



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

(21) PI 0912451-9 A2



(22) Data do Depósito: 12/08/2009

(43) Data da Publicação Nacional: 18/02/2010

(54) **Título:** APARELHO PARA DETERMINAR UM SINAL DE ÁUDIO ESPACIAL CONVERTIDO.

(51) **Int. Cl.:** G10H 1/00; G10L 19/14; H04S 3/02.

(30) **Prioridade Unionista:** 02/02/2009 EP 09001398.8; 13/08/2008 US 61/088,513; 25/08/2008 US 61/091,682.

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

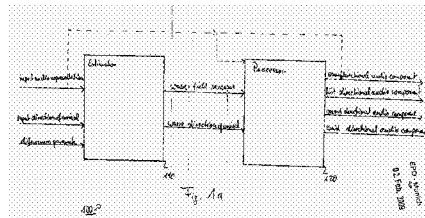
(72) **Inventor(es):** DEL GALDO, GIOVANNI; KUECH, FABIAN.; KALLINGER, MARKUS.; PULKKI, VILLE.; LAITINEN, MIKKO-VILLE.; SCHULTZ-AMLING, RICHARD..

(86) **Pedido PCT:** PCT EP2009005859 de 12/08/2009

(87) **Publicação PCT:** WO 2010/017978 de 18/02/2010

(85) **Data da Fase Nacional:** 14/02/2011

(57) **Resumo:** APARELHO PARA DETERMINAR UM SINAL DE ÁUDIO ESPACIAL CONVERTIDO Um aparelho (100) para determinar um sinal de áudio espacial convertido, tendo um componente de áudio onidirecional (W) e pelo menos um componente de áudio direcional, de um sinal de áudio espacial de entrada, o sinal de áudio espacial de entrada tendo uma representação de áudio de entrada (W) e uma direção de chegada de entrada (vazio) . O aparelho (100) compreende um estimador (110) para estimar uma representação de onda (W) compreendendo uma medida de campo de onda e uma medida de direção de chegada de onda, com base na representação de áudio de entrada (W) e direção de chegada de entrada (vazio) . O aparelho (100) compreende ainda um processador (120) para processar a medida de campo de onda e a medida de direção de chegada de onda para obter o componente de áudio onidirecional (W) e pelo menos um componente direcional (x; y; z).



Aparelho para Determinar um Sinal de Áudio Espacial Convertido

Descrição

A presente invenção está no campo de processamento de áudio, especialmente processamento de áudio espacial e conversão de formatos de áudio espacial diferentes.

Codificação de áudio DirAC (DirAC = Codificação de Áudio Direcional) é um método para a reprodução e processamento de áudio espacial. Sistemas convencionais aplicam DirAC em reprodução de qualidade alta bidimensional e tridimensional de som gravado, aplicações de teleconferência, microfones direcionais, e upmix de estéreo para surround, conforme V. Pulkki e C. Faller, "Directional audio coding: Filterbank and STFT-based design", na 120ª Convenção AES, 20-23 de Maio de 2006, Paris, França, Maio de 2006, V. Pulkki and C. Faller, Directional audio coding in spatial sound reproduction and stereo upmixing, in AES 28th International Conference, Pitea, Sweden, June 2006, V. Pulkki, "Spatial sound reproduction with directional audio coding", *Journal of the Audio Engineering Society*, 55(6):503-516, Junho de 2007, Jukka Ahonen, V. Pulkki e Tapio Lokki, "Teleconference application and B-format microphone array for directional audio coding", na 30ª Conferência Internacional AES.

Outras aplicações convencionais usando DirAC são, por exemplo, o formato de codificação universal e cancelamento de ruído. Na DirAC, algumas propriedades direcionais de som são analisadas em bandas de frequência dependentes do tempo. Os dados de análise são transmitidos com dados de áudio e sintetizados para diferentes objetivos. A análise é comumente feita usando sinais de

formato B, embora teoricamente DirAC não esteja limitada a este formato. Formato B, conforme Michael Gerzon, "Surround Sound psychoacoustics", no *Wireless World*, volume 80, páginas 483-486, dezembro de 1974, foi desenvolvido dentro do trabalho sobre

5 Ambisonics, um sistema desenvolvido pelos pesquisadores britânicos nos anos 70 para trazer o som surround de salas de concerto para salas de estar. O formato B consiste de quatro sinais, isto é $w(t)$, $x(t)$, $y(t)$, e $z(t)$. O primeiro corresponde à pressão medida por um microfone omnidirecional, enquanto os três últimos são

10 leituras de pressão de microfones tendo padrões de captação no formato de oito direcionados para os três eixos de um sistema de coordenada cartesiana. Os sinais $x(t)$, $y(t)$, e $z(t)$ são proporcionais aos componentes de vetor de velocidade de partícula direcionado para x , y e z , respectivamente.

15 A corrente DirAC ~~consiste de 1-4 canais de áudio~~ com metadados direcionais. Em teleconferência e alguns outros casos, a corrente consiste de apenas um único canal de áudio com metadados, denominado de corrente Dirac mono. Esta é uma maneira bastante compacta de descrever áudio espacial, visto que apenas um

20 único canal de áudio necessita ser transmitido com informação de lado, que, por exemplo, provê boa separação espacial entre os interlocutores. Entretanto, nestes casos alguns tipos de som, tais como cenários de som reverberado ou ambiente podem ser reproduzidos com qualidade limitada. Para obter uma qualidade

25 melhor nestes casos, canais de áudio adicionais necessitam serem transmitidos.

A conversão do formato B em DirAC é descrita por V. Pulkki, "A method for reproducing natural or modified spatial

impression in multichannel listening," Patente WO 2004/077884 A1, setembro de 2004. Codificação de Áudio Direcional é uma abordagem eficiente para a análise e reprodução de som espacial. DirAC usa uma representação paramétrica de campos de som com base nas características que são relevantes para a percepção de som espacial, isto é a DOA (DOA = Direção de chegada) e difusão do campo de som nas sub-bandas de frequência. Na verdade, DirAC assume que diferenças de tempo interaurais (ITD) e diferenças de nível interaural (ILD) são percebidas corretamente quando a DOA de um campo é corretamente reproduzida, enquanto coerência interaural (IC) é percebida corretamente, se a difusão for reproduzida com precisão. Estes parâmetros, isto é DOA e difusão, representam informação de lado que acompanha um sinal mono que é referido como uma corrente DirAC mono.

A Figura 7 mostra o codificador DirAC, que a partir dos sinais de microfones próprios, computa um canal de áudio mono e informação de lado, isto é, difusão $\Psi(\mathbf{k}, n)$ e direção de chegada $\mathbf{e}_{DOA}(\mathbf{k}, n)$. A Figura 7 mostra um codificador DirAC 200, que é adaptado para computar um canal de áudio mono e informação de lado de sinais de microfones apropriados. Em outras palavras, a Figura 7 ilustra um codificador DirAC 200 para determinar a difusão e direção de chegada de sinais de microfones apropriados. A Figura 7 mostra um codificador DirAC 200 compreendendo uma unidade de estimativa P/U 210, onde $\mathbf{P}(\mathbf{k}, n)$ representa um sinal de pressão e $\mathbf{U}(\mathbf{k}, n)$ representa um vetor de velocidade de partícula. A unidade de estimativa P/U recebe os sinais de microfones como informação de entrada, na qual a estimativa de P/U é baseada. Um estágio de

análise energética 220 permite a estimativa da direção de chegada e parâmetro de difusão da corrente DirAC mono.

Os parâmetros DirAC, como por exemplo, uma representação de áudio mono $\mathbf{W}(\mathbf{k}, \mathbf{n})$, um parâmetro de difusão $\Psi(\mathbf{k}, \mathbf{n})$ e uma direção de chegada (DOA) $\mathbf{e}_{DOA}(\mathbf{k}, \mathbf{n})$, podem ser obtidos de uma representação de frequência-tempo dos sinais de microfones. Portanto, os parâmetros são dependentes de tempo e frequência. No lado da reprodução, esta informação permite uma transmissão espacial precisa. Para recriar o som espacial em uma posição de audição desejada, um arranjo de alto-falantes múltiplos é requerido. Entretanto, sua geometria pode ser arbitrária. Na verdade, os canais de alto-falantes podem ser determinados como uma função dos parâmetros DirAC.

Existem diferenças substanciais entre DirAC e codificação de áudio de canais múltiplos paramétrica, tal como MPEG Surround, conforme Lars Villemoes, Juergen Herre, Jeroen Breebaart, Gerard Hotho, Sascha Disch, Heiko Purnhagen, e Kristofer Kjrling, "MPEG surround: The forthcoming ISO standard for spatial audio coding", na 28ª Conferência Internacional AES, Pitea, Suécia, junho de 2006, embora eles partilhem estruturas de processamento similares. Embora MPEG Surround seja baseado em uma análise de tempo/frequência dos canais de alto-falantes diferentes, DirAC toma uma entrada dos canais de microfones coincidentes, o que efetivamente descreve o campo de som em um ponto. Dessa maneira, DirAC também representa uma técnica de gravação eficiente para áudio espacial.

Outro sistema que trata de áudio espacial é SAOC (SAOC = Codificação de Objeto de Áudio Espacial), conforme Jonas

Engdegard, Barbara Resch, Cornelia Falch, Oliver Hellmuth, Johannes Hilpert, Andreas Hoelzer, Leonid Terentiev, Jeroen Breebaart, Jeroen Koppens, Erik Schuijers, e Werner Oomen, "Spatial audio object (SAOC) the upcoming MPEG standard on parametric object based audio coding", na 12ª Convenção AES, 17-20 de Maio de 2008, Amsterdã, Holanda, 2008, atualmente sob normatização ISO/MPEG. É baseada na máquina de transmissão de MPEG Surround e trata diferentes fontes de som como objetos. Esta codificação de áudio oferece eficiência bastante alta em termos de taxa de bits e provê liberdade inigualável de interação no lado da reprodução. Esta abordagem promete novas características convincentes e funcionalidade em sistemas 'legacy', assim como várias outras aplicações novas.

O objetivo da presente invenção é prover um conceito-melhorado de processamento espacial.

O objetivo é atingido por um aparelho para determinar um sinal de áudio espacial convertido de acordo com a reivindicação 1 e um método correspondente de acordo com a reivindicação 14.

A presente invenção é baseada na descoberta de que um processamento espacial melhorado pode ser atingido, por exemplo, ao converter um sinal de áudio espacial codificado, como uma corrente DirAC mono, em um sinal de formato B. Nas configurações, o sinal de formato B convertido pode ser processado ou transmitido antes de ser adicionado a alguns outros sinais de áudio e codificados de volta para uma corrente DirAC. Configurações podem ter diferentes aplicações, por exemplo, misturar diferentes tipos de correntes DirAC e de formato B,

baseado em DirAC, etc.. Configurações podem introduzir uma operação inversa a WO 2004/077884 A1, isto é a conversão de uma corrente DirAC mono em um formato B.

A presente invenção é baseada na descoberta de que processamento melhorado pode ser atingido, se sinais de áudio forem convertidos em componentes direcionais. Em outras palavras, é a descoberta da presente invenção que processamento espacial melhorado pode ser atingido quando o formato de um sinal de áudio espacial corresponde a componentes direcionais conforme gravados, por exemplo, por um microfone direcional de formato B. Além do mais, é uma descoberta da presente invenção que componentes direcionais ou omnidirecionais de diferentes fontes podem ser processados conjuntamente e com uma eficiência aumentada. Em outras palavras, especialmente quando processando sinais de áudio espaciais de múltiplas fontes de áudio, processamento pode ser executado mais eficientemente, se os sinais das múltiplas fontes de áudio estiverem disponíveis no formato de seus componentes omnidirecionais e direcionais, visto que estes podem ser processados conjuntamente. Nas configurações, portanto, geradores de efeito de áudio ou processadores de áudio podem ser usados mais eficientemente pelo processamento de componentes combinados de fontes múltiplas.

Nas configurações, sinais de áudio espacial podem ser representados como uma corrente DirAC mono denotando uma técnica de distribuição de informação de multimídia DirAC onde os dados de mídia são acompanhados apenas por um canal de áudio na transmissão. Este formato pode ser convertido, por exemplo, em uma corrente de formato B, tendo múltiplos componentes direcionais.

Configurações podem permitir processamento espacial melhorado pela conversão de sinais de áudio espacial nos componentes direcionais.

Configurações podem prover uma vantagem em relação à decodificação DirAC mono, onde apenas um canal de áudio é usado para criar todos os sinais de alto-falante, pelo fato de que processamento espacial adicional é permitido com base em componentes de áudio direcional, que são determinados antes de criar sinais de alto-falante. Configurações podem prover a vantagem de que problemas na criação de sons reverberantes são reduzidos.

Nas configurações, por exemplo, a corrente DirAC pode usar um sinal de áudio estéreo em lugar de um sinal de áudio mono, onde os canais estéreos são L (L= canal estéreo esquerdo) e R (R= canal estéreo direito) e são transmitidos para serem usados na decodificação DirAC. Configurações podem atingir uma qualidade melhor para som reverberante e prover uma compatibilidade direta com sistemas de alto-falantes estéreos, por exemplo.

Configurações podem prover a vantagem de que decodificação DirAC de microfone virtual pode ser habilitada. Detalhes sobre decodificação DirAC de microfone virtual podem ser encontrados em V. Pulkki, "Spatial sound reproduction with directional audio coding," *Journal of the Audio Engineering Society*, 55(6):503-516, junho de 2007. Estas configurações obtêm os sinais de áudio para os alto-falantes colocando microfones virtuais orientados para a posição dos alto-falantes e tendo fontes de som do tipo de ponto, cuja posição é determinada pelos parâmetros DirAC. Configurações podem prover a vantagem de que

pela conversão, combinação linear conveniente de sinais de áudio, podem ser habilitadas.

Configurações da presente invenção serão detalhadas usando as Figuras em anexo, nas quais

5 A Figura 1a mostra uma configuração de um aparelho para determinar um sinal de áudio espacial convertido;

A Figura 1b mostra pressão e componentes de um vetor de velocidade de partícula em um plano Gaussiano para uma onda plana;

10 A Figura 2 mostra outra configuração para converter uma corrente mono DirAC de um sinal de formato B;

A Figura 3 mostra uma configuração para combinar múltiplos sinais de áudio espacial convertidos;

15 As Figuras 4a-4d mostram configurações para combinar múltiplos sinais de áudio espacial baseados em DirAC aplicando diferentes efeitos de áudio;

A Figura 5 ilustra uma configuração de um gerador de efeito de áudio;

20 A Figura 6 mostra uma configuração de um gerador de efeito de áudio aplicando efeitos de áudio múltiplos em componentes direcionais; e

A Figura 7 mostra um estado da técnica do codificador DirAC.

25 A Figura 1a mostra um aparelho 100 para determinar um sinal de áudio espacial convertido, o sinal de áudio espacial convertido tendo um componente omnidirecional e pelo menos um componente direcional (X; Y; Z), de um sinal de áudio espacial de entrada, o sinal de áudio espacial de entrada tendo

uma representação de áudio de entrada (W) e uma direção de chegada de entrada (\emptyset).

O aparelho 100 compreende um estimador 110 para estimar uma representação de onda compreendendo uma medição de campo de onda e uma medição de direção de chegada de onda com base na representação de áudio de entrada (W) e a direção de chegada de entrada (\emptyset). Além do mais, o aparelho 100 compreende um processador 120 para processamento da medição de campo de onda e medição da direção de chegada da onda para obter o componente omnidirecional e o pelo menos um componente direcional. O estimador 110 pode ser adaptado para estimar a representação de onda como uma representação de onda plana.

Nas configurações, o processador pode ser adaptado para prover a representação do áudio de entrada (W) como o componente de áudio omnidirecional (W'). Em outras palavras, o componente de áudio omnidirecional W' pode ser igual à representação de áudio de entrada W . Portanto, de acordo com as linhas pontilhadas na Figura 1a, a representação de áudio de entrada pode desviar do estimador 110, do processador 120, ou de ambos. Em outras configurações, o componente de áudio omnidirecional W' pode ser baseado na intensidade da onda e na direção da onda de chegada sendo processada pelo processador 120 com a representação do áudio de entrada W . Nas configurações, componentes de áudio direcional múltiplos ($X;Y;Z$) podem ser processados, como, por exemplo, um primeiro (X), um segundo (Y) e/ou um terceiro (Z) componentes de áudio direcional correspondendo a diferentes direções espaciais. Nas configurações, por exemplo, três componentes de áudio direcional diferentes

(X;Y;Z) podem ser derivados de acordo com as direções diferentes de um sistema de coordenadas Cartesianas.

O estimador 110 pode ser adaptado para estimar a medição de campo de onda em termos de uma amplitude de campo de onda e uma fase de campo de onda. Em outras palavras, nas configurações, a medição de campo de onda pode ser estimada como quantidade valorizada complexa. A amplitude de campo de onda pode corresponder a uma magnitude de pressão de som e a fase de campo de onda pode corresponder a uma fase de pressão de som em algumas configurações.

Nas configurações a medição de direção de chegada de onda pode corresponder a qualquer quantidade direcional, expressa, por exemplo, por um vetor, um ou mais ângulos, etc. e pode ser derivada de qualquer medição direcional representando um componente de áudio, como por exemplo, um vetor de intensidade, um vetor de velocidade de partícula, etc. A medição de campo de onda pode corresponder a qualquer quantidade física descrevendo um componente de áudio, que pode ser de valor real ou complexo, corresponder a um sinal de pressão, uma amplitude ou magnitude de velocidade de partícula, altura de som, etc. Além do mais, medições podem ser consideradas no domínio de tempo e/ou frequência.

Configurações podem ser baseadas na estimativa de uma representação de onda plana para cada uma das correntes de entrada, que pode ser executada pelo estimador 110 na Figura 1a. Em outras palavras, a medição de campo de onda pode ser modelada usando uma representação de onda plana. Em geral, existem várias descrições exaustivas (isto é, completas) equivalentes de uma onda

plana ou ondas em geral. A seguir, uma descrição matemática será introduzida para computar parâmetros de difusão e direções de chegada ou medições de direção para diferentes componentes. Embora apenas umas poucas descrições se relacionem diretamente com 5 quantidades físicas, como, por exemplo, pressão, velocidade de partícula, etc., potencialmente existe um número infinito de diferentes maneiras de descrever representações de onda, das quais uma deve ser apresentada como um exemplo subsequente, entretanto, não significando ser limitativa de qualquer maneira às 10 configurações da presente invenção. Qualquer combinação pode corresponder à medição de campo de onda e à medição de direção de chegada de onda.

De modo a detalhar adicionalmente descrições potenciais, dois números reais a e b são considerados. A 15 informação contida em a e b pode ser transferida pelo envio de c e d , quando

$$\begin{bmatrix} c \\ d \end{bmatrix} = \Omega \begin{bmatrix} a \\ b \end{bmatrix},$$

onde Ω é uma matriz 2x2 conhecida. O exemplo considera apenas combinações lineares, geralmente qualquer 20 combinação, isto é, também uma combinação não linear é concebível.

Os seguintes escalares são representados por letras minúsculas a , b , c , enquanto vetores de coluna são representados por letras minúsculas em negrito \mathbf{a} , \mathbf{b} , \mathbf{c} . O sobrescrito $(\cdot)^T$ denota a transposição, respectivamente, enquanto 25 $(\cdot)^*$ e $(\cdot)^{\dagger}$ denota conjugação complexa. A anotação de fasor complexa é distinguida daquela temporal. Por exemplo, a pressão $p(\mathbf{t})$, que é um número real e da qual uma medição de campo de onda possível

pode ser derivada, pode ser expressa por meio do fasor P , que é um número complexo e do qual outra medição de campo de onda possível pode ser derivada, por

$$p(t) = \text{Re}\{P e^{j\omega t}\},$$

5 onde $\text{Re}\{ \}$ denota a parte real e $\omega = 2\pi f$ é a frequência angular. Além do mais, letras maiúsculas usadas para quantidades físicas representam fasores a seguir. Para a anotação do exemplo introdutório a seguir, e para evitar confusão, por favor observe que todas as quantidades com "PW" subscrito se
10 referem a ondas planas.

Para uma onda plana monocromática ideal o vetor de velocidade de partícula U_{PW} pode ser anotado como

$$\mathbf{U}_{PW} = \frac{P_{PW}}{\rho_0 c} \mathbf{e}_d = \begin{bmatrix} U_x \\ U_y \\ U_z \end{bmatrix},$$

onde o vetor de unidade \mathbf{e}_d aponta para a direção
15 de propagação da onda, por exemplo, correspondendo a uma medição de direção. Pode ser provado que

$$\left. \begin{aligned} I_a &= \frac{1}{2\rho_0 c} |P_{PW}|^2 \mathbf{e}_d \\ E &= \frac{1}{2\rho_0 c^2} |P_{PW}|^2, \quad (a) \\ \Psi &= 0 \end{aligned} \right|$$

onde I_a denota a intensidade ativa, ρ_0 denota a densidade do ar, c denota a velocidade do som, E denota a energia
20 de campo de som e Ψ denota a difusão.

É interessante observar que desde que todos os componentes de \mathbf{e}_d são números reais, os componentes de \mathbf{U}_{PW} estão todos em fase com P_{PW} . A Figura 1b ilustra um exemplar \mathbf{U}_{PW} e P_{PW} no plano gaussiano. Conforme mencionado, todos os componentes de \mathbf{U}_{PW}

partilham a mesma fase que P_{PW} , isto é θ . Suas magnitudes, por outro lado, são ligadas a

$$\frac{|P_{PW}|}{c} = \sqrt{|U_x|^2 + |U_y|^2 + |U_z|^2} = \|U_{PW}\|$$

Configurações da presente invenção podem prover
 5 um método para converter uma corrente mono DirAC em um sinal de formato B. Uma corrente DirAC mono pode ser representada por um sinal de pressão capturado, por exemplo, por um microfone omnidirecional e por informação de lado. A informação de lado pode compreender medições dependentes de tempo-frequência de difusão e
 10 direção de chegada de som.

Nas configurações o sinal de áudio espacial de entrada pode compreender ainda um parâmetro de difusão Ψ e o estimador 110 pode ser adaptado para estimar a medição de campo de onda adicionalmente baseada no parâmetro de difusão Ψ .

15 A direção de entrada de chegada e a direção de onda de medição de chegada pode se referir a um ponto de referência correspondendo a um local de registro do sinal de áudio espacial de entrada, isto é, em outras palavras todas as direções podem se referir ao mesmo ponto de referência. O ponto de
 20 referência pode ser o local onde um microfone está colocado ou onde microfones direcionais múltiplos estão colocados para gravar um campo de som.

Nas configurações, o sinal de áudio espacial convertido pode compreender um primeiro (X), um segundo (Y) e um
 25 terceiro (Z) componente direcional. O processador 120 pode ser adaptado para processamento adicional da medição de campo de onda e da medição de direção de chegada da onda para obter o primeiro

(X) e/ou o segundo (Y) e/ou o terceiro (Z) componentes direcionais e/ou os componentes de áudio omnidirecionais.

A seguir as anotações e um modelo de dados serão apresentados.

5 Deixemos $\mathbf{p}(t)$ e $\mathbf{u}(t)=[u_x(t), u_y(t), u_z(t)]^T$ serem a pressão e vetor de velocidade de partícula, respectivamente, para um ponto específico no espaço, onde $[\cdot]^T$ denota a transposição. $\mathbf{p}(t)$ pode corresponder a uma representação de áudio e $\mathbf{u}(t)=[u_x(t), u_y(t), u_z(t)]^T$ pode corresponder a componentes
 10 direcionais. Estes sinais podem ser transformados em um domínio de tempo-frequência por meio de um banco de filtro apropriado ou uma STFT (STFT = Transformada de Fourier de Curta Duração) conforme sugerido, por exemplo, por V. Pulkki e C. Faller, "Directional audio coding: Filterbank and STFT-based design", na 120ª Convenção
 15 AES, 20-23 de Maio de 2006, Paris, França, Maio de 2006.

Deixemos $\mathbf{P}(k, n)$ e $\mathbf{U}(k, n)=[U_x(k, n), U_y(k, n), U_z(k, n)]^T$ denotarem os sinais transformados, onde k e n são índices para frequência (ou banda de frequência) e tempo, respectivamente. O vetor de intensidade ativa $\mathbf{I}_a(k, n)$ pode
 20 ser definido como

$$\mathbf{I}_a(k, n) = \frac{1}{2} \text{Re}\{\mathbf{P}(k, n) \cdot \mathbf{U}^*(k, n)\} \quad (1)$$

onde $(^*)$ denota conjugação complexa e $\text{Re}\{\}$ extrai a parte real. O vetor de intensidade ativa pode expressar o fluxo líquido de energia caracterizando o campo de som, conforme F.J. Fahy, "Sound Intensity", Essex: Elsevier Science Publishers Ltd.,
 25 1989.

Deixemos c denotar a velocidade de som no meio considerado e E a energia de campo de som definida por F.J. Fahy.

$$E(k, n) = \frac{\rho_0}{4} \|U(k, n)\|^2 + \frac{1}{4\rho_0 c^2} |P(k, n)|^2, \quad (2)$$

onde $\|\cdot\|$ computa o 2-norm. A seguir, o conteúdo de uma corrente DirAC mono será detalhado.

A corrente DirAC mono pode consistir do sinal mono $p(t)$ ou representação de áudio e de informação de lado, por exemplo, uma medição de direção de chegada. Esta informação de lado pode compreender a direção dependente de tempo-frequência de chegada e uma medição dependendo de tempo-frequência de difusão. O anterior pode ser denotado por $e_{DOA}(k, n)$, que é um vetor de unidade apontando para a direção da qual o som chega, isto é, pode estar modelando a direção de chegada. O último, por difusão, pode ser denotado por

15

$$\Psi(k, n).$$

Nas configurações, o estimador 110 e/ou o processador 120 podem ser adaptados para estimar/processar a DOA de entrada e/ou a medição DOA de onda em termos de um vetor de unidade $e_{DOA}(k, n)$. A direção de chegada pode ser obtida como

20

$$e_{DOA}(k, n) = -e_I(k, n),$$

Onde o vetor de unidade $e_I(k, n)$, indica a direção na qual a intensidade ativa aponta, isto é,

$$\begin{aligned} I_a(k, n) &= \|I_a(k, n)\| \cdot e_I(k, n), \\ e_I(k, n) &= I_a(k, n) / \|I_a(k, n)\|, \end{aligned} \quad (3)$$

respectivamente. Alternativamente nas configurações, a DOA ou medição DOA pode ser expressa em termos de ângulos azimute e de elevação em um sistema de coordenadas

25

esférico. Por exemplo, se $\varphi(\mathbf{k}, n)$ e $\vartheta(\mathbf{k}, n)$ são ângulos azimute e de elevação, respectivamente, então

$$\begin{aligned} \mathbf{e}_{DOA}(\mathbf{k}, n) &= [\cos(\varphi(\mathbf{k}, n)) \cdot \cos(\vartheta(\mathbf{k}, n)), \sin(\varphi(\mathbf{k}, n)) \cdot \cos(\vartheta(\mathbf{k}, n)), \sin(\vartheta(\mathbf{k}, n))] \\ &= [e_{DOA,x}(\mathbf{k}, n), e_{DOA,y}(\mathbf{k}, n), e_{DOA,z}(\mathbf{k}, n)] \end{aligned}$$

5 (4)

onde $\mathbf{e}_{DOA,x}(\mathbf{k}, n)$ é um componente do vetor de unidade $\mathbf{e}_{DOA}(\mathbf{k}, n)$ da direção de chegada de entrada ao longo de um eixo x de um sistema de coordenadas Cartesianas, $\mathbf{e}_{DOA,y}(\mathbf{k}, n)$ é um componente de $\mathbf{e}_{DOA}(\mathbf{k}, n)$ ao longo de um eixo y e $\mathbf{e}_{DOA,z}(\mathbf{k}, n)$ é um
10 componente de $\mathbf{e}_{DOA}(\mathbf{k}, n)$ ao longo de um eixo z.

Nas configurações, o estimador 110 pode ser adaptado para estimar a medição de campo de onda adicionalmente baseada no parâmetro de difusão Ψ , opcionalmente também expresso por $\Psi(\mathbf{k}, n)$ de uma maneira dependente de tempo-frequência. O
15 estimador 110 pode ser adaptado para estimar com base no parâmetro de difusão em termos de

$$\Psi(\mathbf{k}, n) = 1 - \frac{\| \langle \mathbf{I}_o(\mathbf{k}, n) \rangle \|}{c \langle E(\mathbf{k}, n) \rangle}, \quad (5)$$

onde $\langle . \rangle$ indica uma média temporal.

Existem diferentes estratégias para obter $\mathbf{P}(\mathbf{k}, n)$
20 e $\mathbf{U}(\mathbf{k}, n)$ na prática. Uma possibilidade é usar um microfone de formato B, que envia 4 sinais, isto é $\mathbf{w}(\mathbf{t})$, $\mathbf{x}(\mathbf{t})$, $\mathbf{y}(\mathbf{t})$ e $\mathbf{z}(\mathbf{t})$. O primeiro, $\mathbf{w}(\mathbf{t})$, pode corresponder à leitura de pressão de um microfone omnidirecional. Os três últimos podem corresponder a leituras de pressão de microfones tendo padrões de captação no
25 formato de oito, direcionados para três eixos de um sistema de

coordenadas cartesianas. Estes sinais são, também, proporcionais à velocidade de partícula. Portanto, em algumas configurações

$$\begin{aligned} P(k,n) &= W(k,n) \\ U(k,n) &= -\frac{1}{\sqrt{2}\rho_0 c} [X(k,n), Y(k,n), Z(k,n)]^T \end{aligned} \quad (6)$$

onde $W(t)$, $X(t)$, $Y(t)$ e $Z(t)$ são os sinais de
 5 formato B transformados correspondendo ao componente omnidirecional $W(k,n)$ e aos três componentes direcionais $X(k,n)$, $Y(k,n)$ e $Z(k,n)$. Observe que o fator $\sqrt{2}$ em (6) se origina da convenção usada na definição dos sinais de formato B, conforme Michael Gerzon, "Surround sound psychoacoustics", no *Wireless*
 10 *World*, volume 80, páginas 483-486, Dezembro de 1974.

Alternativamente, $P(k,n)$ e $U(k,n)$ podem ser estimados por meio de um arranjo de microfone omnidirecional conforme sugerido por J. Merimaa, "Applications of a 3-D microphone array", na 112ª Convenção AES, Documento 5501, Munique,
 15 maio de 2002. As etapas de processamento descritas acima são também ilustradas na Figura 7.

A Figura 7 mostra um codificador DirAC 200, que é adaptado para computar um canal de áudio mono e informação de lado de sinais de microfone apropriados. Em outras palavras, a Figura 7
 20 ilustra um codificador DirAC 200 para determinar a difusão $\Psi(k,n)$ e direção de chegada $e_{DOA}(k,n)$ de sinais de microfone apropriados. A Figura 7 mostra um codificador DirAC 200 compreendendo uma unidade de estimativa P/U 210. A unidade de estimativa P/U recebe os sinais de microfone como informação de entrada, na qual a
 25 estimativa P/U é baseada. Visto que toda informação está disponível, a estimativa P/U é direta de acordo com as equações acima. Um estágio de análise energética 220 permite a estimativa

da direção de chegada e do parâmetro de difusão da corrente combinada.

Nas configurações, o estimador 110 pode ser adaptado para determinar a medição do campo de onda ou amplitude baseada em uma fração $\beta(\mathbf{k}, n)$ da representação de áudio de entrada $\mathbf{P}(\mathbf{k}, n)$. A Figura 2 mostra as etapas de processamento de uma configuração para computar os sinais de formato B de uma corrente DirAC mono. Todas as quantidades dependem dos índices de tempo e frequência (\mathbf{k}, n) e são parcialmente omitidas a seguir por uma questão de simplicidade.

Em outras palavras, a Figura 2 ilustra outra configuração. De acordo com a Eq. (6), $\mathbf{W}(\mathbf{k}, n)$ é igual à pressão $\mathbf{P}(\mathbf{k}, n)$. Portanto, o problema de sintetização do formato B de uma corrente DirAC mono é reduzido à estimativa do vetor de velocidade de partícula $\mathbf{U}(\mathbf{k}, n)$, visto que seus componentes são proporcionais à $X(\mathbf{k}, n)$, $Y(\mathbf{k}, n)$ e $Z(\mathbf{k}, n)$.

Configurações podem se aproximar da estimativa com base na pressuposição de que o campo consiste de uma onda plana somada a um campo difuso. Portanto, a pressão e velocidade de partícula podem ser expressas como

$$P(\mathbf{k}, n) = P_{pw}(\mathbf{k}, n) + P_{diff}(\mathbf{k}, n) \quad (7)$$

$$\mathbf{U}(\mathbf{k}, n) = \mathbf{U}_{pw}(\mathbf{k}, n) + \mathbf{U}_{diff}(\mathbf{k}, n). \quad (8)$$

onde os subscritos "pw" e "diff" denotam a onda plana e o campo difuso, respectivamente.

Os parâmetros DirAC carregam informação apenas com relação à intensidade ativa. Portanto, o vetor de velocidade de partícula $\mathbf{U}(\mathbf{k}, n)$ é estimado com $\hat{\mathbf{U}}_{pw}(\mathbf{k}, n)$, que é o estimador para

a velocidade de partícula da onda plana apenas. Ele pode ser definido como

$$\hat{U}_{PW}(k, n) = -\frac{1}{\rho_0 c} \beta(k, n) \cdot P(k, n) \cdot e_{DOA}(k, n), \quad (9)$$

onde o número real $\beta(k, n)$ é um fator de ponderação apropriado, que no geral é dependente da frequência e pode exibir uma proporcionalidade inversa à difusão $\Psi(k, n)$. Na verdade, para difusão baixa, isto é, $\Psi(k, n)$ próximo de 0, pode ser pressuposto que o campo é composto de uma onda plana única, de modo que

$$U_{PW}(k, n) \approx -\frac{1}{\rho_0 c} P(k, n) \cdot e_{DOA}(k, n) = \hat{U}_{PW}(k, n) \Big|_{\Psi(k, n) \rightarrow 0}, \quad (10)$$

implicando que $\beta(k, n) = 1$.

Considerando a equação acima e Eq. (6), o componente omnidirecional e/ou o primeiro e/ou segundo e/ou terceiro componentes direcionais podem ser expressos como

$$\begin{aligned} W(k, n) &= P(k, n) \\ X(k, n) &= \sqrt{2} \beta(k, n) \cdot P(k, n) \cdot e_{DOA,x}(k, n) \\ Y(k, n) &= \sqrt{2} \beta(k, n) \cdot P(k, n) \cdot e_{DOA,y}(k, n) \\ Z(k, n) &= \sqrt{2} \beta(k, n) \cdot P(k, n) \cdot e_{DOA,z}(k, n) \end{aligned} \quad (11)$$

onde $e_{DOA,x}(k, n)$ é o componente do vetor de unidade $e_{DOA}(k, n)$ da direção de chegada de entrada junto ao eixo x de um sistema de coordenadas Cartesianas, $e_{DOA,y}(k, n)$ é o componente de $e_{DOA}(k, n)$ junto ao eixo Y e $e_{DOA,z}(k, n)$ é o componente de $e_{DOA}(k, n)$ junto ao eixo z. Na configuração mostrada na Figura 2, a medição de direção de chegada de onda estimada pelo estimador 110 corresponde a $e_{DOA,x}(k, n)$, $e_{DOA,y}(k, n)$ e $e_{DOA,z}(k, n)$ e a medição do campo de onda corresponde a $\beta(k, n) P(k, n)$. O primeiro componente direcional conforme enviado pelo processador 120 pode corresponder a qualquer um dentre $X(k, n)$, $Y(k, n)$ ou $Z(k, n)$ e o segundo

componente direcional, conseqüentemente, a qualquer outro dentre $X(k, n)$, $Y(k, n)$ ou $Z(k, n)$.

A seguir, duas configurações práticas serão apresentadas sobre como determinar o fator $\beta(k, n)$.

5 A primeira configuração objetiva estimar a pressão de uma onda plana primeiramente, isto é $P_{PW}(k, n)$ e, então, dela, derivar o vetor de velocidade de partícula.

Definindo a densidade do ar ρ_0 igual a 1, e derrubando a dependência funcional (k, n) para simplicidade, pode
10 ser escrito

$$\psi = 1 - \frac{\langle |P_{PW}|^2 \rangle_t}{\langle |P_{PW}|^2 \rangle_t + 2c^2 \langle E_{diff} \rangle_t} \quad (12)$$

Providas as propriedades estatísticas de campos difusos, uma aproximação pode ser introduzida por

$$\langle |P_{PW}|^2 \rangle_t + 2c^2 \langle E_{diff} \rangle_t \approx \langle |P|^2 \rangle_t \quad (13)$$

15 onde E_{diff} é a energia do campo difuso. O estimador pode, assim, ser obtido por

$$\langle |P_{PW}| \rangle_t \approx \langle |\bar{P}_{PW}| \rangle_t = \sqrt{1 - \Psi} \langle |P| \rangle_t \quad (14)$$

Para computar as estimativas instantâneas, isto
20 é, para cada telha de tempo-freqüência, os operadores de expectativa podem ser removidos, obtendo

$$\bar{P}_{PW}(k, n) = \sqrt{1 - \Psi(k, n)} P(k, n) \quad (15)$$

Pela exploração da pressuposição de onda plana, a estimativa para a velocidade de partícula pode ser derivada
25 diretamente

$$\hat{U}_{PW}(k, n) = \frac{1}{\rho_0 c} \hat{P}_{PW}(k, n) \cdot e_t(k, n), \quad (16)$$

da qual ocorre que

$$\beta(k, n) = \sqrt{1 - \Psi(k, n)}. \quad (17)$$

Em outras palavras, o estimador 110 pode ser adaptado para estimar a fração $\beta(k, n)$ com base no parâmetro de difusão $\Psi(k, n)$, de acordo com

$$\beta(k, n) = \sqrt{1 - \Psi(k, n)}.$$

e a medição do campo de onda de acordo com

$$\beta(k, n) P(k, n),$$

onde o processador 120 pode ser adaptado para obter a magnitude do primeiro componente direcional $X(k, n)$ e/ou o segundo componente direcional $Y(k, n)$ e/ou o terceiro componente direcional $Z(k, n)$ e/ou o componente de áudio omnidirecional $W(k, n)$ por

$$\begin{aligned} W(k, n) &= P(k, n) \\ X(k, n) &= \sqrt{2} \beta(k, n) \cdot P(k, n) \cdot e_{DOA,x}(k, n) \\ Y(k, n) &= \sqrt{2} \beta(k, n) \cdot P(k, n) \cdot e_{DOA,y}(k, n) \\ Z(k, n) &= \sqrt{2} \beta(k, n) \cdot P(k, n) \cdot e_{DOA,z}(k, n) \end{aligned}$$

onde a direção de onda da medição de chegada é representada pelo vetor de unidade $[e_{DOA,x}(k, n), e_{DOA,y}(k, n), e_{DOA,z}(k, n)]^T$, onde x, y e z indicam as direções de um sistema de coordenadas Cartesianas.

Uma solução alternativa nas configurações pode ser derivada pela obtenção do fator $\beta(k, n)$ diretamente da expressão da difusão $\Psi(k, n)$. Conforme já foi mencionado, a velocidade de partícula $U(k, n)$ pode ser modelada como

$$U(k, n) = \beta(k, n) \cdot \frac{P(k, n)}{\rho_0 c} \cdot e_i(k, n) \quad (18)$$

A Equação (18) pode ser substituída em (5) conduzindo a

$$\Psi(k, n) = 1 - \frac{\frac{1}{\rho_0 c} \left\| \langle |\beta(k, n) \cdot P(k, n)|^2 \cdot e_i(k, n) \rangle \right\|}{c \left\langle \frac{1}{2\rho_0 c^2} |P(k, n)|^2 \cdot (\beta^2(k, n) + 1) \right\rangle} \quad (19)$$

5 Para obter valores instantâneos os operadores de expectativa podem ser removidos e resolvendo para $\beta(k, n)$ resultar

$$\beta(k, n) = \frac{1 - \sqrt{1 - (1 - \Psi(k, n))^2}}{1 - \Psi(k, n)} \quad (20)$$

Em outras palavras, nas configurações o estimador 110 pode ser adaptado para estimar a fração $\beta(k, n)$ com base em 10 $\Psi(k, n)$ de acordo com

$$\beta(k, n) = \frac{1 - \sqrt{1 - (1 - \Psi(k, n))^2}}{1 - \Psi(k, n)}$$

Nas configurações, o sinal de áudio espacial de entrada pode corresponder a um sinal DirAC mono. Configurações podem ser estendidas para processamento de outras correntes. No 15 caso onde a corrente ou o sinal de áudio espacial de entrada não carrega um canal omnidirecional, configurações podem combinar os canais disponíveis para se aproximarem de um padrão de captação omnidirecional. Por exemplo, no caso de uma corrente DirAC estéreo como sinal de áudio espacial de entrada, o sinal de pressão P na 20 Figura 2 pode ser aproximado pela soma dos canais L e R .

A seguir, uma configuração com $\Psi=1$ será iluminada. A Figura 2 ilustra que se a difusão for igual a um para ambas as configurações, o som é direcionado exclusivamente para o canal W visto que β é igual a zero, de modo que os sinais X , Y e

\mathbf{z} , isto é, os componentes direcionais, são também zero. Se $\Psi=1$ constantemente no tempo, o canal de áudio mono pode, dessa maneira, ser direcionado para o canal \mathbf{W} sem quaisquer computações adicionais. A interpretação física disto é que o sinal de áudio é
5 apresentado ao ouvinte como sendo um campo reativo puro, visto que o vetor de velocidade de partícula tem magnitude zero.

Outro caso quando $\Psi=1$ ocorre considerando uma situação onde um sinal de áudio está presente apenas em um ou qualquer subconjunto de sinais dipolos, e não no sinal \mathbf{W} . Na
10 análise de difusão DirAC este cenário é analisado para ter $\Psi=1$ com Eq. 5, visto que o vetor de intensidade tem constantemente o comprimento de zero visto que a pressão P é zero na Eq. (1). A interpretação física disto é também que o sinal de áudio é apresentado ao ouvinte sendo reativo, visto que nesta vez o sinal
15 de pressão é constantemente zero, enquanto o vetor de velocidade de partícula é diferente de zero.

Devido ao fato de que o formato B é inerentemente uma representação independente de definição de alto-falante, configurações podem usar o formato B como um idioma comum falado
20 por diferentes dispositivos de áudio, significando que a conversão de um para outro pode ser possível pelas configurações por meio de uma conversão intermediária no formato B. Por exemplo, configurações podem unir correntes DirAC de diferentes ambientes acústicos gravados com diferentes ambientes de com sintetizados no
25 formato B. A união de correntes DirAC mono a correntes de formato B pode também ser permitida pelas configurações.

Configurações podem permitir a união de sinais de áudio de canais múltiplos em qualquer formato 'surround' com uma

corrente DirAC mono. Além do mais, configurações podem permitir a união de uma corrente DirAC mono com qualquer corrente de formato B. Além do mais, configurações podem permitir a união de uma corrente DirAC mono com uma corrente de formato B.

5 Estas configurações podem prover uma vantagem, por exemplo, na criação de reverberação ou introduzindo efeitos de áudio, como será detalhado subsequente. Na produção de música, reverberadores podem ser usados como dispositivos de efeito que colocam perceptualmente o áudio processado em um espaço
10 virtual. Na realidade virtual, síntese de reverberação pode ser necessária quando fontes virtuais são auralizadas dentro de um espaço fechado, por exemplo, em salas ou halls de concerto.

Quando um sinal para reverberação está disponível, esta auralização pode ser executada pelas
-15- configurações por meio de aplicação de som-seco e som reverberado a diferentes correntes DirAC. Configurações podem usar diferentes abordagens sobre como processar o sinal reverberado no contexto de DirAc, onde configurações podem produzir o som reverberado sendo extremamente difuso ao redor do ouvinte.

20 A Figura 3 ilustra uma configuração de um aparelho 300 para determinar um sinal de áudio espacial convertido combinado, o sinal de áudio espacial convertido combinado tendo pelo menos um primeiro componente combinado e um segundo componente combinado, onde o sinal de áudio espacial convertido
25 combinado é determinado de um primeiro e um segundo sinal de áudio espacial de entrada tendo uma primeira e uma segunda representação de áudio de entrada e uma primeira e uma segunda direção de chegada.

O aparelho 300 compreende uma primeira configuração do aparelho 101 para determinar um sinal de áudio espacial convertido de acordo com a descrição acima, para prover um primeiro sinal convertido tendo um primeiro componente 5 omnidirecional e pelo menos um componente direcional do primeiro aparelho 101. Além do mais, o aparelho 300 compreende outra configuração de um aparelho 102 para determinar um sinal de áudio espacial convertido de acordo com a descrição acima para prover um segundo sinal convertido, tendo um segundo componente 10 omnidirecional e pelo menos um componente direcional do segundo aparelho 102.

De forma geral, configurações não estão limitadas à compreensão de apenas dois dos aparelhos 100, em geral, uma pluralidade dos aparelhos descritos acima pode estar compreendida 15 no aparelho 300, por exemplo, o aparelho 300 pode ser adaptado para combinar uma pluralidade de sinais DirAC.

De acordo com a Figura 3, o aparelho 300 compreende ainda um gerador de efeito de áudio 301 para transmitir o primeiro componente omnidirecional ou o primeiro componente de 20 áudio direcional do primeiro aparelho 101 para obter um primeiro componente transmitido.

Além do mais, o aparelho 300 compreende um primeiro combinador 311 para combinar o primeiro componente transmitido com o primeiro e segundo componentes omnidirecionais, 25 ou para combinar o primeiro componente transmitido com os componentes direcionais do primeiro aparelho 101 e do segundo aparelho 102 para obter o primeiro componente combinado. O aparelho 300 compreende ainda um segundo combinador 312 para

combinar o primeiro e segundo componentes omnidirecionais ou os componentes direcionais do primeiro ou segundo aparelho 101 e 102 para obter o segundo componente combinado.

Em outras palavras, o gerador de efeito de áudio 301 pode transmitir o primeiro componente omnidirecional, de modo que o primeiro combinador 311 pode, então, combinar o primeiro componente omnidirecional transmitido, o primeiro componente omnidirecional e o segundo componente omnidirecional para obter o primeiro componente combinado. O primeiro componente combinado pode, então, corresponder, por exemplo, a um componente omnidirecional combinado. Nesta configuração, o segundo combinador 312 pode combinar o componente direcional do primeiro aparelho 101 e o componente direcional do segundo aparelho para obter o segundo componente combinado, por exemplo, correspondendo a um primeiro componente direcional combinado.

~~Em outras configurações, o gerador de efeito de~~ áudio 301 pode transmitir os componentes direcionais. Nestas configurações, o combinador 311 pode combinar o componente direcional do primeiro aparelho 101, o componente direcional do segundo aparelho 102 e o primeiro componente transmitido para obter o primeiro componente combinado, neste caso correspondendo a um componente direcional combinado. Nesta configuração, o segundo combinador 312 pode combinar o primeiro e segundo componentes omnidirecionais do primeiro aparelho 101 e do segundo aparelho 102 para obter o segundo componente combinado, isto é, um componente omnidirecional combinado.

De acordo com as configurações descritas acima, cada um dos aparelhos pode produzir múltiplos componentes

direcionais, por exemplo, um componente X, Y e Z. Nas configurações, geradores de efeito de áudio múltiplos podem ser usados, o que é indicado na Figura 3 pelas caixas tracejadas 302, 303 e 304. Estes geradores de efeito de áudio opcionais podem 5 gerar componentes transmitidos correspondentes, baseados nos sinais de entrada omnidirecionais e direcionais. Em uma configuração, um gerador de efeito de áudio pode transmitir um componente direcional com base em um componente omnidirecional. Além do mais, o aparelho 300 pode compreender combinadores 10 múltiplos, isto é, combinadores 311, 312, 313 e 314, de modo a combinar um componente combinado omnidirecional e múltiplos componentes direcionais combinados, por exemplo, para as três dimensões espaciais.

Uma das vantagens da estrutura do aparelho 300 é 15 que um máximo de quatro geradores de efeito de áudio é necessário para transmitir, de forma geral, um número ilimitado de fontes de áudio.

Conforme indicado pelos combinadores tracejados 331, 332, 333 e 334 na Figura 3, um gerador de efeito de áudio 20 pode ser adaptado para transmitir uma combinação de componentes direcionais e omnidirecionais dos aparelhos 101 e 102. Em uma configuração o gerador de efeito de áudio 301 pode ser adaptado para transmitir uma combinação dos componentes omnidirecionais do primeiro aparelho 101 e do segundo aparelho 102, ou para 25 transmitir uma combinação dos componentes direcionais do primeiro aparelho 101 e do segundo aparelho 102 para obter o primeiro componente transmitido. Conforme indicado pelos caminhos

tracejados na Figura 3, combinações de múltiplos componentes podem ser providas para os diferentes geradores de efeito de áudio.

Em uma configuração, todos os componentes omnidirecionais de todas as fontes de som, representados na Figura 3 pelo primeiro aparelho 101 e pelo segundo aparelho 102, podem ser combinados de modo a gerar componentes transmitidos múltiplos. Em cada um dos quatro caminhos mostrados na Figura 3, cada gerador de efeito de áudio pode gerar um componente transmitido a ser adicionado aos componentes direcionais ou omnidirecionais correspondentes das fontes de som.

Além do mais, conforme mostrado na Figura 3, estágios múltiplos de atraso e escalonamento 321 e 322 podem ser usados. Em outras palavras, cada aparelho 101 ou 102 pode ter em seu caminho de saída um estágio de atraso e escalonamento 321 ou 322, de modo a retardar um ou mais de seus componentes de saída. Em algumas configurações, os estágios de atraso e escalonamento podem retardar e escalonar os respectivos componentes omnidirecionais, apenas. De forma geral, estágios de atraso e escalonamento podem ser usados para componentes omnidirecionais e direcionais.

Nas configurações, o aparelho 300 pode compreender uma pluralidade de aparelhos 100 representando fontes de áudio e correspondentemente, uma pluralidade de geradores de efeito de áudio, onde o número de geradores de efeito de áudio é menor que o número de aparelhos correspondendo às fontes de som. Conforme já foi mencionado acima, em uma configuração pode existir até quatro geradores de efeito de áudio, com um número basicamente

ilimitado de fontes de som. Nas configurações, um gerador de efeito de áudio pode corresponder a um reverberador.

A Figura 4a mostra outra configuração de um aparelho 300 em maiores detalhes. A Figura 4a mostra dois aparelhos 101 e 102, cada um enviando um componente de áudio omnidirecional W, e três componentes direcionais X, Y, Z. De acordo com a configuração mostrada na Figura 4a, os componentes omnidirecionais de cada um dos aparelhos 101 e 102 são providos para dois estágios de atraso e escalonamento 321 e 322, que enviam três componentes retardados e escalonados, que são, então, adicionados por combinadores 331, 332, 333 e 334. Cada um dos sinais combinados é, então, transmitido separadamente por um dos quatro geradores de efeito de áudio 301, 302, 303 e 304, que são implementados como reverberadores na Figura 4a. Conforme indicado na Figura 4a, cada um dos geradores de efeito de áudio envia um componente correspondendo a um componente omnidirecional e três componentes direcionais no total. Os combinadores 311, 312, 313 e 314 são, então, usados para combinar os respectivos componentes transmitidos com os componentes originais enviados pelos aparelhos 101 e 102, onde na Figura 4a geralmente pode existir uma multiplicidade de aparelhos 100.

Em outras palavras, no combinador 311, uma versão transmitida dos sinais de saída omnidirecionais combinados de todos os aparelhos, pode ser combinada com os componentes de saída omnidirecionais originais ou não transmitidos. Combinações similares podem ser executadas pelos outros combinadores com relação aos componentes direcionais. Na configuração mostrada na Figura 4a, componentes direcionais transmitidos são criados com

base em versões retardadas e escalonadas dos componentes omnidirecionais.

De forma geral, configurações podem aplicar um efeito de áudio como, por exemplo, uma reverberação eficientemente a uma ou mais correntes DirAC. Por exemplo, pelo menos duas correntes DirAC são inseridas à configuração do aparelho 300, conforme mostrado na Figura 4a. Nas configurações, estas correntes podem ser correntes DirAC reais ou correntes sintetizadas, por exemplo, tomando um sinal mono e adicionando informação de lado como uma direção e difusão. De acordo com a discussão acima, os aparelhos 101, 102 podem gerar até quatro sinais para cada corrente, isto é, W, X, Y e Z. De forma geral, configurações dos aparelhos 101 ou 102 podem prover menos que três componentes direcionais, por exemplo, apenas X, ou X e Y, ou qualquer outra combinação dos mesmos.

Em algumas configurações, os componentes omnidirecionais W podem ser providos para geradores de efeito de áudio, como, por exemplo, reverberadores de modo a criar os componentes transmitidos. Em algumas configurações, para cada uma das correntes DirAC de entrada, os sinais podem ser copiados para as quatro ramificações mostradas na Figura 4a, que podem ser independentemente retardados, isto é, individualmente por aparelho 101 ou 102, quatro independentemente retardados, por exemplo, pelos atrasos τ_w , τ_x , τ_y , τ_z , e escalonados, por exemplo, por fatores de escalonamento γ_w , γ_x , γ_y , γ_z , versões podem ser combinadas antes de serem providas para um gerador de efeito de áudio.

De acordo com as Figuras 3 e 4a, as ramificações de diferentes correntes, isto é, as saídas dos aparelhos 101 e

102, podem ser combinadas para obter quatro sinais combinados. Os sinais combinados podem, então, serem independentemente transmitidos pelos geradores de áudio, por exemplo, reverberadores monofônicos convencionais. Os sinais transmitidos resultantes podem, então, serem somados aos sinais W, X, Y e Z enviados originalmente dos diferentes aparelhos 101 e 102.

Nas configurações, sinais de formato B gerais podem ser obtidos, os quais podem, então, por exemplo, serem reproduzidos com um decodificador de formato B como é, por exemplo, executado em Ambisonics. Em outras configurações os sinais de formato B podem ser codificados como, por exemplo, com o codificador DirAC conforme mostrado na Figura 7, de modo que a corrente DirAC resultante possa, então, ser transmitida, adicionalmente processada ou decodificada com um decodificador DirAC mono convencional. A etapa de decodificação pode corresponder à computação de sinais de alto-falante para reprodução.

A Figura 4b mostra outra configuração de um aparelho 300. A figura 4b mostra os dois aparelhos 101 e 102 com os quatro componentes de saída correspondentes. Na configuração mostrada na Figura 4b, apenas os componentes W omnidirecionais são usados para serem primeiramente individualmente retardados e escalonados nos estágios de atraso e escalonamento 321 e 322 antes de serem combinados pelo combinador 331. O sinal combinado é, então, provido para o gerador de efeito de áudio 301, que é, novamente, implementado como um reverberador na Figura 4b. A saída transmitida do reverberador 301 é, então, combinada com os componentes omnidirecionais originais dos aparelhos 101 e 102 pelo

combinador 311. Os outros combinadores 312, 313 e 314 são usados para combinar os componentes direcionais X, Y e Z dos aparelhos 101 e 102, de modo a obter componentes direcionais combinados correspondentes.

5 Em uma relação com a configuração ilustrada na Figura 4a, a configuração ilustrada na Figura 4b corresponde com a definição de fatores de escalonamento para as ramificações X, Y e Z como zero. Nesta configuração, apenas um gerador de efeito de áudio ou reverberador 301 é usado.

10 No geral, conforme os aparelhos 101, 102 e aparelhos N potencialmente correspondendo a fontes N de som, os estágios N potencialmente de atraso e escalonamento 321 podem simular as distâncias de fontes de som, um atraso curto pode corresponder à percepção de uma fonte de som virtual mais próxima
-15 do ouvinte. A impressão espacial de um ambiente "surround" pode, então, ser criada pelos geradores ou reverberadores de efeito de áudio correspondentes.

 Configurações, conforme ilustrado nas Figuras 3, 4a e 4b podem ser utilizadas para casos quando decodificação DirAC
20 mono é usada para fontes N de som que são, então, conjuntamente reverberadas. Visto que pode ser pressuposto que a saída de um reverberador tem uma saída que é totalmente difusa, isto é, ela também pode ser interpretada como um sinal omnidirecional W. Este sinal pode ser combinado com outros sinais de formato B
25 sintetizados, tais como os sinais de formato B originados de fontes N de áudio, dessa maneira representando o caminho direto para o ouvinte. Quando o sinal de formato B resultante é

adicionalmente codificado e decodificado por DirAC, o som reverberado pode ser disponibilizado pelas configurações.

Na Figura 4c, outra configuração do aparelho 300 é mostrada. Na configuração mostrada na Figura 4c, com base nos 5 sinais omnidirecionais de saída dos aparelhos 101 e 102, componentes transmitidos reverberados direcionais são criados. Portanto, com base na saída omnidirecional, os estágios de atraso e escalonamento 321 e 322 criam componentes individualmente retardados e escalonados, que são combinados pelos combinadores 10 331, 332 e 333. A cada um dos sinais combinados, reverberadores diferentes 301, 302 e 303 são aplicados, que, de forma geral, correspondem a diferentes geradores de efeito de áudio. De acordo com a descrição acima, os componentes omnidirecionais, direcionais e transmitidos correspondentes são combinados pelos combinadores 15 311, 312, 313 e 314, de modo a prover um componente omnidirecional combinado e componentes direcionais combinados.

Em outras palavras, os sinais W ou sinais omnidirecionais para cada corrente são alimentados a três geradores de efeito de áudio, como, por exemplo, reverberadores, 20 conforme mostrado nas figuras. De forma geral, existem também apenas duas ramificações dependendo do fato de um sinal de som bidimensional ou tridimensional ser gerado. Após os sinais de formato B serem obtidos, as correntes podem ser decodificadas por meio de um decodificador de microfone virtual DirAC. O último é 25 descrito em detalhe por V. Pulkki, "Spatial Sound Reproduction With Directional Audio Coding", *Journal of the Audio Engineering Society*, 55 (6): 503-516.

Com este decodificador, os sinais de alto-falante $D_p(\mathbf{k}, n)$ podem ser obtidos como uma combinação linear dos sinais W , X , Y e Z , por exemplo, de acordo com

$$D_p(k, n) = G(k, n) \left[W(k, n) \sqrt{2} + X(k, n) \cos(\alpha_p) \cos(\beta_p) + Y(k, n) \sin(\alpha_p) \cos(\beta_p) + Z(k, n) \sin(\beta_p) \right]$$

5 onde α_p e β_p são os ângulos azimute e de elevação do p-ésimos alto-falante. O termo $G(\mathbf{k}, n)$ é um ganho de colocação de uma fonte de som em um campo estéreo ("panning") dependente da direção de chegada e da configuração do alto-falante.

Em outras palavras, a configuração mostrada na
 10 Figura 4c pode prover os sinais de áudio para os alto-falantes correspondendo aos sinais de áudio que podem ser obtidos pela colocação de microfones virtuais orientados para a posição dos alto-falantes e tendo fontes de som do tipo de ponto, cuja posição é determinada pelos parâmetros DirAC. Os microfones virtuais podem
 15 ter padrões de captação formatados como cardióides, como dipolos, ou como qualquer padrão direcional de primeira ordem.

Os sons reverberados podem, por exemplo, serem eficientemente usados como X e Y na soma do formato B. Estas configurações podem ser aplicadas a layouts de alto-falantes
 20 horizontais tendo qualquer número de alto-falantes, sem criar uma necessidade de mais reverberadores.

Conforme discutido anteriormente, decodificação DirAC mono tem limitações na qualidade de reverberação, onde nas configurações a qualidade pode ser melhorada com decodificação
 25 DirAC de microfone virtual, que se beneficia também de sinais dipolos em uma corrente de formato B.

A criação apropriada de sinais de formato B para reverberar um sinal de áudio para decodificação de microfone virtual DirAC pode ser executada nas configurações. Um conceito simples e efetivo que pode ser usado pelas configurações é

5 direcionar diferentes canais de áudio para diferentes sinais dipolos, por exemplo, para canais **X** e **Y**. As configurações podem implementar isto por meio de dois reverberadores que produzem canais de áudio mono incoerentes do mesmo canal de entrada, tratando suas saídas como canais de áudio dipolo de formato B **X** e

10 **Y**, respectivamente, conforme mostrado na Figura 4c para os componentes direcionais. Visto que os sinais não são aplicados a **W**, eles serão analisados para serem totalmente difusos na codificação DirAC subsequente. Também, qualidade aumentada para reverberação pode ser obtida na decodificação de microfone virtual

15 DirAC, visto que os canais=dipolos contêm som diferentemente reverberado. Configurações podem, aqui, gerar uma percepção "mais ampla" e mais "envolvente" de reverberação do que com decodificação DirAC mono. Configurações podem, portanto, usar um máximo de dois reverberadores nos layouts de alto-falantes

20 horizontais, e três para layouts de alto-falantes 3-D na reverberação baseada em DirAC descrita.

As configurações podem não estar limitadas a reverberação de sinais, mas podem aplicar quaisquer outros efeitos de áudio que objetivem, por exemplo, uma percepção totalmente

25 difusa de som. Similar às configurações descritas acima, o sinal de formato B reverberado pode ser somado a outros sinais de formato B sintetizados nas configurações, tais como aqueles

originados de fontes N de áudio, dessa maneira representando um caminho direto para o ouvinte.

Ainda outra configuração é mostrada na Figura 4d. A Figura 4d mostra uma configuração similar àquela da Figura 4a, entretanto, nenhum estágio de atraso ou escalonamento 321 ou 322 está presente, isto é, os sinais individuais nas ramificações são apenas reverberados. A configuração ilustrada na Figura 4d pode também ser vista como sendo similar à configuração ilustrada na figura 4a com os atrasos e escalonamentos ou ganhos antes dos reverberadores serem ajustados como 0 e 1, respectivamente, entretanto, nesta configuração os reverberadores 301, 302, 303 e 304 não são pressupostos como sendo arbitrários e independentes. Na configuração ilustrada na Figura 4d, os quatro geradores de efeito de áudio são pressupostamente dependentes entre si tendo uma estrutura específica.

Cada um dos geradores de efeito de áudio ou reverberadores pode ser implementado como um "tapped-delay line" como será detalhado subseqüentemente com a ajuda da Figura 5. Os atrasos e ganhos ou escalas podem ser escolhidos apropriadamente de uma maneira tal que cada um dos "taps" modela um eco distinto cuja direção, atraso e energia podem ser definidos conforme desejado.

Nesta configuração, o i -ésimo eco pode ser caracterizado por um fator de ponderação, por exemplo, com referência a um som DirAC ρ_i , um atraso τ , e uma direção de chegada θ , e ϕ , correspondendo à elevação e azimute, respectivamente.

Os parâmetros dos reverberadores podem ser definidos como segue

$$\tau_w = \tau_x = \tau_y = \tau_z = \tau_i$$

$$\gamma_w = \rho_i, \text{ para o reverberador } W,$$

$$5 \quad \gamma_x = \rho_i \cos(\theta) \cdot \cos(\theta_i), \text{ para o reverberador } X,$$

$$\gamma_y = \rho_i \sin(\theta) \cdot \cos(\theta_i), \text{ para o reverberador } Y,$$

$$\gamma_z = \rho_i \sin(\theta), \text{ para o reverberador } Z.$$

Em algumas configurações os parâmetros físicos de cada eco podem ser extraídos de processos aleatórios ou tomados de
 10 uma resposta de impulso espacial de sala. O último poderia, por exemplo, ser medido ou simulado com uma ferramenta de acompanhamento de raio.

No geral, configurações podem, aqui, prover a vantagem de que o número de geradores de efeito de áudio é
 15 independente do número de fontes.

A Figura 5 ilustra uma configuração usando um esquema conceitual de um efeito de áudio mono, como, por exemplo, usado dentro de um gerador de efeito de áudio, que é estendido dentro do contexto DirAC. Por exemplo, um reverberador pode ser
 20 efetuado de acordo com este esquema. A Figura 5 mostra uma configuração de um reverberador 500. A Figura 5 mostra, em princípio, uma estrutura de filtro FIR (FIR = Resposta de Impulso Finito). Outras configurações podem usar filtros IIR (IIR = Resposta de Impulso Infinito), também. Um sinal de entrada é
 25 retardado pelos estágios de atraso K rotulados como 511 a 51K. As cópias retardadas K , para as quais os atrasos são denotados por τ_1 a τ_k do sinal, são, então, amplificadas pelos amplificadores 521 a

52K com fatores de amplificação γ_1 a γ_k antes de serem somados no estágio de soma 530.

A Figura 6 mostra outra configuração com uma extensão da cadeia de processamento da Figura 5 dentro do contexto DirAC. A saída do bloco de processamento pode ser um sinal de formato B. A Figura 6 mostra uma configuração onde múltiplos estágios de soma 560, 562 e 564 são utilizados resultando nos três sinais de saída W , X e Y . De modo a estabelecer diferentes combinações, as cópias de sinal retardado podem ser escalonadas diferentemente antes de serem adicionadas nos três estágios de adição diferentes 560, 562 e 564. Isto é executado pelos amplificadores adicionais 531 a 53K e 541 a 54K. Em outras palavras, a configuração 600 mostrada na Figura 6 executa reverberação para diferentes componentes de um sinal de formato B baseado em uma corrente DirAC mono. Três cópias reverberadas diferentes do sinal são geradas usando três filtros FIR diferentes sendo estabelecidos através de diferentes coeficientes de filtro ρ_1 a ρ_k e η_1 a η_k .

A configuração a seguir pode se aplicar a um reverberador ou efeito de áudio que pode ser modelado como na Figura 5. Um sinal de entrada opera através de um "tapped-delay line" simples, onde cópias múltiplas são somadas. O i -ésimo de ramificações K é retardado e atenuado, por τ_i e γ_i , respectivamente.

Os fatores γ e τ podem ser obtidos dependendo do efeito de áudio desejado. No caso de um reverberador, estes fatores imitam a resposta de impulso da sala que deve ser

simulada. De qualquer maneira, sua determinação não é iluminada e eles são, dessa maneira, pressupostamente providos.

Uma configuração é ilustrada na Figura 6. O esquema da Figura 5 é estendido, de modo que mais duas camadas são obtidas. Nas configurações, para cada ramificação um ângulo de chegada θ pode ser designado, obtido de um processo estocástico. Por exemplo, θ pode ser a realização de uma distribuição uniforme na faixa $[-\pi, \pi]$. A i -ésima ramificação é multiplicada com os fatores η_i e ρ_i , que podem ser definidos como

$$10 \quad \eta_i = \text{sen}(\theta_i) \quad (21)$$

$$\rho_i = \text{cos}(\theta_i) \quad (22)$$

Nas configurações, o i -ésimo eco pode ser percebido como vindo de θ_i . A extensão para 3D é direta. Neste caso, mais uma camada necessita ser adicionada, e um ângulo de elevação necessita ser considerado. Após o sinal de formato B ter sido gerado, isto é, W, X, Y , e possivelmente Z , combinações com outros sinais de formato B podem ser executadas. Então, ele pode ser enviado diretamente para um decodificador DirAC de microfone virtual, ou após codificação DirAC a corrente DirAC mono pode ser enviada para um decodificador DirAC mono.

Configurações podem compreender um método para determinar um sinal de áudio espacial convertido, o sinal de áudio espacial convertido tendo um primeiro componente de áudio direcional e um segundo componentes de áudio direcional, de um sinal de áudio espacial de entrada, o sinal de áudio espacial de entrada tendo uma representação de áudio de entrada e uma direção de entrada de chegada. O método compreende uma etapa de estimar uma representação de onda compreendendo uma medição de campo de

onda e uma medição de direção de chegada de onda baseada na representação de áudio de entrada e na direção de chegada de entrada. Além do mais, o método compreende uma etapa de processamento da medição de campo de onda e medição de direção de chegada de onda para obter o primeiro componente direcional e o segundo componente direcional.

Nas configurações, um método para determinar um sinal de áudio espacial convertido pode ser compreendido de uma etapa de obtenção de uma corrente DirAC mono que é para ser convertida no formato B. Opcionalmente, W pode ser obtido de P , quando disponível. Se não, uma etapa de aproximação de W como uma combinação linear dos sinais de áudio disponíveis pode ser executada. Subseqüentemente, uma etapa de computação do fator β como um fator de ponderação dependente de frequência e tempo inversamente proporcional à difusão pode ser executada, por exemplo, de acordo com

$$\beta(k,n) = \sqrt{1-\Psi(k,n)} \quad \text{or} \quad \beta(k,n) = \frac{1 - \sqrt{1 - (1-\Psi(k,n))^2}}{1-\Psi(k,n)}$$

O método pode compreender ainda uma etapa de computação dos sinais X , Y e Z de P , β e e_{DOA} .

Para casos nos quais $\Psi=1$, a etapa de obtenção de W de P pode ser substituída pela obtenção de W de P com X , Y , e Z sendo zero, obtendo pelo menos um sinal dipolo X , Y , ou Z de P ; W sendo zero, respectivamente. Configurações da presente invenção podem executar processamento de sinal no domínio do formato B, resultando na vantagem de que processamento de sinal avançado pode ser executado antes de os sinais de alto-falante serem gerados.

Dependendo de certos requisitos de implementação dos métodos da invenção, os métodos inventivos podem ser implementados em hardware e software. A implementação pode ser executada usando um meio de armazenagem digital, e especificamente 5 uma memória "flash", um disco, um DVD ou um CD tendo sinais de controle eletronicamente legíveis armazenados nos mesmos, que cooperam com um sistema de computador programável, de modo que os métodos inventivos sejam executados. De forma geral, a presente invenção é, portanto, um código de programa de computador com um 10 código de programa armazenado em um veículo legível por máquina, o código de programa sendo operacional para executar os métodos inventivos quando o programa de computador opera em um computador ou processador. Em outras palavras, os métodos inventivos são, portanto, um programa de computador tendo um código de programa 15 para executar pelo menos um dos métodos inventivos, quando o programa de computador opera em um computador.

Reivindicações

1. Aparelho (100) para determinar um sinal de áudio espacial convertido, o sinal de áudio espacial convertido tendo um componente de áudio omnidirecional (W') e pelo menos um
5 componente de áudio direcional ($X; Y; Z$), de um sinal de áudio espacial de entrada, o sinal de áudio espacial de entrada tendo uma representação de áudio de entrada (W) e uma direção de chegada de entrada (\emptyset), compreendendo

um estimador (110) para estimar uma representação
10 de onda, a representação de onda compreendendo uma medição de campo de onda e uma medição de direção de chegada de onda, com base na representação de áudio de entrada (W) e na direção de chegada de entrada (\emptyset); e

um processador (120) para processar a medição de
15 campo de onda e a medição de direção de chegada de onda para obter o componente omnidirecional (W') e pelo menos um componente direcional ($X; Y; Z$).

2. Aparelho (100), de acordo com a reivindicação 1, caracterizado pelo fato de que o estimador (110) é adaptado
20 para estimar a medição do campo de onda em termos de uma amplitude de campo de onda e de fase de campo de onda.

3. Aparelho (100), de acordo com uma das reivindicações 1 ou 2, caracterizado pelo fato de que o sinal de áudio espacial de entrada compreende ainda um parâmetro de difusão
25 Ψ e onde o estimador (110) é adaptado para estimar a medição do campo de onda baseada adicionalmente no parâmetro de difusão Ψ .

4. Aparelho (100), de acordo com uma das reivindicações de 1 a 3, caracterizado pelo fato de que a direção

x de um sistema de coordenadas Cartesianas, $e_{DOA,y}(k,n)$ é um componente de $e_{DOA}(k,n)$ ao longo do eixo Y e $e_{DOA,z}(k,n)$ é um componente de $e_{DOA}(k,n)$ ao longo do eixo z.

8. Aparelho (100), de acordo com uma das reivindicações 6 ou 7, caracterizado pelo fato de que o estimador (110) é adaptado para estimar uma fração $\beta(k,n)$ baseada no parâmetro de difusão $\Psi(k,n)$, de acordo com

$$\beta(k,n) = \sqrt{1 - \Psi(k,n)}$$

9. Aparelho (100), de acordo com uma das reivindicações 6 ou 7, caracterizado pelo fato de que o estimador (110) é adaptado para estimar uma fração $\beta(k,n)$ baseada em $\Psi(k,n)$, de acordo com

$$\beta(k,n) = \frac{1 - \sqrt{1 - (1 - \Psi(k,n))^2}}{1 - \Psi(k,n)}$$

10. Aparelho (100), de acordo com uma das reivindicações 1 a 9, caracterizado pelo fato de que o sinal de áudio espacial de entrada corresponde a um sinal de áudio codificado por DirAC e o sinal de áudio espacial convertido corresponde a um sinal de formato B.

11. Aparelho (300) para determinar um sinal de áudio espacial convertido combinado, o sinal de áudio espacial convertido combinado tendo pelo menos um primeiro componente combinado e um segundo componente combinado, de um primeiro e um segundo sinal de áudio espacial de entrada, o primeiro sinal de áudio espacial de entrada tendo uma primeira representação de áudio de entrada e uma primeira direção de chegada, o segundo sinal de entrada espacial tendo uma segunda representação de áudio de entrada e uma segunda direção de chegada, compreendendo:

um primeiro aparelho (101) de acordo com uma das reivindicações 1 a 10, para prover um primeiro sinal convertido, tendo um primeiro componente omnidirecional e pelo menos um componente direcional do primeiro aparelho (101);

5 um segundo aparelho (102) de acordo com uma das reivindicações 1 a 10, para prover um segundo sinal convertido, tendo um segundo componente omnidirecional e pelo menos um componente direcional do segundo aparelho (102);

10 um gerador de efeito de áudio (301) para transmitir o primeiro componente omnidirecional ou o componente direcional do primeiro aparelho (101) para obter um primeiro componente transmitido;

um primeiro combinador (311) para combinar o primeiro componente transmitido, o primeiro componente omnidirecional e o segundo componente omnidirecional, ou para combinar o primeiro componente transmitido, o componente direcional do primeiro aparelho (101), e o componente direcional do segundo aparelho (102) para obter o primeiro componente combinado; e

20 um segundo combinador (312) para combinar o componente direcional do primeiro aparelho (101) e o componente direcional do segundo aparelho (102), ou para combinar o primeiro componente omnidirecional e o segundo componente omnidirecional para obter o segundo componente combinado.

25 12. Aparelho (300), de acordo com a reivindicação 11, caracterizado pelo fato de que o gerador de efeito de áudio (301) é adaptado para transmitir uma combinação do primeiro componente omnidirecional e do segundo componente omnidirecional,

ou para transmitir uma combinação do componente direcional do primeiro aparelho (101) e o componente direcional do segundo aparelho (102) para obter o primeiro componente transmitido.

13. Aparelho (300), de acordo com qualquer uma das reivindicações 11 ou 12, compreendendo ainda um primeiro estágio de atraso e escalonamento (321) para retardar e/ou escalonar o primeiro componente omnidirecional e/ou o componente direcional do primeiro aparelho (101), e/ou o segundo estágio de atraso e escalonamento (322) para retardar e/ou escalonar o segundo componente omnidirecional e/ou componente direcional do segundo aparelho (102).

14. Aparelho (300), de acordo com uma das reivindicações 11 a 13, compreendendo uma pluralidade de aparelhos (100) de acordo com uma das reivindicações 1 a 10 para converter uma pluralidade de sinais de áudio espaciais de entrada, o aparelho (300) compreendendo ainda uma pluralidade de geradores de efeito de áudio, caracterizado pelo fato de que o número de geradores de efeito de áudio é menor que o número de aparelhos (100) de acordo com uma das reivindicações 1 a 10.

15. Aparelho (300), de acordo com uma das reivindicações 11 a 14, caracterizado pelo fato de que o gerador de efeito de áudio (301) é adaptado para reverberar o primeiro componente omnidirecional ou o componente direcional do primeiro aparelho (101) para obter o primeiro componente transmitido.

16. Método para determinar um sinal de áudio espacial convertido, o sinal de áudio espacial convertido tendo um componente de áudio omnidirecional (W') e pelo menos um componente de áudio direcional ($X;Y;Z$), de um sinal de áudio espacial de

entrada, o sinal de áudio espacial de entrada tendo uma representação de áudio de entrada (W) e uma direção de entrada de chegada (\emptyset), compreendendo as etapas de:

5 estimar uma representação de onda compreendendo
uma medição de campo de onda e uma medição de direção de chegada
de onda baseada na representação de áudio de entrada (W) e na
10 direção de chegada de entrada (\emptyset); e

 processamento da medição de campo de onda e da
medição de direção de chegada de onda para obter o componente
10 omnidirecional e pelo menos um componente direcional (X;Y;Z).

17. Programa de computador tendo um código de programa para executar o método da reivindicação 16, quando o código de programa opera em um processador de computador.

02 Feb. 2009

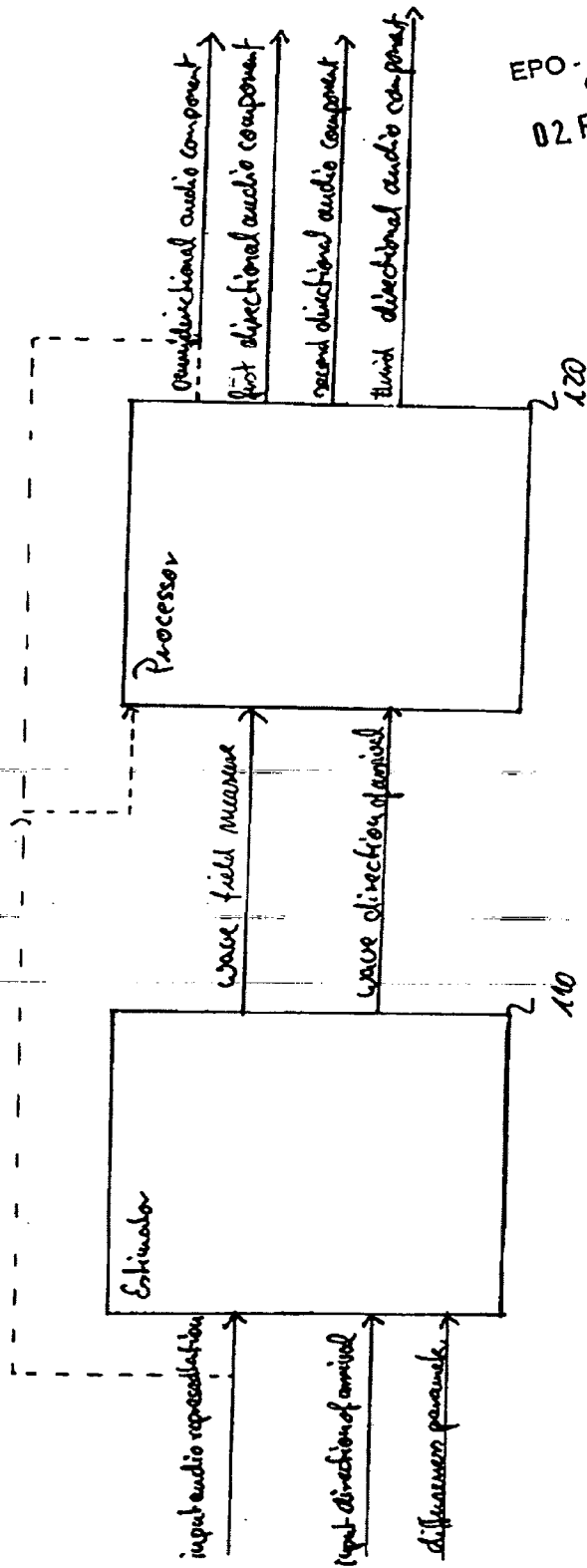


Fig. 1a

100

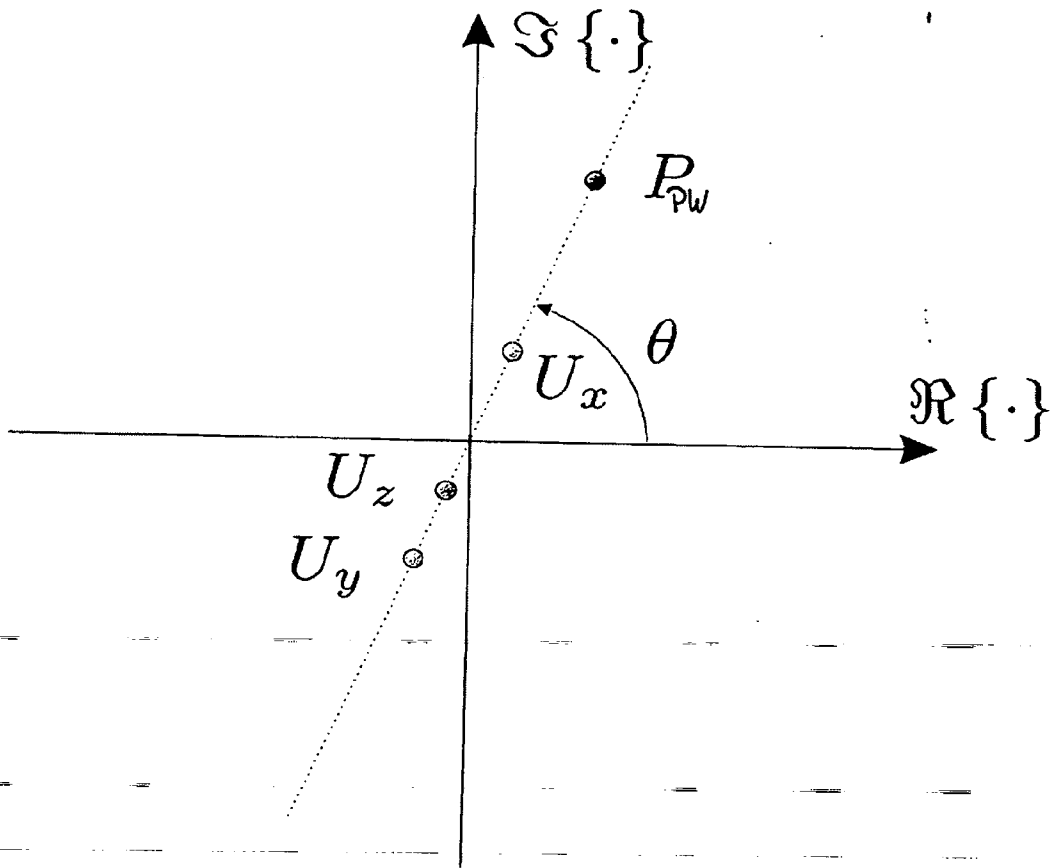


Fig. 16

Dir AC \longrightarrow B-Format

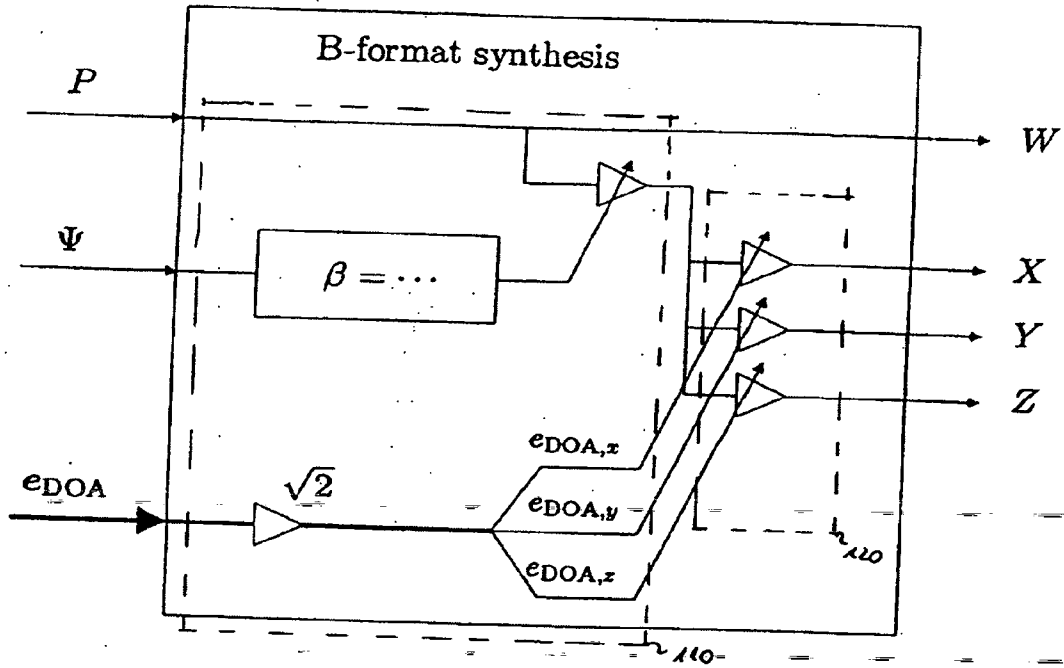


Fig. 2

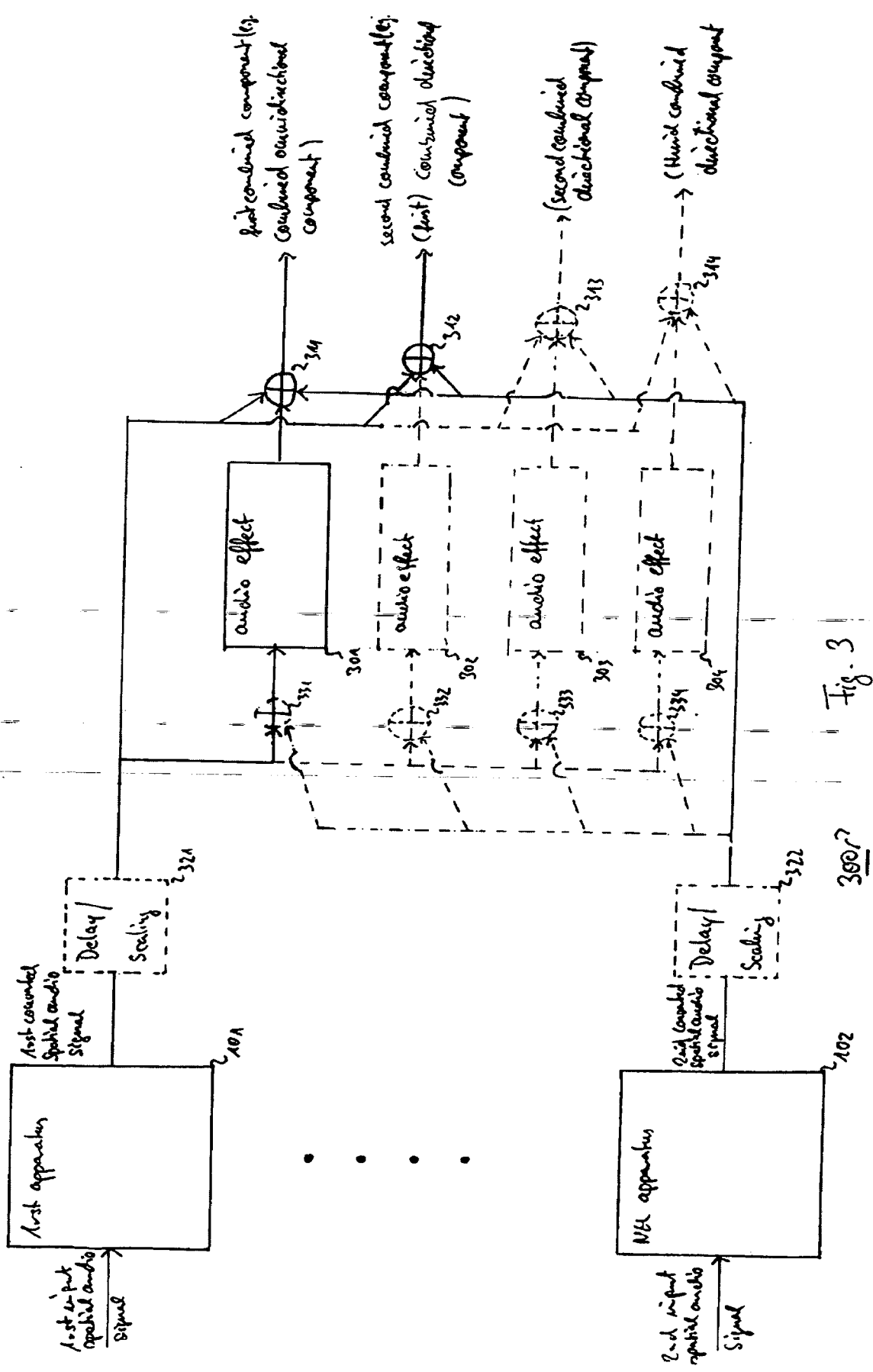
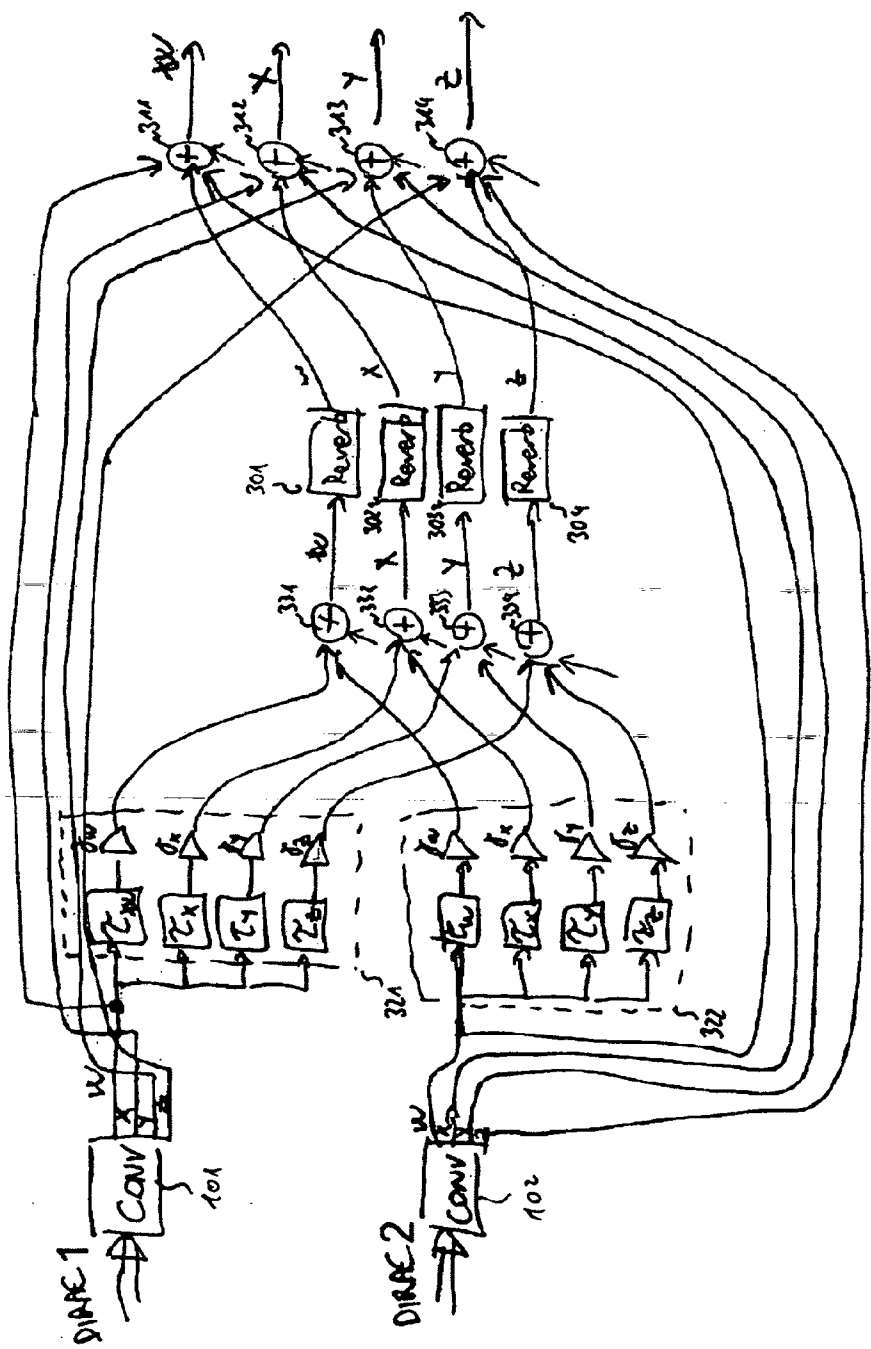


Fig. 3

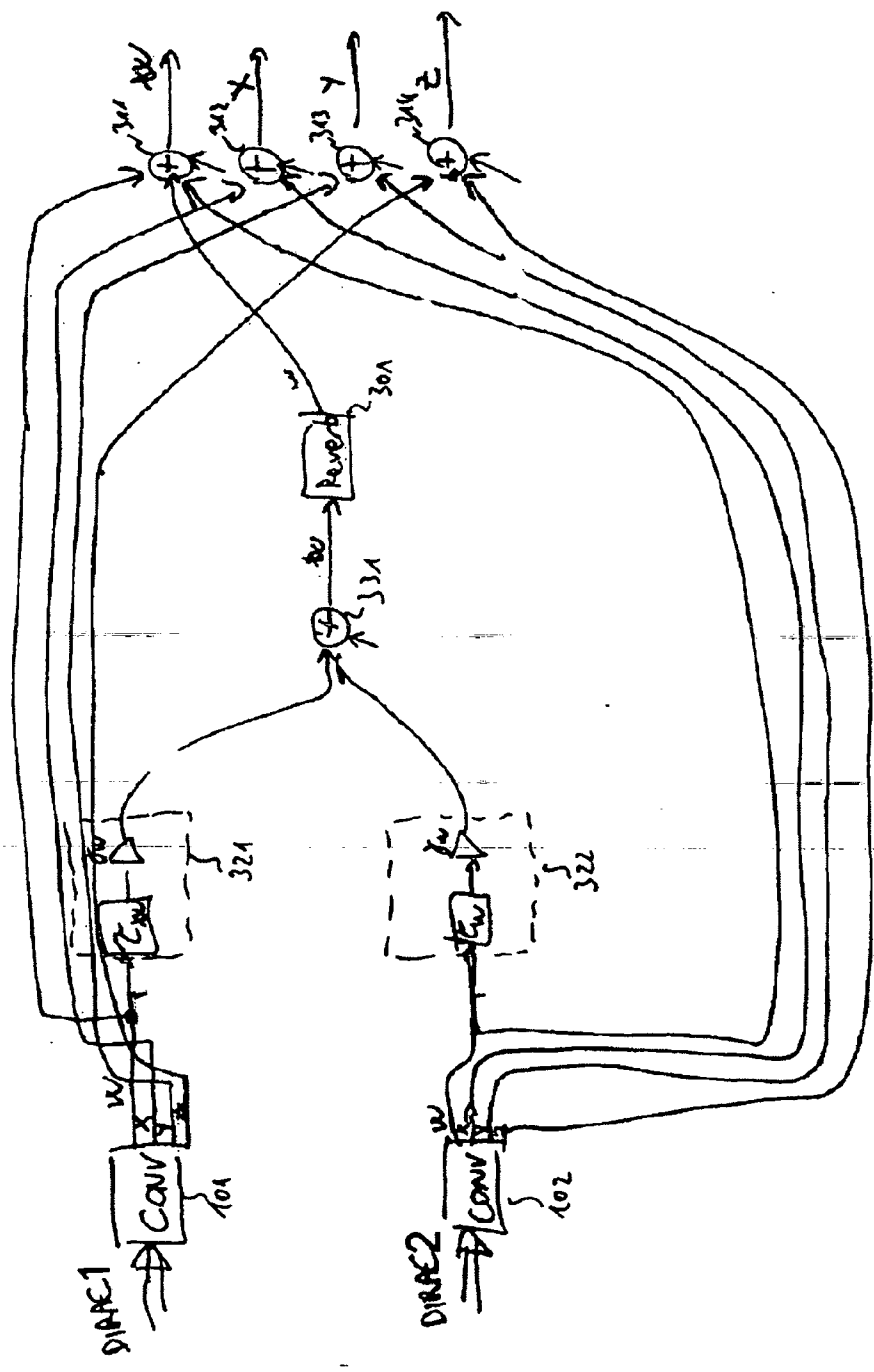
300?



300?

DIRAC N

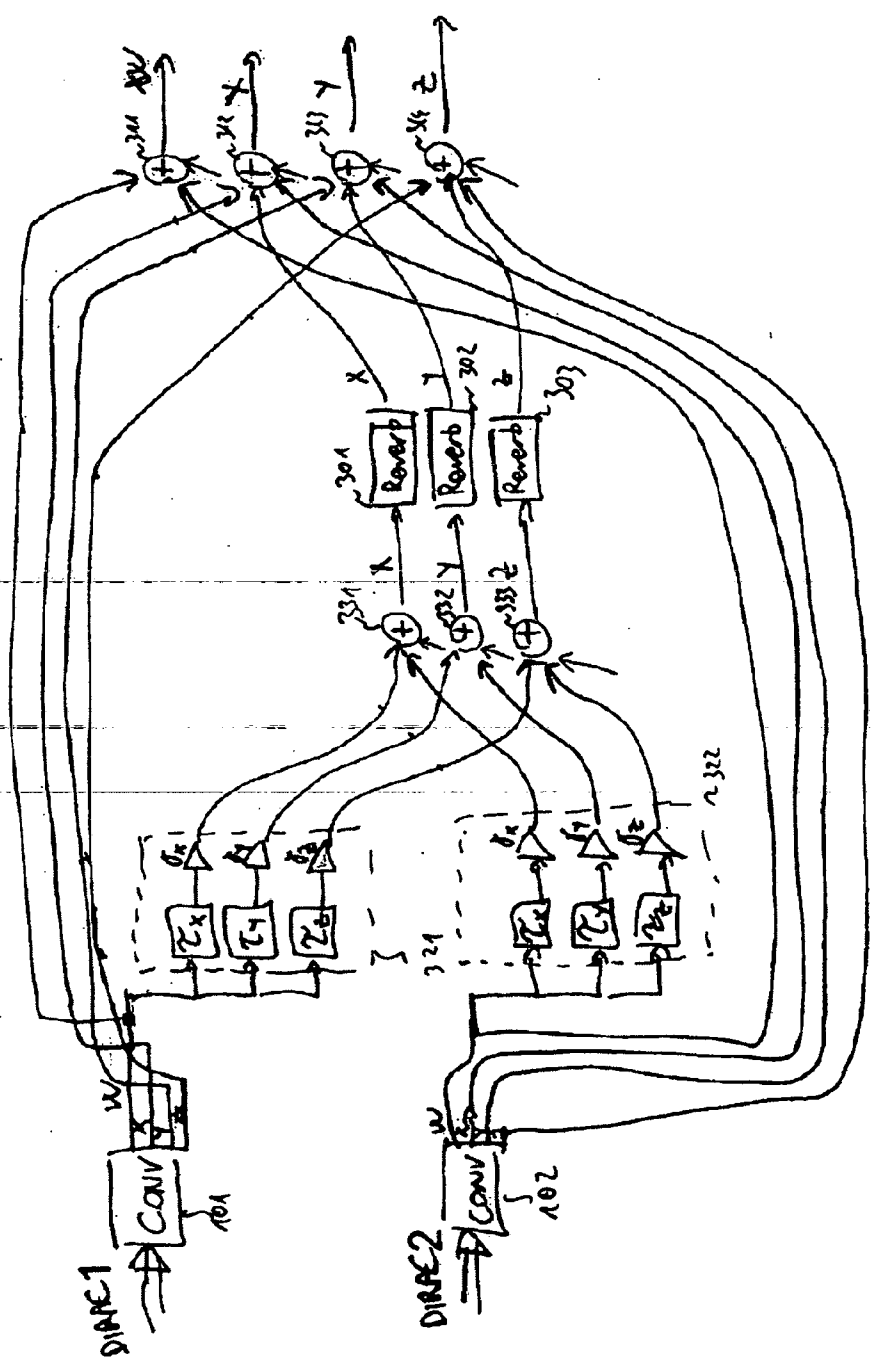
Fig 4a



300.7

DIRACN

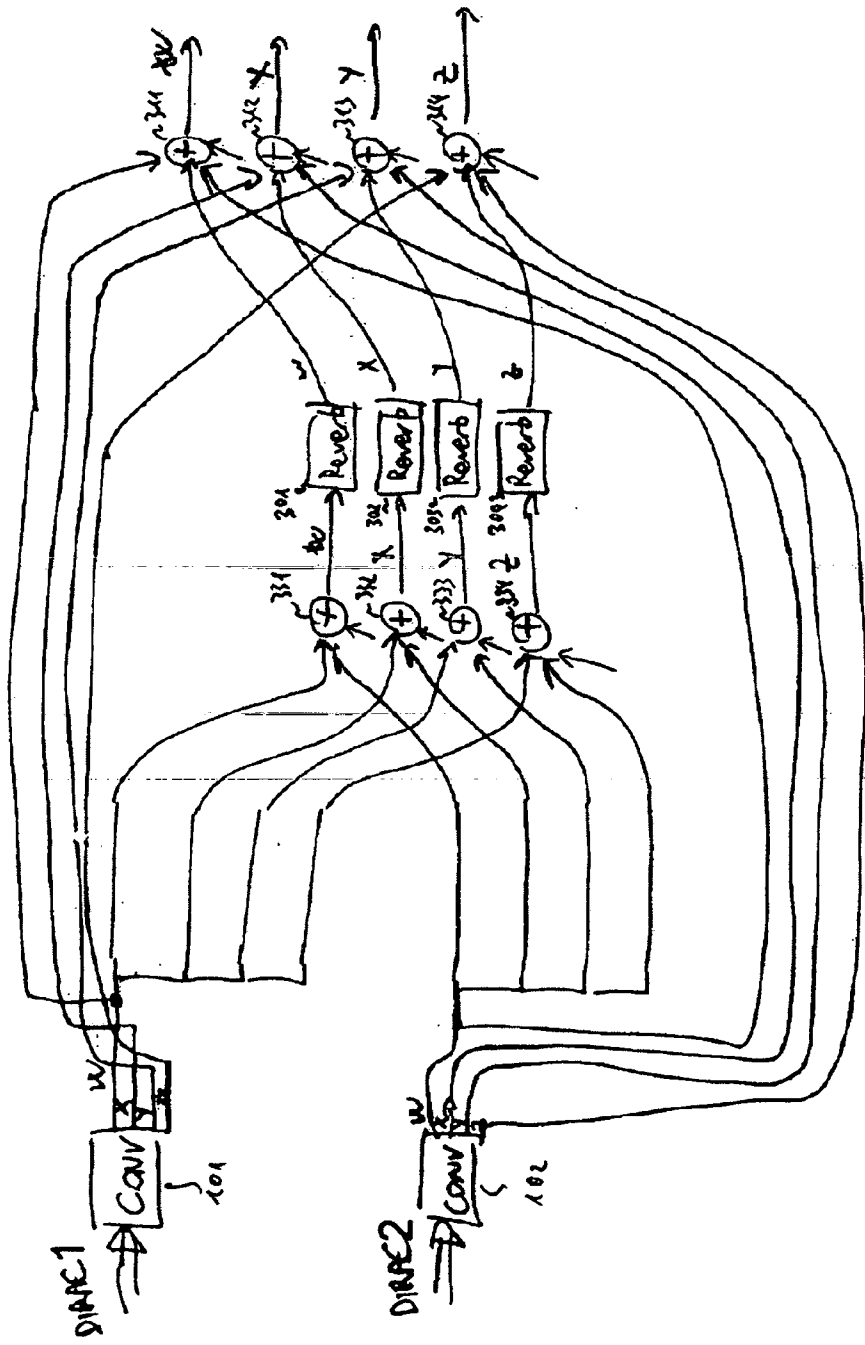
Fig 46



3007

DIRACN

Fig. 4c



DIRACN \rightarrow $\sqrt{300}$

Fig. 4d

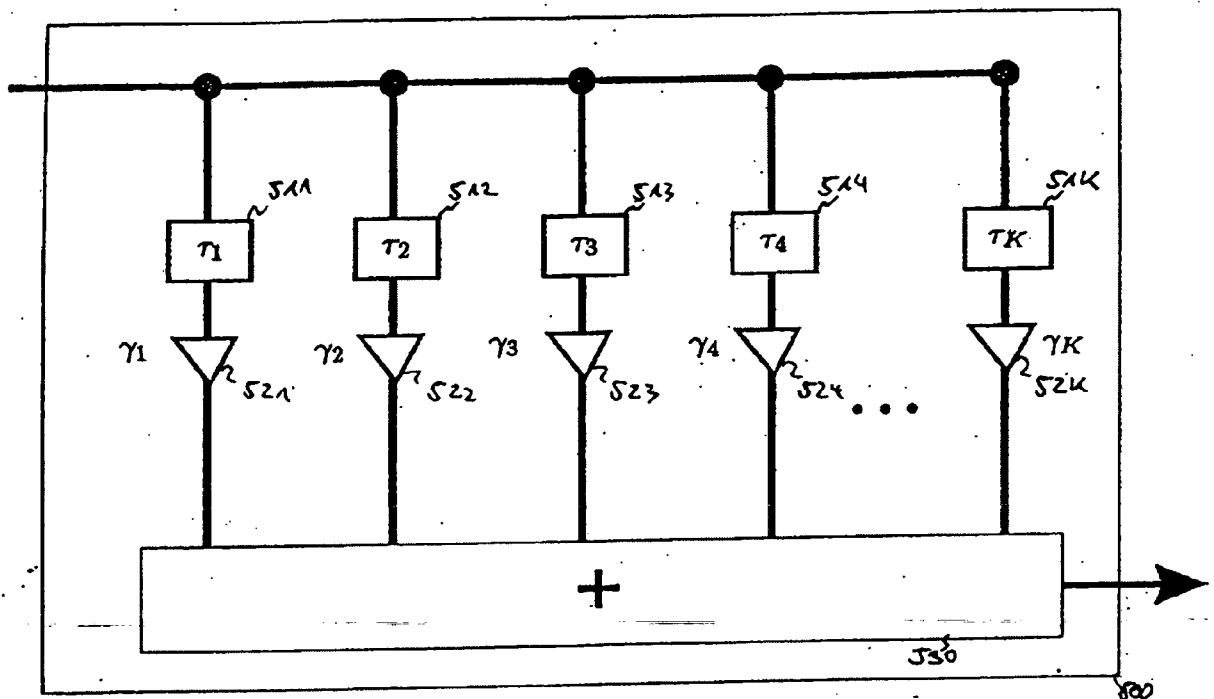


Fig. 5

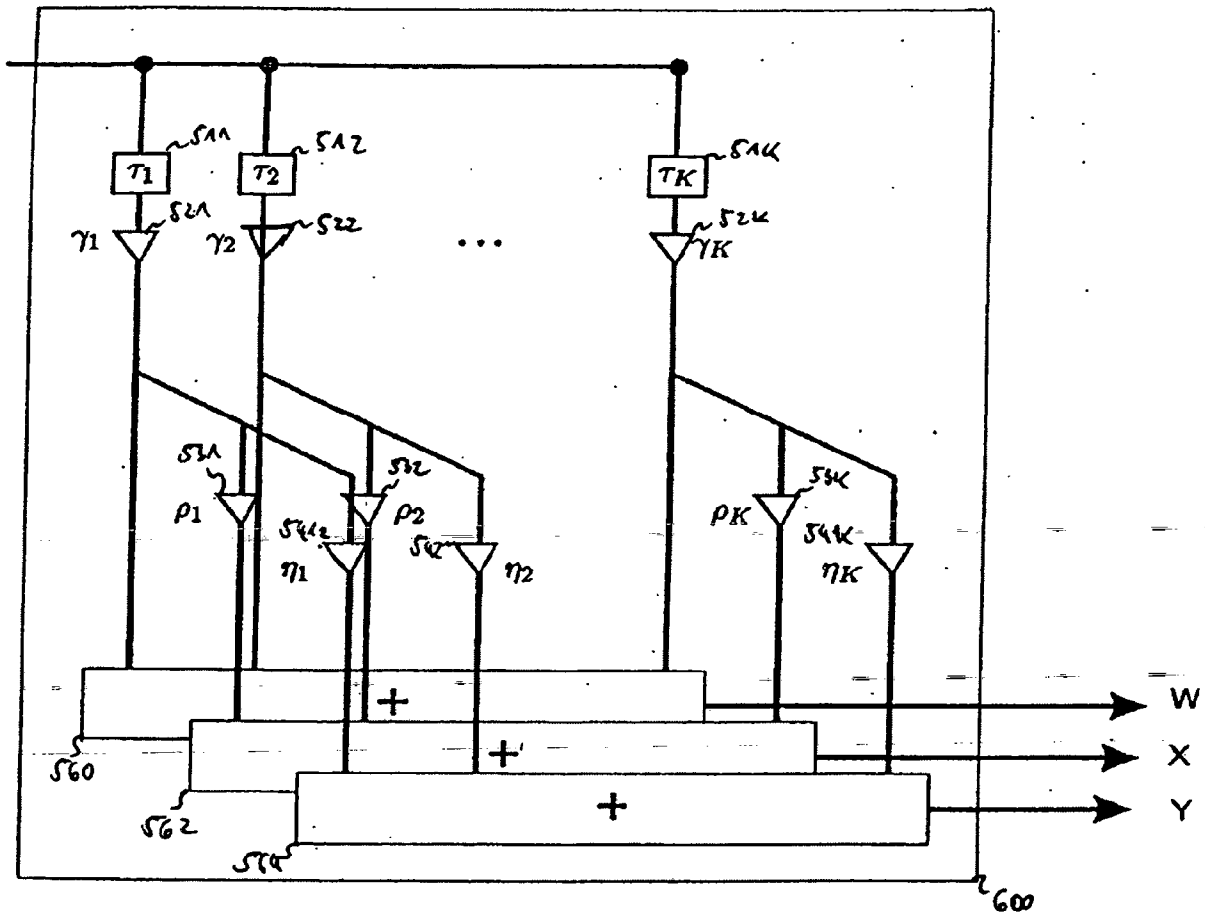


Fig. 6

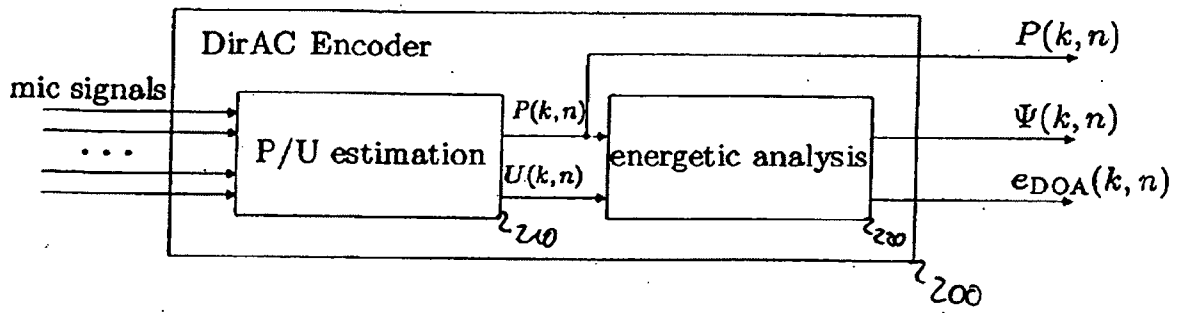


Fig. 2

**Aparelho para Determinar um Sinal de Áudio
Espacial Convertido**

Resumo

Aparelho (100) para determinar um sinal de áudio espacial convertido, o sinal de áudio espacial convertido tendo um componente de áudio omnidirecional (W') e pelo menos um componente de áudio direcional, de um sinal de áudio espacial de entrada, o sinal de áudio espacial de entrada tendo uma representação de áudio de entrada (W) e uma direção de chegada de entrada (\emptyset). O aparelho (100) compreende um estimador (110) para estimar uma representação de onda (W) compreendendo uma medição de campo de onda e uma medição de direção de chegada de onda baseada na representação de áudio de entrada (W) e na direção de chegada de entrada (\emptyset). O aparelho (100) compreende ainda um processador (120) para processar a medição de campo de onda e a medição de direção de chegada de onda para obter o componente de áudio omnidirecional (W') e o pelo menos um componente direcional ($X; Y; Z$).