



República Federativa do Brasil  
Ministério do Desenvolvimento, Indústria  
e do Comércio Exterior  
Instituto Nacional da Propriedade Industrial.

(21) **PI0709276-8 A2**



(22) Data de Depósito: 08/03/2007  
(43) Data da Publicação: 12/07/2011  
(RPI 2114)

(51) *Int.Cl.:*  
H04S 3/02 2006.01

(54) Título: **PROCESSO E DISPOSITIVO DE ESPACIALIZAÇÃO SONORA BINAURAL EFICAZ NO DOMÍNIO TRANSFORMADO**

(30) Prioridade Unionista: 28/03/2006 FR 0602685

(73) Titular(es): France Telecom

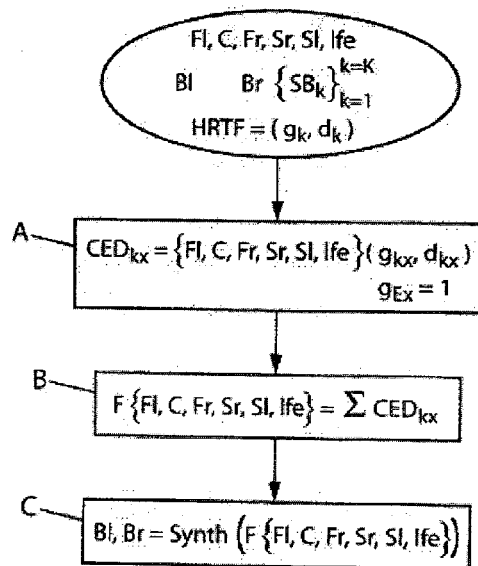
(72) Inventor(es): David Virette, Marc Emerit, Pierrick Philippe

(74) Procurador(es): Montauray Pimenta, Machado & Lioce S/C Ltda

(86) Pedido Internacional: PCT FR2007050894 de 08/03/2007

(87) Publicação Internacional: WO 2007/110519 de 04/10/2007

(57) **Resumo:** PROCESSO E DISPOSITIVO DE ESPACIALIZAÇÃO SONORA BINAURAL EFICAZ NO DOMÍNIO TRANSFORMADO. A invenção refere-se a um processo e um sistema de espacialização sonora de um primeiro conjunto de um número superior ou igual a um de canais áudio codificados em um número de sub-bandas de frequências (SB<sub>k</sub>) e decodificados em um domínio transformado (F<sub>1</sub>, C, Fr, Sr, S<sub>1</sub>, lfe) em um segundo conjunto de um número superior ou igual a dois (B<sub>1</sub>, Br) de canais sonoros no domínio temporal, a partir de filtros de modelação convertidos em um ganho e um retardo, aplicáveis no domínio transformado, efetuando: uma filtragem (A) por retardo do sinal de sub-banda por aplicação de pelo menos um ganho e de um retardo para gerar, a partir de cada um dos canais codificados espacialmente, uma componente equalizada e retardada, a adição (B) de um sub-conjunto de componentes equalizadas e retardadas para criar um número de sinais filtrados no domínio transformado correspondente ao número superior ou igual a dois; a síntese (C) de cada um dos sinais filtrados no domínio transformado por um filtro de síntese para obter o segundo conjunto de um número superior ou igual a dois de canais sonoros de restituição (B<sub>1</sub>, Br) no domínio temporal.



**"PROCESSO E DISPOSITIVO DE ESPACIALIZAÇÃO SONORA BINAURAL  
EFICAZ NO DOMÍNIO TRANSFORMADO"**

A invenção é relativa à espacialização, dita rendimento 3D, de sinais áudio comprimidos.

5 Uma operação desse tipo é, por exemplo, executada quando da descompressão de um sinal comprimido áudio 3D, por exemplo, representado em um certo número de canais, para um número de canais diferentes, dois, por exemplo, para permitir a restituição dos efeitos 3D áudio em um  
10 auscultador.

Assim, o termo "binaural" visa à restituição em um auscultador estereofônico de um sinal sonoro com, no entanto, efeitos de espacialização. A invenção não se limita, no entanto, à técnica precitada e se aplica,  
15 notadamente, a técnicas derivadas do "binaural", tais como as técnicas de restituição ditas técnicas TRANSAURAL®, isto é, em alto-falantes distantes. TRANSAURAL® é uma marca comercial depositada pela sociedade COOPER BAUCK CORPORATION. Tais técnicas podem então utilizar uma  
20 "anulação de diafonia" ("cross-talk cancellation" em inglês), a qual consiste em anular os caminhos acústicos cruzados, de maneira que um som, assim tratado depois emitido pelos alto-falantes, possa só ser percebido por apenas um dos dois ouvidos de um ouvinte.

25 Em consequência, a invenção é igualmente relativa à transmissão e à restituição de sinais áudio multicanais e à sua conversão para um dispositivo de restituição, transdutor, imposto pelo equipamento de um usuário. É, por exemplo, o caso para a restituição de uma cena sonora 5.1  
30 por um auscultador áudio, ou por um par de alto-falantes.

A invenção é igualmente relativa à restituição, no campo de um jogo ou gravação em vídeo, por exemplo, de

uma ou várias amostras sonoras armazenadas em arquivos, com vistas à sua espacialização.

Dentre as técnicas conhecidas no domínio da espacialização sonora binaural, diferentes abordagens foram  
5 propostas.

Em particular, a síntese binaural bicanal consiste, com referência à figura 1a, em filtrar o sinal das diferentes fontes sonoras  $S_i$  que se deseja posicionar, na restituição, em uma posição no espaço, por intermédio de  
10 funções de transferência acústicas esquerda HRTF-l e direita HRTF-r no domínio frequencial correspondente à direção apropriada, definida em coordenadas polares ( $\theta_1$ ,  $\phi_1$ ). As funções de transferência HRTF, para "Head Related Transfer Functions" em inglês, precitadas são as funções de  
15 transferência acústica da cabeça do ouvinte entre as posições do espaço e o conduto auditivo. Designa-se adicionalmente por "HRIR" para "Head Related Impulse Response" sua forma temporal. Essas funções podem adicionalmente comportar um efeito de sala.

20 Obtém-se, para cada fonte sonora  $S_i$ , dois sinais esquerdo e direito que são então adicionados aos sinais esquerdo e direito provenientes da espacialização das outras fontes sonoras, para dar finalmente os sinais L e R difundidos nos ouvidos esquerdo e direito do ouvinte.

25 O número de filtros, ou funções de transferência, necessários é então  $2.N$  para uma síntese binaural estática e  $4.N$  para uma síntese binaural dinâmica,  $N$  designando o número de fontes sonora ou de fluxo áudio a espacializar.

30 *Trabalhos intitulados "A model of head-related transfer functions based on principal components analysis and minimum - phase reconstruction" conduzidos por D. Kistler e F. L. Wightman, publicados no J. Acoust. Soc. Am. 91(3): p 1637-1647 (1992) e por A. Kulkarni 1995 "IEEE*

ASSP Workshop on Applications of Signal Processing to Audio and Acoustics" IEEE catalog number: 95<sup>TH</sup>8144, permitiram verificar que as fases das HRTF podem se decompor na soma de dois termos, um correspondendo ao retardo interaural e o  
5 outro igual à fase mínima associada ao módulo da HRTF.

Assim, para uma função de transferência HRTF expressa sob a forma:

$$H(f) = |H(f)|e^{-j\varphi(f)}$$

$$\varphi(f) = \varphi_{\text{retardo}}(f) + \varphi_{\text{min}}(f)$$

10  $\varphi_{\text{retardo}}(f) = 2\pi f\tau$  corresponde ao retardo interaural;

$$\varphi_{\text{min}}(f) = H(\log(|H(f)|))$$

A implementação dos filtros binaurais se faz, em geral, sob a forma de dois filtros com fase mínima e de um  
15 retardo puro, correspondendo à diferença dos retardos esquerdo e direito aplicados no ouvido mais afastado da fonte. Esse retardo é, em geral, implementado com o auxílio de uma linha de retardo.

O filtro de fase mínima é um filtro de resposta  
20 impulsional finita e pode ser executado no domínio temporal ou frequencial. Filtros de resposta impulsional infinita podem ser buscados para aproximar o módulo dos filtros HRTF de fase mínima.

No que diz respeito à binauralização, situa-se,  
25 com referência à figura 1b, no campo não limitativo de uma cena sonora espacializada no modo 5.1, com vistas à restituição desta no auscultador áudio de um ser humano HB.

Cinco alto-falantes C: Centro, Lf: Left front, Rf: Right front, Sl: Surround left, Sr: Surround right,  
30 produzem, cada um, um som que é percebido pelo ser humano HB nos dois receptores que são seus ouvidos. Modelam-se as transformações sofridas pelo som por uma função de

filtragem representando a modificação que este som sofre quando de sua propagação entre o alto-falante que restitui este som e um ouvido dado.

Em particular, o som emanando do alto-falante  $L_f$  afeta o ouvido esquerdo LE através de um filtro HRTF A, mas este mesmo som alcança o ouvido direito RE modificado por um filtro HRTF B.

A posição dos alto-falantes em relação ao indivíduo HB precitados pode ser simétrica ou não.

Cada ouvido recebe então a contribuição dos 5 alto-falantes sob a forma modelada a seguir:

Ouvido esquerdo LE:  $B_l = A_{L_f} + CC + B_{R_f} + D_{S_l} + E_{S_r}$ ,

Ouvido direito RE:  $B_r = A_{R_f} + CC + B_{L_f} + D_{S_r} + E_{S_l}$ ,

onde  $B_l$  é o sinal binauralizado para o ouvido esquerdo LE e  $B_r$  é o sinal binauralizado para o ouvido direito RE.

Os filtros A, B, C, D e E são modelados, o mais freqüente, por filtros digitais lineares e são necessárias, então, na configuração representada na figura 1b, 10 funções de filtragem a aplicar, as quais podem ser reduzidas a 5, levando em conta simetrias.

De maneira conhecida como tal, as operações de filtragem precitadas podem ser realizadas no domínio frequencial, por exemplo, graças a uma convolução rápida executada no domínio de Fourier. Utiliza-se, então, uma transformada de Fourier rápida FFT, para "Fast Fourier Transform" em inglês, para executar a binauralização de forma eficaz.

Os filtros HRTF A, B, C, D e E podem ser simplificados sob a forma de um equalizador de frequência e de um retardo. O filtro HRTF A pode ser realizado sob a

forma de um simples equalizador, visto que se trata de uma trajetória direta, enquanto que o filtro HRTF B inclui um retardo suplementar. De maneira clássica, os filtros HRTF podem ser decompostos em um filtro de fase mínima e um  
5 retardo puro. O retardo para o ouvido mais próximo da fonte pode ser tomado igual a zero.

A operação de reconstrução por decodificação espacial de uma cena sonora 3D áudio, a partir de um número reduzido de canais transmitidos, tal como representada na  
10 figura 1c, é igualmente conhecida do estado da técnica. A configuração representada na figura 1c é aquela relativa à decodificação de uma via sonora codificada dispondo de parâmetros de localização no domínio frequencial, a fim de reconstruir uma cena sonora espacializada 5.1.

A reconstrução precitada é efetuada por um decodificador espacial por sub-bandas frequenciais, tal como representado na figura 1c. O sinal áudio codificado m sofre 5 etapas de tratamento de espacialização, que são comandadas por parâmetros ou coeficientes complexos de  
20 espacialização CLD e ICC calculados pelo codificador que permitem, pelo viés de operações de descorrelação e de correção de ganho, reconstruir de forma realista a cena sonora composta de seis canais, os cinco canais representados na figura 1b, aos quais é acrescentado um  
25 canal de baixa freqüência lfe.

Quando se deseja proceder a uma binauralização dos canais sonoros provenientes de um decodificador espacial tal como representado na figura 1c, se está, de fato, restrito, no momento atual, a implementar um  
30 tratamento de acordo com o esquema representado na figura 1d.

Com referência ao esquema precitado, parecia necessário realizar a transformação dos canais sonoros dos

quais se dispõe no domínio temporal, antes de proceder à binauralização do sinal. Esta operação de retorno no domínio temporal é simbolizada pelos blocos sintetizadores "Sint" que executam a operação de transformação frequência-tempo para cada um dos canais provenientes do decodificador espacial (SD). A filtragem por filtros HRTF pode, em seguida, ser realizada pelos filtros A, B, C, D, E, com ou sem aplicação do esquema equalizado, correspondendo a uma filtragem clássica.

10 Uma variante de binauralização dos canais sonoros de um decodificador espacial pode consistir igualmente, assim como representado na figura 1e, em converter cada canal sonoro entregue pelo decodificador áudio no domínio temporal por um sintetizador "Sint" e depois em executar a  
15 operação de decodificação espacial e de binauralização, ou espacialização, no domínio frequencial de Fourier, após transformação por FFT.

Nesta hipótese, cada módulo OTT correspondente a uma matriz de coeficientes de decodificação deve então ser  
20 convertido no domínio de Fourier, ao preço de uma aproximação, uma vez que as operações não são efetuadas no mesmo domínio. Além disso, a complexidade é ainda acrescida, uma vez que a operação de síntese "Sint" é seguida de três transformações FFT.

25 Assim, para binauralizar uma cena sonora proveniente de um decodificador espacial, não existe outra possibilidade senão realizar:

- ou 6 transformações tempo-frequência, se for desejado realizar a binauralização fora do decodificador  
30 espacial;

- ou uma operação de síntese seguida de 3 transformações de Fourier, FFT, se for desejado realizar a operação no domínio FFT.

A rigor, uma outra solução pode consistir em efetuar a filtragem HRTF diretamente no domínio das sub-bandas, assim como representado na figura 1f.

No entanto, nesta hipótese, as filtragens HRTF são complexas de realizar, uma vez que estas últimas impõem a utilização de filtros de sub-bandas, cujo comprimento mínimo é fixo e que devem levar em consideração o fenômeno de suavização espectral das sub-bandas.

A economia introduzida pela redução de operações de transformação é compensada negativamente pela explosão do número de operações necessárias para a filtragem, em razão da execução destas operações no domínio PQMF para Pseudo Quadrature Mirror Filter em inglês.

A presente invenção tem por objetivo remediar aos numerosos inconvenientes das técnicas anteriores precitadas de espacialização sonora das cenas áudio 3D, notadamente de transauralização ou de binauralização de cenas áudio 3D.

Em particular, um objetivo da presente invenção é a execução de uma filtragem específica de sinais ou canais áudio codificados espacialmente no domínio das sub-bandas frequenciais de uma decodificação espacial, a fim de limitar o número de transformações duas a duas, reduzindo as operações de filtragem ao mínimo, mas conservando uma boa qualidade de espacialização fonte, notadamente em transauralização ou binauralização.

De acordo com aspecto particularmente remarcável da presente invenção, a execução da filtragem específica precitada se apóia na colocação sob a forma equalizador-retardo dos filtros de espacialização, transaural ou binaural, para uma aplicação direta de uma filtragem por equalização-retardo no domínio das sub-bandas.

Um outro objetivo da presente invenção é a obtenção de uma qualidade de rendimento 3D muito próxima

daquela obtida a partir de filtros de modelação tais como filtros HRTF de origem, pela única adjunção de um tratamento espacial transaural de complexidade muito baixa, em seguida a uma decodificação espacial clássica no domínio transformado.

Um objetivo da presente invenção é, enfim, uma nova técnica de espacialização fonte aplicável não somente ao rendimento transaural ou binaural de um som monofônico, mas igualmente a vários sons monofônicos e notadamente aos canais múltiplos de sons estéreo 5.1, 6.1, 7.1, 8.1 ou superiores.

A presente invenção tem, assim, por objeto um processo de espacialização sonora de uma cena áudio comportando um primeiro conjunto compreendendo um número superior ou igual à unidade de canais áudio codificados espacialmente em um número de sub-bandas de frequências determinado, e decodificados em um domínio transformado, em um segundo conjunto compreendendo um número superior ou igual a dois de canais sonoros de restituição no domínio temporal, a partir de filtros de modelação da propagação acústica dos sinais áudio do primeiro conjunto de canais.

Em conformidade com a invenção, este processo é remarcável pelo fato de que, para cada filtro de modelação convertido sob a forma de pelo menos um ganho e de um retardo aplicáveis no domínio transformado, ele consiste em efetuar pelo menos, para cada sub-banda frequencial do domínio transformado:

- uma filtragem por equalização-retardo do sinal de sub-banda, por aplicação de um ganho respectivamente de um retardo no sinal de sub-banda, para gerar, a partir dos canais codificados espacialmente, uma componente equalizada e retardada de um valor determinado na sub-banda frequencial considerada,

- uma adição de um sub-conjunto de componentes equalizadas e retardadas, para criar um número de sinais filtrados no domínio transformado correspondente ao número do segundo conjunto, superior ou igual a dois, de canais sonoros de restituição no domínio temporal,

- uma síntese de cada um dos sinais filtrados no domínio transformado por um filtro de síntese, para obter o segundo conjunto de número superior ou igual a dois de sinais sonoros de restituição no domínio temporal.

O processo objeto da invenção é igualmente remarcável pelo fato de que a filtragem por equalização-retardo do sinal de sub-banda inclui pelo menos a aplicação de uma defasagem e, quando necessário, de um retardo puro por memorização, para uma pelo menos das sub-bandas de frequências.

O processo objeto da invenção é igualmente remarcável pelo fato de que inclui uma filtragem por equalização-retardo em um domínio transformado híbrido, comportando uma etapa suplementar de corte em frequência de sub-bandas suplementares, com ou sem decimação.

O processo objeto da invenção é, enfim, remarcável pelo fato de que, para converter cada filtro de modelação em um valor de ganho respectivamente de retardo no domínio transformado, ele consiste pelo menos em associar como valor de ganho a cada sub-banda um valor real definido como a média do módulo do filtro de modelação nesta sub-banda e em associar como valor de retardo a cada sub-banda um valor de retardo correspondente ao retardo de recepção entre o ouvido esquerdo e o ouvido direito para diferentes posições.

A presente invenção tem correlativamente por objeto um dispositivo de espacialização sonora de uma cena áudio comportando um primeiro conjunto compreendendo um

número, superior ou igual à unidade, de canais áudio codificados espacialmente em um número de sub-bandas de frequências determinado, e decodificados em um domínio transformado, em um segundo conjunto comportando um número superior ou igual a dois de canais sonoros de restituição no domínio temporal, a partir de filtros de modelação da propagação acústica dos sinais áudio do primeiro conjunto de canais.

Em conformidade com a invenção, este dispositivo é remarcável pelo fato de que, para cada sub-banda frequencial de um decodificador espacial no domínio transformado, este dispositivo compreende, além deste decodificador espacial:

- um módulo de filtragem por equalização-retardo do sinal de sub-banda por aplicação de pelo menos um ganho respectivamente de um retardo no sinal de sub-banda, para gerar, a partir de cada um dos canais áudio codificados espacialmente, uma componente equalizada e retardada de um valor de retardo determinado na sub-banda de frequências considerada,

- um módulo de adição de um sub-conjunto de componentes equalizadas e retardadas para criar um número de sinais filtrados no domínio transformado correspondendo ao número do segundo conjunto superior ou igual a dois dos canais sonoros de restituição no domínio temporal,

- um módulo de síntese de cada um dos sinais filtrados no domínio transformado para obter o segundo conjunto compreendendo um número superior ou igual a dois dos canais sonoros de restituição no domínio temporal.

O processo e o dispositivo objetos da invenção encontram aplicação na indústria eletrônica dos aparelhos de áudio e/ou vídeo com alta fidelidade, na indústria dos jogos de áudio-vídeo executados localmente ou em linha.

Eles serão melhor compreendidos pela leitura do relatório e pela observação dos desenhos logo a seguir, nos quais, além das figuras 1a a 1f relativas à técnica anterior,

5                   - a figura 2a representa um organograma ilustrativo das etapas de implementação do processo de espacialização sonora objeto da invenção;

                  - a figura 2b representa, a título ilustrativo, uma variante de implementação do processo objeto da  
10 invenção representado na figura 2a, obtido por criação de sub-bandas suplementares, na ausência de decimação;

                  - a figura 2c representa, a título ilustrativo, uma variante de implementação do processo objeto da invenção representado na figura 2a obtido por criação de  
15 sub-bandas suplementares, na presença de decimação;

                  - a figura 3a representa, a título ilustrativo, um estágio, para uma sub-banda de frequências de um decodificador espacial, de um dispositivo de espacialização sonora objetos da invenção;

20                   - a figura 3b representa, a título ilustrativo, um detalhe de implementação de um filtro por equalização-retardo permitindo a implementação do dispositivo objeto da invenção representado na figura 3a;

                  - a figura 4, representa, a título ilustrativo, um  
25 exemplo de implementação do dispositivo objeto da invenção, no qual o cálculo dos filtros de equalização-retardo é deslocalizado.

Uma descrição mais detalhada do processo de espacialização sonora de uma cena áudio em conformidade com  
30 o objeto da presente invenção será agora dada em conexão com a figura 2a e as figuras seguintes.

O processo objeto da invenção se aplica a uma cena áudio tal como uma cena áudio 3D representada por um

primeiro conjunto compreendendo um número  $N$  de canais áudio codificados espacialmente superior ou igual à unidade,  $N \geq 1$ , em um número de sub-bandas de frequências determinado e decodificado em um domínio transformado.

5 O domínio transformado se estende de um domínio frequencial transformado, tal como domínio de Fourier, domínio PQMF ou de qualquer domínio híbrido provenientes desses últimos por criação de sub-bandas de frequências suplementares, submetidas ou não a um procedimento de  
10 decimação temporal.

Em conseqüência, os canais áudio codificados espacialmente constitutivos do primeiro conjunto  $N$  de canais são representados de maneira não limitativa pelos canais  $F_l$ ,  $F_r$ ,  $S_l$ ,  $S_r$ ,  $C$ , lfe precedentemente descritos no  
15 relatório e correspondendo a um modo de decodificação de uma cena áudio 3D no domínio transformado correspondente, assim como descrito precedentemente no relatório. Este modo não é outro que o modo 5.1 precedentemente mencionado.

Além disso, estes sinais são decodificados no  
20 domínio transformado precitado de acordo com um número de sub-bandas determinado próprias para a decodificação, o conjunto das sub-bandas sendo sinalizado  $(SB_k)_{k=1}^{k=K}$ .

$k$  designa a ordenação (rang) da sub-banda considerada.

25 O processo objeto da invenção permite transformar o conjunto dos canais áudio codificados espacialmente precedentemente citados em um segundo conjunto comportando um número, superior ou igual a dois, de canais sonoros de restituição no domínio temporal, os canais sonoros de  
30 restituição sendo sinalizados  $B_l$  e  $B_r$  para os canais binaurais esquerdo respectivamente direito, de maneira não limitativa no quadro da figura 2a. Compreende-se, em

particular, que no lugar de dois canais binaurais, o processo objeto da invenção se aplica a qualquer número de canais superior a dois, permitindo, por exemplo, a restituição sonora em tempo real da cena áudio 3D, assim como representado e descrito no relatório em conexão com a figura 1b.

De acordo com um aspecto marcante do processo objeto da invenção, este é implementado a partir de filtros de modelação da propagação acústica dos sinais áudio do primeiro conjunto de canais áudio codificados espacialmente, tendo em conta uma conversão sob a forma de pelo menos um ganho e de um retardo aplicáveis no domínio transformado, assim como será descrito ulteriormente no relatório. De maneira não limitativa, os filtros de modelação serão designados filtros HRTF na seqüência da descrição.

A conversão precitada é sinalizada para cada filtro HRTF considerado para uma sub-banda  $SB_k$  de ordenação  $k$  para estabelecer um valor de ganho  $g_k$  e de retardo  $d_k$  correspondente, a conversão precedente sendo então sinalizada, assim como representada na figura 2a  $HRTF = (g_k, d_k)$ .

Levando em conta a conversão precitada, o processo objeto da invenção consiste, para cada sub-banda frequencial do domínio transformado de ordenação  $k$ , em efetuar uma filtragem na etapa A por equalização-retardo do sinal de sub-banda por aplicação de um ganho  $g_k$  respectivamente de um retardo  $d_k$  no sinal de sub-banda, para gerar a partir dos canais codificados espacialmente precitados, isto é, os canais  $Fl, C, Fr, Sr, Sl$  e  $lfe$ , uma componente equalizada e retardada de um valor de retardo determinado na sub-banda de frequência  $Sb_k$  considerada de ordenação  $k$ .

Na figura 2a, a operação de filtragem por equalização-retardo é sinalizada de maneira simbólica  $CED_{kx} = \{Fl, C, Fr, Sr, Sl, lfe\} (g_{kx}, d_{kx})$ .

Na relação simbólica precitada,  $FEB_{kx}$  designa cada componente equalizada e retardada obtida por aplicação do ganho  $g_{kx}$  e do retardo  $d_{kx}$  em cada um dos canais áudio codificados espacialmente, isto é, os canais Fl, C, Fr, Sr, Sl e lfe.

Em conseqüência e na relação simbólica precitada,  $x$ , para a sub-banda de ordenação  $k$  correspondente, pode tomar de fato os valores Fl, C, Fr, Sr, Sl, lfe.

A etapa A é então seguida no domínio transformado de uma etapa B de adição de um sub-conjunto de componentes equalizadas e retardadas para criar um número de sinais filtrados no domínio transformado correspondendo ao número  $N'$  do segundo conjunto, superior ou igual a 2, de canais sonoros de restituição no domínio temporal.

Na etapa B da figura 2a, a operação de adição é dada pela relação simbólica:

$$F\{Fl, C, Fr, Sr, Sl, lfe\} = \sum CED_{kx}.$$

Na relação simbólica precitada,  $F\{Fl, C, Fr, Sr, Sl, lfe\}$  designa o sub-conjunto dos sinais filtrados no domínio transformado obtido por soma de um sub-conjunto de componentes equalizadas e retardadas  $CED_{kx}$ .

A título de exemplo não limitativo e para fixar as idéias, para um primeiro conjunto comportando um número de canais áudio codificados espacialmente  $N=6$ , correspondendo a um modo 5.1, o sub-conjunto de componentes equalizadas e retardadas pode consistir em adicionar cinco destas componentes equalizadas e retardadas para cada ouvido para obter o número  $N'$  igual a 2 de sinais filtrados no domínio transformado, assim como será descrito de maneira mais detalhada, ulteriormente no relatório.

A etapa de adição B precitada é então seguida de uma etapa C de síntese de cada um dos sinais filtrados no domínio transformado por um filtro de síntese para obter o segundo conjunto de número  $N'$  superior ou igual a dois de 5 sinais sonoros de restituição no domínio temporal.

Na etapa C da figura 2a, a operação correspondente de síntese é representada pela relação simbólica:

$$B_l, B_r = \text{Sint} (F\{F_l, C, F_r, S_r, S_l, lfe\})$$

10 De uma maneira geral, indica-se que o processo objeto da invenção pode ser aplicado a qualquer cena 3D áudio composta de  $N$  variando de 1 ao infinito de vias ou canais áudio codificados de forma espacial para  $N'$  variando de 2 ao infinito de canais sonoros de restituição.

15 No que diz respeito à etapa de soma representada na etapa B da figura 2a, indica-se que esta consiste, de maneira mais específica, em adicionar um sub-conjunto de componentes retardadas de forma diferente pelos diferentes retardos para gerar as  $N'$  componentes para cada sub-banda.

20 De maneira mais específica, indica-se que a filtragem por equalização-retardo do sinal em sub-banda inclui pelo menos a aplicação de uma defasagem completada, quando necessário, por um retardo puro por memorização, para a uma pelo menos das sub-bandas de frequência.

25 A noção de aplicação de um retardo puro é simbolizada na etapa A da figura 2a pela relação  $g_{Ex} = 1$ , a qual representa a ausência de equalização para o conjunto dos canais áudio de índice  $x$  na sub-banda de ordenação  $k = E$ , o valor 1 indicando uma transmissão sem modificação da 30 amplitude de cada um dos canais áudio codificados espacialmente.

O domínio transformado pode, assim como mencionado precedentemente no relatório, corresponder a um

domínio transformado híbrido assim como será descrito em conexão com a figura 2b no caso onde nenhuma decimação em frequência é aplicada na sub-banda correspondente.

Com referência à figura 2b precitada, a filtragem por equalização-retardo representada na etapa A da figura 5 2a é então executada em três sub-etapas A1, A2, A3 representadas na figura 2b.

Nestas condições, a etapa A comporta uma etapa suplementar de corte de frequência de sub-bandas 10 suplementares sem decimação, para aumentar o número de valores de ganho aplicados e, assim, a precisão de frequência, seguida de uma etapa de reagrupamento de sub-bandas suplementares, às quais foram aplicados os valores de ganho precitados.

As operações de corte de frequência depois de reagrupamento são representadas nas sub-etapas A1 e A2 da 15 figura 2b.

A etapa dos cortes de frequência é representada na sub-etapa A1 pela relação:

$$20 \quad \text{HRTF} \equiv \{g_{kz}, dkz\}_{z=1}^{z=Z}$$

A etapa de reagrupamento é representada na sub-etapa A2 pela relação:

$$[\text{GCEB}_{kz}]_1^z x = \{F1, C, F1, Sr, Sl, lfe\} (g_{kz})$$

Na sub-etapa A1, compreende-se que os valores de 25 ganho e de retardo para a sub-banda de ordenação k considerada são subdivididos em Z valores de ganho correspondentes, um valor de ganho  $g_{kz}$  para cada sub-banda suplementar, e na sub-etapa 12, compreende-se que o reagrupamento das sub-bandas suplementares é efetuado a 30 partir dos canais áudio codificados correspondentes para o índice x correspondente ao qual foi aplicado o valor de ganho  $g_{kz}$  na sub-banda suplementar considerada.

Na relação precedente,  $[GCE_{kz}]_{z=1}^{z=Z} \mathbf{x}$  designa o reagrupamento das sub-bandas suplementares às quais foram aplicados os valores de ganho para as sub-bandas suplementares consideradas.

5 A sub-etapa  $A_2$  é então seguida de uma sub-etapa  $A_3$  consistindo em aplicar o retardo às sub-bandas suplementares reagrupadas e, em particular, aos canais de áudio codificados espacialmente de índice  $x$  correspondente por intermédio do retardo  $d_{kx}$  de maneira parecida à etapa A da figura 2a.

10 A operação correspondente é sinalizada pela relação:

$$CE_{kz} \mathbf{x} = [GCE_{kz}]_{z=1}^{z=Z} \mathbf{x} (d_{kx}).$$

Além disso, o processo objeto da invenção pode consistir igualmente em efetuar uma filtragem por equalização-retardo em um domínio transformado híbrido comportando uma etapa suplementar de corte de frequência de sub-bandas suplementares com decimação, assim como representado na figura 2c.

20 Nesta hipótese, a etapa  $A'_1$  da figura 2c é idêntica à etapa  $A_1$  da figura 2b, para executar a criação das sub-bandas suplementares com decimação.

Nesta hipótese, a operação de decimação na etapa  $A'_1$  da figura 2c é executada no domínio temporal.

25 A etapa  $A'_1$  é então seguida de uma etapa  $A'_2$  correspondente a um reagrupamento das sub-bandas suplementares às quais foram aplicados os valores de ganho precitados tirados da decimação.

30 A etapa  $A'_2$  de reagrupamento é ela própria precedida ou seguida da aplicação do retardo  $d_{kx}$  assim representada pela flecha dupla de interversão das etapas  $A'_2$  e  $A'_3$ .

Compreende-se, em particular, que quando a aplicação do retardo é efetuada anteriormente ao reagrupamento, o retardo é aplicado diretamente nos sinais das sub-bandas suplementares anteriormente ao reagrupamento.

No que diz respeito à conversão de cada filtro HRTF em um valor de ganho e de retardo no domínio transformado, esta operação pode consistir, vantajosamente, em associar, como valor de ganho a cada sub-banda de ordenação  $k$ , um valor real definido como a média do módulo do filtro HRTF correspondente e em associar, como valor de retardo a cada sub-banda de ordenação  $k$ , um valor de retardo correspondente ao retardo de propagação entre o ouvido esquerdo e o ouvido direito de um ouvinte para diferentes posições.

Assim, a partir de um filtro HRTF, é possível calcular, de forma automática, os ganhos e os prazos de retardo aplicados em sub-banda. A partir da resolução frequencial do banco de filtro HRTF, associa-se a cada uma das sub-bandas  $SB_k$  um valor de retardo correspondente ao retardo de propagação entre o ouvido esquerdo e o ouvido direito de um ouvinte para diferentes posições.

Assim, a partir de um filtro HRTF, pode-se calcular, de forma automática, os ganhos e os prazos de retardo a aplicar em sub-banda.

A partir da resolução frequencial do banco de filtro, associa-se a cada uma das bandas um valor real. A título de exemplo não limitativo, é possível a partir do módulo do filtro HRTF, calcular, para cada sub-banda, a média do módulo do filtro HRTF precitado. Uma tal operação é similar a uma análise em banda de octavo ou de Bark dos filtros HRTF. Da mesma maneira, determina-se o retardo a ser aplicado para os canais indiretos, isto é, os valores

de retardo que são aplicáveis mais particularmente aos canais cujo retardo não é mínimo. Existem numerosos métodos para determinar de maneira automática os retardos interaurais ainda designados ITD para "Interaural Time Difference" e que correspondem aos retardos entre o ouvido esquerdo e o ouvido direito, para diferentes posições do ouvinte. Pode-se utilizar, a título de exemplo não limitativo, o método do limiar descrito por S. Busson na tese de doutorado da Université de la Méditerranée Est-Marseille II, 2006, intitulada "Individualização de índices acústicos para a síntese binaural". O princípio dos métodos de estimação do retardo interaural do tipo limiar é de determinar o tempo de chegada, ou ainda o retardo inicial da onda no ouvido direito  $T_d$  e no ouvido esquerdo  $T_g$ . O retardo interaural é dado pela relação ITD limiar =  $T_d - T_g$ .

O método mais frequente estima o tempo de chegada como o instante onde o filtro temporal HRIR ultrapassa um limiar dado. Por exemplo, o tempo de chegada pode corresponder ao tempo para o qual a resposta do filtro HRIR alcança 10% de som máximo.

Um exemplo de implementação específico no domínio transformado PQMF será agora dado logo a seguir.

De uma maneira geral, indica-se que a aplicação de um ganho no domínio PQMF complexo consiste em multiplicar o valor de cada amostra do sinal de sub-banda, representado por um valor complexo, pelo valor de ganho formado por um número real.

De fato, é bem conhecido que o uso de um domínio transformado PQMF complexo permite aplicar os ganhos se franqueando problemas de suavização de espectro gerados pela sub-amostragem inerente aos bancos de filtros. Cada

sub-banda  $SB_k$  de cada canal se vê assim afetada de um ganho determinado.

Além disso, a aplicação de um retardo no domínio transformado PQMF consiste pelo menos, para cada amostra do sinal em sub-banda, representado por um valor complexo, em introduzir uma rotação no plano complexo por multiplicação desta amostra por um valor exponencial complexo função da ordenação da sub-banda considerada, da taxa de sub-amostragem na sub-banda considerada e de um parâmetro de retardo ligado à diferença de retardo interaural de um ouvinte.

A rotação no plano complexo é então seguida de um retardo temporal puro da amostra após rotação. Este retardo temporal puro é uma função da diferença do retardo interaural de um ouvinte e da taxa de sub-amostragem na sub-banda considerada.

De maneira prática, indica-se que os retardos precitados são aplicados nos sinais resultantes, isto é, os sinais equalizados e, em particular, nos sub-conjuntos destes sinais ou canais que não se beneficiam de uma trajetória direta.

Em particular, a rotação é efetuada sob a forma de uma multiplicação complexa por um valor exponencial da forma:

$$\exp(-j\pi(k+0,5)d/M)$$

e por um retardo puro implementado por uma linha de retardo, por exemplo, realizando a operação:

$$y(k, n) = x(k, n-D)$$

Nas relações precedentes:

- exp é a função exponencial;
- j é tal que  $j*j=-1$ ;
- k a ordenação da sub-banda  $SB_k$  considerada;

-  $M$  é a taxa de sub-amostragem na sub-banda considerada,  $M$  quer ser tomado igual a 64, por exemplo;

-  $y(k, n)$  é o valor da amostra de saída após aplicação do retardo puro na amostra temporal de ordenação  $n$  da sub-banda  $SB_k$  de ordenação  $k$ , isto é, a amostra  $x(k, n)$  à qual é aplicado o retardo  $B$ ;

-  $d$  e  $D$  nas relações precedentes são tais que eles correspondem à aplicação de um retardo de  $D*M + d$  no domínio temporal não sub-amostrado. O retardo  $D*M + d$  corresponde ao retardo interaural calculado precedentemente.  $d$  pode tomar valores negativos o que permite simular um avanço de fase no lugar de um retardo.

A operação assim realizada induz uma aproximação que é conveniente para o efeito buscado.

Em termos de operações de cálculo, o tratamento implementado consiste então em realizar uma multiplicação complexa entre uma exponencial complexa e de uma amostra de sub-banda formada por um valor complexo.

Um retardo eventual, se o retardo total a ser aplicado for superior ao valor  $M$ , deve ser inserido, mas esta operação não comporta operação aritmética.

O processo objeto da invenção pode igualmente ser implementado em um domínio transformado híbrido. Este domínio transformado híbrido é um domínio frequencial no qual as bandas PQMF são vantajosamente recortadas por um banco de filtros decimado ou não.

Se o banco de filtros for decimado, a decimação se estendendo de uma decimação no tempo, então a introdução de um retardo segue vantajosamente o procedimento incluindo um retardo puro e um defasador.

Se o banco de filtros não for decimado, então o retardo só pode ser aplicado uma única vez quando da síntese. É, de fato, inútil aplicar o mesmo retardo em cada

um dos ramos uma vez que a síntese é uma operação linear, sem sub-amostrador.

A aplicação dos ganhos permanece idêntica, estes sendo simplesmente mais numerosos, assim como descrito  
5 precedentemente em conexão com a figura 2b, por exemplo, e permitem então proceder ao corte mais preciso de frequência. Um ganho real é então aplicado por sub-banda suplementar.

Enfim, de acordo com uma variante de  
10 implementação, reitera-se o processo de acordo com a invenção para pelo menos dois pares equalização-retardo e somam-se os sinais obtidos para obter os canais sonoros no domínio temporal.

Uma descrição mais detalhada de um dispositivo de  
15 espacialização sonora de uma cena áudio comportando um primeiro conjunto compreendendo um número superior ou igual à unidade de canais áudio codificados espacialmente em um número de sub-bandas de frequências determinado e decodificados em um domínio transformado, em um segundo  
20 conjunto compreendendo um número superior ou igual a dois de canais sonoros de restituição no domínio temporal, em conformidade com o objeto da presente invenção, será agora descrito em conexão com as figuras 3a e 3b.

Assim como mencionado precedentemente, o  
25 dispositivo objeto da invenção é baseado no princípio da conversão sob a forma de pelo menos um ganho e de um retardo aplicável no domínio transformado de filtros de modelação da propagação acústica dos sinais áudio do primeiro conjunto de canais precitado. O dispositivo objeto  
30 da invenção permite a espacialização sonora de uma cena áudio, tal como uma cena áudio 3D, em um segundo conjunto comportando um número, superior ou igual a dois, de canais sonoros de restituição no domínio temporal.

O dispositivo objeto da invenção representado na figura 3a refere-se a um estágio deste dispositivo específico em cada sub-banda  $SB_k$  de ordenação  $k$  de decodificação no domínio transformado.

5 Compreende-se, em particular, que o estágio, para cada sub-banda de ordenação  $k$  representada na figura 3a, é, na realidade, replicado para cada uma das sub-bandas para constituir finalmente o dispositivo de espacialização sonora em conformidade com o objeto da presente invenção.

10 Por convenção, o estágio representado na figura 3a será designado logo a seguir dispositivo de espacialização sonora objeto da invenção.

Com referência à figura precitada, o dispositivo objeto da invenção tal como representado na figura 3a  
15 comporta, adicionalmente ao decodificador espacial representado, comportando os módulos  $OTT_0$  a  $OTT_4$  correspondendo sensivelmente a um decodificador espacial SD da técnica anterior tal como representado na figura 1c, mas no qual se procede adicionalmente, de maneira conhecida  
20 como tal do estado da técnica, a uma soma do canal frontal C e do canal de frequência baixa lfe por um somador S, um módulo 1 de filtragem por equalização-retardo do sinal de sub-banda por aplicação de um ganho respectivamente de um retardo no sinal de sub-banda.

25 Na figura 3a, a aplicação de um ganho é representada em cada um dos canais áudio codificados espacialmente, representados por amplificadores  $l_0$  a  $l_8$ , estes últimos gerando uma componente equalizada a qual pode ser submetida ou não a um retardo por intermédio de  
30 elementos de retardo sinalizados  $l_9$  a  $l_{12}$  para gerar, a partir de cada um dos canais áudio codificados espacialmente, uma componente equalizada e retardada de um

valor de retardo determinado na sub-banda de frequência  $SB_k$ .

Com referência à figura 3a, os ganhos dos amplificadores  $l_0$  a  $l_8$  têm valores arbitrários  $A, B, B, A, C, D, E, E, D$  respectivamente. Além disso, os valores de retardo aplicados pelos módulos de retardo  $l_9$  a  $l_{12}$  têm como valores  $D_f, B_f, D_s, D_s$ . Na figura precitada, a estrutura dos ganhos e retardos introduzidos é simétrica. Uma estrutura não simétrica pode ser implementada sem sair do escopo do objeto da invenção.

O dispositivo objeto da invenção comporta igualmente um módulo 2 de adição de um sub-conjunto de componentes equalizadas e retardadas para criar um número de sinais filtrados no domínio transformado correspondente ao número  $N'$  do segundo conjunto superior ou igual a dois de canais sonoros de restituição no domínio temporal.

Enfim, o dispositivo objeto da invenção comporta um módulo 3 de síntese de cada um dos sinais filtrados no domínio transformado para obter o segundo conjunto compreendendo um número  $N'$  superior ou igual a dois de sinais sonoros de restituição no domínio temporal. O módulo de síntese 3 comporta, assim, no modo de realização da figura 3a, um sintetizador  $3_0$  e  $3_1$  os quais permitem entregar um sinal sonoro de restituição no domínio temporal  $B_l$  para sinal binaural esquerdo, respectivamente  $B_r$  para sinal binural direito.

As componentes equalizadas e retardadas no modo de realização da figura 3a são obtidas da maneira logo a seguir com:

- 30 -  $A[k]$  designando o ganho dos amplificadores  $l_0, l_3$  para a sub-banda  $SB_k$  de ordenação  $k$ ,
- $B[k]$  designa o ganho do amplificador  $l_1, l_2$  representado na figura 3a,

- $C[k]$  designa o ganho do amplificador  $l_4$ ,
- $D[k]$  designa o ganho dos amplificadores  $l_5$   $l_8$ ,
- $E[k]$  designa o ganho dos amplificadores  $l_6$   $l_7$ ,

No que diz respeito aos canais áudio codificados espacialmente e, em particular, estes canais Fl, Fr, Clfe, Sl e Sr para a sub-banda  $SB_k$ , designa-se por  $Fl[k][n]$ ,  $Fr[k][n]$ ,  $Fc[k][n]$ ,  $lfe[k][n]$ ,  $SI[k][n]$ ,  $Sr[k][n]$ , a nona amostra da sub-banda  $SB_k$ . Assim, cada amplificador,  $l_0$  a  $l_8$  entrega as componentes equalizadas sucessivamente:

- 10           -  $A[k]*Fl[k][n]$ ,
- $B[k]*Fl[k][n]$ ,
- $B[k]*Fr[k][n]$ ,
- $A[k]*Fr[k][n]$ ,
- $C[k]*Fc[k][n]$ ,
- 15           -  $D[k]*Sl[k][n]$ ,
- $E[k]*Sl[k][n]$ ,
- $E[k]*Sr[k][n]$ ,
- $D[k]*Sr[k][n]$ .

As operações precedentes, assim como mencionado precedentemente no relatório, são realizadas sob a forma de uma multiplicação real agindo, neste caso, sobre números complexos.

Os retardos introduzidos pelos elementos de retardo  $l_9$ ,  $l_{10}$ ,  $l_{11}$  e  $l_{12}$  são aplicados nas componentes equalizadas precitadas, para gerar as componentes equalizadas e retardadas.

No exemplo representado na figura 3a, estes retardos são aplicados no sub-conjunto que não se beneficia de uma trajetória direta. São, na descrição da figura 3a, os sinais que sofreram as multiplicações pelos ganhos  $B[k]$  e  $E[k]$  aplicadas pelos amplificadores ou multiplicadores  $l_1$ ,  $l_2$  e  $l_6$  e  $l_7$ .

Uma descrição mais detalhada de um filtro ou elemento de filtragem por equalização-retardo constituída, por exemplo, por um amplificador multiplicador  $l_1$  e um elemento retardador  $l_9$  será agora dada em conexão com a  
5 figura 3b.

No que diz respeito à aplicação do ganho, indica-se que o elemento de filtragem, correspondente, representado a figura 3b, comporta um multiplicador digital, isto é, um dos multiplicadores ou amplificadores  
10  $l_0$  a  $l_8$  e representado pelo valor de ganho  $g_{kx}$  na figura 3b, este multiplicador permitindo a multiplicação de qualquer amostra complexa de cada canal áudio codificado de índice  $x$  correspondendo aos canais Fl, Fr, Clfe, Sl ou Sr por um valor real, isto é, o valor de ganho precedentemente  
15 mencionado no relatório.

Além disso, o elemento de filtragem representado na figura 3b comporta pelo menos um multiplicador digital complexo permitindo introduzir uma rotação no plano complexo de qualquer amostra do sinal de sub-banda por um  
20 valor exponencial complexo, o valor  $\exp(-j\phi(k, SS_k))$  onde  $\phi(k, SS_k)$  designa um valor de fase função da taxa de sub-amostragem da sub-banda considerada e da ordenação da sub-banda considerada  $k$ .

Em um modo de realização,  $\phi(k, SS_k) =$   
25  $\phi^*(k+0,5) * d/M$ .

O multiplicador digital complexo é seguido de uma linha de retardo sinalizada L.A.R. introduzindo um retardo puro de cada amostra após rotação, permitindo introduzir um retardo temporal puro função da diferença do retardo  
30 interaural de um ouvinte e da taxa de sub-amostragem  $M$  na sub-banda  $SB_k$  considerada.

Assim, a linha de retardo L.A.R. permite introduzir o retardo na amostra complexa após rotação da forma  $y(k, n) = x(k, n-D)$ .

Enfim, indica-se que os valores de  $d$  e  $D$  são tais que estes valores correspondam à aplicação de um retardo  $D*M+d$  no domínio temporal não amostrado e que o retardo  $D*M+d$  corresponda ao retardo interaural precedentemente mencionado.

Para a implementação do dispositivo objeto da invenção, tal como representado na figura 3a, pode-se observar que o sinal  $Fr[k][n]$  é multiplicado pelo ganho  $B[k]$  e depois retardado, o que, em conformidade com um dos aspectos remarcáveis do objeto da invenção, retorna a multiplicar este sinal por um ganho complexo. O produto do ganho  $B[k]$  e da exponencial complexa pode ser realizado uma vez para qualquer evitando, assim, uma operação complementar para cada amostra  $Fr[k][n]$  sucessiva. As componentes equalizadas e retardadas esquerdas sendo referidas  $L_0$  a  $L_4$  e direitas  $R_0$  a  $R_4$  e representadas no desenho reagrupadas pelos módulos somadores  $2_0$  respectivamente  $2_1$ , verificam-se então as relações a seguir:

Tabela T

	$L0[k][n] = A[k]F1[k][n]$
25	$R0[k][n] = B[k]F1[k][n]$ retardado de $D_f$ amostras
	$R1[k][n] = A[k]Fr[k][n]$
	$L1[k][n] = B[k]Fr[k][n]$ retardado de $D_f$ amostras
	$L2[k][n] = R2[k][n] = C[k](Fc[k][n] + 1fe[k][n])$
	$L3[k][n] = D[k]S1[k][n]$
30	$R3[k][n] = E[k]S1[k][n]$ retardado de $D_s$ amostras
	$R4[k][n] = D[k]Sr[k][n]$
	$L4[k][n] = E[k]Sr[k][n]$ retardado de $D_s$ amostras

Para obter os canais sonoros de restituição no domínio temporal, a saber os canais  $B_l$  esquerdo respectivamente  $B_r$  direito representados na figura 3a, isto é, os canais binauralizados no modo de realização da figura

5 3a, adiciona-se, para cada amostra de ordenação  $n$ , as componentes equalizadas e retardadas espaciais, isto é, a adição das componentes:

$L0[k][n] + L1[k][n] + L2[k][n] + L3[k][n] + L4[k][n]$  para o módulo somador  $2_0$ , e

10  $R0[k][n] + R1[k][n] + R2[k][n] + R3[k][n] + R4[k][n]$  para o módulo somador  $2_1$ .

Os sinais resultantes entregues pelos módulos de soma  $2_0$  e  $2_1$  são, em seguida, passados nos bancos de filtros de síntese  $3_0$  respectivamente  $3_1$  a fim de obter os

15 sinais binauralizados no domínio temporal  $B_l$  respectivamente  $B_r$ .

Os sinais precitados podem, em seguida, alimentar um conversor digital-analógico, a fim de permitir a escuta dos sons esquerdo  $B_l$  e direito  $B_r$  em um auscultador, por

20 exemplo.

A operação de síntese realizada pelos módulos de síntese  $3_0$  e  $3_1$  inclui, no caso em questão, a operação de síntese híbrida tal como descrita precedentemente na descrição.

25 O processo objeto da invenção pode vantajosamente consistir em dissociar as operações de equalização e de retardo, as quais podem se basear em sub-bandas de frequência em número diferente. Em variante, a equalização pode, por exemplo, ser efetuada no domínio híbrido e o

30 retardo no domínio PQMF.

Compreende-se que o processo e o dispositivo objetos da invenção, apesar de descritos para a binauralização de seis canais na direção de um auscultador,

podem igualmente ser aplicados para efetuar a transauralização, isto é, a restituição de um campo sonoro 3D em um par de alto-falantes ou para converter, de forma pouco complexa, uma representação de N canais de áudio ou 5 fontes sonoras provenientes de um decodificador espacial ou de vários decodificadores monofônicos na direção de N' canais de áudio disponíveis no nível da restituição. As operações de filtragens podem então servir para multiplicar quando necessário.

10 A título de exemplo complementar não limitativo, o processo e o dispositivo objetos da invenção podem ser aplicados no caso de um jogo 3D interativo nos sons emitidos pelos diferentes objetos ou fontes sonoras, os quais podem então ser espacializados em função de sua 15 posição relativa em relação ao ouvinte. Amostras sonoras são então comprimidas e armazenadas em diferentes arquivos ou diferentes zonas de memória. Para serem jogados e espacializados, eles são parcialmente decodificados a fim de permanecer no domínio codificado e são filtrados no 20 domínio codificado por filtros binaurais adequados de maneira vantajosa utilizando o processo de escrita em conformidade com o objeto da presente invenção.

De fato, reagrupando as operações de decodificação e de espacialização, a complexidade global do 25 procedimento é fortemente reduzida sem, no entanto, causar a perda de qualidade.

A invenção cobre, enfim, um programa de computador comportando uma seqüência de instruções memorizadas em um suporte de memorização para execução por 30 um computador ou um dispositivo dedicado de espacialização sonora, o qual quando desta execução, executa as etapas de filtragem de adição e de síntese tais como descritas em

ligação com as figuras 2a a 2c e 3a, 3b precedentemente no relatório.

Compreende-se, em particular, que as operações representadas nas figuras precitadas podem vantajosamente ser implementadas em amostras digitais complexas por intermédio de uma unidade central de tratamento, de uma memória de trabalho e de uma memória de programa, não representadas no desenho da figura 3a.

Enfim, o cálculo dos ganhos e dos retardos constituindo os filtros de equalização-retardo pode ser executado de maneira externa ao dispositivo objeto da invenção representado na figura 3a e 3b, assim como será descrito logo a seguir em conexão com a figura 4.

Com referência à figura precitada, considera-se uma primeira unidade de codificação espacial e de codificação com redução de fluxo I, incluindo um dispositivo objeto da invenção, tal como representado na figura 3a, 3b, permitindo operar a codificação espacial precitada a partir de uma cena áudio em modo 5.1, por exemplo, e a transmissão de áudio codificado, de um lado, e de parâmetros espaciais, de outro lado, na direção de uma unidade de decodificação e de decodificação espacial II.

O cálculo dos filtros de equalização-retardo pode então ser efetuado por uma unidade distinta III, a qual a partir dos filtros de modelação, filtros HRTF, calcula os valores de equalização de ganho e de retardo e os transmite à unidade I de codificação espacial e à unidade II de decodificação espacial.

A codificação espacial pode, assim, levar em consideração as HRTF que serão aplicadas para corrigir seus parâmetros espaciais e melhorar o rendimento 3D. Da mesma forma, o codificador com redução de fluxo poderá se servir

destas HRTF para medir os efeitos perceptivos de uma quantificação de frequência.

Do lado da decodificação estão as HRTF transmitidas que serão aplicadas no decodificador espacial, e permitirão, quando necessário, reconstruir as vias restituídas.

Como nos exemplos precedentes, são 2 vias a partir de 5 que serão restituídas, mas outros casos podem incluir a construção de 5 vias a partir de 3 como ilustrado acima. O processo de decodificação espacial procederá então como se segue:

- projeção dos 3 canais recebidos em um conjunto de canais virtuais (superior aos 5 de saída) utilizando as informações espaciais (upmix);
- redução dos canais virtuais aos 5 canais de saída utilizando as HRTF.

Se as HRTF tiverem sido aplicadas ao codificador, então poder-se-á eventualmente suprimir sua contribuição antes upmix para realizar o esquema abaixo.

As HRTF após conversão sob sua forma ganho/retardo podem ser quantificadas de forma privilegiada sob a seguinte forma: codificação em diferencial de seus valores depois da quantificação de suas diferenças: se se chama  $G[k]$  os valores dos ganhos do equalizador, então se transmitirão os valores quantificados:

$$e[k] = G[k+1] - G[k],$$

linearmente ou logaritmicamente.

De maneira mais específica com referência à figura 4 precitada, o procedimento implementado pelo dispositivo e o processo objetos da invenção permite, assim, executar uma espacialização sonora de uma cena de áudio na qual o primeiro conjunto comporta um número determinado de canais áudio espacialmente e, o segundo

conjunto comporta um número inferior de canais sonoros de restituição no domínio temporal. Ele permite, além da decodificação, efetuar uma transformação inversa de um número de canais áudio codificados espacialmente na direção

5 de um conjunto comportando um número superior ou igual de canais sonoros de restituição no domínio temporal.

## REIVINDICAÇÕES

1. Processo de espacialização sonora de uma cena áudio comportando um primeiro conjunto, comportando um número, superior ou igual à unidade, canais áudio  
5 codificados espacialmente em um número de sub-bandas de frequências determinado, e decodificados em um domínio transformado, em um segundo conjunto comportando um número superior ou igual a dois de canais sonoros de restituição no domínio temporal, a partir de filtros de modelação da  
10 propagação acústica dos sinais áudio do primeiro conjunto de canais, caracterizado pelo fato de que, para cada filtro de modelação convertido sob a forma de pelo menos um ganho e de um retardo aplicáveis no domínio transformado, o processo inclui pelo menos, para cada sub-banda frequencial  
15 do domínio transformado:

a filtragem por equalização-retardo do sinal de sub-banda por aplicação de um ganho respectivamente de um retardo no sinal de sub-banda, para gerar, a partir dos canais codificados espacialmente, uma componente equalizada e retardada de um valor de retardo determinado na sub-banda  
20 de frequências considerada;

a adição de um sub-conjunto de componentes equalizadas e retardadas, para criar um número de sinais filtrados no domínio transformado correspondente ao número  
25 do segundo conjunto superior ou igual a dois de canais sonoros de restituição no domínio temporal;

a síntese de cada um dos sinais filtrados no domínio transformado por um filtro de síntese, para obter o segundo conjunto de número superior ou igual a dois de  
30 canais sonoros de restituição no domínio temporal.

2. Processo, de acordo com a reivindicação 1, caracterizado pelo fato de que a filtragem por equalização-retardo do sinal de sub-banda inclui pelo menos a aplicação

de uma defasagem para uma pelo menos das sub-bandas de frequências.

3. Processo, de acordo com a reivindicação 2, caracterizado pelo fato de que a filtragem por equalização  
5 retardo inclui adicionalmente um retardo puro por memorização para uma pelo menos das sub-bandas de frequências.

4. Processo, de acordo com qualquer uma das reivindicações 1 a 3, caracterizado pelo fato de que a  
10 filtragem por equalização-retardo em um domínio transformado híbrido comporta uma etapa suplementar de corte de frequência em sub-bandas suplementares sem decimação, para aumentar o número de valores de ganho aplicados, seguida de uma etapa de reagrupamento das sub-  
15 bandas suplementares às quais foram aplicados os valores de ganho, depois de aplicação do retardo.

5. Processo, de acordo com qualquer uma das reivindicações 1 a 3, caracterizado pelo fato de que a  
20 filtragem por equalização-retardo em um domínio transformado híbrido comporta uma etapa suplementar de corte de frequência em sub-bandas suplementares com decimação, para aumentar o número de valores de ganho aplicados, seguida de uma etapa de reagrupamento das sub-  
25 bandas suplementares às quais foram aplicados os valores de ganho, a etapa de reagrupamento sendo, ela própria, precedida ou seguida da aplicação do retardo.

6. Processo, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que, para converter cada filtro de modelação em um valor de  
30 ganho respectivamente de retardo no domínio transformado, este consiste pelo menos em:

associar como valor de ganho a cada sub-banda um valor real definido como a média do módulo do filtro de modelação;

5 associar como valor de retardo a cada sub-banda um valor de retardo correspondente ao retardo de propagação entre o ouvido esquerdo e o ouvido direito para diferentes posições.

7. Processo, de acordo com qualquer uma das reivindicações 1 a 3 ou 6, com exclusão das reivindicações 10 4 ou 5, caracterizado pelo fato de que a aplicação de um ganho no domínio PQMF consiste em multiplicar o valor de cada amostra do sinal de sub-banda, representada por um valor complexo, pelo valor de ganho formado por um número real.

15 8. Processo, de acordo com qualquer uma das reivindicações 1 a 3 ou 6 ou 7, com exclusão das reivindicações 4 ou 5, caracterizado pelo fato de que a aplicação de um retardo no domínio transformado PQMF consiste pelo menos, para cada amostra do sinal em sub- 20 banda, representada por um valor complexo, em:

introduzir uma rotação no plano complexo por multiplicação desta amostra por um valor exponencial complexo função da ordenação da sub-banda considerada, da taxa de sub-amostragem na sub-banda considerada, e de um 25 parâmetro de retardo ligado à diferença de retardo interaural de um ouvinte;

introduzir um retardo temporal puro da amostra após rotação, o retardo temporal puro sendo uma função da diferença do retardo interaural de um ouvinte e da taxa de 30 sub-amostragem na sub-banda considerada.

9. Processo, de acordo com qualquer uma das reivindicações 1 a 8, caracterizado pelo fato de que para uma espacialização sonora binaural de uma cena áudio na

qual o primeiro conjunto comporta um número de canais áudio codificados espacialmente igual a  $N=6$ , em modo 5.1, o segundo conjunto comporta dois canais sonoros de restituição no domínio temporal, para uma restituição por um auscultador de áudio.

10. Processo, de acordo com qualquer uma das reivindicações 1 a 9, caracterizado pelo fato de que se reitera o processo para pelo menos dois pares equalização-retardo e se somam os sinais obtidos para obter os canais sonoros no domínio temporal.

11. Processo, de acordo com qualquer uma das reivindicações 1 a 9, caracterizado pelo fato de que para uma espacialização sonora de uma cena áudio na qual o primeiro conjunto comporta um número determinado de canais áudio codificados espacialmente e o segundo conjunto comporta um número inferior de canais sonoros de restituição no domínio temporal, esse processo consiste, na decodificação, em efetuar uma transformação inversa de um número de canais áudio codificados espacialmente na direção de um conjunto comportando um número superior ou igual de canais sonoros de restituição no domínio temporal.

12. Processo, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que os valores de ganho e de retardo associados ao filtro de modelação são transmitidos sob a forma quantificada.

13. Dispositivo de espacialização sonora de uma cena de áudio comportando um primeiro conjunto, comportando um número superior ou igual à unidade, canais áudio codificados espacialmente em um número de sub-bandas de frequências determinado, e decodificados em um domínio transformado, em um segundo conjunto comportando um número superior ou igual a dois de canais sonoros de restituição no domínio temporal, a partir de filtros de modelação da

propagação acústica dos sinais áudio do primeiro conjunto de canais, caracterizado pelo fato de que, para cada sub-banda frequencial de um decodificador espacial, no domínio transformado, o dispositivo compreende, além deste  
5 decodificador espacial:

meios de filtragem por equalização-retardo do sinal de sub-banda por aplicação de pelo menos um ganho respectivamente de um retardo no sinal de sub-banda, para gerar, a partir de cada um dos canais áudio codificados  
10 espacialmente, uma componente equalizada e retardada de um valor de retardo determinado na sub-banda de frequências considerada;

meios de adição de um sub-conjunto de componentes equalizadas e retardadas, para criar um número de sinais  
15 filtrados no domínio transformado correspondente ao número do segundo conjunto superior ou igual a dois de canais sonoros de restituição no domínio temporal;

meios de síntese de cada um dos sinais filtrados no domínio transformado, para obter o segundo conjunto de  
20 número superior ou igual a dois de canais sonoros de restituição no domínio temporal.

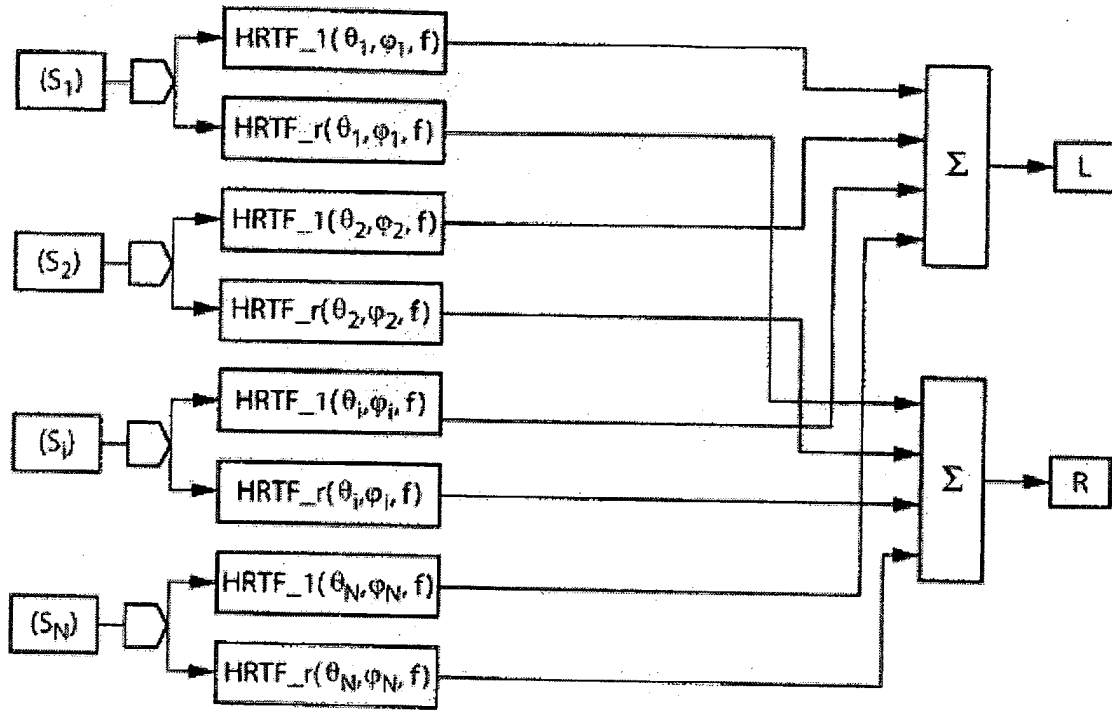
14. Dispositivo, de acordo com a reivindicação 13, caracterizado pelo fato de que os meios de filtragem por aplicação de um ganho comportam um multiplicador  
25 digital de toda amostra complexa de cada canal áudio codificado espacialmente por um valor real.

15. Dispositivo, de acordo com a reivindicação 13 ou 14, caracterizado pelo fato de que os meios de filtragem por aplicação de um retardo comportam pelo menos um  
30 multiplicador digital complexo, permitindo introduzir uma rotação no plano complexo de toda amostra do sinal em sub-banda por um valor exponencial complexo, função da ordenação da sub-banda considerada, da taxa de sub-

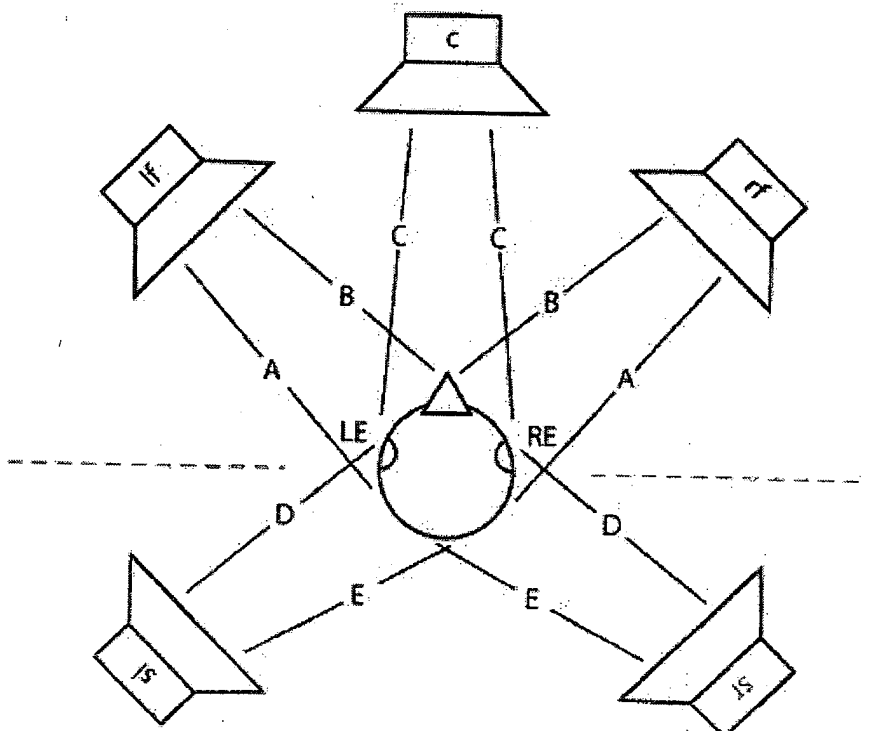
amostragem na sub-banda considerada e de um parâmetro de retardo ligado à diferença de retardo interaural de um ouvinte.

16. Dispositivo, de acordo com a reivindicação 5 15, caracterizado pelo fato de que os meios de filtragem comportam adicionalmente uma linha de retardo puro de cada amostra após rotação, permitindo introduzir um retardo temporal puro função da diferença do retardo interaural de um ouvinte e da taxa de sub-amostragem na sub-banda 10 considerada.

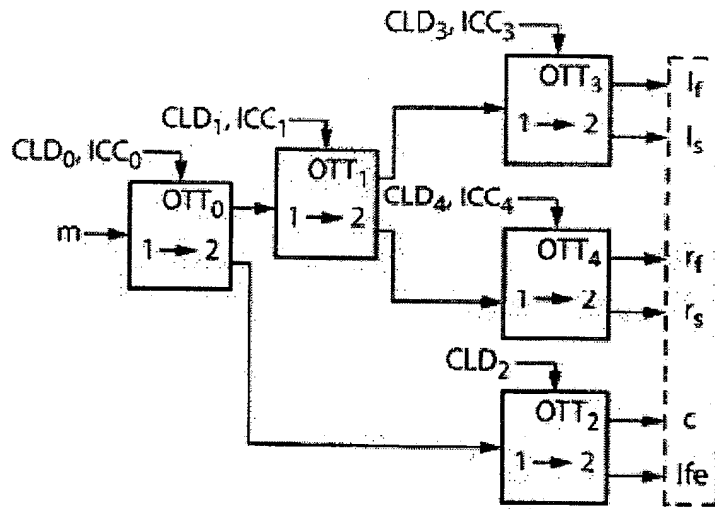
17. Programa de computador comportando uma seqüência de instruções memorizadas em um suporte de memorização para execução por um computador ou um dispositivo dedicado, caracterizado pelo fato de que quando 15 desta execução, o programa executa as etapas de filtragem, de adição e de síntese de acordo com qualquer uma das reivindicações 1 a 12.



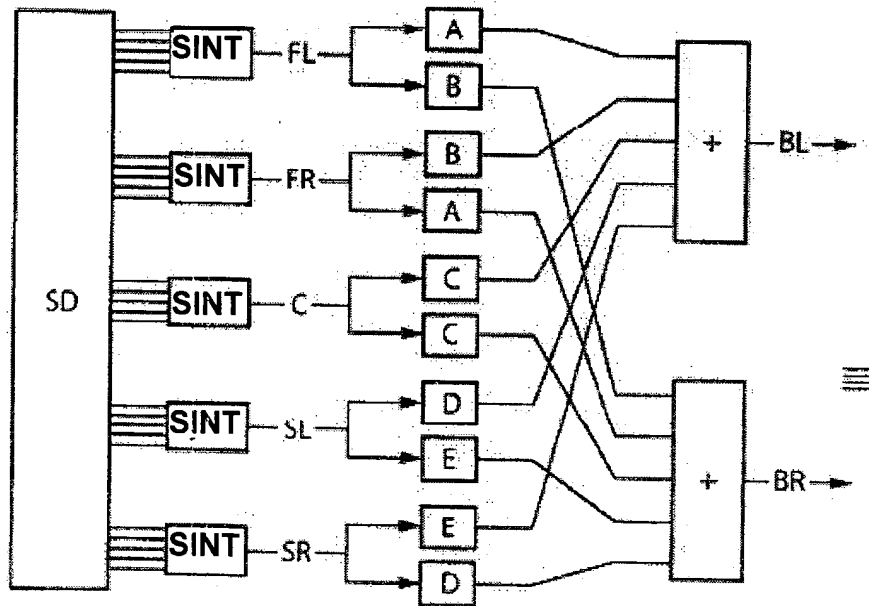
**FIG. 1a**  
(TÉCNICA ANTERIOR)



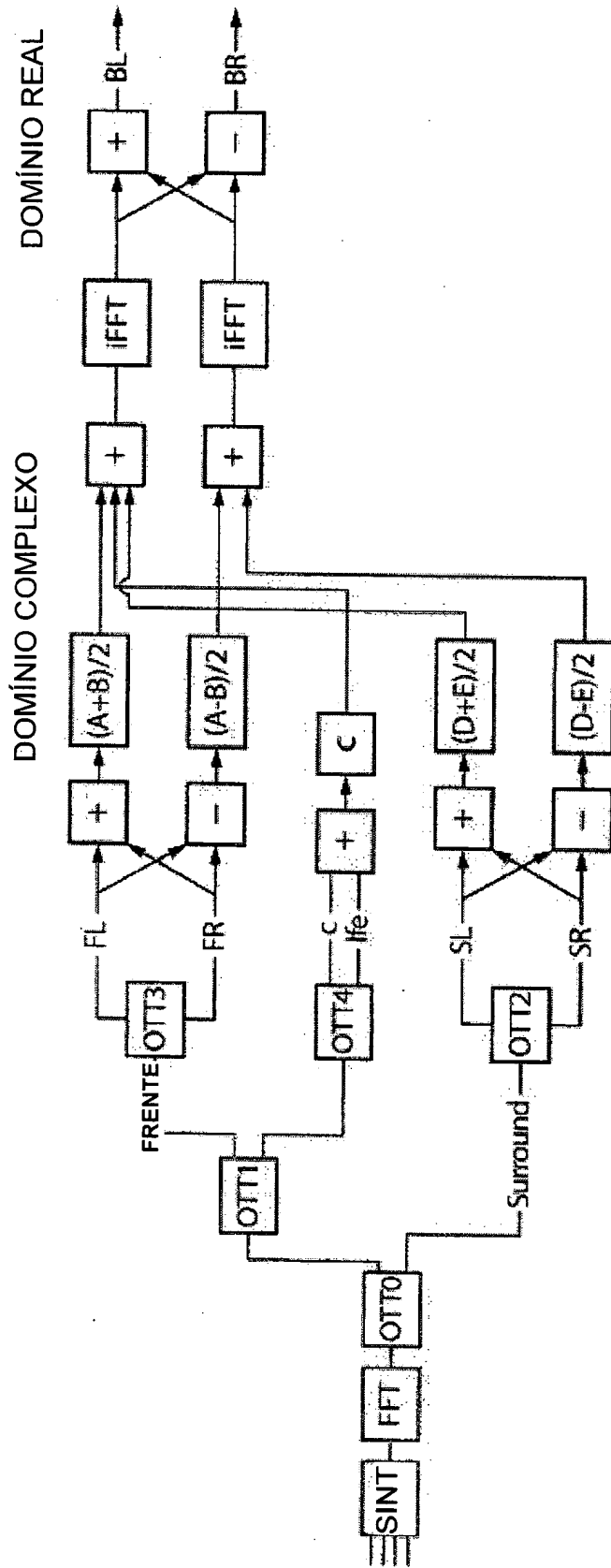
**FIG. 1b**  
(TÉCNICA ANTERIOR)



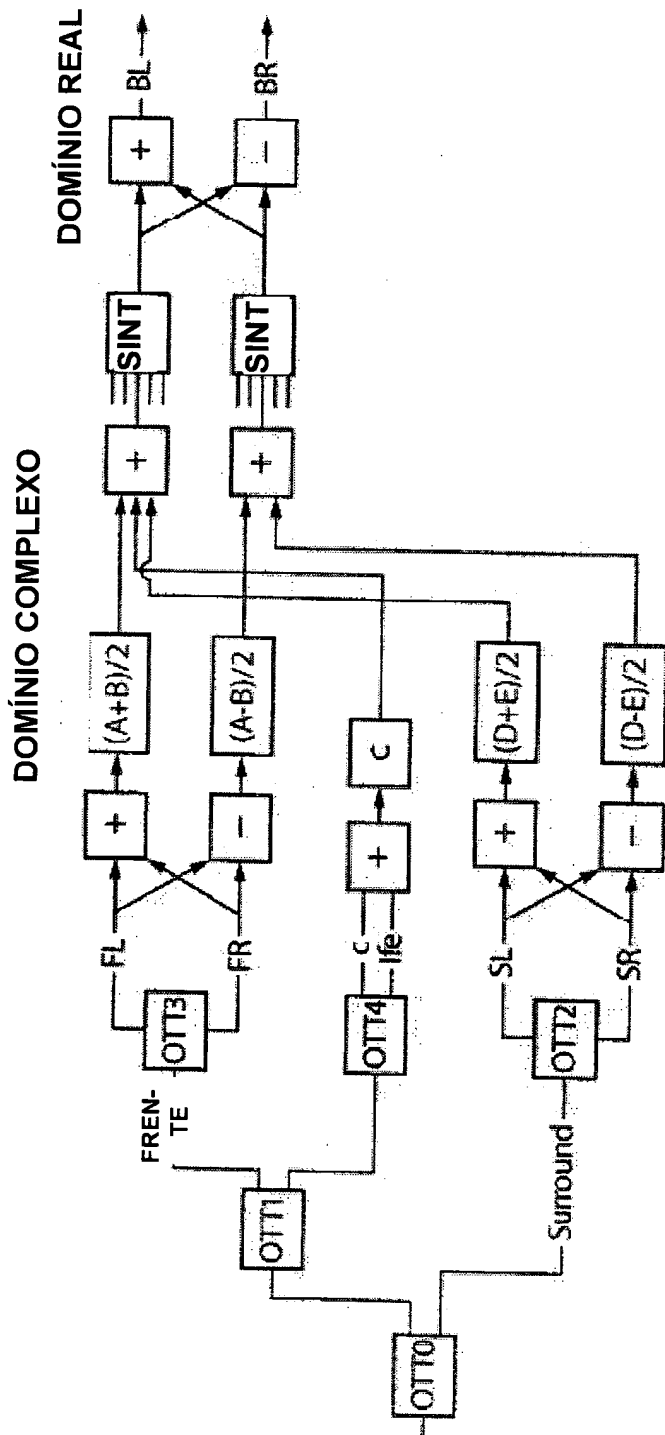
**FIG. 1c**  
(TÉCNICA ANTERIOR)



**FIG. 1d**  
(TÉCNICA ANTERIOR)



**FIG. 1e**  
(TÉCNICA ANTERIOR)



**FIG. 1f**  
(TÉCNICA ANTERIOR)

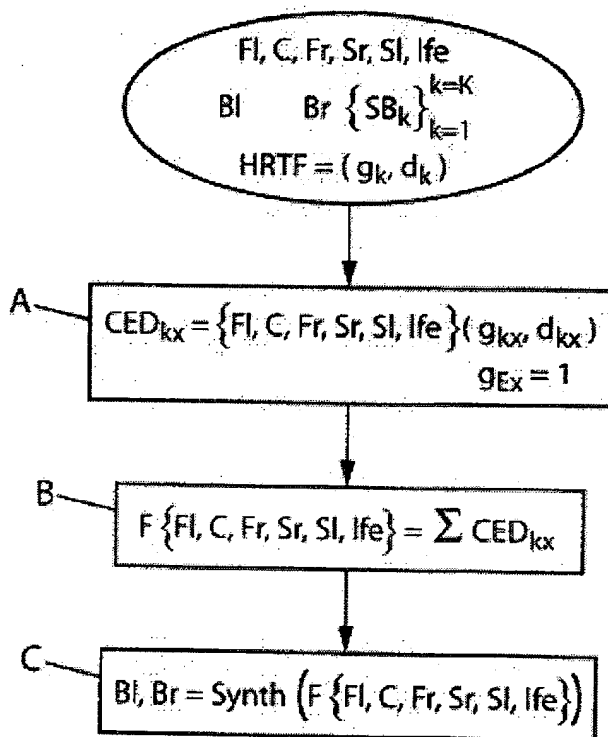


FIG. 2a

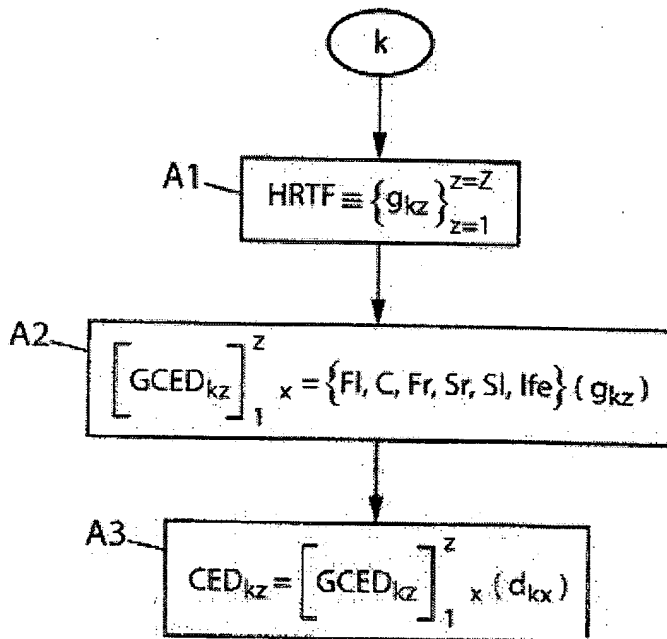


FIG. 2b

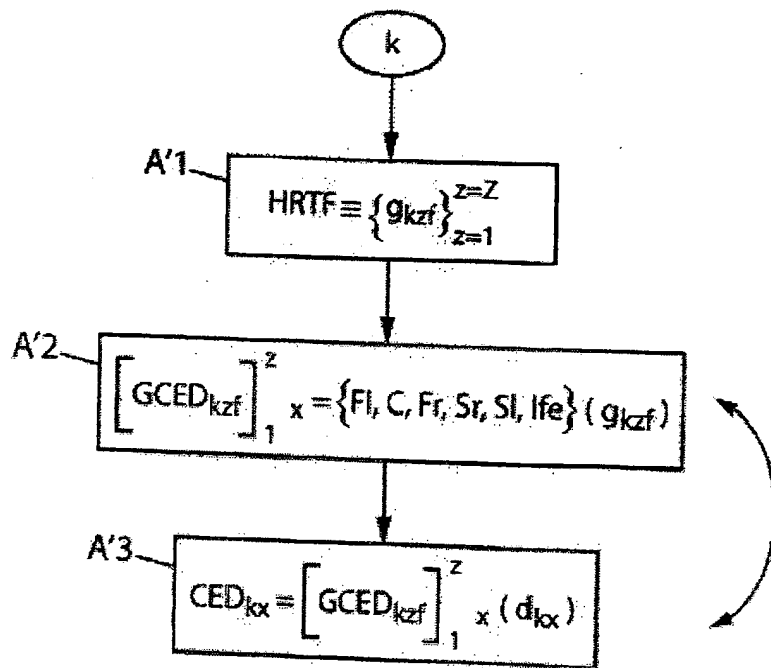


FIG. 2c

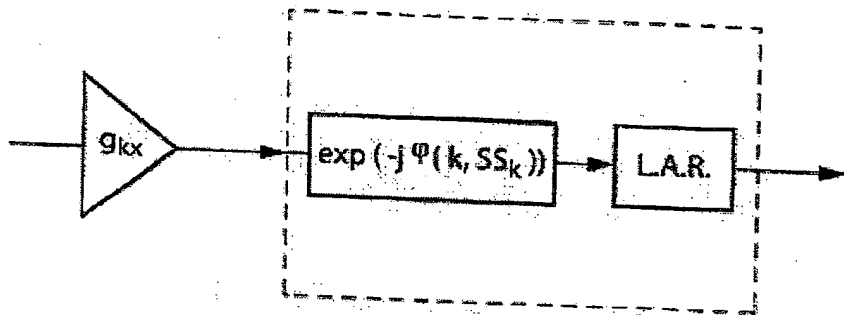


FIG. 3b

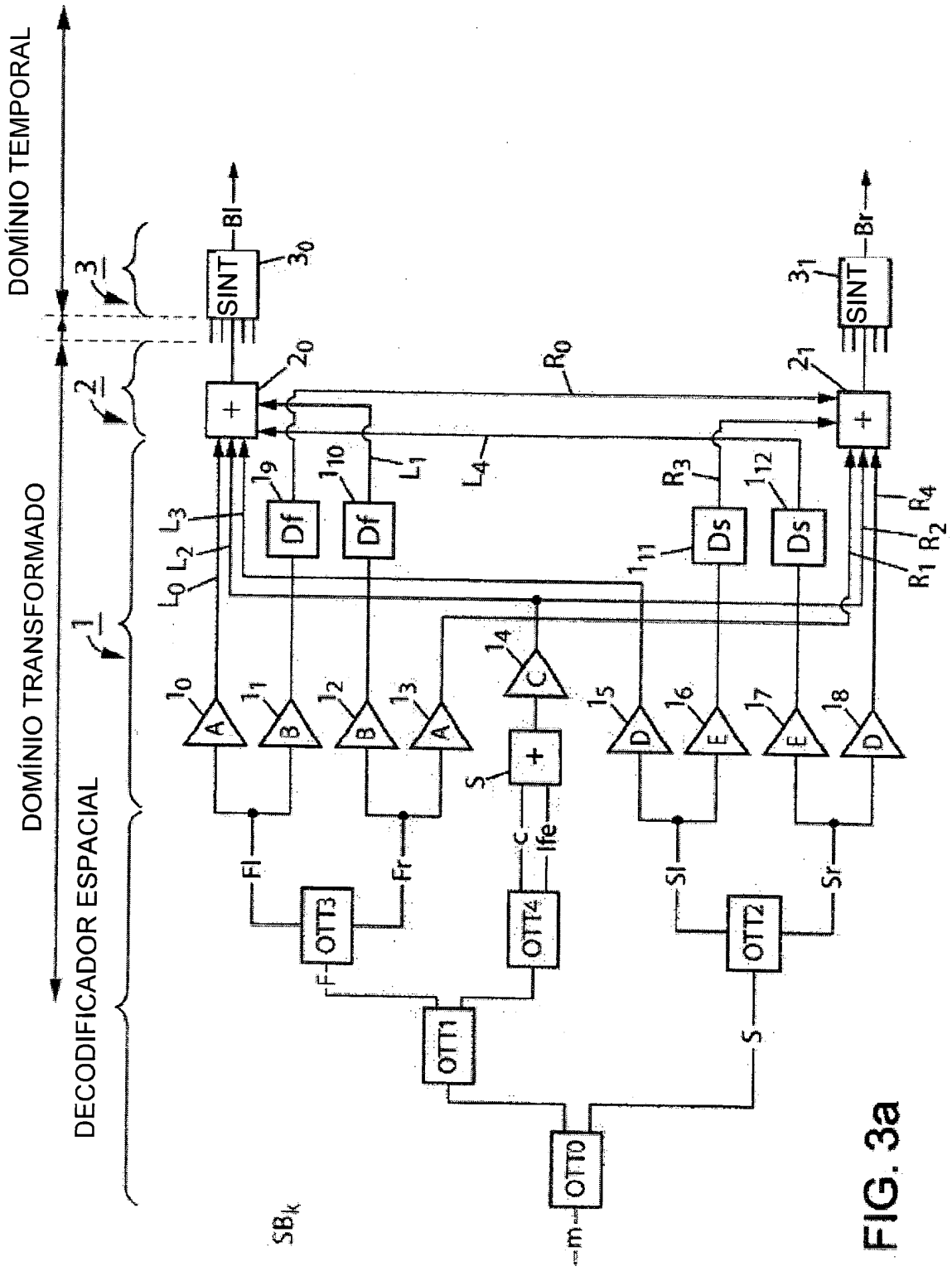


FIG. 3a

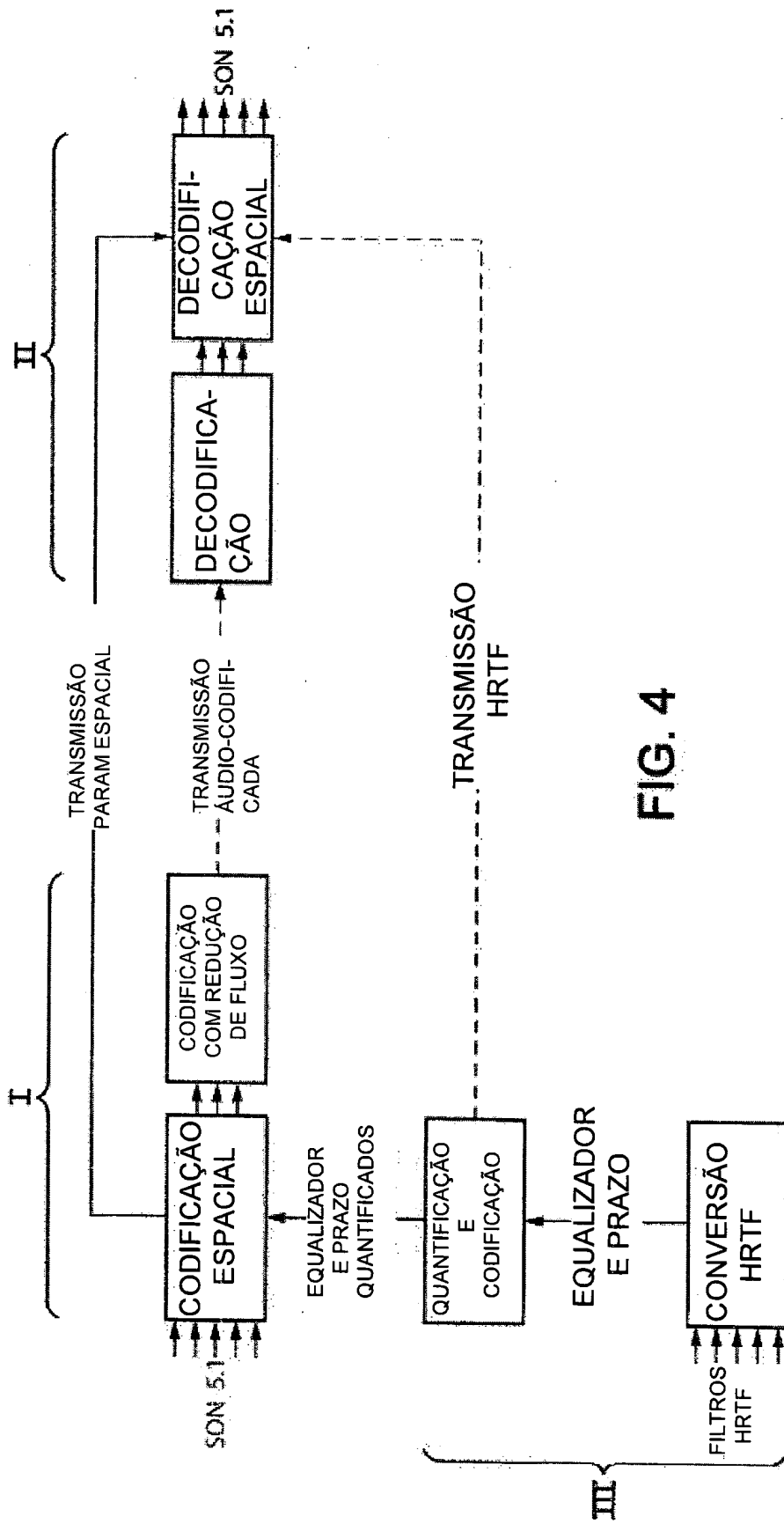


FIG. 4

RESUMO**"PROCESSO E DISPOSITIVO DE ESPACIALIZAÇÃO SONORA BINAURAL  
EFICAZ NO DOMÍNIO TRANSFORMADO"**

A invenção refere-se a um processo e um sistema  
5 de espacialização sonora de um primeiro conjunto de um  
número superior ou igual a um de canais áudio codificados  
em um número de sub-bandas de frequências (SBk) e  
decodificados em um domínio transformado (Fl, C, Fr, Sr,  
Sl, Ife) em um segundo conjunto de um número superior ou  
10 igual a dois (Bl, Br) de canais sonoros no domínio  
temporal, a partir de filtros de modelação convertidos em  
um ganho e um retardo, aplicáveis no domínio transformado,  
efetuando: uma filtragem (A) por retardo do sinal de sub-  
banda por aplicação de pelo menos um ganho e de um retardo  
15 para gerar, a partir de cada um dos canais codificados  
espacialmente, uma componente equalizada e retardada, a  
adição (B) de um sub-conjunto de componentes equalizadas e  
retardadas para criar um número de sinais filtrados no  
domínio transformado correspondente ao número superior ou  
20 igual a dois; a síntese (C) de cada um dos sinais filtrados  
no domínio transformado por um filtro de síntese para obter  
o segundo conjunto de um número superior ou igual a dois de  
canais sonoros de restituição (Bl, Br) no domínio temporal.