todas as funcionalidades dos métodos descritos aqui. Em algumas aplicações, um arranjo de portas lógicas programáveis pode cooperar com um microprocessador para realizar um dos métodos descritos aqui.

Geralmente, os métodos são preferivelmente realizados por qualquer aparelho de hardware.

As aplicações acima descritas são meramente ilustrativas para os princípios da presente invenção. Deve ser 5 entendido que modificações e variações das disposições e detalhes descritos aqui serão evidentes a outros especialistas na técnica. É a intenção, portanto, serem limitadas apenas pelo escopo das reivindicações de patente pendente e não pelos detalhes específicos apresentados em forma de descrição e explicação das 10 aplicações aqui contidas.

LITERATURA:

[1] C. Faller, "Multiple-Loudspeaker Playback of Stereo Signals", *Journal of the Audio Engineering Society*, Vol. 54, No. 11, pp. 1051-1064, June 2006.

15 [2] V. Pulkki, "Spatial Sound Reproduction with Directional Audio Coding", *Journal of the Audio Engineering Society*, Vol. 55, No. 6, pp. 503-516, June 2007.

[3] C. Tournery, C. Faller, F. Küch, J. Herre, "Converting Stereo Microphone Signals Directly to MPEG Surround", 20 128th AES Convention, May 2010.

[4] J. Breebaart, S. van de Par, A. Kohlrausch and E. Schuijers, "Parametric Coding of Stereo Audio," *EURASIP Journal on Applied Signal Processing*, Vol. 2005, No. 9, pp. 1305-1322, 2005.

25 [5] J. Herre, K. Kjörling, J. Breebaart, C. Faller, S. Disch, H. Purnhagen, J. Koppens, J. Hilpert, J. Rödén, W. Oomen, K. Linzmeier and K. S. Chong, "MPEG Surround - The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio

Coding", Journal of the Audio Engineering Society, Vol. 56, No. 11, pp. 932-955, November 2008.

[6] J. Vilkamo, V. Pulkki, "Directional Audio Coding: Virtual Microphone-Based Synthesis and Subjective Evaluation", Journal of the Audio Engineering Society, Vol. 57, No. 9, pp. 709-724, September 2009.

[7] Golub, G.H. and Van Loan, C.F., "Matrix computations", Johns Hopkins Univ Press, 1996.

[8] R. Rebonato, P. Jäckel, "The most general methodology to create a valid correlation matrix for risk management and option pricing purposes", Journal of Risk, Vol. 2, No. 2, pp. 17-28, 2000.

**REIVINDICAÇÕES**

1. Um aparelho para gerar um sinal de saída de áudio com dois ou mais canais de saída de áudio de um sinal de entrada de áudio com dois ou mais canais de entrada de áudio, compreendendo:

um provedor (110) para fornecer as primeiras propriedades de covariância do sinal de entrada de áudio, e

um processador de sinal (120) para gerar o sinal de saída de áudio aplicando uma regra de mixagem em, pelo menos, dois de dois canais de entrada de áudio ou mais,

caracterizado pelo processador de sinal (120) ser configurado para determinar a regra de mixagem com base nas primeiras propriedades de covariância do sinal de entrada de áudio e com base nas segundas propriedades de covariância do sinal de saída de áudio, as segundas propriedades de covariância diferentes das primeiras propriedades de covariância.

2. Um aparelho de acordo com a reivindicação 1, caracterizado pelo provedor (110) ser adaptado para fornecer as primeiras propriedades de covariância, em que as primeiras propriedades de covariância têm um primeiro estado para uma primeira posição de tempo-frequência e em que as primeiras propriedades de covariância têm um segundo estado, diferente do primeiro estado, para uma segunda posição de tempo-frequência, diferente da primeira posição de tempo-frequência.

3. Um aparelho de acordo com a reivindicação 1 ou 2, caracterizado pelo processador de sinal (120) ser adaptado para determinar a regra de mixagem com base nas segundas propriedades de covariância, em que as segundas propriedades de

covariância têm um terceiro estado para uma terceira posição de tempo-frequência e, em que as segundas propriedades de covariância têm um quarto estado, diferente do terceiro estado para uma quarta posição de tempo-frequência, diferente da terceira posição de tempo-frequência.

4. Um aparelho de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo processador de sinal (120) ser adaptado para gerar o sinal de saída de áudio aplicando a regra de mixagem de modo que cada um dos dois canais de saída de áudio ou mais depende de cada um dos dois canais de entrada de áudio ou mais.

5. Um aparelho de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo processador de sinal (120) ser adaptado para determinar a regra de mixagem de modo que uma medição de erro seja reduzida.

6. Um aparelho de acordo com a reivindicação 5, caracterizado pelo processador de sinal (120) ser adaptado para determinar a regra de mixagem de modo que a regra de mixagem depende de

$$\| \mathbf{y}_{\text{ref}} - \mathbf{y} \|^2$$

em que

$$\mathbf{y}_{\text{ref}} = \mathbf{Q}\mathbf{x} ,$$

em que  $\mathbf{x}$  é o sinal de entrada de áudio, em que  $\mathbf{Q}$  é uma matriz de mapeamento e em que  $\mathbf{y}$  é o sinal de saída de áudio.

7. Um aparelho de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo processador de sinal (120) ser configurado para determinar a regra de mixagem determinando as segundas propriedades de covariância, em que o

processador de sinal (120) é configurado para determinar as segundas propriedades de covariância com base nas primeiras propriedades de covariância.

8. Um aparelho de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo processador de sinal (120) ser adaptado para determinar uma matriz de mixagem como a regra de mixagem, em que o processador de sinal (120) é adaptado para determinar a matriz de mixagem com base nas primeiras propriedades de covariância e com base nas segundas propriedades de covariância.

9. Um aparelho de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo provedor (110) ser adaptado para fornecer as primeiras propriedades de covariância determinando uma primeira matriz de covariância do sinal de entrada de áudio e em que o processador de sinal (120) é configurado para determinar a regra de mixagem com base em uma segunda matriz de covariância do sinal de saída de áudio como as segundas propriedades de covariância.

10. Um aparelho de acordo com a reivindicação 9, caracterizado pelo provedor (110) ser adaptado para determinar a primeira matriz de covariância, de modo que cada valor da diagonal da primeira matriz de covariância indique uma energia de um dos canais de entrada de áudio e de modo que cada valor da primeira matriz de covariância, que não é um valor da diagonal indique uma correlação intercanal entre um primeiro canal de entrada de áudio e um diferente segundo canal de entrada de áudio.

11. Um aparelho de acordo com a reivindicação 9 ou 10, caracterizado pelo processador de sinal (120) ser

configurado para determinar a regra de mixagem com base na segunda matriz de covariância, em que cada valor da diagonal da segunda matriz de covariância indica uma energia de um dos canais de saída de áudio e em que cada valor da segunda matriz de covariância, que não é um valor da diagonal, indica uma correlação intercanal entre um primeiro canal de saída de áudio e um segundo canal de saída de áudio.

12. Um aparelho de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo processador de sinal (120) ser adaptado para determinar uma matriz de mixagem como a regra de mixagem, em que o processador de sinal (120) é adaptado para determinar a matriz de mixagem com base nas primeiras propriedades de covariância e com base nas segundas propriedades de covariância, em que o provedor (110) é adaptado para fornecer as primeiras propriedades de covariância determinando uma primeira matriz de covariância do sinal de entrada de áudio e em que o processador de sinal (120) é configurado para determinar a regra de mixagem com base em uma segunda matriz de covariância do sinal de saída de áudio como as segundas propriedades de covariância, em que o processador de sinal (120) é adaptado para determinar a matriz de mixagem de modo que:

$$\mathbf{M} = \mathbf{K}_y \mathbf{P} \mathbf{K}_x^{-1},$$

de modo que ,

$$\mathbf{K}_x \mathbf{K}_x^T = \mathbf{C}_x,$$

$$\mathbf{K}_y \mathbf{K}_y^T = \mathbf{C}_y$$

em que **M** é a matriz de mixagem, em que **C<sub>x</sub>** é a primeira matriz de covariância, em que **C<sub>y</sub>** é a segunda matriz de covariância, em que **K<sub>x</sub><sup>T</sup>** é uma primeira matriz transposta de uma

primeira matriz decomposta  $\mathbf{K}_x$ , em que  $\mathbf{K}_y^T$  é uma segunda matriz transposta de uma segunda matriz decomposta  $\mathbf{K}_y$ , em que  $\mathbf{K}_x^{-1}$  é uma matriz inversa da primeira matriz decomposta  $\mathbf{K}_x$  e em que  $\mathbf{P}$  é uma primeira matriz unitária.

13. Um aparelho de acordo com a reivindicação 12, caracterizado pelo processador de sinal (120) ser adaptado para determinar a matriz de mixagem de modo que

$$\mathbf{M} = \mathbf{K}_y \mathbf{P} \mathbf{K}_x^{-1},$$

em que

$$\mathbf{P} = \mathbf{V} \mathbf{\Lambda} \mathbf{U}^T,$$

em que  $\mathbf{U}^T$  é uma terceira matriz transposta de uma segunda matriz unitária  $\mathbf{U}$ , em que  $\mathbf{V}$  é uma terceira matriz unitária, em que  $\mathbf{\Lambda}$  é uma matriz de identidade anexada com zeros, em que

$$\mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y,$$

em que  $\mathbf{Q}^T$  é uma quarta matriz transposta da matriz de mapeamento  $\mathbf{Q}$ ,

em que  $\mathbf{V}^T$  é uma quinta matriz transposta da terceira matriz unitária  $\mathbf{V}$  e em que  $\mathbf{S}$  é uma matriz diagonal.

14. Um aparelho de acordo com a reivindicação 1, caracterizado pelo processador de sinal (120) ser adaptado para determinar uma matriz de mixagem como a regra de mixagem, em que o processador de sinal (120) é adaptado para determinar a matriz de mixagem com base nas primeiras propriedades de covariância e com base nas segundas propriedades de covariância,

em que o provedor (110) é adaptado para fornecer as primeiras propriedades de covariância determinando uma primeira matriz de covariância do sinal de entrada de áudio e,

em que o processador de sinal (120) é configurado para determinar a regra de mixagem com base em uma segunda matriz de covariância do sinal de saída de áudio como as segundas propriedades de covariância,

em que o processador de sinal (120) é adaptado para determinar a regra de mixagem modificando pelo menos alguns valores da diagonal de uma matriz diagonal  $\mathbf{S}_x$  quando os valores da matriz diagonal  $\mathbf{S}_x$  são zero ou menores do que um valor limite, de modo que os valores sejam maiores do que ou iguais ao valor limite,

em que a matriz diagonal depende da primeira matriz de covariância.

15. Um aparelho de acordo com a reivindicação 14, caracterizado pelo processador de sinal (120) ser configurado para modificar, pelo menos, alguns valores diagonais da matriz diagonal  $\mathbf{S}_x$ , em que  $\mathbf{K}_x = \mathbf{U}_x \mathbf{S}_x \mathbf{V}_x^T$ , e em que  $\mathbf{C}_x = \mathbf{K}_x \mathbf{K}_x^T$ , em que  $\mathbf{C}_x$  é a primeira matriz de covariância, em que  $\mathbf{S}_x$  é a matriz diagonal, em que  $\mathbf{U}_x$  é uma segunda matriz,  $\mathbf{V}_x^T$  é uma terceira matriz transposta e em que  $\mathbf{K}_x^T$  é uma quarta matriz transposta da quinta matriz  $\mathbf{K}_x$ , e em que  $\mathbf{V}_x$  e  $\mathbf{U}_x$  são matrizes unitárias.

16. Um aparelho de acordo com a reivindicação 14 ou 15, caracterizado pelo processador de sinal (120) ser adaptado para gerar o sinal de saída de áudio aplicando a matriz de mixagem em, pelo menos, dois dos dois canais de entrada de áudio ou mais para obter um sinal intermediário e adicionando um sinal residual ao sinal intermediário para obter o sinal de saída de áudio.

17. Um aparelho de acordo com a reivindicação 14 ou 15, caracterizado pelo processador de sinal (120) ser adaptado

para determinar a matriz de mixagem com base em uma matriz de ganho diagonal  $\mathbf{G}$  e uma matriz intermediária  $\hat{\mathbf{M}}$ , de modo que  $\mathbf{M}' = \mathbf{G}\hat{\mathbf{M}}$ , em que a matriz de ganho diagonal tem o valor

$$\mathbf{G}(i,i) = \sqrt{\frac{\mathbf{C}_y(i,i)}{\hat{\mathbf{C}}_y(i,i)}}$$

onde  $\hat{\mathbf{C}}_y = \hat{\mathbf{M}}\mathbf{C}_x\hat{\mathbf{M}}^T$ ,

em que  $\mathbf{M}'$  é a matriz de mixagem, em que  $\mathbf{G}$  é a matriz de ganho diagonal, em que  $\mathbf{C}_y$  é a segunda matriz de covariância e em que  $\hat{\mathbf{M}}^T$  é uma quinta matriz transposta da matriz intermediária  $\hat{\mathbf{M}}$ .

18. Um aparelho de acordo com a reivindicação 1, caracterizado pelo processador de sinal (120) compreender:

um módulo de formulação da matriz de mixagem (420; 530; 630; 730; 830; 1030) para gerar uma matriz de mixagem como a regra de mixagem com base nas primeiras propriedades de covariância e,

um módulo de utilização da matriz de mixagem (430; 540; 640; 740; 840; 1040) para aplicar a matriz de mixagem no sinal de entrada de áudio para gerar o sinal de saída de áudio.

19. Um aparelho de acordo com a reivindicação 18, caracterizado pelo provedor (110) compreender um módulo de análise da matriz de covariância (410; 705; 805; 1005) para fornecer a entrada das propriedades de covariância do sinal de entrada de áudio de entrada para obter um resultado de análise como as primeiras propriedades de covariância e,

em que o módulo de formulação da matriz de mixagem (420; 530; 630; 730; 830; 1030) é adaptado para gerar a matriz de mixagem com base no resultado de análise.

20. Um aparelho de acordo com a reivindicação 18 ou 19, caracterizado pelo módulo de formulação da matriz de mixagem (420; 530; 630; 730; 830; 1030) ser adaptado para gerar a matriz de mixagem com base em um critério de erro.

21. Um aparelho de acordo com qualquer uma das reivindicações 18 a 20,

caracterizado pelo processador de sinal (120) compreender, ainda, um módulo de determinação de dados espaciais (520; 620) para determinar os dados de informação da configuração compreendendo os dados espaciais surround, dados de correlação intercanal ou dados do nível de sinal de áudio e,

em que o módulo de formulação da matriz de mixagem (420; 530; 630; 730; 830; 1030) é adaptado para gerar a matriz de mixagem com base nos dados de informação da configuração.

22. Um aparelho de acordo com qualquer uma das reivindicações 18 a 20,

caracterizado pelo processador de sinal (120) compreender, além disso, um módulo de formulação da matriz de covariância alvo (730; 1018) para gerar uma matriz de covariância alvo com base no resultado de análise e,

em que o módulo de formulação da matriz de mixagem (420; 530; 630; 730; 830; 1030) é adaptado para gerar uma matriz de mixagem com base na matriz de covariância alvo.

23. Um aparelho de acordo com a reivindicação 22, caracterizado pelo módulo de formulação da matriz de covariância alvo (1018) ser configurado para gerar a matriz de covariância alvo com base em uma configuração do alto-falante.

24. Um aparelho de acordo com a reivindicação 18 a 19, caracterizado pelo processador de sinal (120) compreender, ainda, um módulo de melhoria (815) para obter os dados de correlação intercanal de saída com base nos dados de correlação intercanal de entrada, diferentes dos dados de correlação intercanal de entrada e,

em que o módulo de formulação da matriz de mixagem (420; 530; 630; 730; 830; 1030) é adaptado para gerar a matriz de mixagem com base nos dados de correlação intercanal de saída.

25. Um método para gerar um sinal de saída de áudio com dois canais de saída de áudio ou mais de um sinal de entrada de áudio com dois canais de entrada de áudio ou mais, compreendendo:

fornecer as primeiras propriedades de covariância do sinal de entrada de áudio e,

gerar o sinal de saída de áudio aplicando uma regra de mixagem em, pelo menos, dois de dois canais de entrada de áudio ou mais,

caracterizado pela regra de mixagem ser determinada com base nas primeiras propriedades de covariância do sinal de entrada de áudio e com base nas segundas propriedades de covariância do sinal de saída de áudio diferentes das primeiras propriedades de covariância.

26. Memória lida por computador caracterizada por compreender instruções que quando executadas realiza o método da reivindicação 25.

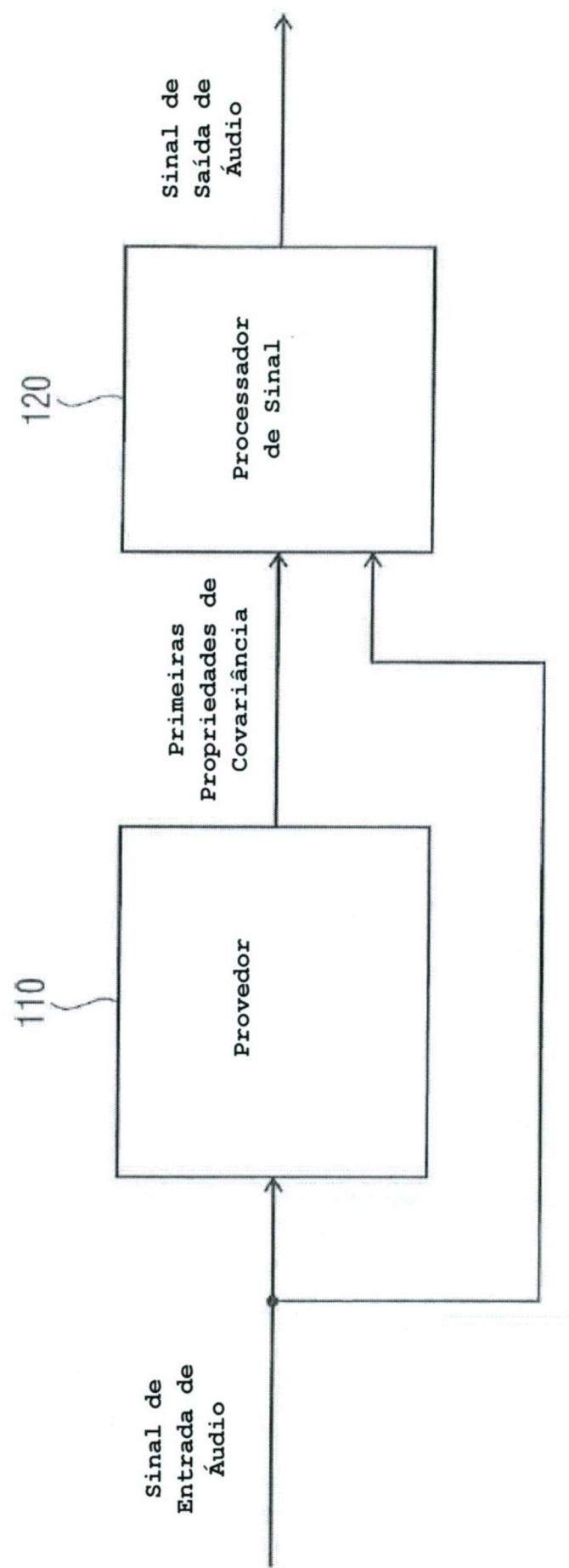


FIG 1

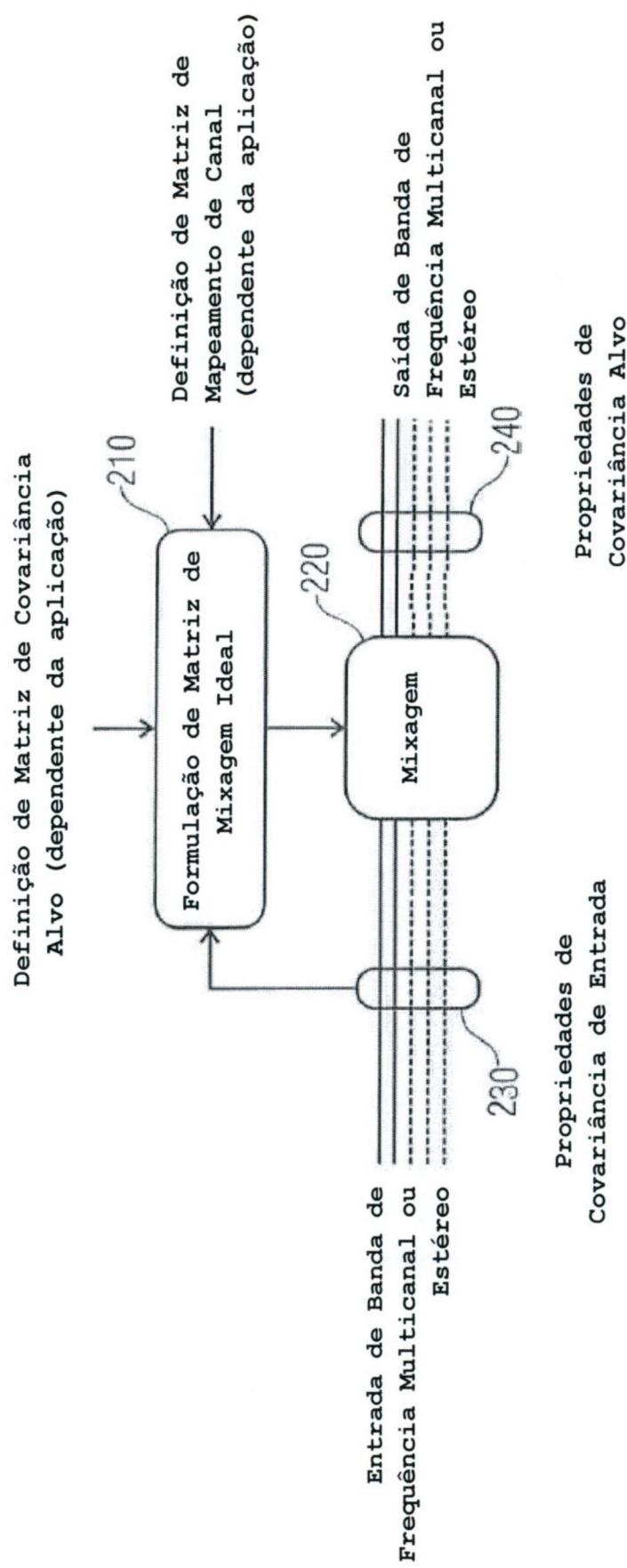


FIG 2

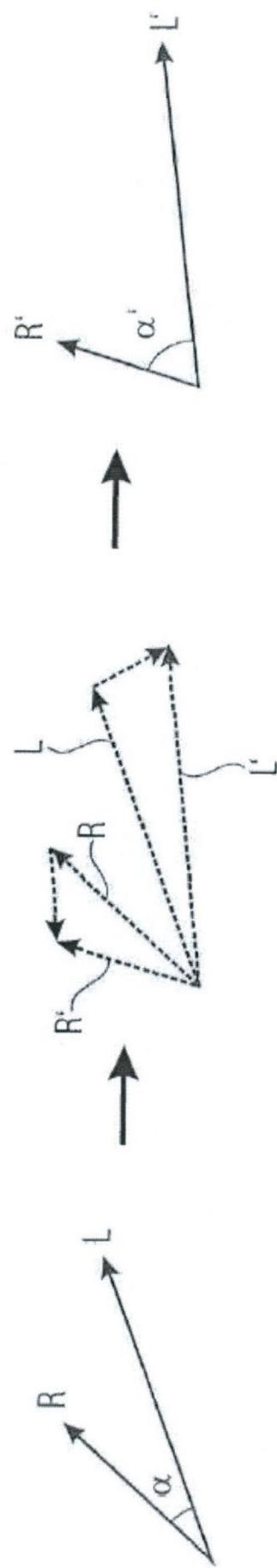


FIG 3

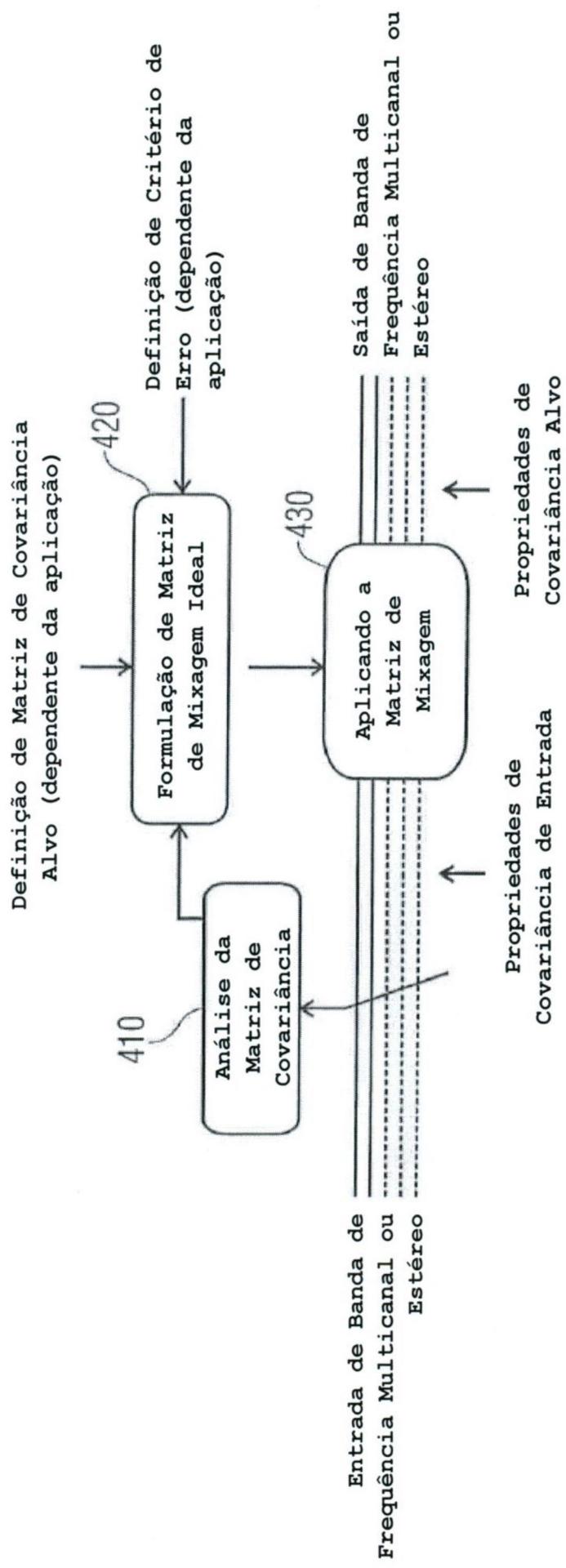


FIG 4

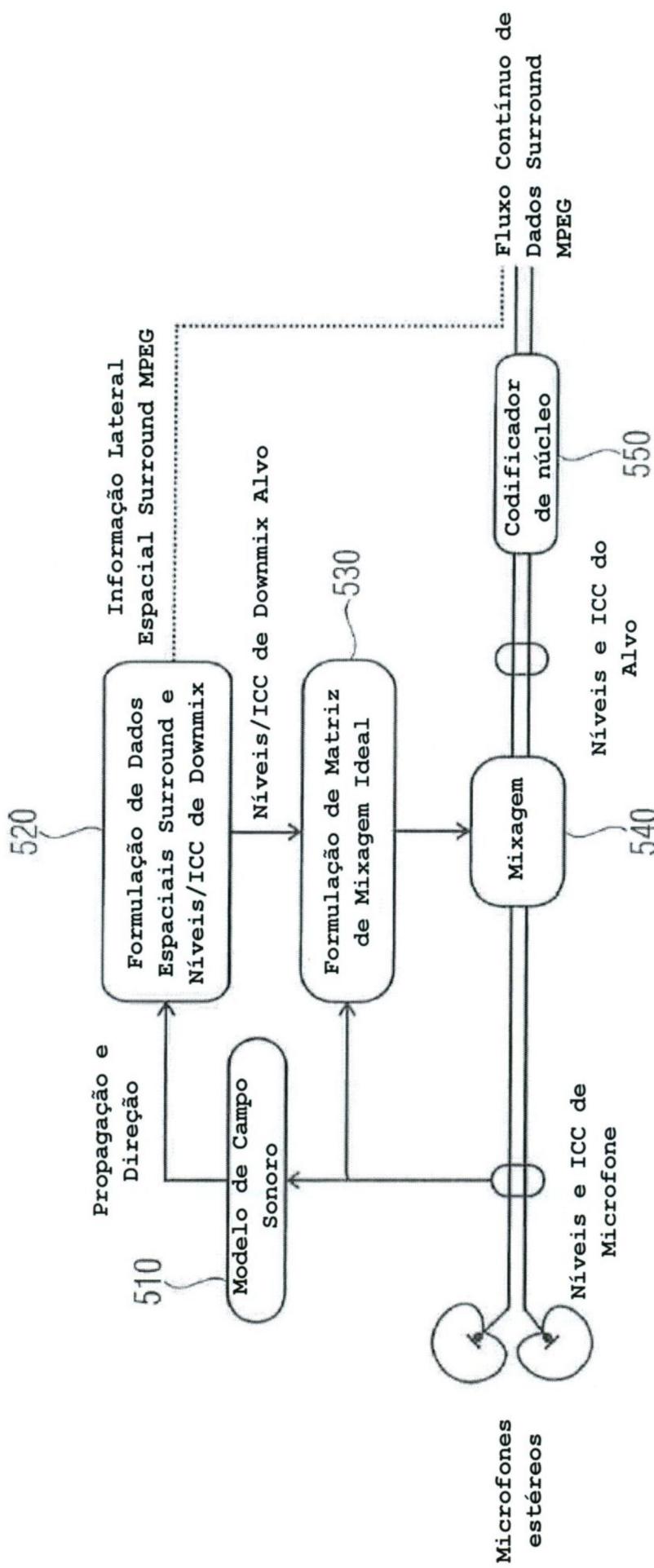


FIG 5

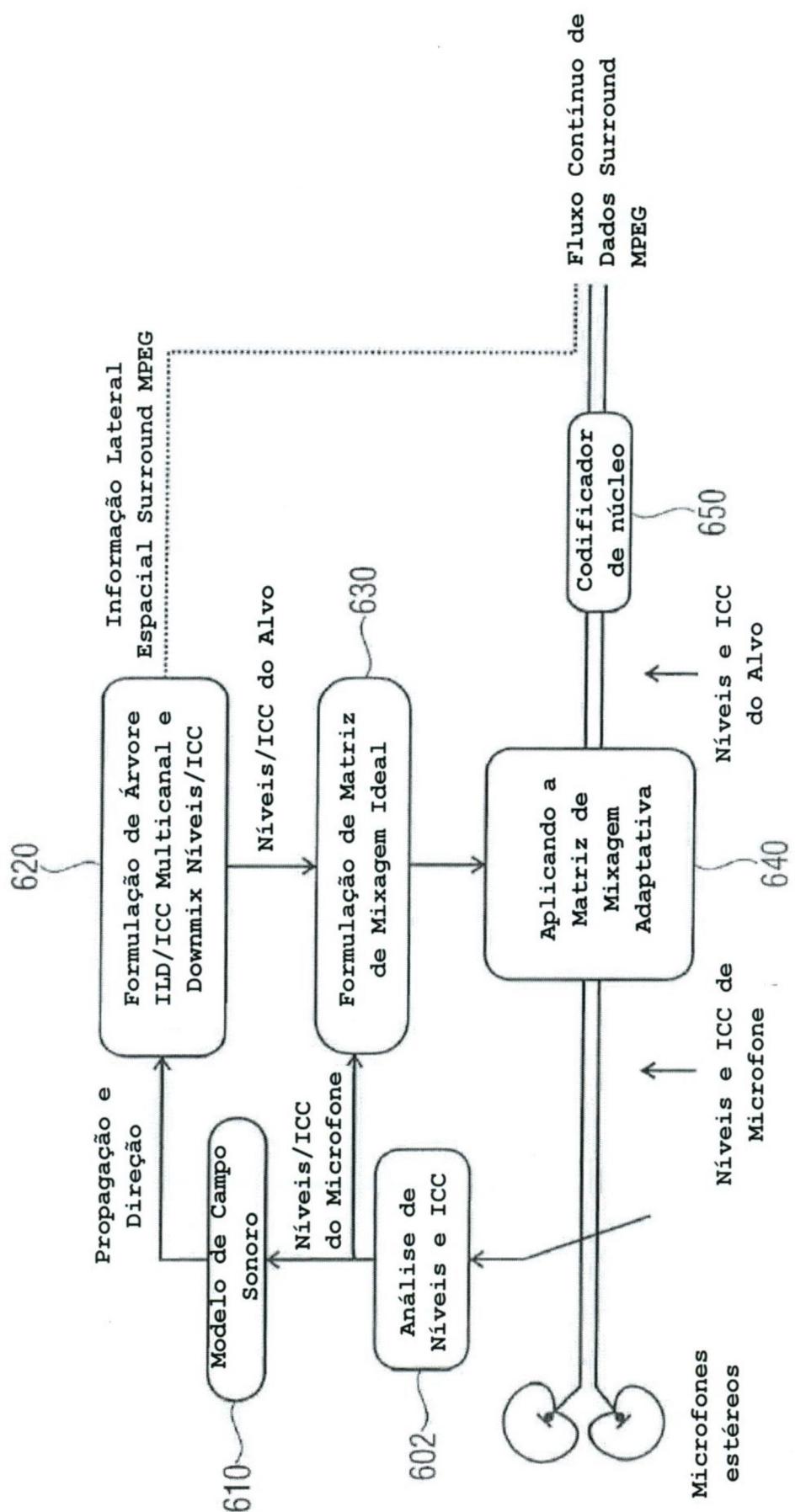


FIG 6

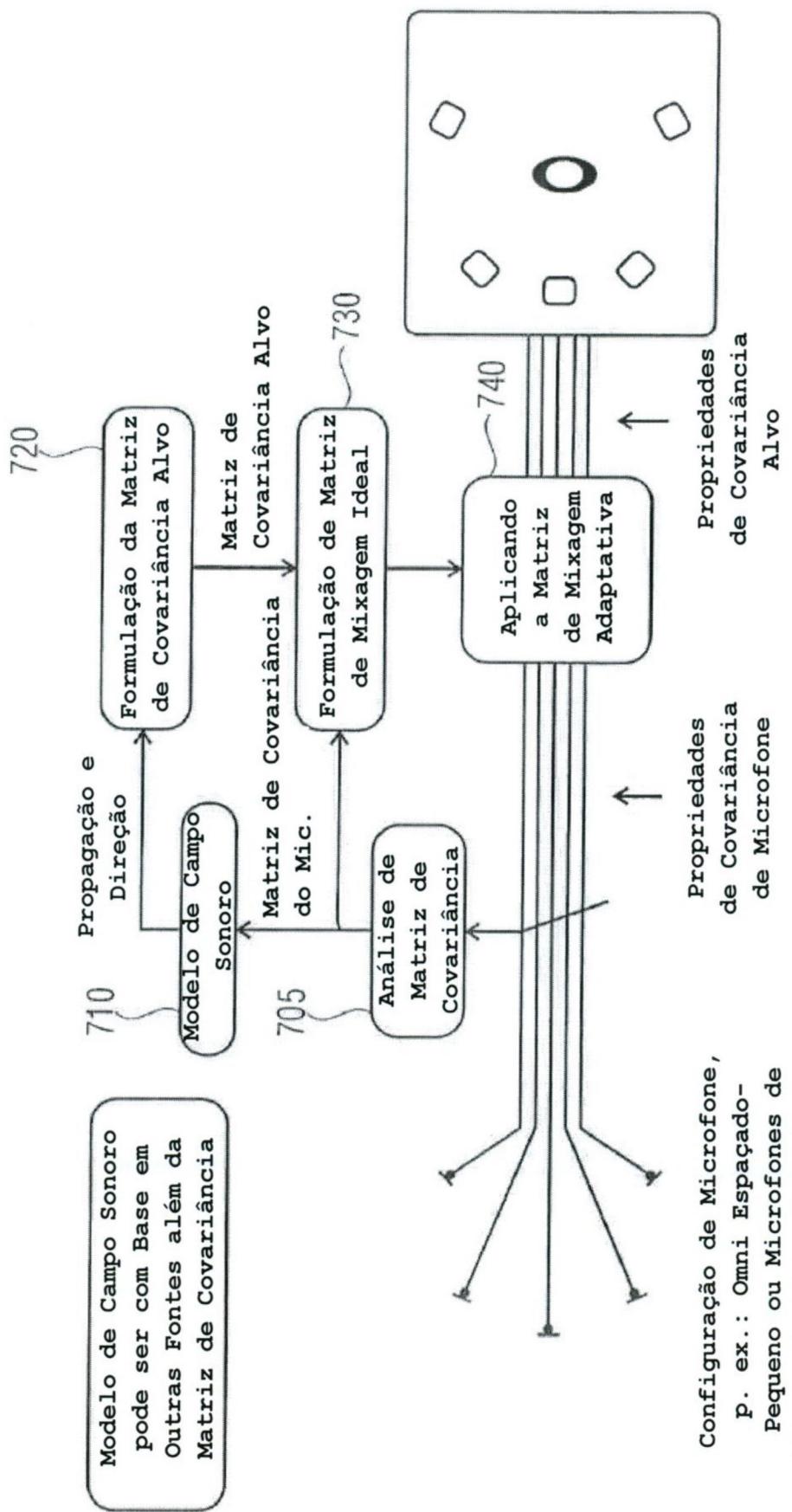


FIG 7

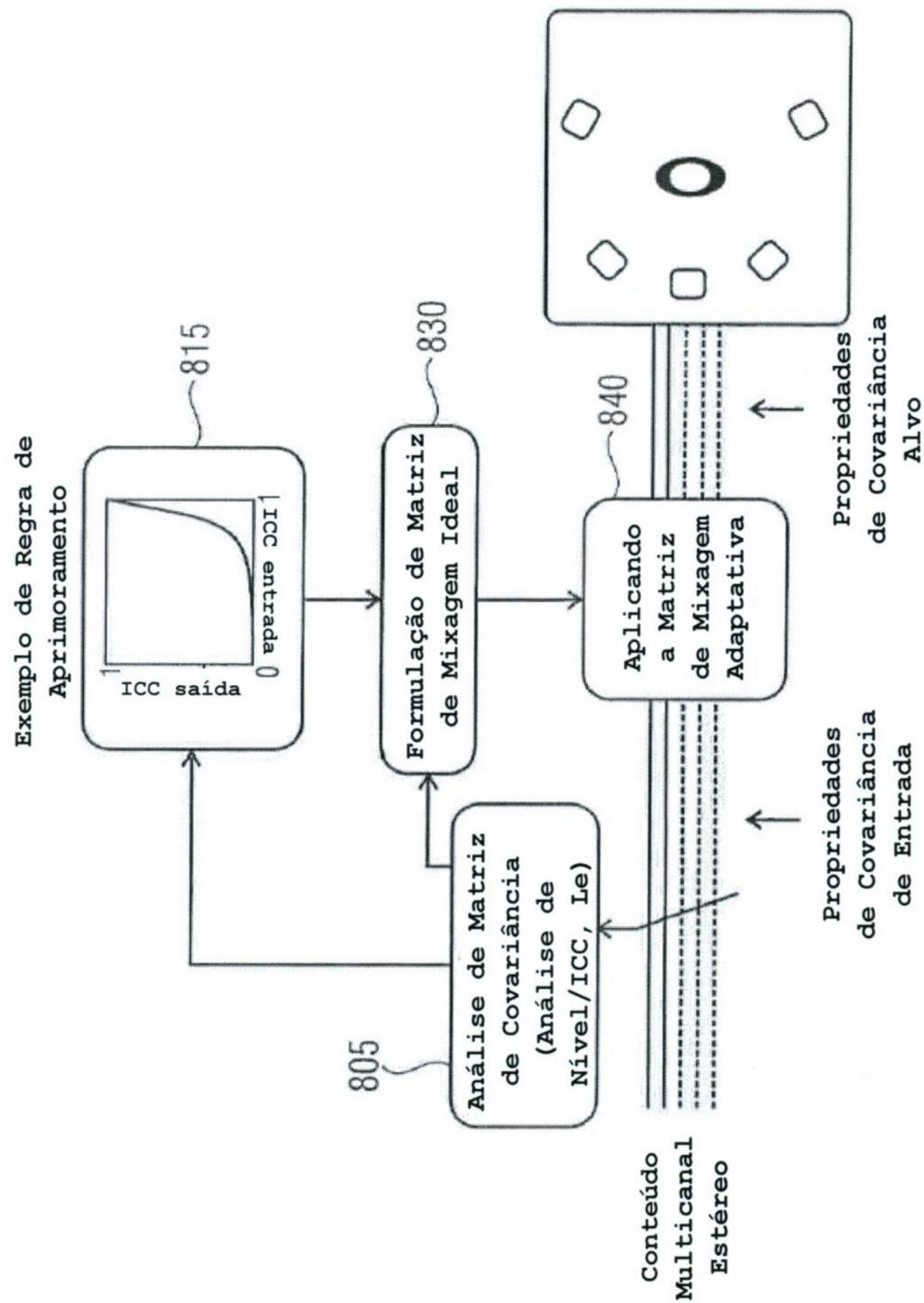


FIG 8

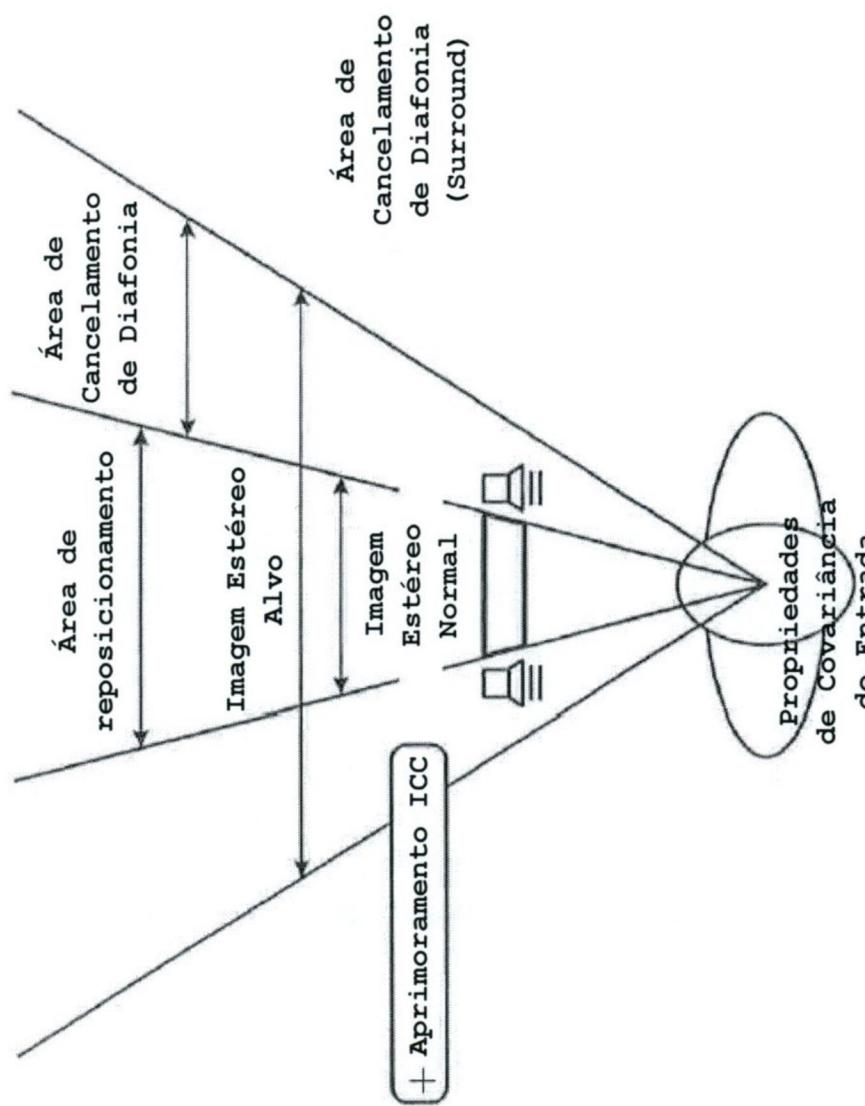


FIG 9

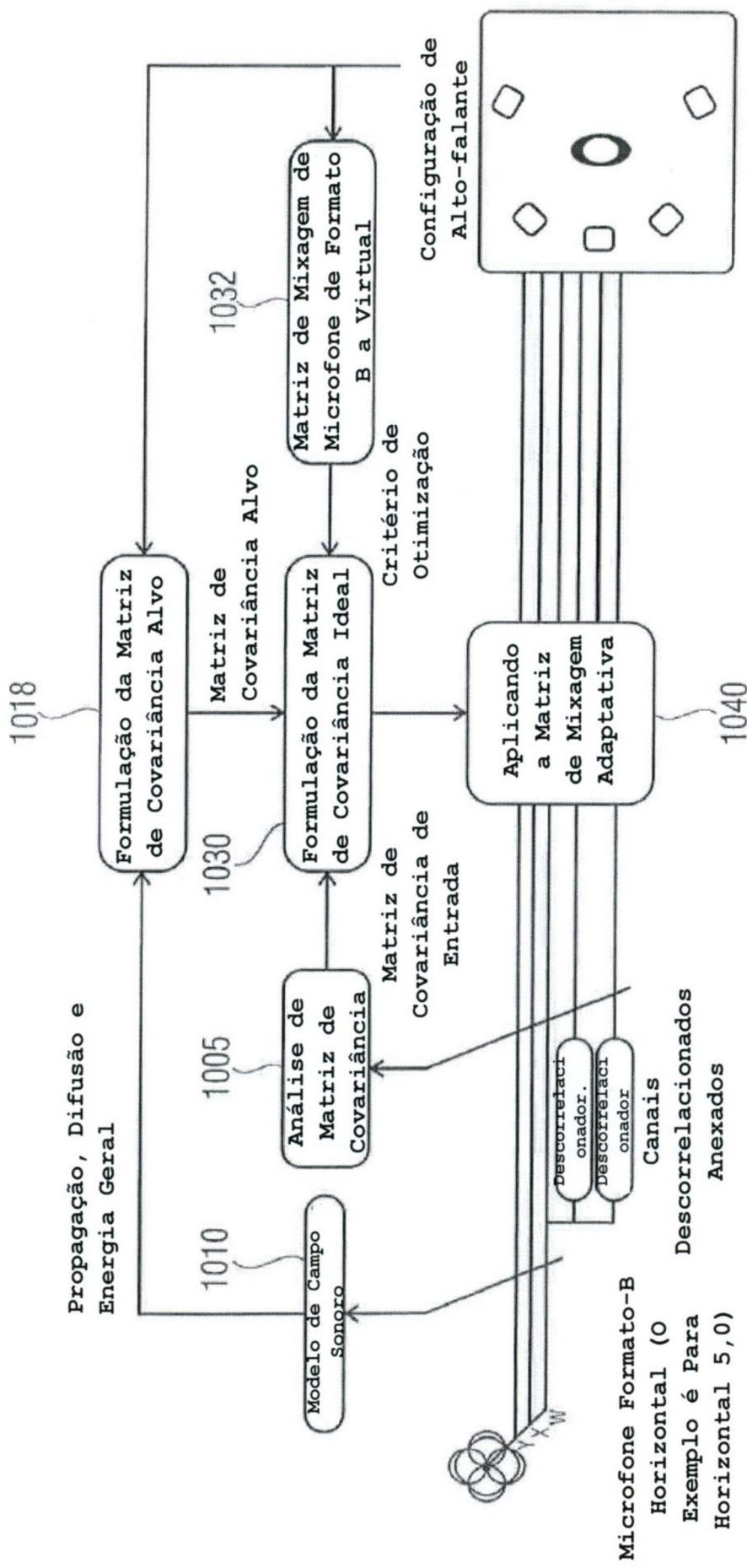


FIG 10

Contexto	$C_x$
Descorrelação: ICC de Entrada Alta	$\begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$
Descorrelação: ICC de Entrada Muito Alta	$\begin{bmatrix} 1 & 0.97 \\ 0.97 & 1 \end{bmatrix}$
Upmixing Estéreo	$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$
Upmixing de 5,0 a 7,0	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$
Downmixing: com a Mesma Coerência Não-Zero	$\begin{bmatrix} 1 & 0 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 & 0.5 \\ 0 & 0 & 1 & 0 & 0 \\ 0.5 & 0 & 0 & 1 & 0 \\ 0 & 0.5 & 0 & 0 & 1 \end{bmatrix}$
Descorrelação: Coincidente Cardioide Espaçado-Igual 5,0	$\begin{bmatrix} 1 & 0.86 & 0.64 & 0.64 & 0.86 \\ 0.86 & 1 & 0.86 & 0.64 & 0.64 \\ 0.64 & 0.86 & 1 & 0.86 & 0.64 \\ 0.64 & 0.64 & 0.86 & 1 & 0.86 \\ 0.86 & 0.64 & 0.64 & 0.86 & 1 \end{bmatrix}$
Descorrelação: Hipercardioide Coincidente de Layout Padrão 5,0	$\begin{bmatrix} 1 & 0.65 & 0.91 & 0.43 & -0.22 \\ 0.65 & 1 & 0.91 & -0.22 & 0.43 \\ 0.91 & 0.91 & 1 & 0.07 & 0.07 \\ 0.43 & -0.22 & 0.07 & 1 & -0.22 \\ -0.22 & 0.43 & 0.07 & -0.22 & 1 \end{bmatrix}$

Tabela 1

FIG 11	FIG 11A	FIG 11B	FIG 11C

FIG 11A

Q	$C_y$
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.71 & 0.71 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0 \\ 0 & 0 & 0 & 0 & 0.71 \\ 0 & 0 & 0 & 0.71 & 0 \\ 0 & 0 & 0 & 0 & 0.71 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0.5 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 0.71 & 1 & 0 \\ 0 & 1 & 0.71 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 2.5 & 0.5 \\ 0.5 & 2.5 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

Tabela 1

FIG 11	FIG 11A	FIG 11B	FIG 11C

FIG 11B

M	C <sub>r</sub>
$\begin{bmatrix} 1.5 & -0.75 \\ -0.75 & 1.5 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
$\begin{bmatrix} 2.1 & -1.4 \\ -1.4 & 2.1 \end{bmatrix}$	$\begin{bmatrix} 0.31 & -0.31 \\ -0.31 & 0.31 \end{bmatrix}$
$\begin{bmatrix} 0.33 & -0.17 \\ -0.17 & 0.33 \\ 0.47 & 0.47 \\ 0.33 & -0.17 \\ -0.17 & 0.33 \end{bmatrix}$	$\begin{bmatrix} 0.33 & 0.08 & -0.24 & -0.17 & 0.08 \\ 0.08 & 0.33 & -0.24 & 0.08 & -0.17 \\ -0.24 & -0.24 & 0.67 & -0.24 & -0.24 \\ -0.17 & 0.08 & -0.24 & 0.33 & 0.08 \\ 0.08 & -0.17 & -0.24 & 0.08 & 0.33 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 0 & -0.25 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0 & -0.25 \\ 0 & 0 & 0 & -0.25 & 0 & 0.25 & 0 \\ 0 & 0 & 0 & 0 & -0.25 & 0 & 0.25 \end{bmatrix}$
$\begin{bmatrix} 0.84 & 0.02 & 0.61 & 0.84 & 0.02 \\ 0.02 & 0.84 & 0.61 & 0.02 & 0.84 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
$\begin{bmatrix} 1.7 & -0.53 & -0.05 & -0.05 & -0.53 \\ -0.53 & 1.7 & -0.53 & -0.05 & -0.05 \\ -0.05 & -0.53 & 1.7 & -0.53 & -0.05 \\ -0.05 & -0.05 & -0.53 & 1.7 & -0.53 \\ -0.53 & -0.05 & -0.05 & -0.53 & 1.7 \end{bmatrix}$	$\begin{bmatrix} 0.4 & -0.32 & 0.12 & 0.12 & -0.32 \\ -0.32 & 0.4 & -0.32 & 0.12 & 0.12 \\ 0.12 & -0.32 & 0.4 & -0.32 & 0.12 \\ 0.12 & 0.12 & -0.32 & 0.4 & -0.32 \\ -0.32 & 0.12 & 0.12 & -0.32 & 0.4 \end{bmatrix}$
$\begin{bmatrix} 2 & -0.51 & -0.83 & -0.53 & 0.41 \\ -0.51 & 2 & -0.83 & -0.41 & -0.53 \\ -0.83 & -0.83 & 2.1 & 0.04 & 0.04 \\ -0.53 & 0.41 & 0.04 & 1.2 & -0.07 \\ 0.41 & -0.53 & 0.04 & -0.07 & 1.2 \end{bmatrix}$	$\begin{bmatrix} 0.58 & -0.2 & -0.34 & -0.23 & 0.19 \\ -0.2 & 0.58 & -0.34 & 0.19 & -0.23 \\ -0.34 & -0.34 & 0.62 & 0.03 & 0.03 \\ -0.23 & 0.19 & 0.03 & 0.11 & -0.11 \\ 0.19 & -0.23 & 0.03 & -0.11 & 0.11 \end{bmatrix}$

Tabela 1

FIG 11	FIG 11A	FIG 11B	FIG 11C

FIG 11C

Listagem 1: Implementação de Matlab do Método Proposto

```

1 Função [M, Cr] = formular_M_e_Cr (Cx, Cy, Q, indicador)
2 % indicador = 0: Uso esperado de resíduos
3 % indicador = 1: Fixar energias no lugar
4 Lambda = olho (comprimento (Cy), comprimento (Cx));
5
6 % Decomposição de Cy
7 [U_Cy,S_Cy]=svd(Cy);
8 Ky=U_Cy*sqrt(S_Cy);
9
10 % Decomposição de Cx
11 [U_Cx,S_Cx]=svd(Cx);
12 Kx=U_Cx*sqrt(S_Cx);
13
14 % SVD de Kx
15 Ux=U_Cx;
16 Sx=sqrt(S_Cx);
17 % Vx = matriz de identidade
18
19 % Uma regularização simples do inverso
20 Sx_diag=diag(Sx)
21 Limite = Max (Sx_diag)*0,2;
22 Sx_hat_diag = max (Sx_diag, limite);
23

```

FIG 12	FIG 12A
	FIG 12B

FIG 12A

```

24 % para Formular Kx^-1 regularizado
25 Kx_hat_inverso = diag (1./Sx_hat_diag)*Ux';
26
27 % para Formular P ideal
28 [U,S,V]=svd(Kx'*Q'*Ky);
29 P=V*lambda*U';
30
31 % Utilizando o que temos de M ideal
32 M = Ky * P * Kx_hat_inverso;
33
34 % para Formular matriz de covariância residual
35 Cy_hat = M*Cx*M';
36 Cr=Cy-Cy_hat;
37
38 % Uso de compensação de energia ao invés de resíduos
39 Se indicador == 1
40     Ajuste = diag (Cy)./diag (Cy_hat +1 e -20);
41     G = diag (sqrt (ajuste));
42
43     Cr = 'desnecessário'
44 Fim

```

FIG 12

FIG 12A

FIG 12B

FIG 12B