



República Federativa do Brasil

Ministério do Desenvolvimento, Indústria,
Comércio e Serviços

Instituto Nacional da Propriedade Industrial



(11) BR 112014028439-3 B1

(22) Data do Depósito: 06/05/2013

(45) Data de Concessão: 14/02/2023

(54) Título: MÉTODO E APARELHO PARA COMPRIMIR UM SINAL AMBISSÔNICO DE ORDEM SUPERIOR (AOS), MÉTODO E APARELHO PARA DESCOMPRIMIR UM SINAL AMBISSÔNICO DE ORDEM SUPERIOR (AOS) COMPRIMIDO, E REPRESENTAÇÃO DE SINAL AOS

(51) Int.Cl.: H04S 3/02; G10L 19/008.

(52) CPC: H04S 3/02; H04S 3/008; G10L 19/008; H04S 2420/11.

(30) Prioridade Unionista: 14/05/2012 EP 12305537.8.

(73) Titular(es): DOLBY INTERNATIONAL AB.

(72) Inventor(es): ALEXANDER KRÜGER; SVEN KORDON; JOHANNES BOEHM; JOHANN-MARKUS BATKE.

(86) Pedido PCT: PCT EP2013059363 de 06/05/2013

(87) Publicação PCT: WO 2013/171083 de 21/11/2013

(85) Data do Início da Fase Nacional: 14/11/2014

(57) Resumo: MÉTODO E APARELHO PARA COMPRIMIR E DESCOMPRIMIR UMA REPRESENTAÇÃO DE SINAL AMBISONICS DE ORDEM SUPERIOR. Ambisonics de Ordem Superior (AOS) representa um campo sonoro completo nas proximidades de um ponto perfeito, independente da configuração de alto-falante. A resolução espacial elevada requer um número elevado de coeficientes de AOS. Na invenção, direções sonoras dominantes são estimadas e a representação do sinal de AOS é decomposta em sinais direcionais dominantes no domínio do tempo e informações direcionais correlatas, e um componente ambiental no domínio da AOS, seguido pela compressão do componente ambiental através da redução de sua ordem. O componente ambiental de ordem reduzida é transformado para o domínio espacial, e é codificado perceptivamente junto com os sinais direcionais. No lado do receptor, os sinais direcionais codificados e o componente ambiental codificado de ordem reduzida são perceptivamente descomprimidos, os sinais ambientais perceptivamente descomprimidos são transformados para uma representação do domínio da AOS de ordem reduzida, seguido pela extensão da ordem. A representação de AOS total é recomposta a partir dos sinais direcionais, das informações direcionais correspondentes, e do componente de AOS ambiental de ordem original.

“MÉTODO E APARELHO PARA COMPRIMIR UM SINAL AMBISSÔNICO DE ORDEM SUPERIOR (AOS), MÉTODO E APARELHO PARA DESCOMPRIMIR UM SINAL AMBISSÔNICO DE ORDEM SUPERIOR (AOS) COMPRIMIDO, E REPRESENTAÇÃO DE SINAL AOS”.

[001]A invenção se refere a um método e a um aparelho para comprimir e descomprimir uma representação de sinal Ambissônico de Ordem Superior, em que componentes direcionais e ambientais são processados de maneira diferente.

Antecedentes

[002]Ambissônico de Ordem Superior (AOS) oferece a vantagem de capturar um campo sonoro completo nas proximidades de um local específico no espaço tridimensional, sendo esse local denominado “ponto perfeito” (*sweet spot*). Essa representação de AOS independe de uma configuração de alto-falante específica, contrariamente às técnicas baseadas em canal, como estéreo ou *surround*. Mas esta flexibilidade ocorre à custa de um processo de decodificação necessário para reproduzir a representação de AOS em uma configuração de alto-falante particular.

[003]AOS se baseia na descrição de amplitudes complexas da pressão do ar para números de onda angular individual k para as posições x nas proximidades de uma posição do ouvinte desejado, o qual, sem perda da generalidade, pode ser presumido como sendo a origem de um sistema esférico de coordenadas, utilizando uma expansão truncada da Harmônica Esférica (HE). A resolução espacial desta representação aumenta com uma ordem máxima crescente N da expansão. Infelizmente, o número de coeficientes de expansão 0 cresce quadraticamente com a ordem N , isto é, $0 = (N + 1)^2$. Por exemplo, representações de AOS típicas que utilizam a ordem $N = 4$ requerem coeficientes de AOS $0 = 25$. Dada uma taxa de amostragem desejada f_s e o número N_b de bits por amostra, a taxa de bit

total para a transmissão de uma representação de sinal AOS é determinada por $0.f_s.N_b$, e a transmissão de uma representação de sinal AOS de ordem $N = 4$ com uma taxa de amostragem de $f_s=48\text{kHz}$ empregando $N_b = 16$ bits por amostra está resultando em uma taxa de bit de 19,2 MBits/s. Desse modo, a compressão das representações do sinal de AOS é altamente desejável.

[004]Um panorama das abordagens de compressão de áudio espacial existente pode ser encontrado no pedido de patente EP 10306472.1 ou em I. Elfitri, B. Günel, A.M. Kondoz, "Multichannel Audio Coding Based on Analysis by Synthesis", Proceedings of the IEEE, vol.99, no.4, pp.657-670, Abril de 2011.

[005]As técnicas a seguir são mais relevantes em relação à invenção.

[006]Sinais de formato B, que são equivalentes às representações Ambissônico de primeira ordem, podem ser comprimido utilizando a Codificação de Áudio Direcional (DirAC) conforme descrito in V. Pulkki, "Spatial Sound Reproduction with Directional Audio Coding", Journal of Audio Eng. Society, vol.55 (6), pp.503-516, 2007. Em uma versão sugerida para aplicações em teleconferência, o sinal no formato B é codificado em um único sinal omnidirecional, bem como informações laterais na forma de uma única direção e de um parâmetro de redundância por banda de frequência. No entanto, a redução drástica da taxa de dados resultante implica em menor qualidade do sinal obtido na reprodução. Ademais, a DirAC é limitada à compressão das representações Ambissônico de primeira ordem, que sofrem com a baixíssima resolução espacial.

[007]Os métodos conhecidos para a compressão das representações de AOS com $N>1$ são raríssimos. Um deles realiza a codificação direta de sequências de coeficiente de AOS individuais empregando o codec perceptivo Codificação de Áudio Avançada (AAC), de

acordo com E. Hellerud, I. Burnett, A. Solvang, U. Peter Svensson, "Encoding High Order Ambissônico with AAC", 124th AES Convention, Amsterdam, 2008. No entanto, o problema inerente com tal abordagem é a codificação perceptiva dos sinais que jamais são ouvidos. Os sinais de reprodução reconstruídos são normalmente obtidos por uma soma ponderada das sequências de coeficiente de AOS. Por essa razão é grande a probabilidade de desmascaramento do ruído de codificação perceptiva quando a representação de AOS descomprimida é fornecida em uma configuração de alto-falante particular. Em termos mais técnicos, o principal problema para o desmascaramento do ruído de codificação perceptiva são as correlações cruzadas altas entre as sequências de coeficientes de AOS individuais. Como os sinais de ruído codificados nas sequências de coeficiente de AOS individuais normalmente não são correlacionadas entre si, pode ocorrer uma sobreposição construtiva do ruído de codificação perceptiva, enquanto, ao mesmo tempo, as sequências de coeficiente de AOS livres de ruído são canceladas na sobreposição. Outro problema é que as correlações cruzadas mencionadas levam à eficiência reduzida dos codificadores perceptivos.

[008]A fim de minimizar a extensão destes efeitos, o documento EP 10306472.1 sugere transformar a representação de AOS em uma representação equivalente no domínio espacial antes da codificação perceptiva. Os sinais do domínio espacial correspondem aos sinais direcionais convencionais, e corresponderiam aos sinais de alto-falante se os alto-falantes estivessem posicionados exatamente nas mesmas direções que aquelas assumidas para a transformada do domínio espacial.

[009]A transformada para o domínio espacial reduz as correlações cruzadas entre os sinais do domínio espacial individuais. Entretanto, as correlações cruzadas não são completamente eliminadas. Um exemplo para

as correlações cruzadas relativamente altas é um sinal direcional, cuja direção recai entre as direções adjacentes cobertas pelos sinais do domínio espacial.

[010]Outra desvantagem do documento EP 10306472.1 e no citado artigo de Hellerud et al. é que o número de sinais codificados perceptivamente é $(N + 1)^2$, onde N é a ordem da representação de AOS. Portanto, a taxa de dados para a representação de AOS comprimida cresce quadraticamente com a ordem de Ambissônico.

[011]O processamento de compressão da invenção executa uma decomposição de uma representação do campo sonoro de AOS em um componente direcional e em um componente ambiental. Em particular, para a computação do componente direcional do campo do som, descreve-se abaixo um novo processamento para a estimativa de diversas direções sonoras dominantes.

[012]Quanto aos métodos existentes para estimativa de direção baseados em Ambissônico, o artigo de Pulkki já mencionado descreve um método associado à codificação DirAC para a estimativa da direção, com base na representação do campo do som no formato B. A direção é obtida do vetor de intensidade média, que aponta para a direção do fluxo de energia do campo sonoro. Uma alternativa baseada no formato B é sugerida em D. Levin, S. Gannot, E.A.P. Habets, "Direction-of-Arrival Estimation using Acoustic Vector Sensors in the Presence of Noise", IEEE Proc. d ICASSP, pp.105-108, 2011. A estimativa de direção é realizada iterativamente buscando a direção que fornece a energia máxima de um sinal de saída formador de feixe dirigido para aquela direção.

[013]No entanto, ambas as abordagens são limitadas ao formato B para a estimativa de direção, que sofre com uma resolução espacial relativamente baixa. Uma desvantagem adicional é que a estimativa está

restrita a uma única direção dominante.

[014]As representações de AOS oferecem uma resolução espacial aprimorada e, portanto, permite uma melhor estimativa das diversas direções dominantes. Os métodos existentes que realizam uma estimativa das diversas direções com base nas representações do campo sonoro de AOS são bastante raros. Uma abordagem baseada na percepção compressiva é sugerida em N. Epain, C. Jin, A. van Schaik, "The Application of Compressive Amostragem to the Analysis and Synthesis of Spatial Sound Fields", 127th Convention of the Audio Eng. Soc, Nova York, 2009, em in A. Wabnitz, N. Epain, A. van Schaik, C Jin, "Time Domain Reconstruction of Spatial Sound Fields Using Compressed Sensing", IEEE Proc. of the ICASSP, pp.465-468, 2011. A ideia principal é assumir que o campo sonoro é espacialmente esparso, isto é, consistir apenas de um pequeno número de sinais direcionais. Em seguida à alocação de um número elevado de direções de teste sobre a esfera, um algoritmo de otimização é utilizado a fim de encontrar o menor número possível de direções de teste junto com os sinais direcionais correspondentes, de modo que são bem descritos pela representação de AOS dada. Este método fornece uma resolução espacial em relação à que é efetivamente fornecida pela dada representação de AOS, já que evita a dispersão espacial resultante de uma ordem limitada da representação de AOS dada. No entanto, o desempenho do algoritmo depende intensamente de a presunção de dispersão ser satisfeita. Em particular, a abordagem falha, se o campo sonoro contiver quaisquer componentes ambientais menores adicionais, ou se a representação de AOS for afetada pelo ruído que ocorrerá quando computado de gravações em múltiplos canais.

[015]Ademais, um método especialmente intuitivo é transformar a representação de AOS dada para o domínio espacial conforme descrito em

B. Rafaely, "Plane-wave decomposition of the sound Field on a sphere by spherical convolution", J. Acoust. Soc. Am., vol.4, no.116, pp.2149-2157, Outubro de 2004, e em seguida buscar a máxima nas energias direcionais. A desvantagem desta abordagem é que a presença de componentes ambientais prejudica a distribuição da energia direcional e a um deslocamento da máxima das energias direcionais em comparação à ausência de qualquer componente ambiental.

Invenção

[016]Um problema a ser solucionado pela invenção é fornecer uma compressão para os sinais AOS através da qual a resolução espacial elevada da representação de sinal AOS ainda é mantida. Este problema é solucionado pelos métodos revelados nas reivindicações 1 e 2. Os aparelhos que utilizam estes métodos são revelados nas reivindicações 3 e 4.

[017]A invenção aborda a compressão das representações de AOS Ambissônico de Ordem Superior dos campos sonoros. No presente relatório descritivo, o termo 'AOS' denota a representação Ambissônico de Ordem Superior em tal condição, bem como um sinal de áudio correspondentemente codificado ou representado. As direções sonoras dominantes são estimadas e a representação de sinal AOS é decomposta em diversos sinais direcionais dominantes no domínio do tempo e informações direcionais correlatas, e em um componente ambiental no domínio AOS, seguido pela compressão do componente ambiental através da redução de sua ordem. Depois dessa decomposição, o componente AOS ambiental de ordem reduzida é transformado para o domínio espacial, e é codificado perceptivamente junto com os sinais direcionais.

[018]No lado do receptor ou do decodificador, os sinais direcionais codificados e o componente ambiental codificado de ordem reduzida são perceptivamente descomprimidos. Os sinais ambientais perceptivamente

descomprimidos são transformados para uma representação do domínio AOS de ordem reduzida, seguido pela extensão de ordem. A representação de AOS total é recomposta a partir dos sinais direcionais e das informações direcionais correspondentes e do componente AOS ambiental de ordem original.

[019]Em termos de vantagens, o componente do campo sonoro ambiental pode ser representado com exatidão suficiente por uma representação de AOS com uma ordem inferior à ordem original, e a extração dos sinais direcionais dominantes garante que, após a compressão e descompressão, ainda se obtém uma elevada resolução espacial.

[020]Em princípio, o método da invenção é adequado para comprimir uma representação do sinal Ambissônico de Ordem Superior AOS, sendo que o dito método inclui as etapas de:

- estimar as direções dominantes, em que a dita estimativa de direção dominante depende de uma distribuição da energia direcional dos componentes AOS energeticamente dominantes;

- decompor ou decodificar a representação de sinal AOS em diversos sinais direcionais dominantes no domínio do tempo e informações direcionais correlatas, e um componente ambiental residual no domínio AOS, em que o dito componente ambiental residual representa a diferença entre a dita representação de sinal AOS e uma representação dos ditos sinais direcionais dominantes;

- comprimir o dito componente ambiental residual reduzindo sua ordem em relação à sua ordem original;

- transformar o dito componente AOS ambiental residual de ordem reduzida para o domínio espacial;

- codificar perceptivamente os ditos sinais direcionais dominantes e o dito componente AOS ambiental residual transformado.

[021]Em princípio, o método da invenção é adequado para descomprimir uma representação do sinal Ambissônico de Ordem Superior AOS que foi comprimido pelas etapas de:

- estimar as direções dominantes, em que a dita estimativa de direção dominante depende de uma distribuição da energia direcional dos componentes AOS energeticamente dominantes;
- decompor ou decodificar a representação de sinal AOS em diversos sinais direcionais dominantes no domínio do tempo e informações direcionais correlatas, e um componente ambiental residual no domínio AOS, em que o dito componente ambiental residual representa a diferença entre a dita representação de sinal AOS e uma representação dos ditos sinais direcionais dominantes;
- comprimir o dito componente ambiental residual reduzindo sua ordem em relação à sua ordem original;
- transformar o dito componente AOS ambiental residual de ordem reduzida para o domínio espacial;
- codificar perceptivamente os ditos sinais direcionais dominantes e o dito componente AOS ambiental residual transformado, sendo que o método inclui as etapas de:
 - decodificar perceptivamente os ditos sinais direcionais dominantes codificados perceptivamente e o dito componente AOS ambiental residual transformado codificado perceptivamente;
 - submeter à transformada inversa o dito componente AOS ambiental residual transformado decodificado perceptivamente de modo a obter uma representação do domínio AOS;
 - executar uma extensão de ordem do dito componente AOS ambiental residual submetido à transformada inversa de modo a estabelecer um componente AOS ambiental de ordem original;

- compor os ditos sinais direcionais dominantes decodificados perceptivamente, a dita informação de direção e o dito componente AOS ambiental estendido de ordem original de modo a obter uma representação de sinal AOS.

[022]Em princípio, o aparelho da invenção é adequado para comprimir uma representação do sinal Ambissônico de Ordem Superior AOS, sendo que o aparelho inclui:

- meio que é adaptado para estimar as direções dominantes, em que a dita estimativa de direção dominante depende de uma distribuição da energia direcional dos componentes AOS energeticamente dominantes;

- meio que é adaptado para decompor ou decodificar a representação de sinal AOS em diversos sinais direcionais dominantes no domínio do tempo e informações direcionais correlatas, e um componente ambiental residual no domínio AOS, em que o dito componente ambiental residual representa a diferença entre a dita representação de sinal AOS e uma representação dos ditos sinais direcionais dominantes;

- meio que é adaptado para comprimir o dito componente ambiental residual reduzindo sua ordem em relação à sua ordem original;

- meio que é adaptado para transformar o dito componente AOS ambiental residual de ordem reduzida para o domínio espacial;

- meio que é adaptado para codificar perceptivamente os ditos sinais direcionais dominantes e o dito componente AOS ambiental transformado residual.

[023]Em princípio o aparelho da invenção é adequado para descomprimir uma representação do sinal Ambissônico de Ordem Superior AOS que foi comprimido pelas etapas de:

- estimar as direções dominantes, em que a dita estimativa de direção dominante depende de uma distribuição da energia direcional dos

componentes AOS energeticamente dominantes;

- decompor ou decodificar a representação de sinal AOS em diversos sinais direcionais dominantes no domínio do tempo e informações direcionais correlatas, e um componente ambiental residual no domínio AOS, em que o dito componente ambiental residual representa a diferença entre a dita representação de sinal AOS e uma representação dos ditos sinais direcionais dominantes;

- comprimir o dito componente ambiental residual reduzindo sua ordem em relação à sua ordem original;

- transformar o dito componente AOS ambiental residual de ordem reduzida para o domínio espacial;

- codificar perceptivamente os ditos sinais direcionais dominantes e o dito componente AOS ambiental residual transformado, sendo que o aparelho inclui:

- meio que é adaptado para decodificar perceptivamente os ditos sinais direcionais dominantes codificados perceptivamente e o dito componente AOS ambiental residual transformado codificado perceptivamente;

- meio que é adaptado para submeter à transformada inversa o dito componente AOS ambiental residual transformado decodificado perceptivamente de modo a obter uma representação do domínio AOS;

- meio que é adaptado para executar uma extensão de ordem do dito componente AOS ambiental residual submetido à transformada inversa de modo a estabelecer um componente AOS ambiental de ordem original;

- meio que é adaptado para compor os ditos sinais direcionais dominantes decodificados perceptivamente, a dita informação de direção e o dito componente AOS ambiental estendido de ordem original de modo a obter uma representação de sinal AOS.

[024] Outras modalidades vantajosas da invenção são reveladas nas respectivas reivindicações dependentes.

Desenhos

[025] Modalidades exemplificativas da invenção são descritas fazendo referência aos desenhos em anexo, que mostram na:

[026] Figura 1 - Função de dispersão normalizada $v_N(\theta)$ para diferentes ordens Ambissônico N e para os ângulos $\theta \in [0, \pi]$;

[027] Figura 2 - Diagrama de bloco do processamento de compressão de acordo com a invenção;

[028] Figura 3 - Diagrama de bloco do processamento de descompressão de acordo com a invenção.

Modalidades exemplificativas

[029] Os sinais Ambissônico descrevem os campos sonoros em áreas livres de fonte utilizando a expansão da Harmônica Esférica (HS). A viabilidade desta descrição pode ser atribuída à propriedade física em que o comportamento temporal e espacial da pressão sonora é essencialmente determinado pela equação da onda.

Equação da onda e expansão da Harmônica Esférica

[030] Para uma descrição mais detalhada de Ambissônico, a seguir assume-se um sistema esférico de coordenadas, onde um ponto no espaço $\mathbf{x} = (r, \theta, \phi)^T$ é representado por um raio $r > 0$ (isto é, a distância até a origem da coordenada), um ângulo de inclinação $\theta \in [0, \pi]$ medido a partir do eixo polar z, e um ângulo do azimute $\phi \in [0, 2\pi]$ medido no plano $x=y$ a partir do eixo x. Neste sistema esférico de coordenadas, a equação da onda para a pressão sonora $p(t, x)$ em uma área livre de fonte conectada, onde t denota tempo, é dada pelo livro de Earl G. Williams, "Fourier Acoustics", vol.93 of Applied Mathematical Sciences, Academic Press, 1999:

$$[031] \frac{1}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial p(t, x)}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial p(t, x)}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 p(t, x)}{\partial \phi^2} \right] - \frac{1}{c_s^2} \frac{\partial^2 p(t, x)}{\partial t^2} = 0 \quad (1)$$

com c_s indicando a velocidade do som. Em consequência disso, a transformada de Fourier da pressão sonora em relação ao tempo

$$P(\omega, \mathbf{x}) := \mathcal{F}_t\{p(t, \mathbf{x})\} \quad (2)$$

$$:= \int_{-\infty}^{\infty} p(t, \mathbf{x}) e^{-i\omega t} dt , \quad (3)$$

onde i denota a unidade imaginária, pode ser expandida na série de HS de acordo com o livro de Williams:

$$P(kc_s, (r, \theta, \phi)^T) = \sum_{n=0}^{\infty} \sum_{m=-n}^n p_n^m(kr) Y_n^m(\theta, \phi) . \quad (4)$$

[032] Cabe notar que esta expansão é válida para todos os pontos \mathbf{x} em uma área livre de fonte conectada, que corresponde à região de convergência da série.

[033] Na equação (4), k denota o número da onda angular definido por

$$k := \frac{\omega}{c_s} \quad (5)$$

$e^{p_n^m(kr)}$ indica os coeficientes de expansão HS, que dependem somente do produto kr .

[034] Ademais, $Y_n^m(\theta, \phi)$ são as funções de HS de ordem n e o grau m :

$$Y_n^m(\theta, \phi) := \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi} , \quad (6)$$

onde $P_n^m(\cos\theta)$ denotam as funções de Legendre associadas e $(\cdot)!$ indica o fatorial.

[035] As funções de Legendre associadas para índices de grau não negativo m são definidas pelas polinomiais de Legendre $P_n(x)$ por

$$P_n^m(x) := (-1)^m (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x) \quad for \ m \geq 0 . \quad (7)$$

[036] Para índices de grau negativo, isto é, $m < 0$, as funções de Legendre associadas são definidas por

$$P_n^m(x) := (-1)^m \frac{(n+m)!}{(n-m)!} P_n^{-m}(x) \quad for \ m < 0 . \quad (8)$$

[037] As polinomiais de Legendre $P_n(x)$ ($n \geq 0$), por sua vez, podem ser definidas utilizando a Fórmula de Rodrigues como

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n . \quad (9)$$

[038] Na técnica anterior, por exemplo, em M. Poletti, "Unified Description of Ambissônico using Real and Complex Spherical Harmonics", Proceedings of the Ambissônico Symposium 2009, 25-27 junho de 2009, Graz, Áustria, também existem definições das funções de HS que divergem daquelas na equação (6) por um fator de $(-1)^m$ para os índices de grau negativo m .

[039] Como alternativa, a transformada de Fourier da pressão sonora em relação ao tempo pode ser expressa utilizando funções reais de HS $S_n^m(\theta, \phi)$ como

$$P(kc_s, (r, \theta, \phi)^T) = \sum_{n=0}^{\infty} \sum_{m=-n}^n q_n^m(kr) S_n^m(\theta, \phi) . \quad (10)$$

[040] Na literatura, existem várias definições de funções reais de HS (vide, por exemplo, o artigo de Poletti citado). Uma possível definição, que é aplicada em todo este documento, é dada por

$$S_n^m(\theta, \phi) := \begin{cases} \frac{(-1)^m}{\sqrt{2}} [Y_n^m(\theta, \phi) + Y_n^{m*}(\theta, \phi)] & \text{for } m > 0 \\ Y_n^m(\theta, \phi) & \text{for } m = 0 \\ \frac{(-1)}{\sqrt{2}} [Y_n^m(\theta, \phi) - Y_n^{m*}(\theta, \phi)] & \text{for } m < 0 \end{cases} \quad (11)$$

onde $(\cdot)^*$ denota uma conjugação complexa. Uma expressão alternativa é obtida inserindo-se a equação (6) na equação (11):

$$S_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) \operatorname{trg}_m(\phi) ,$$

com

$$\begin{cases} (-1)^m \sqrt{2} \cos(m\phi) & \text{for } m > 0 \\ 1 & \text{for } m = 0 \\ -(-1)^m \sqrt{2} \sin(m\phi) & \text{for } m < 0 \end{cases}$$

[041] Embora as funções reais de HS sejam de valor real por definição, isso não se reproduz para os coeficientes de expansão correspondentes $q_n^m(kr)$ em geral.

[042] As funções de HS complexas estão relacionadas às funções reais de HS como segue:

$$Y_n^m(\theta, \phi) = \begin{cases} \frac{q_n^m(kr)}{\sqrt{2}} [S_n^m(\theta, \phi) + iS_n^{-m}(\theta, \phi)] & \text{for } m > 0 \\ S_n^0(\theta, \phi) & \text{for } m = 0 \\ \frac{1}{i\sqrt{2}} [S_n^m(\theta, \phi) + iS_n^{-m}(\theta, \phi)] & \text{for } m < 0 \end{cases} \quad (14)$$

[043] As funções complexas de HS $Y_n^m(\theta, \phi)$ bem como as funções reais de HS $S_n^m(\theta, \phi)$ com o vetor de direção $\Omega = (\theta, \phi)^T$ formam uma base ortonormal para as funções de valor complexo do quadrado integrável na esfera unitária S^2 no espaço tridimensional e, portanto, obedece as condições

$$\int_{S^2} Y_n^m(\Omega) Y_{n'}^{m'*}(\Omega) d\Omega = \int_0^{2\pi} \int_0^\pi Y_n^m(\theta, \phi) Y_{n'}^{m'*}(\theta, \phi) \sin\theta d\theta d\phi \\ = \delta_{n-n'} \delta_{m-m'}, \quad (15)$$

$$\int_{S^2} S_n^m(\Omega) S_{n'}^{m'}(\Omega) d\Omega = \delta_{n-n'} \delta_{m-m'}, \quad (16)$$

onde δ denota a função delta de Kronecker. O segundo resultado pode ser derivado utilizando a equação (15) e a definição de harmônicas esféricas reais na equação (11).

Problema interno e os coeficientes Ambissônico

[044] O propósito do Ambissônico é uma representação de um campo sonoro nas proximidades da origem da coordenada. Sem perda da generalidade, esta região de interesse é aqui assumida como sendo uma bola de raio R centralizado na origem da coordenada, o que é especificado pelo conjunto $\{x | 0 \leq r \leq R\}$. Uma assunção essencial para a representação é que esta bola supostamente não contém fontes sonoras. Descobrir a representação do campo sonoro no interior desta bola denomina-se “problema interno”, cf. o livro de Williams já mencionado.

[045] Pode-se mostrar que, para o problema interno, os coeficientes de expansão das funções de HS $p_n^m(kr)$ podem ser expressas como

$$p_n^m(kr) = a_n^m(k) j_n(kr), \quad (17)$$

onde $j_n(\cdot)$ denotam as funções esféricas de Bessel de primeira ordem. A partir da equação (17) ocorre que a informação completa sobre o campo sonoro está contida nos coeficientes $a_n^m(k)$, denominados coeficientes Ambissônico.

[046] Do mesmo modo, os coeficientes da expansão das funções reais de HS $q_n^m(kr)$ podem ser fatorados como

$$q_n^m(kr) = b_n^m(k)j_n(kr) , \quad (18)$$

onde os coeficientes $b_n^m(k)$ são denominados coeficientes Ambissônico em relação à expansão utilizando as funções de HS de valor real. Eles estão relacionados a $a_n^m(k)$ através de:

$$b_n^m(k) = \begin{cases} \frac{1}{\sqrt{2}}[(-1)^m a_n^m(k) + a_n^{-m}(k)] & \text{for } m > 0 \\ a_n^0(k) & \text{for } m = 0 \\ \frac{1}{\sqrt{m}}[a_n^m(k) - (-1)^m a_n^{-m}(k)] & \text{for } m < 0 \end{cases} . \quad (19)$$

Decomposição da onda plana

[047] O campo sonoro no interior de uma bola livre de fonte sonora centralizada na origem da coordenada pode ser expresso por uma sobreposição de um número infinito de ondas planas de diferentes números da onda angular k , que colidem na bola a partir de todas as direções possíveis, cf. o já citado artigo de Rafaely "Plane-wave decomposition...". Assumindo que a amplitude complexa de uma onda plana com número da onda angular k a partir da direção Ω_0 seja dada por $D(k, \Omega_0)$, é possível mostrar de modo similar com o uso da equação (11) e da equação (19) que os coeficientes Ambissônico correspondentes em relação à expansão das funções reais de HS são dados por

$$b_{n, \text{plane wave}}^m(k; \Omega_0) = 4\pi i^n D(k, \Omega_0) S_n^m(\Omega_0) . \quad (20)$$

[048] Consequentemente, os coeficientes Ambissônico para o campo sonoro resultante da sobreposição de um número infinito de ondas planas do número da onda angular k são obtidos pela integração da equação (20) em todas as direções possíveis $\Omega_0 \in S^2$:

$$b_n^m(k) = \int_{S^2} b_{n, \text{plane wave}}^m(k; \Omega_0) d\Omega_0 \quad (21)$$

$$= 4\pi i^n \int_{S^2} D(k, \Omega_0) S_n^m(\Omega_0) d\Omega_0 . \quad (22)$$

[049] A função $D(k, \Omega_0)$ é denominada "densidade de amplitude" e é assumida como sendo o quadrado integrável sobre a esfera unitária S^2 . É possível expandi-la na série de funções reais de HS como

$$D(k, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n c_n^m(k) S_n^m(\Omega) , \quad (23)$$

onde os coeficientes de expansão $c_n^m(k)$ são iguais à integral que

ocorre na equação (22), isto é,

$$c_n^m(k) = \int_{S^2} D(k, \Omega) S_n^m(\Omega) d\Omega . \quad (24)$$

[050] Inserindo a equação (24) na equação (22) pode-se observar que os coeficientes Ambissônico $b_n^m(k)$ são uma versão escalonada dos coeficientes de expansão $c_n^m(k)$, isto é,

$$b_n^m(k) = 4\pi i^n c_n^m(k) . \quad (25)$$

[051] Ao aplicar a transformada inversa de Fourier em relação ao tempo aos coeficientes Ambissônico escalonados $c_n^m(k)$ e à função da densidade de amplitude $D(k, \Omega_0)$, as quantidades de domínio do tempo correspondentes

$$\tilde{c}_n^m(t) := \mathcal{F}_t^{-1} \left\{ c_n^m \left(\frac{\omega}{c_s} \right) \right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} c_n^m \left(\frac{\omega}{c_s} \right) e^{i\omega t} d\omega \quad (26)$$

$$d(t, \Omega) := \mathcal{F}_t^{-1} \left\{ D \left(\frac{\omega}{c_s}, \Omega \right) \right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} D \left(\frac{\omega}{c_s}, \Omega \right) e^{i\omega t} d\omega \quad (27)$$

são obtidas. Então, no domínio do tempo, a equação (24) pode ser formulada como

$$\tilde{c}_n^m(t) = \int_{S^2} d(t, \Omega) S_n^m(\Omega) d\Omega . \quad (28)$$

[052] O sinal direcional do domínio do tempo $d(t, \Omega)$ pode ser representado por uma expansão da função de HS real de acordo com

$$d(t, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \tilde{c}_n^m(t) S_n^m(\Omega) . \quad (29)$$

[053] Usando o fato de que as funções de HS $S_n^m(\Omega)$ são de valor real, seu conjugado complexo pode ser expresso por

$$d^*(t, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \tilde{c}_n^{m*}(t) S_n^m(\Omega) . \quad (30)$$

[054] Assumindo the sinal do domínio do tempo $d(t, \Omega)$ seja de valor real, isto é, $d(t, \Omega) = d^*(t, \Omega)$, a partir da comparação da equação (29) com a equação (30) os coeficientes $\tilde{c}_n^{m*}(t)$ são de valor real naquele caso, isto é, $\tilde{c}_n^m(t) = \tilde{c}_n^{m*}(t)$.

[055] Os coeficientes $\tilde{c}_n^{m*}(t)$ serão denominados a seguir como coeficientes Ambissônico do domínio do tempo escalonados.

[056] A seguir, assume-se ainda que a representação do campo

sonoro é dada por esses coeficientes, que serão descritos mais detalhadamente na seção abaixo que trata da compressão.

[057] Observa-se que a representação de AOS do domínio do tempo pelos coeficientes $\tilde{c}_n^{m*}(t)$ utilizados para o processamento de acordo com a invenção é equivalente a uma representação de AOS de domínio de frequência correspondente $c_n^m(k)$. Portanto, a compressão e descompressão descritas podem ser realizadas equivalentemente no domínio de frequência com as respectivas modificações mínimas das equações.

Resolução espacial com ordem finita

[058] Na prática, o campo sonoro nas proximidades da origem da coordenada é descrito utilizando apenas um número finito de coeficientes Ambissônico $c_n^m(k)$ de ordem $n \leq N$. Computar a função da densidade de amplitude a partir da série truncada de funções de HS de acordo com

$$D_N(k, \Omega) := \sum_{n=0}^N \sum_{m=-n}^n c_n^m(k) S_n^m(\Omega) \quad (31)$$

introduz um tipo de dispersão espacial comparada à função da densidade de amplitude verdadeira $D(k, \Omega)$, cf. o já citado artigo "Plane-wave decomposition...". Isso pode ser realizado computando a função da densidade de amplitude para uma única onda plana a partir da direção Ω_0 com o uso da equação (31):

$$D_N(k, \Omega) = \sum_{n=0}^N \sum_{m=-n}^n \frac{1}{4\pi i^n} \cdot b_{n, \text{plane wave}}^m(k; \Omega_0) S_n^m(\Omega) \quad (32)$$

$$= D(k, \Omega_0) \sum_{n=0}^N \sum_{m=-n}^n S_n^m(\Omega_0) S_n^m(\Omega) \quad (33)$$

$$= D(k, \Omega_0) \sum_{n=0}^N \sum_{m=-n}^n Y_n^{m*}(\Omega_0) Y_n^m(\Omega) \quad (34)$$

$$= D(k, \Omega_0) \sum_{n=0}^N \frac{2n+1}{4\pi} P_n(\cos\theta) \quad (35)$$

$$= D(k, \Omega_0) \left[\frac{N+1}{4\pi(\cos\theta-1)} (P_{N+1}(\cos\theta) - P_N(\cos\theta)) \right] \quad (36)$$

$$= D(k, \Omega_0) v_N(\theta) \quad (37)$$

com

$$v_N(\theta) := \frac{N+1}{4\pi(\cos\theta-1)} (P_{N+1}(\cos\theta) - P_N(\cos\theta)) , \quad (38)$$

onde Θ denota o ângulo entre os dois vetores que apontam para as direções Ω e Ω_0 satisfazendo a propriedade

$$\cos\theta = \cos\theta \cos\theta_0 + \cos(\phi - \phi_0) \sin\theta \sin\theta_0 . \quad (39)$$

[059] Na equação (34), são empregados os coeficientes Ambissônico para uma onda plana dada na equação (20), enquanto nas equações (35) e (36) alguns teoremas matemáticos são explorados, cf. o já mencionado artigo "Plane-wave decomposition...". A propriedade na equação (33) pode ser mostrada utilizando a equação (14).

[060] Comparando equação (37) à função da densidade de amplitude verdadeira

$$D(k, \Omega) = D(k, \Omega_0) \frac{\delta(\Theta)}{2\pi} , \quad (40)$$

onde $\delta(\cdot)$ denota a função delta Dirac, a dispersão espacial torna-se óbvia a partir da substituição da função delta Dirac escalonada pela função de dispersão $v_N(\Theta)$ que, depois de normalizada por seu valor máximo, é ilustrada na Figura 1 para diferentes ordens Ambissônico N e ângulos $\Theta \in [0, \pi]$.

[061] Como o primeiro zero de $v_N(\Theta)$ está localizado aproximadamente a $\frac{\pi}{N}$ para $N \geq 4$ (vide o já mencionado artigo "Plane-wave decomposition..."), o efeito da dispersão é reduzido (e, desse modo, a resolução espacial é aprimorada) com ordem de Ambissônico crescente N .

[062] Para $N \rightarrow \infty$, a função de dispersão $v_N(\Theta)$ converge para a função delta Dirac escalonada. Isso pode ser observado se a relação de completude para as polinomiais de Legendre

$$\sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(x) P_n(x') = \delta(x - x') \quad (41)$$

for utilizada junto com a equação (35) para expressar o limite de $v_N(\Theta)$ para $N \rightarrow \infty$ como

$$\lim_{N \rightarrow \infty} v_N(\Theta) = \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(\cos\Theta) \quad (42)$$

$$= \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(\cos\Theta) P_n(1) \quad (43)$$

$$= \frac{1}{2\pi} \delta(\cos\Theta - 1) \quad (44)$$

$$= \frac{1}{2\pi} \delta(\Theta) . \quad (45)$$

[063] Ao definir o vetor das funções reais de HS da ordem $n \leq N$ por

$$\mathbf{S}(\Omega) := (S_0^0(\Omega), S_1^{-1}(\Omega), S_1^0(\Omega), S_1^1(\Omega), S_2^{-2}(\Omega), S_N^N(\Omega))^T \in \mathbb{R}^O , \quad (46)$$

onde $0 = (N + 1)^2$ e onde $(\cdot)^T$ denota transposição, a comparação a equação (37) com a equação (33) mostra que a função de dispersão pode ser expressa através do produto escalar de dois vetores de HS reais como $v_N(\Theta) = \mathbf{S}^T(\Theta)\mathbf{S}(\Theta_0)$. (47).

[064]A dispersão pode ser equivalentemente expressa no domínio do tempo como

$$d_N(t, \Omega) := \sum_{n=0}^N \sum_{m=-n}^n \tilde{c}_n^m(t) S_n^m(\Omega) \quad (48)$$

$$= d(t, \Omega_0) v_N(\Theta) . \quad (49)$$

Amostragem

[065]Para algumas aplicações é desejável determinar os coeficientes Ambissônico do domínio do tempo escalonados $\tilde{c}_n^m(t)$ a partir das amostras da função da densidade de amplitude do domínio do tempo $d(t, \Omega)$ em um número finito J de direções discretas Ω_j . A integral na equação (28) é então aproximada por uma soma finita de acordo com B. Rafaely, "Analysis and Design of Spherical Microphone Arrays", IEEE Transactions on Speech and Audio Processing, vol.13, nº 1, pp.135-143, Janeiro de 2005:

$$\tilde{c}_n^m(t) \approx \sum_{j=1}^J g_j \cdot d(t, \Omega_j) S_n^m(\Omega_j) , \quad (50)$$

onde g_j denotam alguns pesos de amostragem corretamente escolhidos. Em oposição ao artigo "Analysis and Design...", a aproximação (50) se refere a uma representação do domínio do tempo que usa funções reais de HS invés de uma representação do domínio de frequência utilizando funções de HS complexas. Uma condição necessária para que a aproximação (50) seja exata é que a densidade de amplitude seja da ordem harmônica limitada N , significando que

$$\tilde{c}_n^m(t) = 0 \quad for \quad n > N . \quad (51)$$

[066]Se esta condição não for satisfeita, a aproximação (50) sofre de erros de *aliasing* espacial, cf. B. Rafaely, "Spatial Aliasing in Spherical Microphone Arrays", IEEE Transactions on Signal Processing, vol.55, nº 3,

pp.1003-1010, Março de 2007. Uma segunda condição necessária requer que os pontos de amostragem e os pesos correspondentes satisfaçam as condições correspondentes dados no artigo "Analysis and Design...":

$$\sum_{j=1}^J g_j S_{n'}^{m'}(\Omega_j) S_n^m(\Omega_j) = \delta_{n-n'} \delta_{m-m'} \quad \text{for } m, m' \leq N. \quad (52)$$

[067]As condições (51) e (52) em conjunto são suficientes para a amostragem exata.

[068]A condição da amostragem (52) consiste em um conjunto de equações lineares, que podem ser formuladas compactamente utilizando uma única equação de matriz como

$$\Psi G \Psi^H = I, \quad (53)$$

onde Ψ indica a matriz de modo definida por

$$\Psi := [S(\Omega_1) \quad \dots \quad S(\Omega_J)] \in \mathbb{R}^{O \times J} \quad (54)$$

e G denota a matriz com os pesos em sua diagonal, isto é,

$$G := \text{diag}(g_1, \dots, g_J). \quad (55)$$

[069]A partir da equação (53) pode-se observar que uma condição necessária para que a equação (52) se mantenha é que o número J de pontos de amostragem satisfaça $J \geq O$. Coletando os valores da densidade de amplitude do domínio do tempo nos pontos de amostragem J no vetor

$$\mathbf{w}(t) := (D(t, \Omega_1), \dots, D(t, \Omega_J))^T, \quad (56)$$

e definindo o vetor dos coeficientes Ambissônico do domínio do tempo escalonados por

$$\mathbf{c}(t) := (\tilde{c}_0^0(t), \tilde{c}_1^{-1}(t), \tilde{c}_1^0(t), \tilde{c}_1^1(t), \tilde{c}_2^{-2}(t), \dots, \tilde{c}_O^0(t))^T, \quad (57)$$

ambos os vetores são relacionados através da expansão das funções de HS (29). Esta relação fornece o seguinte sistema de equações lineares: $\mathbf{w}(t) = \Psi^H \mathbf{c}(t)$.

[070]Utilizando a notação do vetor introduzido, a computação dos coeficientes Ambissônico do domínio do tempo escalonados a partir dos valores das amostras da função da densidade de amplitude do domínio do

tempo pode ser redigida como

$$\mathbf{c}(t) \approx \Psi \mathbf{G} \mathbf{w}(t) . \quad (59)$$

[071] Dada uma ordem de Ambissônico fixa N , em geral não é possível computar um número $J \geq 0$ de pontos de amostragem Ω , e os pesos correspondentes, de modo que a equação (52) da condição de amostragem se mantém. No entanto, se os pontos de amostragem forem selecionados de modo que a condição da amostragem esteja bem aproximada, então a classe da matriz de modo Ψ é 0 e seu número de condição é baixo. Neste caso, a pseudoinversa $\Psi^+ := (\Psi \Psi^H)^{-1} \Psi \Psi^+$ (60) da matriz de modo Ψ existe e uma aproximação razoável do vetor do coeficiente Ambissônico do domínio do tempo escalonado $\mathbf{c}(t)$ a partir do vetor das amostras da função da densidade de amplitude do domínio do tempo amostras é dada por

$$\mathbf{c}(t) \approx \Psi^+ \mathbf{w}(t) . \quad (61)$$

[072] Se $J = 0$, e a classe da matriz de modo for 0, então sua pseudoinversa coincide com sua inversa desde

$$\Psi^+ = (\Psi \Psi^H)^{-1} \Psi = \Psi^{-H} \Psi^{-1} \Psi = \Psi^{-H} . \quad (62)$$

[073] Se adicionalmente a equação da condição de amostragem (52) for satisfeita, então $\Psi^{-H} = \Psi \mathbf{G}$ (63) se mantém e ambas as aproximações (59) e (61) são equivalentes e exatas.

[074] O vetor $\mathbf{w}(t)$ pode ser interpretado como um vetor dos sinais do domínio do tempo espacial. A transformada do domínio AOS para o domínio espacial pode ser realizada, por exemplo, utilizando a equação (58). Este tipo de transformada é denominada “Transformada Harmônica Esférica” (THE) no presente relatório descritivo e é utilizada quando o componente AOS ambiental de ordem reduzida é transformado para o domínio espacial. Assume-se implicitamente que os pontos de amostragem espaciais Ω_j para a THE satisfazem aproximadamente a condição de amostragem na equação

(52) com $g_j \approx \frac{4\pi}{n}$ for $j = 1, \dots, J$ and that $J = 0$.

[075]Sob essas assunções, a matriz da THE satisfaç $\Psi^H \approx \frac{4\pi}{0} \Psi^{-1}$.

Caso o escalonamento absoluto para a THE não seja importante, a constante — pode ser negligenciada.

Compressão

[076]Esta invenção está relacionada à compressão de uma dada representação de sinal AOS. Como mencionado acima, a representação de AOS é decomposta em um número previamente definido de sinais direcionais dominantes no domínio do tempo e em um componente ambiental no domínio AOS, seguido pela compressão da representação de AOS do componente ambiental através da redução de sua ordem. Esta operação explora a assunção, que é sustentada pelos testes de audição, de que o componente do campo sonoro ambiental pode ser representado com exatidão suficiente por uma representação de AOS com uma ordem inferior. A extração dos sinais direcionais dominantes garante que, após essa compressão e uma descompressão correspondente, uma elevada resolução espacial é obtida.

[077]Depois da decomposição, o componente AOS ambiental de ordem reduzida é transformado para o domínio espacial, e é codificado perceptivamente junto com os sinais direcionais conforme descrito na seção de Modalidades exemplificativas do pedido de patente EP 10306472.1.

[078]O processamento de compressão inclui duas etapas sucessivas, que são representadas na Figura 2. As definições exatas dos sinais individuais estão descritas na seção que se segue intitulada Detalhes da compressão.

[079]Na primeira etapa ou estágio mostrada na Figura 2a, em um estimador da direção dominante 22 as direções dominantes são estimadas e é realizada a decomposição do sinal Ambissônico $C(l)$ em um componente

direcional e um componente residual ou ambiental, onde l denota o índice do quadro. O componente direcional é calculado em uma etapa ou estágio de computação do sinal direcional 23, e assim a representação Ambissônico é convertida nos sinais do domínio do tempo representados por um conjunto de D sinais direcionais convencionais $X(l)$ com direções correspondentes $\bar{\Omega}_{\text{DOM}}(l)$. O componente ambiental residual é calculado em uma etapa ou estágio de computação do componente AOS ambiental 24, e é representado pelos coeficientes de domínio AOS $C_A(l)$.

[080]Na segunda etapa mostrada na Figura 2b, uma codificação perceptiva dos sinais direcionais $X(l)$ e do componente AOS ambiental $C_A(V)$ é realizada da seguinte maneira:

- Os sinais direcionais do domínio do tempo convencionais $X(l)$ podem ser comprimidos individualmente em um codificador perceptivo 27 utilizando qualquer técnica de compressão perceptiva conhecida.
- A compressão do componente do domínio AOS ambiental $C_A(l)$ é realizada em duas subetapas ou estágios.

[081]A primeira subetapa ou estágio 25 realiza a redução da ordem original de Ambissônico N para N_{RED} , por exemplo, $N_{\text{RED}} = 2$, resultando no componente AOS ambiental $C_{A,\text{RED}}(l)$. Aqui, é explorada a assunção de que o componente do campo sonoro ambiental pode ser representado com exatidão por AOS com uma ordem inferior. A segunda subetapa ou estágio 26 baseia-se na compressão descrita no pedido de patente EP 10306472.1. Os $O_{\text{RED}} = (N_{\text{RED}} + 1)^2$ AOS sinais $C_{A,\text{RED}}(l)$ do componente do campo sonoro ambiental, que foram computados na subetapa/estágio 25, são transformados em O_{RED} sinais equivalentes $W_{A,\text{RED}}(l)$ no domínio espacial através da aplicação de uma Transformada Harmônica Esférica, resultando em sinais de domínio do tempo convencionais que podem ser inseridos em um banco de codecs perceptivos paralelos 27. Qualquer codificação

perceptiva ou compressão técnica conhecida pode ser aplicada. Os sinais direcionais codificados $\bar{X}(l)$ e os sinais do domínio espacial codificados de ordem reduzida $\bar{W}_{A,RED}(l)$ são emitidos e podem ser transmitidos ou armazenados.

[082]Como vantagem, a compressão perceptiva de todos os sinais do domínio do tempo $X(l)$ e $W_{A,RED}(l)$ pode ser executada conjuntamente em um codificador perceptivo 27 a fim de aprimorar a eficiência geral da codificação explorando as correlações intercanais potencialmente remanescentes.

Descompressão

[083]O processamento de descompressão para um sinal recebido ou reproduzido é representado na Figura 3. Assim como o processamento de compressão, inclui duas etapas sucessivas.

[084]Na primeira etapa ou estágio mostrado na Figura 3a, em uma decodificação perceptiva 31, realiza-se uma decodificação ou descompressão perceptiva dos sinais direcionais codificados $\bar{X}(l)$ e dos sinais do domínio espacial codificados de ordem reduzida $\bar{W}_{A,RED}(l)$, onde $\bar{X}(l)$ representa o componente e $\bar{W}_{A,RED}(l)$ representa o componente AOS ambiental. Os sinais do domínio espacial decodificados ou descomprimidos perceptivamente $\bar{W}_{A,RED}(l)$ são transformados em um transformador inverso da harmônica esférica 32 em uma representação do domínio AOS da $\hat{C}_{A,RED}(l)$ da ordem N_{RED} por meio de uma transformada inversa da Harmônica Esférica. Depois disso, em uma etapa ou estágio de extensão de ordem 33, uma representação apropriada de AOS $\hat{C}_A(l)$ da ordem N é estimada a partir de $\hat{C}_{A,RED}(l)$ por extensão de ordem.

[085]Na segunda etapa ou estágio mostrada na Figura 3b, a representação de AOS total $\hat{C}(l)$ é recomposta em um montador do sinal de AOS 34 a partir dos sinais direcionais $\bar{X}(l)$ e das informações direcionais

correspondentes $\bar{\Omega}_{\text{DOM}}(l)$, bem como a partir do componente AOS ambiental de ordem original $\hat{C}_A(l)$.

[086] Redução da taxa de dados obtêivel

[087] Um problema solucionado pela invenção é a redução considerável da taxa de dados em comparação aos métodos de compressão existentes para as representações de AOS. A seguir, discute-se a taxa de compressão obtêivel em comparação com a representação de AOS não comprimido. A taxa de compressão resulta da comparação da taxa de dados requerida para a transmissão de um sinal de AOS não comprimido $C(l)$ de ordem N com a taxa de dados requerida para a transmissão de uma representação de sinal comprimido que consiste em D sinais direcionais codificados perceptivamente $X(l)$ com as direções correspondentes $\bar{\Omega}_{\text{DOM}}(l)$ e N_{RED} sinais codificados perceptivamente do domínio espacial $W_{A,\text{RED}}(l)$ representando o componente AOS ambiental.

[088] Para a transmissão do sinal de AOS não comprimido $C(l)$, uma taxa de dados de $0 \cdot f_s \cdot N_b$ é requerida. Por outro lado, a transmissão de D sinais direcionais codificados perceptivamente $X(l)$ requer uma taxa de dados de $D \cdot f_{b,\text{COD}}$, onde $f_{b,\text{COD}}$ denota a taxa de bit dos sinais codificados perceptivamente. Do mesmo modo, a transmissão dos N_{RED} sinais codificados perceptivamente do domínio espacial $W_{A,\text{RED}}(l)$ sinais requer uma taxa de bit de $0_{\text{RED}} \cdot f_{b,\text{COD}}$.

[089] As direções $\bar{\Omega}_{\text{DOM}}(l)$ são assumidas como sendo computadas com base em uma taxa bem menor do que a taxa de amostragem f_s , isto é, são assumidas como sendo fixas pela duração de um sinal quadro consistindo de amostras B , por exemplo, $B = 1200$ para uma taxa de amostragem de $f_s = 48\text{kHz}$, e a parcela correspondente da taxa de dados pode ser desprezada para a computação da taxa de dados total do sinal de AOS comprimido.

[090] Portanto, a transmissão da representação comprimida requer uma taxa de dados de aproximadamente $(D + O_{RED}) \cdot f_{b,COD}$.

Consequentemente, a taxa de compressão r_{COMPR} é

$$r_{COMPR} \approx \frac{O \cdot f_S \cdot N_b}{(D + O_{RED}) \cdot f_{b,COD}} . \quad (64)$$

[091] Por exemplo, a compressão de uma representação de AOS de ordem $N = 4$ empregando uma taxa de amostragem $f_S = 48\text{kHz}$ e $N_b = 16$ bits por amostra para uma representação com $D = 3$ direções dominantes utilizando uma ordem de AOS reduzida $N_{RED} = 2$ e uma taxa de bit de 64 kbits/s resultará em uma taxa de compressão de $r_{COMPR} \approx 25$. A transmissão da representação comprimida requer uma taxa de aproximadamente 768 kbits/s.

[092] Probabilidade reduzida para a ocorrência do desmascaramento do ruído da codificação

[093] Conforme explicação na seção de Antecedentes, a compressão perceptiva dos sinais do domínio espacial descrita no pedido de patente EP 10306472.1 sofre com as correlações cruzadas remanescentes entre os sinais, o que pode levar ao desmascaramento do ruído da codificação perceptiva. De acordo com a invenção, os sinais direcionais dominantes são primeiramente extraídos da representação do campo sonoro de AOS antes de serem codificados perceptivamente. Isso significa que, quando se compõe a representação de AOS, após a decodificação perceptiva, o ruído da codificação possui exatamente a mesma direitividade espacial dos sinais direcionais. Em particular, as contribuições do ruído da codificação, bem como do sinal direcional para qualquer direção arbitrária, são descritas deterministicamente descrito pela função da dispersão espacial explicada na seção Resolução espacial com ordem finita. Em outras palavras, em qualquer momento instantâneo, o vetor dos coeficientes AOS que representam o ruído da codificação é exatamente um múltiplo do vetor dos

coeficientes AOS que representam o sinal direcional. Sendo assim, uma soma ponderada arbitrariamente dos coeficientes AOS com ruído não resultará no desmascaramento do ruído de codificação perceptiva.

[094]Ademais, o componente ambiental de ordem reduzida é processado exatamente como sugere o documento EP 10306472.1, contudo, como, por definição, os sinais do domínio espacial do componente ambiental exibem uma correlação especialmente baixa entre si, a probabilidade para o desmascaramento do ruído da codificação perceptiva é baixa.

Estimativa de direção aprimorada

[095]A estimativa de direção da invenção depende da distribuição da energia direcional do componente AOS energeticamente dominante. A distribuição da energia direcional é computada a partir matriz de correlação de classe reduzida da representação de AOS, que é obtida pela decomposição de valor próprio da matriz de correlação da representação de AOS.

[096]Comparada à estimativa de direção utilizada no artigo já mencionado "Plane-wave decomposition...", oferece a vantagem de ser mais precisa, pois se concentra no componente AOS energeticamente dominante invés de utilizar a representação de AOS completa, já que a estimativa de direção reduz o borramento espacial da distribuição de energia direcional.

[097]Comparada à estimativa de direção sugerida nos artigos já mencionados "The Application of Compressive Sampling to the Analysis and Synthesis of Spatial Sound Fields" e "Time Domain Reconstruction of Spatial Sound Fields using Compressed Sensing", oferece a vantagem de ser mais robusta. A razão é que a decomposição da representação de AOS no componente direcional e ambiental praticamente não consegue ser alcançada com perfeição, de modo que ainda resta uma pequena

quantidade do componente ambiental no componente direcional. Então, os métodos de amostragem compressiva como os desses dois artigos são falhos em fornecer estimativas de direção razoável em decorrência de sua elevada sensibilidade à presença dos sinais ambientais.

[098]Como vantagem, a estimativa de direção da invenção não demonstra esse problema.

Aplicações alternativas da decomposição da representação de AOS

[099]A decomposição descrita da representação de AOS em diversos sinais direcionais com informações direcionais correlatas e um componente ambiental no domínio AOS pode ser utilizada para uma renderização similar à DirAc adaptativa ao sinal da representação de AOS de acordo com a sugerida no já mencionado artigo de Pulkki "Spatial Sound Reproduction with Directional Audio Coding".

[0100]Cada componente AOS pode ser renderizado distintamente porque as características físicas dos dois componentes são diferentes. Por exemplo, os sinais direcionais podem ser renderizados para os alto-falantes utilizando técnicas de posicionamento de sinal como o Posicionamento da Amplitude Baseada em Vetor (*Vetor Based Amplitude Panning* - VBAP), cf. V. Pulkki, "Virtual Sound Source Positioning using Vector Base Amplitude Panning", *Journal of Audio Eng. Society*, vol. 45, nº 6, pp.456-466, 1997. O componente AOS ambiental pode ser renderizado utilizando as técnicas de renderização de AOS convencionais conhecidas.

[0101]Essa renderização não se restringe à representação Ambissônico da ordem ' 1 ' e, assim, pode ser observada como uma extensão da renderização similar à DirAC para as representações de AOS da ordem $N > 1$.

[0102]A estimativa das diversas direções a partir de uma representação de sinal AOS pode ser utilizada para qualquer tipo

relacionado da análise do campo sonoro.

[0103]As seções a seguir descrevem em mais profundidade as etapas de processamento do sinal.

Compressão

Definição do formato de entrada

[0104]Como entrada, os coeficientes de AOS do domínio do tempo escalonado $\tilde{c}_n^m(t)$ definidos na equação (26) são assumidos como sendo amostrados em uma taxa $f_s = \frac{1}{T_s}$.

[0105]Define-se um vetor $c(j)$ para ser formado de todos os coeficientes pertencentes ao tempo de amostragem $t = jT_s$, $j \in \mathbb{Z}$, de acordo com

$$c(j) := [\tilde{c}_0^0(jT_s), \tilde{c}_1^{-1}(jT_s), \tilde{c}_1^0(jT_s), \tilde{c}_1^1(jT_s), \tilde{c}_2^{-2}(jT_s), \dots, \tilde{c}_N^N(jT_s)]^T \in \mathbb{R}^o . \quad (65)$$

Quadros

[0106]Os vetores de entrada $c(j)$ dos coeficientes de AOS escalonados são organizados em quadros na etapa ou estágio de quadros 21 em quadros não sobrepostos de comprimento B de acordo com

$$\mathbf{C}(l) := [c(lB + 1) \ c(lB + 2) \ \dots \ c(lB + B)] \in \mathbb{R}^{o \times B} . \quad (66)$$

[0107]Assumindo uma taxa de amostragem de $f_s = 48\text{kHz}$, um comprimento de quadro adequado é de $B = 1200$ amostras, correspondendo a uma duração de um quadro de 25ms.

Estimativa das direções dominantes

[0108]Para a estimativa das direções dominantes a seguinte matriz de correlação é computada.

$$\mathbf{B}(l) := \frac{1}{L_B} \sum_{l'=0}^{L_B-1} \mathbf{C}(l - l') \mathbf{C}^T \quad (l - l') \in \mathbb{R}^{o \times o} . \quad (67)$$

[0109]O somatório em relação ao quadro atual l e $L-1$ quadros anteriores indica que a análise direcional é baseada em grupos de quadros sobrepostos longos com $L.B$ amostras, isto é, para cada quadro atual o conteúdo de quadros adjacentes é levado em consideração. Isso contribui

para a estabilidade da análise direcional por duas razões: quadros mais longos estão resultando em um número maior de observações, e as estimativas da direção são suavizada devido aos quadros em sobreposição.

[0110]Assumindo $f_s = 48\text{kHz}$ e $B = 1200$, um valor razoável para L é 4, correspondendo a uma duração geral de um quadro de 100ms.

[0111]Em seguida, uma decomposição de valor próprio da matriz de correlação $B(l)$ é determinada de acordo com $\mathbf{B}(l) = \mathbf{V}(l)\Lambda(l)\mathbf{V}^T(l)$, (68) em que a matriz $\mathbf{V}(l)$ é formadas de vetores próprios $\mathbf{V}_i(l)$, $1 \leq i \leq O$, como $\mathbf{V}(l) := [\mathbf{v}_1(l) \ \mathbf{v}_2(l) \ \dots \ \mathbf{v}_O(l)] \in \mathbb{R}^{O \times O}$ (69) e a matriz $\Lambda(l)$ é uma matriz diagonal com os valores próprios correspondentes $\lambda_i(l)$, $1 \leq i \leq O$, em sua diagonal:

$$\Lambda(l) := \text{diag}(\lambda_1(l), \lambda_2(l), \dots, \lambda_O(l)) \in \mathbb{R}^{O \times O} . \quad (70)$$

[0112]Assume-se que os valores próprios são indexados em ordem não ascendente, isto é, $\lambda_1(l) \geq \lambda_2(l) \geq \dots \geq \lambda_O(l)$. (71) Depois disso, o conjunto do índice $\{1, \dots, \tilde{J}(l)\}$ de valores próprios dominantes é computado. Uma possibilidade para administrar isso é definir uma razão de energia direcional-ambiental da largura de banda mínima desejada DAR_{MIN} e então determinar $\tilde{J}(l)$ de modo que

$$10\log_{10}\left(\frac{\lambda_i(l)}{\lambda_1(l)}\right) \geq -\text{DAR}_{\text{MIN}} \quad \forall i \leq \tilde{J}(l) \quad \text{and} \quad 10\log_{10}\left(\frac{\lambda_i(l)}{\lambda_1(l)}\right) > -\text{DAR}_{\text{MIN}} \\ \text{para } i = \tilde{J}(l) + 1 . \quad (72)$$

[0113]Uma escolha razoável para DAR_{MIN} é 15dB. O número de valores próprios dominantes é ainda mais restringido para não ser maior que D a fim de se concentrar em no máximo D direções dominantes. Isso é obtido com a substituição do conjunto do índice

$$\{1, \dots, \tilde{J}(l)\} \text{ by } \{1, \dots, J(l)\}, \text{ where } J(l) := \max(\tilde{J}(l), D) . \quad (73)$$

[0114]Em seguida, a aproximação de $J(l)$ classe de $\mathbf{B}(l)$ é obtida por $\mathbf{B}_J(l) := \mathbf{V}_J(l)\Lambda_J(l)\mathbf{V}_J^T(l)$, onde (74)

$$\mathbf{V}_J(l) := [\mathbf{v}_1(l) \quad \mathbf{v}_2(l) \quad \dots \quad \mathbf{v}_{J(l)}(l)] \in \mathbb{R}^{O \times J(l)} \quad , \quad (75)$$

$$\mathbf{\Lambda}_J(l) := \text{diag}(\lambda_1(l), \lambda_2(l), \dots, \lambda_{J(l)}(l)) \in \mathbb{R}^{J(l) \times J(l)} \quad . \quad (76)$$

[0115] Esta matriz deve conter as contribuições dos componentes direcionais dominantes para $\mathbf{B}(l)$.

Depois disso, computa-se o vetor

$$\sigma^2(l) := \text{diag}(\mathbf{\Xi}^T \mathbf{B}_J(l) \mathbf{\Xi}) \in \mathbb{R}^Q \quad (77)$$

$$= (\mathbf{S}_1^T \mathbf{B}_J(l) \mathbf{S}_1, \dots, \mathbf{S}_Q^T \mathbf{B}_J(l) \mathbf{S}_Q)^T \quad (78)$$

onde $\mathbf{\Xi}$ denota uma matriz de modo em relação a um número elevado de direções de teste quase igualmente distribuídas $\Omega_q := (\theta_q, \phi_q)$, $1 \leq q \leq Q$, onde $\theta_q \in [0, \pi]$ denota o ângulo de inclinação $\theta \in [0, \pi]$ medido desde o eixo polar z e $\phi_q \in [-\pi, \pi[$ denota o ângulo do azimute medido no plano x=y desde o eixo x.

[0116] A matriz de modo $\mathbf{\Xi}$ é definida por $\mathbf{\Xi} := [\mathbf{S}_1 \quad \mathbf{S}_2 \quad \dots \quad \mathbf{S}_Q] \in \mathbb{R}^{O \times Q}$ (79) com $\mathbf{S}_q := [S_0^0(\Omega_q), S_1^{-1}(\Omega_q), S_1^0(\Omega_q), S_1^{-1}(\Omega_q), S_2^{-2}(\Omega_q), \dots, S_N^N(\Omega_q)]^T$ (80)

para $1 \leq q \leq Q$.

[0117] Os $\sigma_q^2(l)$ elementos de $\sigma^2(l)$ são aproximações das energias de ondas planas, correspondendo aos sinais direcionais dominantes, partindo das direções em colisão. A explicação teórica para isso é fornecida na seção abaixo intitulada Explicação do algoritmo de busca da direção.

[0118] A partir de $\sigma^2(l)$, um número $\tilde{D}(l)$ de direções dominantes $\Omega_{\text{CURRDOM}, \tilde{d}}(l)$,

[0119] $1 \leq \tilde{d} \leq \tilde{D}(l)$, para a determinação dos componentes do sinal direcional é computado. O número de direções dominantes está assim limitado a satisfazer $\tilde{D}(l) \leq D$ a fim de garantir uma taxa de dados constante. No entanto, se permitida uma taxa de dados variável, o número de direções dominantes pode ser adaptado para a cena sonora corrente.

[0120] Uma possibilidade de computar as $\tilde{D}(l)$ direções dominantes é estabelecer a primeira direção dominante com a energia máxima, isto é,

$\Omega_{\text{CURRDOM},1}(l) = \Omega_{q_1}$ com $q_1 := \text{argmax}_{q \in \mathcal{M}_1} \sigma_q^2(l)$ e $\mathcal{M}_1 := \{1, 2, \dots, Q\}$. Assumindo que o máximo da energia é criado por um sinal direcional dominante, e considerando o fato de que o uso de uma representação de AOS de ordem finita N resulta em uma dispersão espacial dos sinais direcionais (cf. o já mencionado artigo "Plane-wave decomposition..."), é possível concluir que nas proximidades direcionais de $\Omega_{\text{CURRDOM},1}(l)$ deve haver componentes de energia pertencentes ao mesmo sinal direcional. Como a dispersão do sinal espacial pode ser expressa pela função $v_N(\Theta_{q,q_1})$ (vide a equação (38)), onde $\Theta_{q,q_1} := \angle(\Omega_q, \Omega_{q_1})$ denota o ângulo entre Ω_q e $\Omega_{\text{CURRDOM},1}(l)$, a energia pertencente ao sinal direcional declina de acordo com $v_N^2(\Theta_{q,q_1})$. Portanto, é razoável excluir todas as direções nas proximidades direcionais de Ω_{q_1} com $\Theta_{q,1} \leq \Theta_{\text{MIN}}$ para a busca de outras direções dominantes. A distância Θ_{MIN} pode ser escolhida como o primeiro zero de $v_N(x)$, que é aproximadamente dado por $\frac{\pi}{N}$ para $N \geq 4$. A segunda direção dominante é então estabelecida para isso com a energia máxima nas direções remanescentes $\Omega_q \in \mathcal{M}_2$ com $\mathcal{M}_2 := \{q \in \mathcal{M}_1 \mid \Theta_{q,1} > \Theta_{\text{MIN}}\}$. As direções dominantes remanescentes são determinadas de maneira análoga.

[0121] O número $\tilde{D}(l)$ de direções dominantes pode ser determinado considerando as energias $\sigma_{q_{\tilde{D}}}(l)$ atribuídas às direções dominantes individuais $\Omega_{q_{\tilde{D}}}$ e usando o caso em que a razão $\sigma_{q_1}^2(l)/\sigma_{q_{\tilde{D}}}(l)$ excede o valor de uma razão direcional-ambiental desejada DAR_{MIN} . Isso significa que $\tilde{D}(l)$ satisfaz

$$10 \log_{10} \left(\frac{\sigma_{q_1}^2(l)}{\sigma_{q_{\tilde{D}(l)}}^2(l)} \right) \leq \text{DAR}_{\text{MIN}} \wedge \left[10 \log_{10} \left(\frac{\sigma_{q_1}^2(l)}{\sigma_{q_{\tilde{D}(l)+1}}^2(l)} \right) > \text{DAR}_{\text{MIN}} \vee \tilde{D}(l) = D \right]. \quad (81)$$

[0122] O processamento geral para a computação de todas as direções dominantes pode ser realizada da seguinte maneira:

Algoritmo 1 Busca de direções dominantes dada a distribuição de energia na esfera

PowerFlag = verdadeiro

$$\tilde{d} = 1$$

$$\mathcal{M}_1 = \{1, 2, \dots, Q\}$$

repetir

$$q_{\tilde{d}} = \operatorname{argmax}_{q \in \mathcal{M}_{\tilde{d}}} \sigma_q^2(l)$$

$$\text{Se } \left[\tilde{d} > 1 \wedge 10 \log_{10} \left(\frac{\sigma_{q_1}^2(l)}{\sigma_{q_{\tilde{d}}(l)}^2} \right) > \text{DAR}_{\text{MIN}} \right] \text{então}$$

PowerFlag = falso

Do contrário

$$\Omega_{\text{CURRDOM}, \tilde{d}}(l) = \Omega_{q_{\tilde{d}}}$$

$$\mathcal{M}_{\tilde{d}+1} = \{q \in \mathcal{M}_{\tilde{d}} \mid \angle(\Omega_q, \Omega_{q_{\tilde{d}}}) > \theta_{\text{MIN}}\}$$

$$\tilde{d} = \tilde{d} + 1$$

Terminar se

até [$d > D \vee \text{PowerFlag} = \text{falso}$]

$$\tilde{D}(l) = \tilde{d} - 1$$

[0123] Em seguida, as direções $\Omega_{\text{CURRDOM}, \tilde{d}}(l)$, $1 \leq \tilde{d} \leq \tilde{D}(l)$,

obtidas no quadro atual são suavizadas com as direções dos quadros anteriores, resultando nas direções suavizadas $\bar{\Omega}_{\text{DOM}, d}(l)$, $1 \leq d \leq D$. Esta operação pode ser subdividida em duas partes sucessivas:

(a) As direções dominantes correntes $\Omega_{\text{CURRDOM}, \tilde{d}}(l)$, $1 \leq \tilde{d} \leq \tilde{D}(l)$, são atribuídas às direções suavizadas $\bar{\Omega}_{\text{DOM}, d}(l-1)$, $1 \leq d \leq D$ do quadro anterior. A função de atribuição $f_{\mathcal{A}, l}: \{1, \dots, \tilde{D}(l)\} \rightarrow \{1, \dots, D\}$ é determinada de modo que a soma dos ângulos entre direções atribuídas

$$\sum_{\tilde{d}=1}^{\tilde{D}(l)} \angle(\Omega_{\text{CURRDOM}, \tilde{d}}(l), \bar{\Omega}_{\text{DOM}, f_{\mathcal{A}, l}(\tilde{d})}(l-1)) \quad (82)$$

é minimizada. Esse problema de atribuição pode ser solucionado utilizando o algoritmo húngaro bastante conhecido, cf. H.W. Kuhn, "The Hungarian method for the attribution problem", Naval research logistics quarterly 2, nº 1-2, pp.83-97, 1955.

[0124] Os ângulos entre as direções correntes $\Omega_{\text{CURRDOM}, \tilde{d}}(l)$ e as direções inativas (vide abaixo para explicação do termo "direção inativa") do quadro anterior $\bar{\Omega}_{\text{DOM}, d}(l-1)$ são estabelecidas para $2\theta_{\text{MIN}}$. Esta operação

tem como efeito o fato de que as direções correntes $\Omega_{\text{CURRDOM},\tilde{d}}(l)$, que estão mais próximas do que $2\theta_{\text{MIN}}$ das direções previamente ativas $\bar{\Omega}_{\text{DOM},d}(l-1)$, estão tentando ser atribuídas a elas. Se a distância exceder $2\theta_{\text{MIN}}$, a direção corrente correspondente é assumida como pertencendo a um novo sinal, o que significa que estar favorecido para a atribuição de uma direção previamente inativa $\bar{\Omega}_{\text{DOM},d}(l-1)$.

[0125]Nota: quando se permite uma maior latência do algoritmo de compressão total, a atribuição das estimativas de direção sucessivas pode ser executada de forma mais robusta. Por exemplo, mudanças abruptas de direção podem ser mais identificadas sem misturá-las com os valores atípicos resultantes dos erros de estimativa.

(b) As direções suavizadas $\bar{\Omega}_{\text{DOM},d}(l-1)$, $1 \leq d \leq D$ são computadas utilizando a atribuição da etapa (a). A suavização se baseia na geometria esférica e não na geometria euclidiana. Para cada uma das direções dominantes correntes $\Omega_{\text{CURRDOM},\tilde{d}}(l)$, $1 \leq \tilde{d} \leq D$, a suavização é realizada ao longo do arco menor do maior círculo que cruza os dois pontos sobre a esfera, os quais são especificados pelas direções $\Omega_{\text{CURRDOM},\tilde{d}}(l)$ e $\bar{\Omega}_{\text{DOM},d}(l-1)$. Explicitamente, os ângulos do azimute e da inclinação são suavizados independentemente computando a média ponderada exponencialmente com um fator de suavização α_{Ω} . Para o ângulo de inclinação isso resulta na seguinte operação de suavização:

$$\bar{\theta}_{\text{DOM},f_{\mathcal{A},l}(\tilde{d})}(l) = (1 - \alpha_{\Omega}) \cdot \bar{\theta}_{\text{DOM},f_{\mathcal{A},l}(\tilde{d})}(l-1) + \alpha_{\Omega} \cdot \theta_{\text{DOM},\tilde{d}}(l) ,$$

$$1 \leq \tilde{d} \leq \tilde{D}(l) . \quad (83)$$

[0126]Para o ângulo do azimute, a suavização deve ser modificada para que se obtenha a suavização correta na transição de $\pi - \varepsilon$ para $-\pi$, $\varepsilon > 0$, e na transição na direção oposta. Isso pode ser levado em consideração computando primeiramente o módulo do ângulo da diferença 2π como

$$\Delta_{\phi,[0,2\pi],\tilde{d}}(l) := [\phi_{\text{DOM},\tilde{d}}(l) - \bar{\phi}_{\text{DOM},f_{\mathcal{A},l}(\tilde{d})}(l-1)] \bmod 2\pi , \quad (84) \text{ que}$$

é convertido para o intervalo $[-\pi, \pi[$ por

$$\Delta_{\phi,[-\pi, \pi], \tilde{d}}(l) := \begin{cases} \Delta_{\phi, [0, 2\pi], \tilde{d}}(l) & \Delta_{\phi, [0, 2\pi], \tilde{d}}(l) < \pi \\ \Delta_{\phi, [0, 2\pi], \tilde{d}}(l) - 2\pi & \text{para } \Delta_{\phi, [0, 2\pi], \tilde{d}}(l) \geq \pi \end{cases} \quad (85)$$

[0127] O módulo do ângulo do azimute dominante suavizado 2π é determinado como

$$\bar{\phi}_{\text{DOM}, [0, 2\pi], \tilde{d}}(l) := [\bar{\phi}_{\text{DOM}, \tilde{d}}(l-1) + \alpha_{\Omega} \cdot \Delta_{\phi, [-\pi, \pi], \tilde{d}}(l)] \bmod 2\pi \quad (86)$$

e é finalmente convertido para recair no intervalo $[-\pi, \pi[$ por

$$\bar{\phi}_{\text{DOM}, \tilde{d}}(l) = \begin{cases} \bar{\phi}_{\text{DOM}, [0, 2\pi], \tilde{d}}(l) & \bar{\phi}_{\text{DOM}, [0, 2\pi], \tilde{d}}(l) < \pi \\ \bar{\phi}_{\text{DOM}, [0, 2\pi], \tilde{d}}(l) - 2\pi & \text{para } \bar{\phi}_{\text{DOM}, [0, 2\pi], \tilde{d}}(l) \geq \pi \end{cases} \quad (87)$$

[0128] No caso $\tilde{D}(l) < D$, existem direções $\bar{\Omega}_{\text{DOM}, d}(l-1)$ do quadro anterior que não obtém uma direção dominante corrente atribuída. O conjunto do índice correspondente é denotado por

$$\mathcal{M}_{\text{NA}}(l) := \{1, \dots, D\} \setminus \{f_{\mathcal{A}, l}(\bar{d}) \mid 1 \leq \bar{d} \leq D\} \quad (88)$$

[0129] As respectivas direções são copiadas do último quadro, isto é, $\bar{\Omega}_{\text{DOM}, d}(l) = \bar{\Omega}_{\text{DOM}, d}(l-1)$ para $d \in \mathcal{M}_{\text{NA}}(l)$. (89)

[0130] As direções que não forem atribuídas a um número previamente definido de quadros são denominadas inativas.

[0131] Depois disso, o conjunto do índice das direções ativas denotado por $\mathcal{M}_{\text{ACT}}(l)$ é computado. Sua cardinalidade é denotada por $D_{\text{ACT}}(l) := |\mathcal{M}_{\text{ACT}}(l)|$.

[0132] Em seguida, todas as direções suavizadas são concatenadas em uma única matriz de direção como

$$\bar{\Omega}_{\text{DOM}}(l) := [\bar{\Omega}_{\text{DOM}, 1}(l) \quad \bar{\Omega}_{\text{DOM}, 2}(l) \quad \dots \quad \bar{\Omega}_{\text{DOM}, D}(l)] \quad (90)$$

Computação de sinais direcionais

[0133] A computação dos sinais direcionais se baseia na correspondência de modo. Em particular, realiza-se uma busca para aqueles sinais direcionais cuja representação de AOS resulta na melhor aproximação de dado sinal de AOS. Como as mudanças de direções entre quadros

sucessivos podem levar a uma descontinuidade dos sinais direcionais, as estimativas dos sinais direcionais para quadros em sobreposição podem ser computadas, seguida pela suavização dos resultados de quadros sucessivamente sobrepostos utilizando uma função de janela apropriada. A suavização, no entanto, introduz a latência de um único quadro.

[0134]A estimativa detalhada dos sinais direcionais é explicada a seguir:

[0135]Em primeiro lugar, a matriz de modo baseada nas direções suavizadas ativas é computada de acordo com

$$\mathbf{S}_{\text{ACT}}(l) := [\mathbf{S}_{\text{DOM},d_{\text{ACT},1}}(l) \ \mathbf{S}_{\text{DOM},d_{\text{ACT},2}}(l) \ \dots \ \mathbf{S}_{\text{DOM},d_{\text{ACT},D_{\text{ACT}}(l)}}(l)] \in \mathbb{R}^{O \times D_{\text{ACT}}(l)} \quad (91)$$

Com $\mathbf{S}_{\text{DOM},d}(l) :=$

$$[S_0^0(\bar{\Omega}_{\text{DOM},d}(l)), S_1^{-1}(\bar{\Omega}_{\text{DOM},d}(l)), S_1^0(\bar{\Omega}_{\text{DOM},d}(l)), \dots, S_N^N(\bar{\Omega}_{\text{DOM},d}(l))]^T \in \mathbb{R}^O,$$

(92)

em que $d_{\text{ACT},j}$, $1 \leq j \leq D_{\text{ACT}}(l)$ denota os índices das direções ativas.

[0136]Em seguida, é computada uma matriz $\mathbf{X}_{\text{INST}}(l)$ que contém as estimativas não suavizadas de todos os sinais direcionais para o ($l-1$)-ésimo e (l)-ésimo quadro:

$$\mathbf{X}_{\text{INST}}(l) := [\mathbf{x}_{\text{INST}}(l, 1) \ \mathbf{x}_{\text{INST}}(l, 2) \ \dots \ \mathbf{x}_{\text{INST}}(l, 2B)] \in \mathbb{R}^{D \times 2B} \quad (93)$$

Com

$$\mathbf{x}_{\text{INST}}(l, j) = [x_{\text{INST},1}(l, j), x_{\text{INST},2}(l, j), \dots, x_{\text{INST},D}(l, j)]^T \in \mathbb{R}^D, 1 \leq j \leq 2B. \quad (94)$$

[0137]Isso é realizado em duas etapas. Na primeira etapa, as amostras do sinal direcional nas fileiras correspondentes às direções inativas são definidas em zero, isto é,

$$x_{\text{INST},d}(l, j) = 0 \quad \forall 1 \leq j \leq 2B, \text{ se } d \notin \mathcal{M}_{\text{ACT}}(l) \quad . \quad (95)$$

[0138]Na segunda etapa, as amostras do sinal direcional correspondentes às direções ativas são obtidas primeiramente organizando-as em uma matriz de acordo com

$$\mathbf{X}_{\text{INST,ACT}}(l) := \begin{bmatrix} x_{\text{INST},d_{\text{ACT},1}}(l, 1) & x_{\text{INST},d_{\text{ACT},1}}(l, 2B) \\ \vdots & \ddots \\ x_{\text{INST},d_{\text{ACT},D_{\text{ACT}}(l)}}(l, 1) & x_{\text{INST},d_{\text{ACT},D_{\text{ACT}}(l)}}(l, 2B) \end{bmatrix} . \quad (96)$$

[0139] Esta matriz é então computada de modo a minimizar a norma euclidiana do erro $\mathbf{\Xi}_{\text{ACT}}(l)\mathbf{X}_{\text{INST,ACT}}(l) - [\mathbf{C}(l-1) \quad \mathbf{C}(l)]$. (97)

A solução é dada por

$$\mathbf{X}_{\text{INST,ACT}}(l) = [\mathbf{\Xi}_{\text{ACT}}^T(l)\mathbf{\Xi}_{\text{ACT}}(l)]^{-1}\mathbf{\Xi}_{\text{ACT}}^T(l)[\mathbf{C}(l-1) \quad \mathbf{C}(l)] . \quad (98)$$

[0140] As estimativas dos sinais direcionais $x_{\text{INST},d}(l, j)$, $1 \leq d \leq D$, são colocadas em janela por uma função de janela apropriada $w(j)$:

$$x_{\text{INST,WIN},d}(l, j) := x_{\text{INST},d}(l, j) \cdot w(j), \quad 1 \leq j \leq 2B \quad (99)$$

[0141] Um exemplo para a função de janela é dado pela janela de Hamming periódica definida por

$$w(j) := \begin{cases} K_w \left[0.54 - 0.46 \cos\left(\frac{2\pi j}{2B+1}\right) \right] & \text{for } 1 \leq j \leq 2B \\ 0 & \text{else} \end{cases} , \quad (100)$$

onde K_w denota um fator de escalonamento que é determinado de modo que a soma das janelas deslocadas equivale a '1'. Os sinais direcionais suavizados para o $(l-1)$ -ésimo quadro são computados pela sobreposição apropriada de estimativas não suavizadas em janela de acordo com

$$x_d((l-1)B + j) = x_{\text{INST,WIN},d}(l-1, B + j) + x_{\text{INST,WIN},d}(l, j) . \quad (101)$$

[0142] As amostras de todos os sinais direcionais suavizados para $(l-1)$ -ésimo quadro são arranjadas na matriz $\mathbf{X}(l-1)$ como

$$\mathbf{X}(l-1) := [\mathbf{x}((l-1)B + 1) \quad \mathbf{x}((l-1)B + 2) \quad \dots \quad \mathbf{x}((l-1)B + B)] \in \mathbb{R}^{D \times B} \quad (102)$$

$$\text{com } \mathbf{x}(j) = [x_1(j), x_2(j), \dots, x_D(j)]^T \in \mathbb{R}^D . \quad (103)$$

Computação do componente AOS ambiental

[0143] O componente AOS ambiental $\mathbf{C}_A(l-1)$ é obtido subtraindo-se o componente AOS direcional total $\mathbf{C}_{\text{DIR}}(l-1)$ da representação de AOS total $\mathbf{C}(l-1)$ de acordo com

$$\mathbf{C}_A(l-1) := \mathbf{C}(l-1) - \mathbf{C}_{DIR}(l-1) \in \mathbb{R}^{O \times B} , \quad (104)$$

onde $\mathbf{C}_{DIR}(l-1)$ é determinado por

$$\begin{aligned} \mathbf{C}_{DIR}(l-1) &:= \mathbf{\Xi}_{DOM}(l-1) \begin{bmatrix} x_{INST,WIN,1}(l-1, B+1) & \dots & x_{INST,WIN,1}(l-1, 2B) \\ \vdots & \ddots & \vdots \\ x_{INST,WIN,D}(l-1, B+1) & \dots & x_{INST,WIN,D}(l-1, 2B) \end{bmatrix} \\ &+ \mathbf{\Xi}_{DOM}(l) \begin{bmatrix} x_{INST,WIN,1}(l, 1) & \dots & x_{INST,WIN,1}(l, B) \\ \vdots & \ddots & \vdots \\ x_{INST,WIN,D}(l, 1) & \dots & x_{INST,WIN,D}(l, B) \end{bmatrix} , \end{aligned} \quad (105)$$

e onde $\mathbf{\Xi}_{DOM}(l)$ denota a matriz de modo baseada em todas as direções suavizadas definidas por

$$\mathbf{\Xi}_{DOM}(l) := [\mathbf{S}_{DOM,1}(l) \quad \mathbf{S}_{DOM,2}(l) \quad \dots \quad \mathbf{S}_{DOM,D}(l)] \in \mathbb{R}^{O \times D} . \quad (106)$$

[0144] Como a computação do componente AOS direcional total também está baseada em uma suavização espacial dos componentes AOS direcionais totais instantâneos sucessivos em sobreposição, o componente AOS ambiental também é obtido com uma latência de um único quadro.

Redução da ordem para o componente AOS ambiental

[0145] Ao expressar $\mathbf{C}_A(l-1)$ através de seus componentes como

$$\mathbf{C}_A(l-1) = \begin{bmatrix} c_{0,A}^0((l-1)B+1) & \dots & c_{0,A}^0((l-1)B+B) \\ \vdots & \ddots & \vdots \\ c_{N,A}^N((l-1)B+1) & \dots & c_{N,A}^N((l-1)B+B) \end{bmatrix} , \quad (107)$$

a redução da ordem é realizada causando a queda de todos os coeficientes AOS $c_{n,A}^m(j)$ com $n > N_{RED}$: 108

$$\mathbf{C}_{A,RED}(l-1) := \begin{bmatrix} c_{0,A}^0((l-1)B+1) & \dots & c_{0,A}^0((l-1)B+B) \\ \vdots & \ddots & \vdots \\ c_{N_{RED}}^0((l-1)B+1) & \dots & c_{N_{RED}}^0((l-1)B+B) \end{bmatrix} , \quad (108)$$

[0146] Transformada Harmônica Esférica para o componente AOS ambiental

[0147] A Transformada Harmônica Esférica é realizada pela multiplicação do componente AOS ambiental de ordem reduzida $\mathbf{C}_{A,RED}(l)$ com a inversa da matriz de modo

$$\mathbf{\Xi}_A := [\mathbf{S}_{A,1} \quad \mathbf{S}_{A,2} \quad \dots \quad \mathbf{S}_{A,O_{RED}}] \in \mathbb{R}^{O_{RED} \times O_{RED}} \quad (109)$$

$$\text{com } \mathbf{S}_{A,d} := [S_0^0(\mathbf{\Omega}_{A,d}), S_1^{-1}(\mathbf{\Omega}_{A,d}), S_1^0(\mathbf{\Omega}_{A,d}), \dots, S_{N_{RED}}^{N_{RED}}(\mathbf{\Omega}_{A,d})]^T \in \mathbb{R}^{O_{RED}} , \quad (110)$$

baseada em O_{RED} que são as direções uniformemente distribuídas $\Omega_{A,d}$,

$$1 \leq d \leq O_{RED} : \mathbf{W}_{A,RED}(l) = (\mathbf{\Xi}_A)^{-1} \mathbf{C}_{A,RED}(l) \quad \cdot (111)$$

Descompressão

Transformada Inversa de Harmônica Esférica

[0148] Os sinais do domínio espacial perceptivamente descomprimidos $\widehat{\mathbf{W}}_{A,RED}(l)$ são transformados para uma representação do domínio AOS $\widehat{\mathbf{C}}_{A,RED}(l)$ da ordem N_{RED} por meio de um Transformada Inversa de Harmônicas Esféricas por

$$\widehat{\mathbf{C}}_{A,RED}(l) = \mathbf{\Xi}_A \widehat{\mathbf{W}}_{A,RED}(l) \quad \cdot (112)$$

Extensão de ordem

[0149] A ordem de Ambissônico da representação de AOS $\widehat{\mathbf{C}}_{A,RED}(l)$ é estendida para N através da anexação de zeros de acordo com

$$\widehat{\mathbf{C}}_A(l) := \begin{bmatrix} \widehat{\mathbf{C}}_{A,RED}(l) \\ \mathbf{0}_{(O-O_{RED}) \times B} \end{bmatrix} \in \mathbb{R}^{O \times B} \quad , \quad (113)$$

onde $\mathbf{0}_{m \times n}$ denota uma matriz zero com m fileiras e n colunas

Composição dos coeficientes AOS

[0150] Os coeficientes de AOS descomprimidos finais são adicionalmente formados pelo componente AOS direcional e ambiental de acordo com $\widehat{\mathbf{C}}(l-1) := \widehat{\mathbf{C}}_A(l-1) + \widehat{\mathbf{C}}_{DIR}(l-1) \quad \cdot (114)$

[0151] Neste estágio, novamente uma latência de um único quadro é introduzida para permitir que o componente AOS direcional seja computado com base na suavização espacial. Fazendo isso, potenciais descontinuidades indesejadas no componente direcional do campo sonoro resultante das mudanças das direções entre quadros sucessivos são evitadas.

[0152] Para computar o componente AOS direcional suavizado, dois quadros sucessivos contendo as estimativas de todos os sinais direcionais

individuais são concatenados em um único quadro longo como $\hat{\mathbf{X}}_{\text{INST}}(l) := [\hat{\mathbf{X}}(l-1) \quad \hat{\mathbf{X}}(l)] \in \mathbb{R}^{D \times 2B}$. (115)

[0153]Cada excerto do sinal individual contido neste quadro longo é multiplicado por uma função de janela, por exemplo, como o da equação (100). Quando se expressa o quadro longo $\hat{\mathbf{X}}_{\text{INST}}(l)$ através de seus componentes por

$$\hat{\mathbf{X}}_{\text{INST}}(l) = \begin{bmatrix} \hat{x}_{\text{INST},1}(l,1) & \ddots & \hat{x}_{\text{INST},1}(l,2B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{INST},D}(l,1) & \ddots & \hat{x}_{\text{INST},D}(l,2B) \end{bmatrix} , \quad (116)$$

a operação de colocação em janela pode ser formulada como a computação dos excertos do sinal em janela $\hat{x}_{\text{INST},\text{WIN},d}(l,j)$, $1 \leq d \leq D$, por

$$\hat{x}_{\text{INST},\text{WIN},d}(l,j) = \hat{x}_{\text{INST},d}(l,j) \cdot w(j), \quad 1 \leq j \leq 2B, \quad 1 \leq d \leq D . \quad (117)$$

[0154]Finalmente, o componente AOS direcional total $C_{\text{DIR}}(l-1)$ é obtido codificando todos os excertos do sinal direcional em janela nas direções apropriadas e posicionando-os em sobreposição:

$$\begin{aligned} \hat{C}_{\text{DIR}}(l-1) &= \mathbf{E}_{\text{DOM}}(l-1) \begin{bmatrix} \hat{x}_{\text{INST},\text{WIN},1}(l-1, B+1) & \ddots & \hat{x}_{\text{INST},\text{WIN},1}(l-1, 2B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{INST},\text{WIN},D}(l-1, B+1) & \ddots & \hat{x}_{\text{INST},\text{WIN},D}(l-1, 2B) \end{bmatrix} \\ &+ \mathbf{E}_{\text{DOM}}(l) \begin{bmatrix} \hat{x}_{\text{INST},\text{WIN},1}(l, 1) & \ddots & \hat{x}_{\text{INST},\text{WIN},1}(l, B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{INST},\text{WIN},D}(l, 1) & \ddots & \hat{x}_{\text{INST},\text{WIN},D}(l, B) \end{bmatrix} . \end{aligned} \quad (118)$$

Explicação do algoritmo de busca de direção

[0155]A seguir, o motivo é explicado após o processamento de busca de direção descrito na seção intitulada Estimativa de direções dominantes. A explicação está fundada em algumas assunções que inicialmente são definidas.

Assunções

[0156]Assume-se que vetor dos coeficientes de AOS $c(j)$, que em geral está relacionado à função da densidade de amplitude do domínio do

tempo $d(j, \Omega)$ através de

$$\mathbf{c}(j) = \int_{S^2} d(j, \Omega) \mathbf{S}(\Omega) d\Omega , \quad (119)$$

obedece ao modelo abaixo:

$$\mathbf{c}(j) = \sum_{i=1}^I x_i(j) \mathbf{S}(\Omega_{x_i}(l)) + \mathbf{c}_A(j) \quad \text{for } lB + 1 \leq j \leq (l + 1)B . \quad (120)$$

[0157]Este modelo declara que o vetor dos coeficientes de AOS $c(j)$,

por um lado, é criado pelos / sinais da fonte direcional dominante $x_i(j)$, $1 \leq i \leq I$, partindo das direções $\Omega_{x_i}(l)$ no l -ésimo quadro. Em particular, assume-se que as direções são fixas pela duração de um único quadro. Assume-se que o número de sinais da fonte dominante / é distintamente menor que o número total de coeficientes AOS 0. Ademais, assume-se que o comprimento do quadro B é distintamente maior que 0. Por outro lado, o vetor $c(j)$ consiste em um componente residual $c_A(j)$, que pode ser considerado como representando o campo sonoro ambiental idealmente isotrópico.

[0158]Assume-se que os componentes do vetor do coeficiente AOS individuais possuam as seguintes propriedades:

- Os sinais da fonte dominante presumidamente possuem média zero, isto é, $\sum_{j=lB+1}^{(l+1)B} x_i(j) \approx 0 \quad \forall 1 \leq i \leq I$, (121)

e presumidamente não possuem correlação entre si, isto é,

$$\frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) x_{i'}(j) \approx \delta_{i-i'} \bar{\sigma}_{x_i}^2(l) \quad \forall 1 \leq i, i' \leq I \quad (122)$$

com $\bar{\sigma}_{x_i}^2(l)$ denotando a energia média do i -ésimo sinal para o l -ésimo quadro.

- Os sinais da fonte dominante presumidamente não possuem correlação com o componente ambiental do vetor do coeficiente AOS, isto é,

$$\frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) \mathbf{c}_A(j) \approx \mathbf{0} \quad \forall 1 \leq i \leq I \quad (123)$$

- O vetor do componente AOS ambiental presumidamente é uma média zero e presumidamente possui a matriz de covariância

$$\Sigma_A(l) := \frac{1}{B} \sum_{j=lb+1}^{(l+1)B} \mathbf{c}_A(j) \mathbf{c}_A^T(j) \quad . \quad (124)$$

- A razão da energia direcional/ambiente $DAR(l)$ de cada quadro Z ,
 $DAR(l) := 10 \log_{10} \left[\frac{\max_{1 \leq i \leq l} \bar{\sigma}_{x_i}^2(l)}{\|\Sigma_A(l)\|^2} \right]$, presumidamente que é aqui definida por é maior que um valor desejado previamente definido DAR_{MIN} , isto é,
 $DAR(l) \geq DAR_{MIN}$. (126)

Explicação da busca de direção

[0159] Para a explicação, considera-se o caso em que a matriz de correlação $B(Z)$ (vide a equação (67)) é computada em função somente das amostras do l -ésimo quadro sem considerar as amostras dos $L-1$ quadros anteriores. Esta operação corresponde à configuração $L = l$. Consequentemente, a matriz de correlação pode ser expressa por
 $B(l) = \frac{1}{B} \mathbf{C}(l) \mathbf{C}^T(l) \quad (127)$
 $= \frac{1}{B} \sum_{j=lb+1}^{(l+1)B} \mathbf{c}(j) \mathbf{c}^T(j) \quad . \quad (128)$

[0160] Substituindo a assunção do modelo na equação (120) para a equação (128) e utilizando as equações (122) e (123) e a definição na equação (124), a matriz de correlação $B(l)$ pode ser aproximada como

$$\begin{aligned} (129) \quad B(l) &= \frac{1}{B} \sum_{j=lb+1}^{(l+1)B} [\sum_{i=1}^l x_i(j) \mathbf{S}(\Omega_{x_i}(l)) + \mathbf{c}_A(j)] [\sum_{i=1}^l x_i(j) \mathbf{S}(\Omega_{x_i}(l)) + \mathbf{c}_A(j)]^T \\ &= \sum_{i=1}^l \sum_{i'=1}^l \mathbf{S}(\Omega_{x_i}(l)) \mathbf{S}^T(\Omega_{x_{i'}}(l)) \frac{1}{B} \sum_{j=lb+1}^{(l+1)B} x_i(j) x_{i'}(j) \\ &\quad + \sum_{i=1}^l \mathbf{S}(\Omega_{x_i}(l)) \frac{1}{B} \sum_{j=lb+1}^{(l+1)B} x_i(j) \mathbf{c}_A^T(j) + \sum_{i=1}^l \frac{1}{B} \sum_{j=lb+1}^{(l+1)B} x_i(j) \mathbf{c}_A(j) \mathbf{S}^T(\Omega_{x_i}(l)) \\ &\quad + \frac{1}{B} \sum_{j=lb+1}^{(l+1)B} \mathbf{c}_A(j) \mathbf{c}_A^T(j) \end{aligned} \quad (130)$$

$$\approx \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) \mathbf{S}(\Omega_{x_i}(l)) \mathbf{S}^T(\Omega_{x_i}(l)) + \Sigma_A(l) . \quad (131)$$

[0161] A partir da equação (131) pode-se observar que $B(l)$ consiste aproximadamente de dois componentes aditivos atribuíveis ao componente AOS direcional e ambiental. Sua aproximação da classe $\mathcal{I}(l)$ fornece uma aproximação do componente AOS direcional, isto é,
 $B_{\mathcal{I}}(l) \approx \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) \mathbf{S}(\Omega_{x_i}(l)) \mathbf{S}^T(\Omega_{x_i}(l))$, (132) que acompanha a equação (126) na razão de energia direcional-ambiental.

[0162] No entanto, cabe destacar que alguma parte de $\Sigma_A(l)$ inevitavelmente vazará para $\mathbf{B}_J(l)$, já que $\Sigma_A(l)$ de maneira geral possui classe integral e, portanto, os subespaços abrangidos pelas colunas das matrizes $\sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) \mathbf{S}(\Omega_{x_i}(l)) \mathbf{S}^T(\Omega_{x_i}(l))$ e $\Sigma_A(l)$ não são ortogonais entre si. Com a equação (132), o vetor $\sigma^2(l)$ na equação (77), que é utilizada para a busca das direções dominantes, pode ser expressa por $\sigma^2(l) = \text{diag}(\mathbf{E}^T \mathbf{B}_J(l) \mathbf{E})$

(133)

$$= \text{diag} \begin{pmatrix} \mathbf{S}^T (\Omega_1) \mathbf{B}_J(l) \mathbf{S}(\Omega_1) & \mathbf{S}^T (\Omega_1) \mathbf{B}_J(l) \mathbf{S}(\Omega_Q) \\ \vdots & \vdots \\ \mathbf{S}^T (\Omega_Q) \mathbf{B}_J(l) \mathbf{S}(\Omega_1) & \mathbf{S}^T (\Omega_Q) \mathbf{B}_J(l) \mathbf{S}(\Omega_Q) \end{pmatrix}$$

$$\text{diag} \begin{pmatrix} \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\Omega_1, \Omega_{x_i})) & \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N(\angle(\Omega_1, \Omega_{x_i})) \\ \vdots & \vdots \\ \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N(\angle(\Omega_Q, \Omega_{x_i})) v_N(\angle(\Omega_{x_i}, \Omega_1)) & \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\Omega_Q, \Omega_{x_i})) \end{pmatrix}$$

[0163] Na equação (135) a seguinte propriedade da Harmônica Esférica mostrada na equação (47) foi utilizada:

$$\mathbf{S}^T(\Omega_q) \mathbf{S}(\Omega_{q'}) = v_N(\angle(\Omega_q, \Omega_{q'})) \quad (137)$$

[0164] A equação (136) mostra que os componentes de $\sigma^2(l)$ são aproximações das energias dos sinais originários das direções de teste Ω_q , $1 \leq q \leq Q$.

REIVINDICAÇÕES

1. Método para comprimir uma representação de sinal Ambissônico de Ordem Superior AOS ($\mathbf{c}(l)$), o dito método **CARACTERIZADO** pelo fato de que inclui:

- estimar (22) direções dominantes;
- decompor (23, 24) a representação de sinal AOS em um número de sinais direcionais dominantes ($\mathbf{x}(l)$) no domínio do tempo e informações direcionais correlatas ($\mathbf{\Omega}_{\text{DOM}}(l)$), e um componente ambiental residual no domínio AOS ($\mathbf{c}_A(l)$), em que o dito componente ambiental residual representa a diferença entre a dita representação de sinal AOS ($\mathbf{c}(l)$) e uma representação ($\mathbf{c}_{\text{DIR}}(l)$) dos ditos sinais direcionais dominantes ($\mathbf{x}(l)$);
- comprimir (25) o dito componente ambiental residual reduzindo sua ordem quando comparado à sua ordem original;
- transformar (26) o dito componente AOS ambiental residual ($\mathbf{c}_{\text{A,RED}}(l)$) de ordem reduzida para o domínio espacial;
- codificar perceptivamente (27) os ditos sinais direcionais dominantes e o dito componente AOS ambiental residual transformado.

2. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que vetores de entrada ($\mathbf{c}(j)$) de coeficientes de AOS são enquadrados (21) em quadros não sobrepostos ($\mathbf{c}(l)$), e em que uma duração de quadro é de **25ms**.

3. Método, de acordo com a reivindicação 1 ou 2, **CARACTERIZADO** pelo fato de que a dita estimativa de direções dominantes (22) é baseada em longos grupos de quadros sobrepostos, de modo que, para cada quadro atual, o conteúdo de quadros adjacentes é levado em consideração.

4. Método, de acordo com qualquer uma das reivindicações 1 a 3, **CARACTERIZADO** pelo fato de que, para a dita estimativa de direção dominante, uma distribuição da energia direcional dos componentes AOS

energeticamente dominantes é utilizada.

5. Método para descomprimir uma representação de sinal Ambissônico de Ordem Superior AOS ($\mathbf{C}(l)$) que foi comprimido pelo método conforme definido na reivindicação 1, o dito método **CARACTERIZADO** pelo fato de que inclui as etapas de:

- decodificar perceptivamente (31) os ditos sinais direcionais dominantes codificados perceptivamente ($\mathbf{\tilde{X}}(l)$) e o dito componente AOS ambiental residual transformado codificado perceptivamente ($\mathbf{\tilde{W}}_{A,RED}(l)$);

- transformar inversamente (32) o dito componente AOS ambiental residual transformado decodificado perceptivamente ($\mathbf{\tilde{W}}_{A,RED}(l)$) de modo a obter uma representação de domínio AOS ($\mathbf{\tilde{C}}_{A,RED}(l)$);

- executar (33) uma extensão de ordem do dito componente AOS ambiental residual transformado inversamente de modo a estabelecer um componente AOS ambiental de ordem original ($\mathbf{\hat{C}}_A(l)$);

- compor (34) os ditos sinais direcionais dominantes decodificados perceptivamente ($\mathbf{\tilde{X}}(l)$), as ditas informações direcionais ($\mathbf{\bar{\Omega}}_{DOM}(l)$) e o dito componente AOS ambiental estendido de ordem original ($\mathbf{\hat{C}}_A(l)$) de modo a obter uma representação de sinal AOS ($\mathbf{\hat{C}}(l)$).

6. Aparelho para comprimir uma representação de sinal Ambissônico de Ordem Superior AOS ($\mathbf{C}(l)$), o dito aparelho **CARACTERIZADO** pelo fato de que inclui:

- meio (22) adaptado para estimar direções dominantes;
- meio (23, 24) adaptado para decompor a representação de sinal AOS em um número de sinais direcionais dominantes ($\mathbf{X}(l)$) no domínio do tempo e informações direcionais correlatas ($\mathbf{\bar{\Omega}}_{DOM}(l)$), e um componente ambiental residual no domínio AOS ($\mathbf{C}_A(l)$), em que o dito componente ambiental residual representa a diferença entre a dita representação de sinal

AOS ($C(l)$) e uma representação ($C_{DIR}(l)$) dos ditos sinais direcionais dominantes ($X(l)$);

- meio (25) adaptado para comprimir o dito componente ambiental residual reduzindo sua ordem quando comparado à sua ordem original;

- meio (26) adaptado para transformar o dito componente AOS ambiental residual ($C_{A,RED}(l)$) de ordem reduzida para o domínio espacial;

- meio (27) adaptado para codificar perceptivamente os ditos sinais direcionais dominantes e o dito componente AOS ambiental residual transformado.

7. Aparelho, de acordo com a reivindicação 6, **CARACTERIZADO** pelo fato de que vetores de entrada ($c(j)$) de coeficientes de AOS são enquadrados (21) em quadros não sobrepostos ($C(l)$), e em que uma duração de quadro é de **25ms**.

8. Aparelho, de acordo com a reivindicação 6 ou 7, **CARACTERIZADO** pelo fato de que a dita estimativa de direções dominantes (22) é baseada em longos grupos de quadros sobrepostos, de modo que, para cada quadro atual, o conteúdo de quadros adjacentes é levado em consideração.

9. Aparelho, de acordo com qualquer uma das reivindicações 6 a 8, **CARACTERIZADO** pelo fato de que, para a dita estimativa de direção dominante, uma distribuição de energia direcional dos componentes AOS energeticamente dominantes é utilizada.

10. Aparelho para descomprimir uma representação de sinal Ambissônico de Ordem Superior AOS ($C(l)$) que foi comprimido pelo método conforme definido na reivindicação 1, o dito aparelho **CARACTERIZADO** pelo fato de que inclui:

- meio (31) adaptado para decodificar perceptivamente os ditos sinais direcionais dominantes codificados perceptivamente ($X(l)$) e o dito componente AOS ambiental residual transformado codificado perceptivamente ($W_{A,RED}(l)$);

- meio (32) adaptado para transformar inversamente o dito componente AOS ambiental residual transformado decodificado perceptivamente ($\bar{W}_{A,RED}(l)$) de modo a obter uma representação de domínio AOS ($\bar{C}_{A,RED}(l)$);

- meio (33) adaptado para executar uma extensão de ordem do dito componente AOS ambiental residual transformado inversamente de modo a estabelecer um componente AOS ambiental de ordem original ($\bar{C}_A(l)$);

- meio (34) adaptado para compor os ditos sinais direcionais dominantes decodificados perceptivamente ($\bar{X}(l)$), as ditas informações direcionais ($\bar{\Omega}_{DOM}(l)$) e o dito componente AOS ambiental estendido de ordem original ($\bar{C}_A(l)$) de modo a obter uma representação de sinal AOS ($\bar{C}(l)$).

11. Representação de sinal AOS **CARACTERIZADA** por ser comprimida em conformidade com o método conforme definido em qualquer uma das reivindicações 1 a 4.

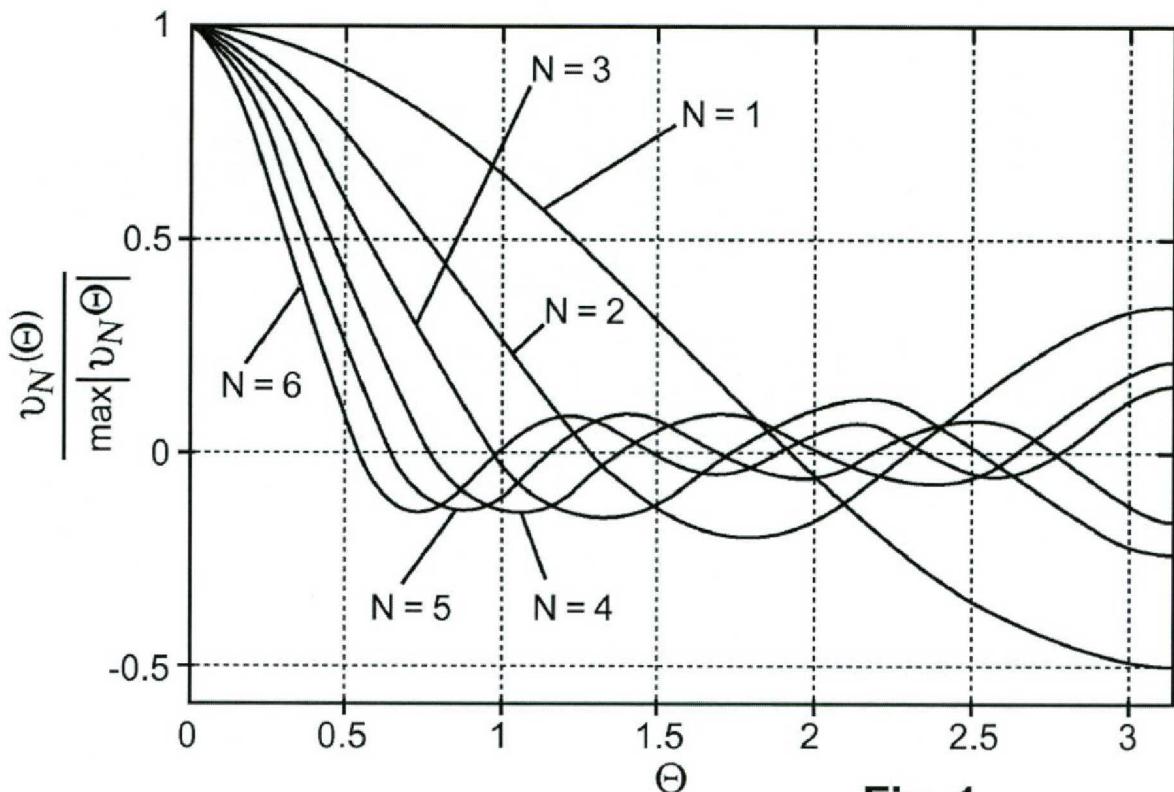


Fig. 1

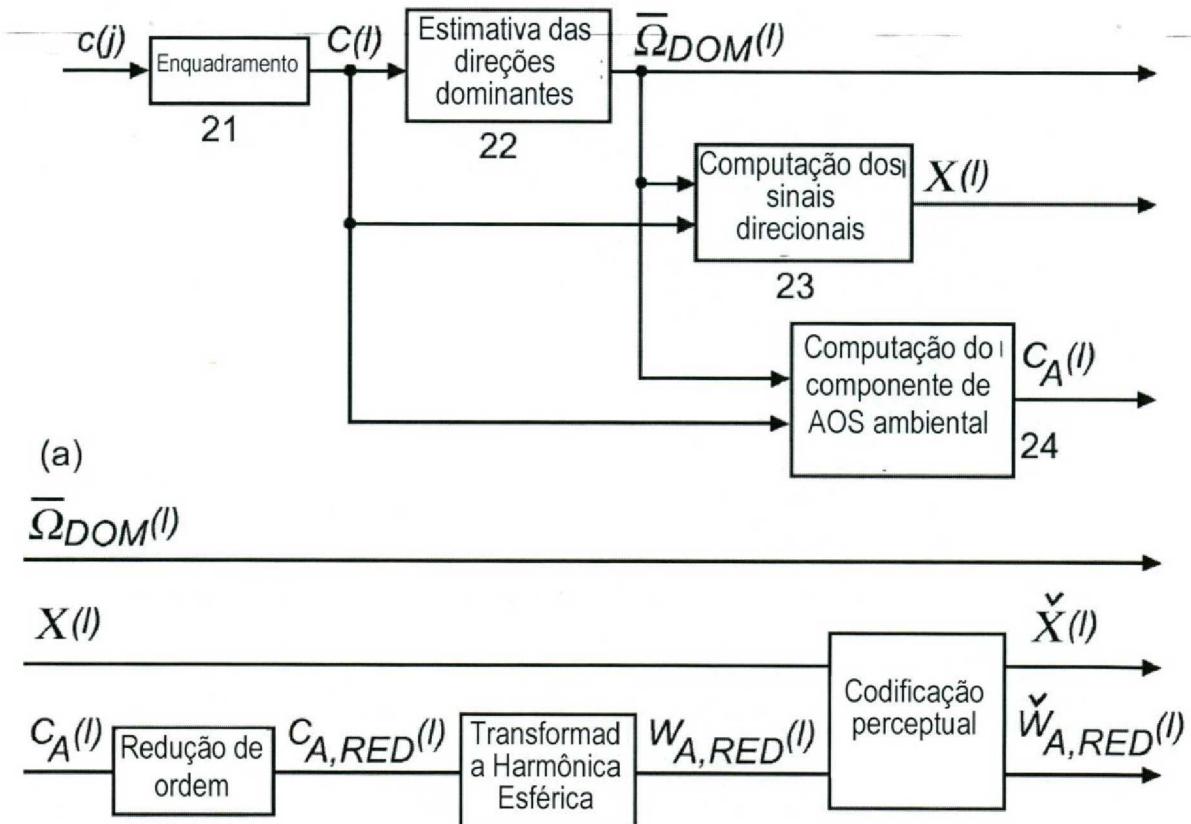


Fig. 2

