



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

(11) PI 1005311-5 B1

(22) Data do Depósito: 05/01/2010

(45) Data de Concessão: 01/12/2020



* B R P I 1 0 0 5 3 1 1 B 1 *

(54) Título: EQUIPAMENTO E MÉTODO PARA MANIPULAR UM SINAL DE ÁUDIO COMPREENDENDO UM EVENTO TRANSIENTE.

(51) Int.Cl.: G10L 19/02; G10L 21/04.

(30) Prioridade Unionista: 05/08/2009 US 61/231,563; 30/01/2009 US 61/148,759; 30/09/2009 EP 09012410.8.

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

(72) Inventor(es): FREDERICK NAGEL; ANDREAS WALTHER; GUILLAUME FUCHS; JÉRÉMIE LECOMTE; HARALD POPP; TILO WIK.

(86) Pedido PCT: PCT EP2010050042 de 05/01/2010

(87) Publicação PCT: WO 2010/086194 de 05/08/2010

(85) Data do Início da Fase Nacional: 29/07/2011

(57) Resumo: APARELHO, MÉTODO E PROGRAMA DE COMPUTADOR PARA MANIPULAR UM SINAL DE ÁUDIO COMPREENDENDO UM EVENTO TRANSIENTE. Aparelho para manipular um sinal de áudio compreendendo um evento transiente que compreende um repositor de sinal transiente configurado para substituir uma parte de sinal transiente, compreendendo o evento transiente do sinal de áudio, com uma parte do sinal de reposição adaptada às características da energia do sinal de uma ou mais partes de sinal transiente do sinal de áudio, ou para as características da energia de sinal da parte de sinal transiente para obter um sinal de áudio com redução de transiente. O equipamento também compreende um processador de sinal configurado para processar o sinal de áudio com redução de transiente para obter a versão processada do sinal de áudio com redução de transiente.

**"EQUIPAMENTO E MÉTODO PARA MANIPULAR UM SINAL DE
ÁUDIO COMPREENDENDO UM EVENTO TRANSIENTE"**

[0001] HISTÓRICO DA INVENÇÃO

[0002] Configurações, de acordo com a invenção, relativas a um equipamento, um método e um programa de computador para manipular um sinal de áudio compreendendo um evento transiente.

[0003] A seguir, serão descritos cenários típicos de aplicação, nos quais podem ser aplicadas as configurações de acordo com a invenção.

[0004] Nos sistemas atuais de processamento de sinal de áudio, os sinais de áudio frequentemente são processados com o uso de técnicas digitais. As partes de sinais específicas como transientes, por exemplo, apresentam exigências especiais mediante o processamento de sinal digital.

[0005] Eventos transientes (ou "transientes") são eventos em um sinal durante os quais a energia do sinal, na banda inteira ou em uma determinada faixa de frequência, é rapidamente alterada, ou seja, sua energia é rapidamente aumentada ou rapidamente diminuída. As características particulares dos transientes específicos (eventos transientes) podem ser encontradas na distribuição da energia do sinal no espectro. Tipicamente, a energia de um sinal de áudio, durante um evento transiente, é distribuída além da frequência inteira, ao passo que em partes de sinal não transiente a energia é normalmente concentrada em uma parte de baixa frequência do sinal de áudio ou em uma ou mais bandas específicas. Isso significa que uma parte de sinal não transiente, que também é denominada parte estacionária

ou de sinal “tonal”, tem um espectro, que é não achatado. Ademais, o espectro da parte de sinal transiente é tipicamente caótico e “não previsível” (por exemplo, ao conhecer um espectro de uma parte de sinal que precede uma parte de sinal transiente), ou seja, a energia do sinal está incluída em um número comparativamente pequeno de linhas espectrais ou bandas espectrais, que são fortemente enfatizadas além de um patamar de ruído de um sinal de áudio. Entretanto, em uma parte transiente, a energia do sinal de áudio será distribuída além de várias bandas diferentes de frequência e, especificamente, será distribuída em uma parte de alta frequência, assim um espectro para a parte transiente do sinal de áudio será comparativamente achatada e tipicamente será mais achatada que um espectro de uma parte tonal do sinal de áudio. No entanto, deve ser observado que há outros tipos de sinais que apresentam um espectro achatado, como, por exemplo, os sinais semelhantes a ruído, cujos sinais não representam um transiente. Contudo, enquanto escaninhos espectrais de sinais semelhantes a ruído apresentam valores de fase não correlacionada ou debilmente correlacionada, frequentemente há uma correlação de fase bastante significativa dos escaninhos espectrais na presença de um transiente.

[0006] Tipicamente, um evento transiente é uma forte alteração em uma representação de domínio de tempo do sinal de áudio, que significa que o sinal incluirá vários componentes de frequência mais alta, quando é realizada uma decomposição de Fourier. Uma característica importante destes vários harmônicos mais altos é que as fases destes harmônicos mais altos estão em uma relação mútua bastante específica, assim a sobreposição de

todos os harmônicos resultará em uma rápida alteração de energia de sinal (quando considerado no domínio de tempo), ou seja, existe uma forte correlação no espectro, na proximidade de um evento transiente. A situação de fase específica entre todos os harmônicos pode ser denominada como “coerência vertical”. Esta “coerência vertical” está relacionada a uma representação de espectrograma de tempo/frequência do sinal, onde um sentido horizontal se refere a uma evolução do sinal além do tempo e onde uma dimensão vertical descreve a dependência além da frequência dos componentes espectrais, em um espectro de tempo curto, além da frequência.

[0007] Se, por exemplo, forem realizadas alterações além dos grandes domínios de tempo, por exemplo, por quantização, as referidas alterações influenciarão o bloco inteiro. Uma vez que os transientes são caracterizados por um aumento curto prazo na energia, esta energia provavelmente será indistinta, quando o bloco é alterado, em toda a região representada pelo bloco.

[0008] O problema também se torna particularmente evidente quando a velocidade de reprodução de um sinal é alterada, ao passo que o tom é mantido, ou quando o sinal é transposto, enquanto a duração original da reprodução é mantida. Isso pode ser obtido usando-se um vocoder de fase ou um método como (P)SOLA (consulte as referências [A1] a [A4] relativas a este assunto). Este último é atingido por meio da reprodução do sinal estendido, acelerado pelo fator de extensão de tempo. Com a representação de sinais tempo-discreto, isso se refere a efetuar o *downsampling* do sinal pelo fator de extensão, ao passo que é mantida a frequência de ilustragem. Os métodos de extensão de tempo, como o vocoder de

fase são, de fato, adequados somente para os sinais estacionários ou quase-estacionários, uma vez que os transientes são “indistintos” no tempo pela dispersão. O vocoder de fase prejudica as chamadas propriedades de coerência vertical (relativas a uma representação de espectrograma de tempo/frequência) do sinal.

[0009] A extensão de tempo dos sinais de áudio desempenha um importante papel no entretenimento e nas artes. Os algoritmos comuns baseiam-se em técnicas de *overlap and add* (OLA), como vocoder de fase (VF), *Synchronous Overlap Add* (SOLA), *Pitch Synchronous Overlap Add* (PSOLA) e *Waveform Similarity Overlap Add* (WSOLA). Embora estes algoritmos sejam capazes de alterar a velocidade de repetição dos sinais de áudio, apesar de preservar o tom original, os transientes não são bem preservados. A extensão de tempo de um sinal de áudio, sem alterar seu tom, com o uso de OLA, exige o processamento separado de transientes e as partes de sinal mantido, para evitar a dispersão transiente [B1] e o *aliasing* do domínio de tempo que frequentemente ocorre com WSOLA e SOLA. Um desafio é apresentado pela tarefa de estender uma combinação de um sinal bastante tonal, como um diapasão e um sinal percussivo, como as castanholas.

[00010] A seguir, serão mencionadas algumas das abordagens convencionais para prover o retrospecto da presente invenção.

[00011] Alguns métodos atuais estendem o tempo dos transientes de modo mais intenso, assim não há ou há apenas uma pequena extensão de tempo na duração do transiente (consulte, por exemplo, as referências [5] a [8]).

[00012] Os seguintes artigos e patentes descrevem métodos

de manipulação de tempo e/ou tom: [A1], [A2], [A3], [A4], [A5], [A6], [A7], [A8].

[00013] Em [B2] um método é proposto que, de modo aproximado, o envelope de um sinal, na versão de tempo estendido, é preservado, bem como suas características espectrais. Esta abordagem prevê um evento percussivo de tempo ampliado para diminuir de modo mais lentamente que o original.

[00014] Diversos métodos amplamente conhecidos permitem um processamento diferenciado de transientes e de componentes de sinal estacionário, por exemplo, a modelação de um sinal como soma de senos, transientes e ruído (S+T+N) [B4, B5]. Para preservar os transientes depois de uma modificação de escala de tempo, as três partes são separadamente estendidas. Esta técnica é capaz de preservar perfeitamente os componentes transientes de sinais de áudio. Entretanto, o som resultante frequentemente é percebido como não natural.

[00015] As abordagens adicionais podem variar a quantidade de extensão de tempo e estabelecê-la em uma, durante o tempo transiente ou bloquear a fase no evento transiente [B3, B6, B7].

[00016] O documento [B8] demonstra como os transientes podem ser preservados na extensão de tempo e frequência com o VF. Na referida abordagem, os transientes são cortados do sinal, antes que ele seja estendido. A retirada de partes transientes resultou em lacunas dentro do sinal, que foram estendidas pelo processo VF. Depois da extensão, os transientes foram novamente adicionados ao sinal, com um *surrounding* que ajustou as lacunas estendidas.

[00017] Em vista do acima apresentado, há necessidade de

um conceito de manipulação de um sinal de áudio compreendendo um evento transiente que é provido para um sinal de saída de qualidade melhorada percebida.

[00018] SUMÁRIO DA INVENÇÃO

[00019] Configuração, de acordo com a invenção, que cria um equipamento para manipular um sinal de áudio compreendendo um evento transiente. O equipamento compreende um repositor de sinal transiente configurado para substituir uma parte de sinal transiente, compreendendo o evento transiente, do sinal de áudio com uma parte do sinal de reposição, adaptado a características da energia do sinal de uma ou mais partes de sinal não transiente do sinal de áudio ou para uma característica da energia do sinal da parte de sinal transiente, para obter um sinal de áudio com redução de transiente. O equipamento compreende ainda um processador de sinal configurado para processar o sinal de áudio com redução de transiente, para obter uma versão processada do sinal de áudio com redução de transiente. O equipamento também compreende um reinsensor de sinal transiente configurado para combinar a versão processada do sinal de áudio com redução de transiente com uma representação de sinal transiente, em uma forma original ou processada, um teor transiente da parte de sinal transiente.

[00020] A configuração acima descrita está baseada no achado de que o processador de sinal provê um sinal de saída de qualidade melhorada, se a parte de sinal transiente é substituída por uma parte do sinal de reposição, uma energia de sinal da qual é adaptada às características da energia do sinal do sinal de áudio original, enquanto reduz ou elimina o evento transiente.

Este conceito evita grandes alterações passo a passo da energia da entrada de sinal para o processador de sinal, que poderia ser causada, simplesmente, pela eliminação da parte de sinal transiente do sinal de áudio, bem como evita também ou, no mínimo, reduz o efeito deletério de um transiente no processador de sinal.

[00021] Assim, por meio da eliminação ou redução do evento transiente no sinal de áudio (para obter o sinal de áudio com redução de transiente) e por meio da limitação de uma alteração da energia de sinal de áudio com redução de transiente, quando comparado ao sinal de entrada de áudio, o processador de sinal recebe um sinal de entrada adequado, assim seu sinal de saída se aproxima de um sinal de saída desejado, na ausência de um evento transiente.

[00022] Em uma configuração preferida, o repositor de sinal transiente está configurado para prover a parte de sinal de reposição (ou parte de sinal com redução de transiente), assim a parte do sinal de reposição representa um sinal de tempo que apresenta uma evolução temporal suavizada, quando comparada a uma parte de sinal transiente e assim, um desvio entre a energia da parte do sinal de reposição e uma energia de uma parte de sinal não transiente, do sinal de áudio que precede a parte de sinal transiente ou após a parte de sinal transiente é menor que um valor de limiar pré-determinado. Neste modo, pode ser obtido que a parte do sinal de reposição atende duas condições, assim denominadas “condição transiente” e “condição de energia”. A condição transiente indica que um evento transiente, que é representado por um passo ou pico em um domínio de tempo, está limitado na intensidade (ou altura do passo ou altura do pico)

dentro da parte do sinal de reposição. A condição de energia indica adicionalmente que o sinal de áudio com redução de transiente (da parte do sinal de reposição) deve ter uma evolução temporal suave da distribuição de energia espectral. As descontinuidades na evolução temporal da distribuição de energia espectral tipicamente resultam na geração de artefatos audíveis. Adequadamente, por meio da limitação das referidas descontinuidades temporais da distribuição de energia espectral, os artefatos audíveis podem ser evitados, o que poderia resultar de uma mera exclusão (sem reposição) de uma parte de sinal transiente, a partir do sinal de áudio de entrada.

[00023] Em uma configuração preferida, o repositor de sinal transiente está configurado para extrapolar os valores de amplitude de uma ou mais partes de sinal que precede a parte de sinal transiente, para obter os valores de amplitude da parte do sinal de reposição. O repositor de sinal transiente também está configurado para extrapolar valores de fase de uma ou mais partes de sinal que precede a parte de sinal transiente para obter valores de fase da parte do sinal de reposição. Com o uso desta abordagem, também pode ser obtida uma evolução de amplitude suave do sinal de áudio com redução de transiente. Ademais, as fases dos componentes espectrais diferentes do sinal de áudio com redução de transiente são bem controladas (por meio de extrapolação), assim o evento transiente, que é caracterizado por valores de fase específica durante a parte de sinal transiente (diferente dos valores de fase de partes de sinal não transiente), é suprimido.

[00024] Ou seja, os valores de fase são reforçados por meio de extrapolação, que são gerados de modo diferente dos

valores de fase que caracterizam o transiente. A extrapolação também provê a vantagem de que o conhecimento das partes de sinal de áudio que precede a parte de sinal transiente seja suficiente para realizar a extrapolação. Entretanto, é naturalmente possível aplicar adicionalmente alguma informação colateral, por exemplo, parâmetros de extrapolação, para realizar a extrapolação.

[00025] Em outra configuração preferida, o reinsersor de sinal transiente (150) está configurado para efetuar o desvanecimento cruzado da versão processada do sinal de áudio com redução de transiente com a representação de sinal transiente, em uma forma original ou processada, um teor transiente da parte de sinal transiente. Neste caso, a versão processada do sinal com redução de transiente pode ser uma versão com tempo estendido do sinal de áudio de entrada. Adequadamente, o transiente pode ser suavemente reinserido em uma versão estendida do sinal de áudio de entrada, ou seja, depois da extensão de (tempo) do sinal de áudio com redução de transiente, os transientes (na forma processada ou não processada) são novamente adicionados ao sinal com um *surrounding* que se ajusta às lacunas estendidas.

[00026] Em outra configuração preferida, o repositor de sinal transiente está configurado para interpolar entre um valor de amplitude, de uma parte de sinal que precede a parte de sinal transiente, e um valor de amplitude de uma parte de sinal que segue a parte de sinal transiente para obter um ou mais valores de amplitude da parte do sinal de reposição. O repositor de sinal transiente está, ademais, configurado para interpolar entre um valor de fase, de uma parte de sinal que precede a parte de sinal transiente, e um valor de fase de uma parte de sinal que segue a

parte de sinal transiente, para obter um ou mais valores de fase da parte do sinal de reposição. Por meio da realização da interpolação, é possível obter uma evolução temporal particularmente suave dos valores de amplitude e de fase. A interpolação da fase também resulta, tipicamente, em uma redução ou cancelamento do evento transiente, conforme os transientes tipicamente compreendem uma distribuição de fase bastante específica na proximidade direta do transiente, cuja distribuição de fase é tipicamente diferente da distribuição de fase em um determinado espaço distante do transiente.

[00027] Em uma configuração preferida, o repositor de sinal transiente está configurado para aplicar um ruído ponderado (por exemplo, um espectro de sinal semelhante a ruído, adaptado às características da energia do sinal de uma ou mais partes de sinal não transiente do sinal de áudio, ou para uma característica da energia do sinal da parte de sinal transiente) para obter os valores de amplitude da parte do sinal de reposição e para aplicar um ruído ponderado para obter os valores de fase da parte do sinal de reposição. É possível, por meio da aplicação de um ruído ponderado, reduzir adicionalmente o transiente, enquanto o impacto na energia é mantido de modo suficientemente pequeno.

[00028] Em uma configuração preferida, o repositor de sinal transiente está configurado para combinar componentes não transientes da parte de sinal transiente com os valores extrapolados ou interpolados para obter a parte do sinal de reposição. Foi descoberto que uma qualidade melhorada do sinal de áudio com redução de transiente (e de sua versão processada, que é obtida usando-se o processador de sinal) pode ser atingida, se

forem mantidos os componentes não transientes da parte de sinal transiente. Por exemplo, os componentes tonais da parte de sinal transiente podem ter somente um impacto limitado no transiente (uma vez que um transiente temporal é tipicamente provocado por um sinal de banda larga que tem uma distribuição de fase específica sobre a frequência). Assim, os componentes não transientes tonais da parte de sinal transiente podem carregar uma informação preciosa que, de fato, pode favorecer um sinal de saída desejável do processador de sinal. Assim, mantendo as referidas partes de sinal - enquanto é reduzido o transiente - é possível favorecer um aperfeiçoamento do sinal de áudio processado.

[00029] Em uma configuração da invenção, o repositor de sinal transiente está configurado para obter partes do sinal de reposição de comprimento variável, conforme o comprimento de uma parte de sinal transiente. Foi descoberto que a qualidade do sinal de áudio pode, às vezes, ser aperfeiçoada por meio da adaptação do comprimento de partes do sinal de reposição para um comprimento variável das partes de sinal transiente. Por exemplo, em alguns sinais, as partes de sinal transiente podem ter uma duração bastante curta. Neste caso, um sinal de áudio processado otimizado pode ser obtido pela reposição de somente uma parte relativamente curta do sinal de áudio de entrada. Assim, o máximo possível de informação (não transiente) do sinal de áudio de entrada original pode ser mantido. Ademais, mantendo também as partes do sinal de reposição curtas (em conformidade com o comprimento da parte de sinal transiente), é possível evitar uma sobreposição das partes do sinal de reposição, em várias situações. Portanto, na maioria dos casos, é possível obter que há uma parte do sinal não

transiente original entre duas partes do sinal de reposição subsequentes. Consequentemente, o sinal de áudio processado é gerado com precisão suficiente, mantendo o máximo possível de informação (não transiente) do sinal de áudio de entrada original.

[00030] Em uma configuração preferida, o processador de sinal está configurado para processar o sinal de áudio com redução de transiente, assim uma parte de sinal temporal dada da versão processada do sinal de áudio com redução de transiente é dependente de uma pluralidade de partes de sinal temporal temporariamente sem sobreposição do sinal de áudio com redução de transiente, ou seja, é preferido que o processador de sinal compreenda a memória temporal ao gerar as partes de sinal da versão processada do sinal de áudio com redução de transiente. O processamento de sinal com o uso de uma memória permite um processo em bloco do sinal de áudio com redução de transiente ou para uma filtragem temporal (por exemplo, filtragem FIR ou filtragem IIR) do sinal de áudio com redução de transiente. Também foi descoberto que o conceito da invenção de reposição das partes de sinal transiente é bastante bem adaptado para o trabalho aliado ao referido processador de sinal. Embora os transientes normalmente possam apresentar um impacto negativo significativo no processador de sinal descrito, realizando um processamento em bloco ou apresentando uma memória temporal, as partes do sinal de reposição da invenção reduzem este efeito deletério do transiente. Embora um transiente normalmente possa apresentar um impacto em múltiplas partes de sinal providas pelo processador de sinal - estendendo-se além dos limites temporais da parte de sinal transiente - o efeito deletério de um transiente é reduzido ou,

inclusive, eliminado por meio do conceito da invenção. Por meio da manutenção de uma evolução temporal suave da energia do sinal com redução de transiente, qualquer degradação pode ser mantida suficientemente suave. Por exemplo, um bloco (do processamento em bloco do processador de sinal), que compreende uma parte do sinal de reposição (por exemplo, além da parte de sinal não transiente original) não é gravemente degradado, uma vez que a parte do sinal de reposição é adaptada à energia para o restante do bloco. Assim, o bloco em sua totalidade somente é discretamente afetado pela eliminação ou redução do evento transiente. Ademais, uma filtragem temporal, que poderia ser negativamente afetada por um evento transiente, bem como por uma eliminação completa (por exemplo, na forma de um reforço igual a zero) da parte do sinal transiente, é mantida praticamente não afetada pela eliminação (ou redução) do transiente devido ao uso de uma parte do sinal de reposição.

[00031] Em uma configuração preferida, o processador de sinal está configurado para realizar um processamento com base em bloco de tempo do sinal de áudio com redução de transiente para obter a versão processada do sinal de áudio com redução de transiente. O repositor de sinal transiente também está configurado para ajustar a duração da parte do sinal a ser substituído pela parte do sinal de reposição com uma resolução temporal que é melhor que a duração de um bloco de tempo ou para substituir uma parte de sinal transiente que apresenta uma duração temporal menor que a duração do bloco de tempo com uma parte do sinal de reposição que apresenta uma duração menor que a duração do bloco de tempo. Assim, a reposição ora sugerida permite um processamento de baixa distorção dos sinais de áudio, inclusive se

o comprimento das partes de transiente eliminadas for diferente do comprimento do bloco de tempo.

[00032] Em uma configuração preferida, o processador de sinal está configurado para processar o sinal de áudio com redução de transiente em um modo dependente da frequência, assim o processamento introduz mudanças de fase dependentes da frequência degradante de transiente no sinal de áudio com redução de transiente. Entretanto, mesmo o referido processamento de sinal degradante de transiente não tem um impacto deletério significativo no sinal de áudio processado, uma vez que os transientes são tipicamente processados separadamente do processamento do sinal de áudio com redução de transiente. Adequadamente, apesar de que um algoritmo de processamento de um sinal degradante transiente possa ser aplicado no processador de sinal, a qualidade dos transientes pode ser mantida com o uso de um processamento separado do transiente e por uma reinserção dos transientes, em uma etapa posterior do processamento.

[00033] Em uma configuração preferida, o repositório de sinal transiente compreende um detector de transientes, onde o detector de transiente está configurado para prover um limite de detecção com variação de tempo para a detecção do transiente no sinal de áudio, assim o limite de detecção segue um envelope do sinal de áudio em uma constante de tempo de ajuste suave. O detector de transiente está configurado para alterar a constante de tempo suave em resposta à detecção de um transiente e/ou na dependência de uma evolução temporal do sinal de áudio. Pelo uso do referido detector de transiente, é possível detectar transientes de intensidades diferentes, inclusive se os

transientes estão estreitamente espaçados no tempo. Por exemplo, o conceito da invenção permite a detecção de um transiente fraco, inclusive se o transiente fraco segue estreitamente um transiente mais forte precedente. Adequadamente, a detecção de transiente para a reposição de transiente pode ser realizada de modo confiável e preciso.

[00034] Em uma configuração preferida, o equipamento compreende um processador transiente configurado para receber uma informação transiente que representa o teor transiente da parte do sinal transiente. Neste caso, o processador transiente pode ser configurado para obter, com base na informação transiente, um sinal transiente processado no qual os componentes tonais são reduzidos. O reinsersor de sinal transiente pode ser configurado para combinar a versão processada do sinal de áudio com redução de transiente a um sinal transiente processado provido pelo processador transiente. Assim, o processamento separado do sinal de áudio com redução de transiente e o componente transiente do sinal de áudio de entrada (representado pela informação transiente) podem ser realizados de modo que uma combinação subsequente das partes de sinal diferente resulte em um sinal de saída geral adequado. Estes componentes de sinal da parte de sinal transiente, que foram processados pelo processador de sinal “principal” (por exemplo, componentes de sinal) não necessitam ser incluídos no processamento separado do transiente. Adequadamente, é possível realizar o compartilhamento adequado do processamento dos componentes de áudio da parte de sinal transiente.

[00035] As configurações adicionais, de acordo com a invenção, criam um método e um programa de computador para

manipular um sinal de áudio compreendendo um evento transiente.

[00036] BREVE DESCRIÇÃO DAS FIGURAS

[00037] As configurações, de acordo com a invenção, serão descritas a seguir tomando-se como referência as figuras anexas, nas quais:

[00038] A Figura 1 ilustra um diagrama de blocos esquemático de um equipamento para manipular um sinal de áudio compreendendo um evento transiente, de acordo com uma configuração da presente invenção;

[00039] A Figura 2 ilustra um diagrama de blocos esquemático de um repositório de sinal transiente, de acordo com uma configuração da presente invenção;

[00040] As Figuras 3a a 3d ilustram diagramas de blocos esquemáticos de um processador de sinal, de acordo com as configurações da presente invenção;

[00041] A Figura 4 ilustra um diagrama de blocos esquemático de um reinseridor de sinal transiente, de acordo com uma configuração da presente invenção;

[00042] A Figura 5a ilustra uma visão geral da implementação de um vocoder a ser utilizado no processador de sinal da Figura 1;

[00043] A Figura 5b ilustra uma implementação de partes (análise) de um processador de sinal da Figura 1;

[00044] A Figura 5c ilustra outras partes (extensão) de um processador de sinal da Figura 1;

[00045] A Figura 6 ilustra uma implementação de transformada de um vocoder de fase a ser utilizada no processador de sinal da Figura 1;

[00046] A Figura 7 ilustra uma representação esquemática da operação de um algoritmo de vocoder de fase com o tamanho da grade de tempo de síntese sendo diferente do tamanho da grade de tempo de análise, por exemplo, por um fator de 2;

[00047] A Figura 8 ilustra uma representação gráfica de uma evolução temporal da amplitude de um sinal de áudio;

[00048] A Figura 9 ilustra uma representação gráfica de um *timing* do processamento de sinal no equipamento da Figura 1;

[00049] A Figura 10 ilustra uma representação gráfica de sinais que podem aparecer em um equipamento, de acordo com a Figura 1;

[00050] A Figura 11 ilustra outra representação gráfica de sinais que podem aparecer em um equipamento, de acordo com a Figura 1;

[00051] A Figura 12 ilustra um fluxograma de um método para manipular um sinal de áudio, de acordo com uma configuração da presente invenção;

[00052] A Figura 13 ilustra uma representação gráfica de uma eliminação de transiente e interpolação, de acordo com a configuração da invenção;

[00053] A Figura 14 ilustra uma representação gráfica de uma extensão de tempo e da reinserção de transiente, de acordo com uma configuração da invenção;

[00054] A Figura 15 ilustra uma representação gráfica de formas de onda de sinal que ocorrem em diferentes etapas ao tratar o transiente da invenção, em uma aplicação de extensão de tempo com o vocoder de fase; e

[00055] A Figura 16 ilustra uma representação gráfica dos

sinais, que estão presentes em diferentes etapas de uma extensão de tempo.

[00056] DESCRIÇÃO DETALHADA DAS CONFIGURAÇÕES

[00057] A seguir, serão descritas algumas configurações, de acordo com a invenção. Uma primeira configuração de um equipamento para manipular um sinal de áudio que compreende um evento transiente será descrita com referência à Figura 1, que ilustra uma visão geral da primeira configuração, também com referência às Figuras 2, 3a a 3c, 4, 5a, 5b, 5c, 6 e 7, que ilustram detalhes dos componentes da primeira configuração e a operação do vocoder de fase (Figura 7). Um sinal transiente é ilustrado na Figura 8, e o seu processamento está ilustrado nas Figuras de 9 a 11. A Figura 12 ilustra um fluxograma de um método respectivo.

[00058] De modo subsequente, a operação de uma segunda configuração de um equipamento para manipular um sinal de áudio, compreendendo um evento transiente, será descrito tomando-se como referência as Figuras 13 a 17.

[00059] Configuração de acordo com a Figura 1

[00060] A Figura 1 ilustra um diagrama de blocos esquemático de um equipamento para manipular um sinal de áudio compreendendo um evento transiente, de acordo com uma configuração da invenção. O equipamento ilustrado na Figura 1 está desenhado em sua totalidade com 100. O equipamento 100 está configurado para receber um sinal de áudio 110 que compreende um evento transiente, e para prover, com base neste, um sinal de áudio processado 120 com um transiente não processado "natural" ou sintetizado. O equipamento 100 compreende um repositor de sinal transiente 130

configurado para substituir uma parte de sinal transiente, compreendendo o evento transiente do sinal de áudio 110, com uma parte do sinal de reposição adaptada a características da energia do sinal, de uma ou mais partes de sinal não transiente do sinal de áudio ou para uma característica da energia do sinal da parte de sinal transiente, para obter um sinal de áudio com redução de transiente 132. Opcionalmente, as características de fase da parte do sinal de reposição podem ser adaptadas a características de fase de uma ou mais partes de sinal não transiente do sinal de áudio. O equipamento 100 compreende ainda um processador de sinal 140 configurado para processar o sinal de áudio com redução de transiente 132, para obter uma versão processada 142 do sinal de áudio com redução de transiente. O equipamento 100 compreende ainda um reinsertor de sinal transiente 150 configurado para combinar a versão processada 142 do sinal de áudio com redução de transiente com um sinal transiente 152 para obter o sinal de áudio processado 120 com transiente não processado "natural" ou sintetizado. O sinal transiente 152 pode representar, em uma forma original ou processada, um teor transiente da parte de sinal transiente, que foi substituída com a parte de sinal de reposição por meio de um repositor de sinal transiente 130.

[00061] O repositor de sinal transiente 130 pode, ainda, de modo opcional, prover uma informação transiente 134 que representa o teor transiente da parte de sinal transiente (que é substituída pela parte do sinal de reposição no sinal de áudio com redução de transiente 132). Adequadamente, a informação transiente 134 pode servir para "salvar" o teor transiente do sinal de áudio 110, que é reduzido ou, inclusive, completamente suprimido no

signal de áudio com redução de transiente 132. A informação transiente 134 pode ser diretamente encaminhada ao reinsensor de signal transiente 150, para atuar como signal transiente 152. Contudo, o equipamento 100 pode compreender ainda um processador de transiente opcional 160, que está configurado para processar a informação transiente 134, para derivar a partir dela o signal transiente 152. Por exemplo, o processador de transiente 160 pode ser configurado para realizar uma transposição de frequência transiente, uma mudança de frequência transiente ou uma síntese transiente.

[00062] O equipamento 100 pode compreender ainda, de modo opcional, um condicionador de signal 170 configurado para condicionar o signal de áudio processado 120 para obter um signal de áudio condicionado para reprodução.

[00063] Em relação à funcionalidade do equipamento 100, pode-se afirmar, de modo geral, que o equipamento 100 permite um processamento separado de um teor de áudio não transiente do signal de áudio 110 (representado pelo signal de áudio com redução de transiente 132) e de um teor de áudio transiente do signal de áudio 110 (representado pela informação transiente 134). Os eventos transientes são reduzidos ou, inclusive, suprimidos, no signal de áudio com redução de transiente 132, assim o processador de signal 140 pode realizar um processamento de signal que poderia degradar os eventos transientes e/ou que poderiam se afetados, de modo deletério, pelos eventos transientes. Entretanto, por meio da substituição de partes de signal transiente por partes do signal de reposição adaptado à energia, o repositor de signal transiente 130 atua para evitar artefatos audíveis, que poderiam ser introduzidos

pelo processador de sinal 140, se as partes de sinal transientes fossem simplesmente estabelecidas em zero.

[00064] Uma impressão de audição adequada também é obtida usando-se uma reinserção transiente por meio do reinsensor de sinal transiente 150. Certamente, a impressão de audição tipicamente poderia ser gravemente degradada, se os eventos transientes fossem simplesmente eliminados. Por esta razão, os transientes são reinseridos no sinal de áudio processado 142. Os transientes reinseridos podem ser idênticos aos transientes eliminados do sinal de áudio 110 pelo repositor de sinal transiente 130. Alternativamente, é possível realizar um processamento dos referidos transientes eliminados (ou substituídos), por exemplo, na forma de uma transposição de frequência ou comutação de frequência. Contudo, em algumas configurações, os transientes reinseridos podem, inclusive, ser sinteticamente gerados, por exemplo, com base nos parâmetros transientes que descrevem o tempo e a intensidade dos transientes a serem reinseridos.

[00065] Detalhes do repositor de sinal transiente

[00066] A seguir, a funcionalidade do repositor de sinal transiente 130 será descrita tomando-se como referência a Figura 2, onde a Figura 2 ilustra um diagrama de blocos esquemático de uma configuração do repositor de sinal transiente 130. O repositor de sinal transiente 130 recebe o sinal de áudio 110 e provê, com base nele, o sinal de áudio com redução de transiente 132.

[00067] Para esta finalidade, o repositor de sinal transiente 130 pode, por exemplo, compreender um detector transiente 130a, que está configurado para detectar um transiente

e para prover uma informação sobre um *timing* do transiente. Por exemplo, o detector de transiente 130a pode prover uma informação 130b que descreve o tempo de início e o tempo de término de uma parte de sinal transiente. Conceitos diferentes para a detecção de transiente são conhecidos na técnica, assim uma descrição detalhada será ora omitida. Entretanto, em alguns casos, o detector transiente 130a pode ser configurado para diferenciar transientes de comprimento diferente, assim o comprimento de uma parte de sinal transiente reconhecido pode variar conforme a forma real do sinal.

[00068] Alternativamente, o repositor de sinal transiente pode compreender um extrator de informação colateral 130c, por exemplo, se uma informação colateral que descreve um *timing* de transientes estiver associada ao sinal de áudio 110. Neste caso, o detector de transiente 130a pode, naturalmente, ser omitido. O extrator de informação colateral 130c pode ainda, opcionalmente, ser configurado para prover um ou mais parâmetros de interpolação, parâmetros de extrapolação e/ou parâmetros de reposição, com base na informação colateral associada ao sinal de áudio 110. O repositor de transiente 130 compreende ainda um repositor de parte transiente 130d, por exemplo, um interpolador de parte transiente ou um extrapolador de parte transiente. O repositor de parte transiente 130e está configurado para receber o sinal de áudio 110 e a informação de tempo transiente 130b (provida pelo detector de transiente 130a ou pelo extrator de informação colateral 130c) e para repor uma parte transiente do sinal de áudio 110 pela parte do sinal de reposição.

[00069] A seguir, serão descritos detalhes referentes à

detecção e à reposição (ou eliminação) de transientes. Em particular, serão discutidos em detalhes métodos diferentes para a eliminação de transientes.

[00070] Os transientes (por exemplo, o início de um instrumento ou sinais percussivos) podem, geralmente, ser descritos como um intervalo curto de tempo durante o qual o sinal se desenvolve rapidamente, de modo imprevisível. Por exemplo, um transiente pode ser detectado (com o uso do detector de transiente 130a) pela avaliação de uma representação de domínio de tempo do sinal de áudio 110. Se a representação de domínio de tempo do sinal de áudio 110 excede um limite (que pode variar conforme o tempo), então é possível indicar a presença de um evento transiente. Uma região temporal que compreende o evento transiente pode ser considerada uma parte de sinal transiente e pode ser descrita pela informação de tempo transiente 130b.

[00071] Uma vez que as partes de sinal (ou seja, transientes, ou intervalos de tempo durante os quais o sinal se desenvolve rapidamente de modo imprevisível), são idealmente não estendidas no tempo, é vantajoso eliminar “um período de tempo transiente” a partir do sinal anterior à extensão de tempo (que pode ser realizado pelo processador de sinal 140). A supressão pode ocorrer durante o todo o período de tempo que é considerado “não estacionário”. Para instrumentos percussivos, este período de tempo é constituído, em sua maior parte, pelo evento sonoro completo (ou seja, batida HiHat simples). Para o início de um instrumento, um envelope denominado ADSR (*Attack Decay Sustain Release*) pode servir para ilustrar o período de tempo transiente.

[00072] A Figura 8 ilustra uma representação gráfica 800

de uma evolução temporal de uma amplitude de sinal. Uma abscissa 810 descreve um tempo e uma ordena 812 descreve uma amplitude. Uma curva 814 descreve uma evolução temporal da amplitude. Conforme pode ser observado a partir da Figura 8, a evolução temporal da amplitude compreende um intervalo de ataque, um intervalo de queda, um intervalo de manutenção e um intervalo de liberação. O intervalo de ataque e o intervalo de queda podem, por exemplo, ser considerados uma “região transiente” ou parte de sinal transiente.

[00073] Contudo, foi descoberto que para o processamento de sinal adicional (por exemplo, no processador de sinal 140), a lacuna no sinal de áudio, que é provocada pela supressão transiente, deve ser preenchida, assim ao ser ouvido o sinal processado (=sinal de síntese) (por exemplo, processado com o uso do processador de sinal 140), há uma sensação auditiva de um sinal contínuo, transiente e livre, sem pausa de interrupção e modulações de amplitude.

[00074] Para o caso específico de aplicação ora descrita, é preferido suprimir todas as partes transientes do sinal original (por exemplo, sinal 110) no sinal de síntese (por exemplo, no sinal 132 provido ao processador de sinal 140 ou, conseqüentemente, no sinal 142 provido pelo processador de sinal 140), onde continuam a existir as partes tonais e componentes de ruído não transientes.

[00075] Sobre este assunto, há várias abordagens que já existem, mas um objetivo deste é que nunca é um sinal de transiente ajustado de alta qualidade (ou transiente purgado). Em relação a esta questão, a referência é feita na publicação [Edler], por exemplo.

[00076] Em relação à eficiência dos métodos de detecção de transiente e à decomposição em vários componentes, como, por exemplo, “transientes+ruído”, podem ser assumidas as seguintes conclusões, a partir de publicações especializadas específicas [Bello] e [Daudet], que provêem um bom panorama geral dos métodos comuns: nenhum dos métodos é claramente superior aos demais; a seleção deve ser orientada pela aplicação respectiva e pela capacidade de processamento disponível.

[00077] Segue-se que a seleção de detecção específica e métodos de decomposição pode influenciar significativamente o resultado do método da invenção. Para os especialistas na técnica, é prontamente possível aplicar qualquer um dos diversos métodos conhecidos, para assim prover a melhor condição possível para o cenário de aplicação respectivo.

[00078] Conceitos para substituição de parte transiente

[00079] Alguns cenários de aplicação tratam da geração de partes de sinal que não necessitam ser avaliadas como “certas” ou “erradas” pela verificação com um sinal de referência, mas somente com base nos respectivos sons ideais gerais. Isso significa que as configurações, de acordo com a invenção, não estão limitadas à separação de partes e para omitir os componentes transientes, mas podem gerar para si mesmos sinais de síntese que apresentam propriedades específicas.

[00080] A geração de sinal de síntese (por exemplo, geração de um sinal com redução de transiente 132 pelo repositório de sinal transiente 132d) pode, portanto, ser uma combinação de decomposição de sinal e de geração de sinal (no sentido de uma interpolação e/ou extrapolação do sinal assumido) durante o

período de tempo transiente. Os componentes não transientes do sinal original podem ser misturados aos componentes interpolados/extrapolados ou podem substituí-los.

[00081] Em algumas configurações, de acordo com a presente invenção, a extrapolação pode ser igual a uma geração de sinal de síntese com o uso de valores anteriores. Adequadamente, a extrapolação pode ser possível em tempo real. Por outro lado, em algumas configurações, a interpolação pode ser igual a uma geração de sinal de síntese com o uso de valores precedentes e subsequentes. Assim, em alguns casos, a interpolação pode exigir uma visão antecipada.

[00082] Para resumir o acima exposto, é possível aplicar diferentes conceitos no repositor de parte transiente 130d para obter o sinal de áudio com redução de transiente 132.

[00083] Por exemplo, o repositor de parte transiente 130d pode ser configurado para reduzir os componentes transientes do sinal de áudio 110, para obter o sinal de áudio com redução de transiente. Neste caso, o repositor de parte transiente 130d pode ser configurado para garantir que a energia suficiente permaneça na parte do sinal de reposição, assumindo o lugar da parte de sinal transiente. Por exemplo, os componentes de frequência, que compreendem uma característica de fase transiente, podem ser eliminados do sinal de áudio 110, ao passo que demais componentes de frequência, que não compreendem a característica de fase transiente, (por exemplo, componentes de frequência tonal) podem ser transferidos da parte de sinal transiente para a parte do sinal de reposição. Adequadamente, é possível garantir que a parte do sinal de reposição compreende energia de sinal suficiente, que

não se desvia muito enfaticamente da energia do sinal das partes de sinal precedente e subsequente.

[00084] Alternativamente, o repositor de parte transiente 130d pode ser configurado para obter a parte do sinal de reposição pela destruição da relação de fase de forma transiente na parte de sinal transiente. Por exemplo, o repositor de parte transiente pode ser configurado para distribuir aleatoriamente ou ajustar (de modo determinístico) a fase de diferentes componentes de frequência da parte de sinal transiente. Adequadamente, a parte do sinal de reposição obtida desta forma pode compreender (no mínimo, de modo aproximado) a mesma energia que a parte de sinal transiente (uma vez que a modificação de fase dos componentes de frequência não altera a energia). Entretanto, a evolução temporal com forma transiente do sinal de tempo descrito pela parte do sinal de reposição pode ser perdida, uma vez que a evolução temporal transiente tem como base uma relação de fase específica de diferentes componentes de frequência, que é destruída.

[00085] Contudo, alternativamente, o repositor de parte transiente 130d pode interpolar, por exemplo, uma evolução temporal da energia em diferentes bandas de frequência, com base em uma parte de sinal não transiente que precede a parte de sinal transiente. Adequadamente, o teor da parte do sinal de reposição pode ser meramente baseado em uma extrapolação do teor de uma parte de sinal não transiente que precede a parte de sinal transiente. Adequadamente, o teor da parte de sinal transiente pode ser completamente desconsiderado.

[00086] Contudo, alternativamente, é possível obter o teor da parte do sinal de reposição, com o uso do repositor de

parte transiente 130d, por meio da interpolação entre um teor de uma parte de sinal não transiente que precede a parte de sinal transiente e uma parte de sinal não transiente que segue a parte de sinal transiente. Novamente, o teor da parte de sinal transiente pode ser completamente desconsiderado. É possível realizar a interpolação, por exemplo, em um domínio tempo-frequência.

[00087] Contudo, alternativamente, uma combinação dos métodos acima descritos pode ser utilizada para obter o teor da parte do sinal de reposição. Por exemplo, um teor não transiente da parte de sinal transiente (extraído, por exemplo, pela eliminação do teor transiente ou pela destruição da relação de fase com formação de transiente) pode ser combinado a um teor de sinal de áudio obtido pela interpolação ou extrapolação de uma ou mais partes de sinal transiente. Como outro exemplo, uma relação de fase de formação de transiente, em uma parte de sinal transiente, pode ser destruída e uma energia da parte de sinal transiente pode ser escalonada para ser adaptada a uma energia de partes de sinal não transientes adjacentes.

[00088] Em vista do acima exposto, é possível afirmar que a parte do sinal de reposição é sintetizada com base somente em partes de sinal não transiente (por exemplo, parte de sinal transiente precedente e/ou seguinte) (sem o uso do teor da parte de sinal transiente), com base somente na parte de sinal transiente ou com base em uma combinação de uma ou mais partes de sinal não transiente e a parte de sinal transiente.

[00089] Conceito adicional para a geração de sinal de áudio com redução de transiente - básico

[00090] A seguir, será descrito um conceito adicional para a geração do sinal de áudio com redução de transiente 132, cujos aspectos podem ser aplicados em qualquer configuração ora descrita. Em relação ao processo de detecção e substituição, é feita referência à patente WO 2007/118533, que foi ora incorporada em sua totalidade, para referência.

[00091] A patente WO 2007/118533 A1 descreve um equipamento e um método para a produção de um sinal de área *surrounding*. Este documento descreve um detector de transiente, que é provido para detectar um período de tempo transiente. O detector de transiente descrito na patente WO 2007/118533 A1 pode, por exemplo, ser utilizado para implementar (ou substituir) o detector de transiente 130a ora descrito. A referida publicação descreve ainda um gerador de sinal de síntese, que produz um sinal de síntese que atende a uma condição transiente e uma condição de continuidade. O gerador de síntese, descrito na patente WO 2007/118533 A1, pode, por exemplo, ser utilizado para implementar o repositório de parte transiente 130d ou pode, inclusive, substituir o repositório de parte transiente 130d. Assim, o conceito descrito na patente WO 2007/118533 A1, para a geração de um sinal de síntese, pode ser utilizado para a geração do sinal de áudio com redução de transiente 132, em algumas configurações da presente invenção.

[00092] Conceito adicional para a geração de sinal de áudio com redução de transiente - extensões

[00093] Conforme a aplicação ora descrita (processamento de um sinal que compreende um transiente, ao passo que mantém uma boa impressão de audição), a alta qualidade de áudio do sinal

resultante é substancialmente mais crítica que na aplicação da patente WO 2007/118533 (*Ambient Signal Generation*), o método descrito na patente WO 2007/118533 é ampliado por algumas etapas, para melhorar a qualidade do sinal de áudio.

[00094] Por exemplo, ademais da extrapolação de amplitude, uma configuração, de acordo com a presente invenção, também pode compreender a extrapolação ou interpolação de valores de fase, para assim obter um sinal de síntese de qualidade melhorada, que não apresenta partes transientes.

[00095] A extrapolação e a interpolação são realizadas, por exemplo, com o uso de uma predição linear ou uma codificação de predição linear (LPC), ou linearmente e/ou com splines ou como o ruído ponderado +.

[00096] Em algumas configurações, a geração acima descrita do sinal de áudio com redução de transiente 132 pode ser particularmente vantajosa quando utilizada em combinação com um vocoder de fase, que pode ser parte do processador de sinal 140 ou que pode constituir o processador de sinal 140. Em algumas configurações, é explorada a propriedade do vocoder de fase - que geralmente é considerada um grande problema [8] - que consiste em que não há relação previsível para os frames precedentes durante os transientes. Em algumas configurações, este fato é explorado, para assim suprimir o transiente no qual o transiente é apagado pelo reforço da relação com os escaninhos precedentes, ou seja, a fase de coeficientes diferentes que descrevem os diferentes escaninhos tempo-frequência da parte do sinal de reposição (por exemplo, na forma de números complexos) são, por exemplo, ajustados pela extrapolação a partir de escaninhos precedentes de

tempo-frequência (de uma parte de sinal não transiente precedente) ou que interpolam entre escaninhos respectivos de tempo-frequência de uma parte de sinal não transiente precedente e uma parte de sinal não transiente seguinte. Na publicação [Maher], um método comparável de interpolação é descrito. O método apresentado em [Maher] não é possível ser executado em tempo real, uma vez que as partes que seguem a lacuna de sinal também são exigidas. Ademais, [Maher] somente descreve o processamento de “picos” em um sinal de áudio (por outro lado, algumas configurações, de acordo com a invenção, processam todas as linhas de frequência) e os componentes de ruído não são tratados de modo explícito, ou seja, em algumas configurações, o conceito descrito em [Maher] para relacionar as lacunas em um sinal de áudio pode ser aplicado com a presente aplicação para obter um sinal de áudio com redução de transiente 132, com base no sinal de áudio de entrada original 110. Em vez de relacionar uma parte “faltante” de um sinal de áudio, a parte identificada como parte de sinal transiente pode ser substituída com o uso do método descrito em [Maher]. Contudo, é possível realizar a interpolação/extrapolação de modo independente, para cada escaninho de frequência. De modo opcional, a amplitude e a fase podem ser interpoladas (por exemplo, separadamente).

[00097] Detector de transiente 130a

[00098] A seguir, serão descritos alguns detalhes referentes ao detector de transiente 130a. Contudo, deve ser observado que várias implementações diferentes do detector de transiente 130a podem ser utilizadas, assim os detalhes a seguir devem ser considerados como exemplos de uma implementação

vantajosa. Em algumas configurações, os limites adaptativos são preferidos para reconhecer os períodos de tempo transiente. Normalmente, os limites adaptativos são versões suavizadas de uma função de detecção, que pode resultar em flutuações principais e, portanto, na não detecção de picos pequenos nas áreas circundantes dos picos grandes. Para detalhes, é feita referência à publicação [Bello]. Este problema pode ser resolvido, por exemplo, por meio da adaptação adequada de constantes de suavização na dependência em uma condição atualmente detectada (região de transiente/região sem transiente) e no desenvolvimento de uma função de detecção (por exemplo, ataque, queda).

[00099] A seguir, serão fornecidas algumas referências bibliográficas relativas aos aspectos acima mencionados: [Edler], [Bello], [Goodwin], [Walther], [Maher], [Daudet].

[000100] Extrator de parte transiente 130e

[000101] Ademais das funcionalidades acima descritas, o repositório de sinal transiente 130 pode ainda compreender um extrator de parte transiente 130e, cujo extrator de parte transiente 130e pode ser configurado para receber o sinal de áudio 110 (ou, no mínimo, sua parte de sinal transiente) e para prover a informação transiente 134. O extrator de parte transiente 130e pode ser configurado para prover a informação transiente 134, em qualquer forma possível, por exemplo, na forma de um sinal de tempo de parte de sinal transiente, na forma de uma representação de domínio de frequência-tempo de parte de sinal transiente ou na forma de parâmetros transientes (por exemplo, uma informação de tempo transiente e/ou uma informação de intensidade de transiente e/ou uma informação de queda transiente e/ou qualquer outra

informação transiente adequada).

[000102] Em particular, o extrator de parte transiente 130e pode ser configurado para prover a informação transiente 134 somente para as partes de sinal que foram eliminadas a partir do sinal de áudio 110 para obter o sinal de áudio com redução de transiente 132, para manter a taxa de dados razoavelmente pequena.

[000103] Alternativas de implementação para o processador de sinal 140 – Panorama

[000104] A seguir, serão descritos diferentes conceitos básicos para a implementação do processador de sinal 140. A Figura 3a ilustra uma implementação preferida do processador de sinal 140 da Figura 1. Esta implementação compreende um analisador seletivo de frequência 310 e um dispositivo de processamento seletivo de frequência subsequentemente conectado 312 que está implementado para assim fornecer uma influência negativa na “coerência vertical” do sinal de áudio original. Um exemplo para este processamento seletivo de frequência é estender o sinal no tempo ou encurtar o sinal no tempo, onde a extensão ou encurtamento é aplicado em um modo seletivo de frequência assim, por exemplo, o processamento introduz mudanças de fase no sinal de áudio processado, que são diferentes para diferentes bandas de frequência. Estas mudanças de fase podem, por exemplo, ser introduzidas para que os transientes sejam degradados. O processador de sinal 140, ilustrado na Figura 3a, pode, ainda, de modo opcional, compreender um combinador de frequência 314, que está configurado para combinar diferentes componentes de frequência do sinal de áudio processado, provido pelo processamento seletivo de frequência 312 em um sinal único (por

exemplo, sinal domínio-tempo).

[000105] Tanto o analisador seletivo de frequência 310, que pode dividir o sinal de áudio com redução de transiente 132 em uma pluralidade de componentes de frequência (por exemplo, coeficientes espectrais de valores complexos) quanto o combinador de frequência 314, que pode ser configurado para obter a representação de domínio de tempo do sinal de áudio processado 142, com base em uma pluralidade de coeficientes espectrais de valores complexos para diferentes bandas de frequência, podem ser configurados para realizar um processamento em blocos. Por exemplo, o analisador seletivo de frequência 310 pode processar um bloco (por exemplo, janelado) de amostras do sinal de áudio 132, para obter um conjunto de coeficientes espectrais de valores complexos representando o conteúdo do áudio do bloco de amostras de sinais de áudio. De maneira semelhante, o combinador de frequência 314 opcional pode receber um conjunto de coeficientes de valores complexos (por exemplo, um para cada banda de frequência fora de uma pluralidade de bandas de frequência) e para prover, nesta base, uma representação de domínio de tempo sobre um intervalo limitado de tempo que compreende uma pluralidade de amostras de domínio de tempo.

[000106] Outro processamento de sinal preferido está ilustrado na Figura 3b, no contexto de um processamento de vocoder de fase. Geralmente, um vocoder de fase compreende um analisador de sub-banda/transformada 320, um processador subsequentemente conectado 322 para realizar um processamento seletivo de frequência de uma pluralidade de sinais de saída providos pelo analisador 320, e subsequentemente um combinador de sub-

banda/transformada 324 que combina os sinais processados pelo processador 322 para, finalmente, obter um sinal processado 142 no domínio de tempo em uma saída 326. O sinal processado 142 no domínio de tempo, novamente, é um sinal de largura de banda completa para um sinal de filtro passa-baixo, contanto que a largura de banda do sinal processado 142 seja maior que a largura de banda representada por um único ramo entre o item 322 e 324, uma vez que o combinador de sub-banda/transformada 324 realiza uma combinação de sinais seletivos de frequência.

[000107] Os detalhes adicionais sobre este vocoder de fase serão discutidos abaixo, em relação às Figuras 5a, 5b, 5c e 6.

[000108] A Figura 3c ilustra outra implementação possível do processador de sinal 140. Conforme pode ser observado, o sinal de áudio com redução de transiente 132 pode, inclusive, ser processado no domínio de tempo, em algumas configurações. Tipicamente, o processamento de domínio de tempo 330 pode compreender uma memória, desse modo um transiente no sinal 132 apresentaria um impacto de maior duração no sinal de áudio processado 142. Em alguns casos, o sinal de áudio com redução de transiente 132 poderia provocar uma resposta transiente no sinal de áudio processado 142, que é significativamente mais longa (por exemplo, mais longa por um fator de 2, ou inclusive por um fator de 5, ou ainda por um fator de 10) que a duração do transiente (ou que a duração da parte do sinal transiente). Neste caso, os transientes no sinal de áudio 132 poderiam significativamente degradar, em um modo não desejável, o sinal de áudio processado 142, por exemplo, pela produção de ecos audíveis. Ademais, uma exclusão completa de uma parte de sinal transiente poderia também

apresentar um impacto de longa duração no sinal de áudio processado 142, uma vez que a exclusão completa de uma parte de sinal transiente provoca um transiente em si.

[000109] Implementação do processador de sinal com o uso de um vocoder - implementação de banco de filtros

[000110] A seguir, com referência às Figuras 5 e 6, estão ilustradas as implementações preferidas para um vocoder, que pode ser utilizado para uma implementação do processador de sinal 140, ou que pode ser parte do processador de sinal 140. A Figura 5a ilustra uma implementação de banco de filtros de um vocoder de fase, onde um sinal de áudio de entrada (por exemplo, sinal de áudio com redução de transiente 132) é alimentado em um sinal de áudio de entrada 500 e um sinal de áudio processado (por exemplo, sinal de áudio processado 142) é obtido em uma saída 510. Em particular, cada canal do banco de filtros esquemático ilustrado, na Figura 5a, inclui um filtro de passa-faixas 501 e um oscilador a jusante 502. Os sinais de saída de todos os osciladores, a partir de cada canal, são combinados por um combinador, que é, por exemplo, implementado como um adicionador e indicado em 503, para obter o sinal de saída na saída 510. Cada filtro 501 é implementado, para assim prover, de um lado, um sinal de amplitude e, de outro lado, um sinal de frequência. O sinal de amplitude e o sinal de frequência são sinais de tempo que ilustram um desenvolvimento da amplitude em um filtro 501 ao longo do tempo, ao passo que o sinal de frequência representa um desenvolvimento da frequência do sinal filtrado por um filtro 501.

[000111] Uma organização esquemática do filtro 501 está ilustrada na Figura 5b. Cada filtro 501 da Figura 5a pode ser

organizado conforme ilustrado na Figura 5b, onde, contudo, somente as frequências f_i fornecidas aos dois misturadores de entrada 551 e ao adicionador 552 são diferentes conforme o canal. Os sinais de saída do misturador são ambos filtrados pelo passa-baixo por passa-baixo 553, onde os sinais passa-baixo são diferentes na medida do possível, uma vez que foram gerados por sinais do oscilador local, que estão fora de fase em 90° . O filtro de passa-baixo mais alto 553 provê um sinal de quadratura 554, ao passo que o filtro mais baixo 553 provê um sinal em fase 555. Estes dois sinais, ou seja, I e Q, são fornecidos para um transformador de coordenadas 556, que gera uma representação de fase de magnitude a partir da representação retangular. O sinal de magnitude ou sinal de amplitude, respectivamente, da Figura 5a ao longo do tempo é liberado em uma saída 557. O sinal de fase é fornecido para um *phase unwrapper* 558. Na saída do elemento 558, não há mais um valor de fase presente, que está sempre entre 0 e 360° , mas um valor de fase que aumenta linearmente. Este valor de fase “unwrapped” é fornecido para um conversor de fase/frequência 559, que pode, por exemplo, ser implementado como um formador de diferença de fase simples, que subtrai uma fase de um ponto anterior no tempo, a partir de uma fase em um ponto corrente no tempo para obter um valor de frequência para o ponto corrente no tempo. Este valor de frequência é adicionado para o valor de frequência constante f_i do canal de filtros i para obter um valor de frequência temporariamente variante na saída 560. O valor de frequência na saída 560 tem um componente direto = f_i e um componente alternativo = desvio de frequência, pelo qual uma frequência atual do sinal no canal de filtro se desvia da

frequência média f_i .

[000112] Assim, conforme ilustrado nas Figuras 5a e 5b, o vocoder de fase atinge uma separação da informação espectral e da informação de tempo. A informação espectral está no canal especial ou na frequência f_i , que provê a parte direta da frequência para cada canal, ao passo que a informação de tempo está contida no desvio de frequência ou na magnitude ao longo do tempo, respectivamente.

[000113] A Figura 5c ilustra uma manipulação que pode ser realizada no vocoder, no local do vocoder apresentado em gráfico nas linhas tracejadas, na Figura 5a.

[000114] Para escalonamento de tempo, por exemplo, os sinais de amplitude $A(t)$, em cada canal ou frequência dos sinais $f(t)$, em cada sinal, podem ser destruídos ou interpolados, respectivamente. Para fins de transposição, conforme seja útil para a presente invenção, um interpolação, ou seja, uma extensão temporal ou distribuição de sinais $A(t)$ e $f(t)$ é realizada para obter sinais distribuídos $A'(t)$ e $f'(t)$, onde a interpolação é controlada por um fator de distribuição. Pela interpolação da variação de fase, ou seja, o valor antes da adição da frequência constante pelo adicionador 552, a frequência de cada oscilador individual 502, na Figura 5a, não é alterada. Contudo, a alteração temporal do sinal de áudio geral é diminuída, ou seja, pelo fator 2. O resultado é um tom de distribuição temporária que apresenta o tom original, ou seja, a onda fundamental original com seus harmônicos.

[000115] Para a transposição de frequência, é possível utilizar o seguinte conceito. Pela realização do processamento de

signal ilustrado na Figura 5c, onde o referido processamento é executado em cada canal de banda de filtro, na Figura 5a, e pela destruição do sinal temporal resultante em um decimador, o sinal de áudio pode ser reduzido novamente para sua duração original, ao passo que todas as frequências são simultaneamente duplicadas. Isso leva a uma transposição de tom pelo fator 2, onde, contudo, um sinal de áudio é obtido, o qual tem o mesmo comprimento que o sinal de áudio original, ou seja, o mesmo número de amostras.

[000116] Implementação do processador de sinal com o uso de um vocoder - implementação de transformada

[000117] Como uma alternativa à implementação do banco de filtros ilustrado na Figura 5a, uma implementação de transformada de um vocoder de fase também pode ser utilizada, conforme apresentado na Figura 6. Assim, o sinal de áudio 132 é alimentado em um processador FFT ou, de modo mais geral, em um Processador de Transformada de Fourier de Tempo Curto 600, como uma sequência de amostras de tempo. O processador FFT 600 está esquematicamente implementado na Figura 6 para realizar um janelamento de tempo de um sinal de áudio, para assim, por meio de uma FFT, calcular a magnitude e a fase do espectro, onde este cálculo é realizado por espectros sucessivos que estão relacionados aos blocos do sinal de áudio, que são enfaticamente sobrepostos.

[000118] Em um caso extremo, para cada nova amostra de sinal de áudio um novo espectro pode ser calculado, onde um novo espectro pode ser calculado também, por exemplo, somente a cada vigésima amostra nova. Esta distância a nas amostras entre dois espectros é preferencialmente dada por um controlador 602. O controlador 602 é ainda implementado para alimentar um processador

IFFT 604, que é implementado para operar em uma operação de sobreposição. Em particular, o processador IFFT 604 é implementado para, assim, realizar um Transformada de Fourier de tempo curto inversa, por meio da realização de um IFFT por espectro, com base na magnitude e na fase de um espectro modificado, para assim, realizar uma operação de adição de sobreposição, a partir da qual é obtido o sinal de tempo resultante. A operação de adição de sobreposição elimina os efeitos da janela de análise.

[000119] Uma distribuição do sinal de tempo é atingida pela distância b entre dois espectros, conforme são processados pelo processador IFFT 604, sendo superior à distância a entre os espectros, na geração de espectros FFT. A idéia básica é distribuir o sinal de áudio pelas FFTs inversas, simplesmente sendo ainda espaçadas pelas FFTs de análise. Como resultado, as alterações temporais no sinal de áudio sintetizado ocorrem mais lentamente que no sinal de áudio original.

[000120] Contudo, sem uma nova escala de fase no bloco 606, isso poderia levar a artefatos. Quando, por exemplo, um único escaninho de frequência isolado é considerado, para o qual os valores de fase sucessivos por 45° são implementados, isso implica que o sinal dentro do banco de filtros aumente na fase com uma taxa de $1/8$ de um ciclo, ou seja, por 45° por intervalo de tempo, onde o intervalo de tempo aqui é o intervalo de tempo entre as FFTs sucessivas. Se, agora, as FFTs inversas são mais espaçadas entre si, isso significa que um aumento de fase de 45° ocorre durante um intervalo de tempo maior. Isso significa que, devido à mudança de fase, ocorre uma incompatibilidade no processo de adição de sobreposição subsequente, levando a um cancelamento de

signal não desejado. Para eliminar este artefato, a fase é novamente escalonada exatamente pelo mesmo fator, por meio do qual o sinal de áudio é distribuído no tempo. A fase de cada valor espectral FFT é, assim, aumentada pelo fator b/a , assim esta incompatibilidade é eliminada.

[000121] Apesar de na configuração ilustrada na Figura 5c a distribuição pela interpolação dos sinais de controle de amplitude/frequência ser atingida por um oscilador de sinal na implementação de banco de filtros da Figura 5a, a distribuição na Figura 6 é atingida pela distância entre dois espectros IFFT, sendo superior a distância entre dois espectros FFT, ou seja, superior a a , onde, contudo, para a prevenção de artefato é executada novamente um escalonamento de fase, de acordo com b/a .

[000122] Em relação à descrição detalhada dos vocoders de fase, a referência à feita aos seguintes documentos:

[000123] *"The phase Vocoder: A tutorial"*, Mark Dolson, *Computer Music Journal*, vol. 10, no. 4, pp. 14 -- 27, 1986, ou *"New phase Vocoder techniques for pitch-shifting, harmonizing and other exotic effects"*, L. Laroche und M. Dolson, *Proceedings 1999 IEEE Workshop on applications of signal processing to audio and acoustics*, New Paltz, New York, October 17 - 20, 1999, pages 91 to 94; *"New approached to transient processing interphase vocoder"*, A. Röbel, *Proceeding of the 6th international conference on digital audio effects (DAFx-03)*, London, UK, September 8-11, 2003, pages DAFx-1 to DAFx-6; *"Phase-locked Vocoder"*, Meller Puckette, *Proceedings 1995, IEEE ASSP, Conference on applications of signal processing to audio and acoustics* ou solicitação de patente norte-americana número 6.549.884.

[000124] A seguir, um exemplo da funcionalidade do vocoder de fase com base em transformada será brevemente descrito, com referência à Figura 7. A Figura 7 ilustra uma representação esquemática da operação de um algoritmo de vocoder de fase com tamanho da grade de tempo de síntese sendo diferente do tamanho da grade de tempo de análise, por exemplo, por um fator de 2.

[000125] O algoritmo do vocoder de fase (VF) é utilizado para modificar a duração de um sinal, sem alterar seu tom [B9]. Ele divide um sinal nos denominados grãos, que denota cortes janelados do sinal, com tipicamente um comprimento na faixa de dez milissegundos. Os grãos são reorganizados em um processo *overlap-and-add* (OLA), com um tamanho da grade de tempo de síntese que difere do tamanho da grade de tempo de análise. Para estender o sinal por um fator ou dois, por exemplo, o tamanho da grade de tempo de síntese é duas vezes o tamanho da grade de tempo de análise. A Figura 7 ilustra o algoritmo.

[000126] Reinsersor de sinal transiente

[000127] A seguir, uma implementação preferida do reinsersor de sinal transiente 150, ilustrada na Figura 1, será descrita com referência à Figura 4.

[000128] O reinsersor de sinal transiente 150 compreende, como um componente principal, um combinador de sinal 150a. O combinador de sinal 150a está configurado para receber o sinal de áudio processado 142 e o sinal transiente 152, bem como para prover, com base nele, o sinal de áudio processado 120. O combinador de sinal 150a pode, por exemplo, ser configurado para realizar uma substituição de comutação de uma parte do sinal de áudio processado 142 por uma parte do sinal transiente 152.

Entretanto, em uma configuração preferida, o combinador de sinal 150a pode ser configurado para formar um desvanecimento cruzado entre o sinal de áudio processado 142 e o sinal transiente 152, assim há uma transição suave entre os referidos sinais 142, 152 no sinal de áudio processado 120.

[000129] Entretanto, um reinsensor de sinal transiente 150 pode ser configurado para determinar um coeficiente de inserção ideal. Por exemplo, o reinsensor de sinal transiente 150 pode compreender um calculador 150b para calcular o comprimento da parte de reinserção transiente. O cálculo deste comprimento da parte de reinserção transiente pode, por exemplo, ser importante se o comprimento da parte transiente substituída (conforme determinado, por exemplo, pelo detector de transiente 130a) for variável conforme as características de sinal. No caso de que o sinal de áudio processado 142 compreenda um comprimento diferente (número diferente de amostras por segundo ou número diferente de amostras gerais), quando comparado ao sinal de áudio de entrada original 110, um fator de extensão ou um fator de compressão pode ser considerado pelo calculador 150b para determinar o comprimento da parte de reinserção transiente. Uma discussão detalhada desta variação de comprimento é provida abaixo, com referência às Figuras 10 e 11.

[000130] O reinsensor de sinal transiente 150 pode, ainda, compreender um calculador 150c para calcular uma posição de reinserção. Em alguns casos, o cálculo da posição de reinserção pode considerar uma extensão ou uma compressão do sinal de áudio processado 142. Em alguns casos, é preferido que a referida relação entre um sinal de áudio não transiente e um teor de sinal

transiente (por exemplo, relação temporal) no sinal de áudio processado 120 seja, no mínimo, aproximadamente, idêntica à relação temporal do referido teor de áudio não transiente e do referido teor de áudio transiente, no sinal de áudio de entrada original 110. Entretanto, ademais do pré-processamento da posição de reinserção de sinal de áudio adequado, é possível realizar um ajuste preciso da referida posição de reinserção. Por exemplo, o calculador 150c, para calcular as posições de reinserção, pode ser configurado para ler o sinal de áudio processado 142 e o sinal transiente 152, bem como para determinar um tempo de reinserção, por exemplo, com base em uma comparação do sinal de áudio processado 142 com o sinal transiente 152. Os detalhes referentes ao cálculo possível da posição de reinserção serão discutidos abaixo, tomando-se como referência os exemplos ilustrados nas Figuras 10 e 11.

[000131] Possível relação de *timing*

[000132] A seguir, serão descritos detalhes referentes a uma possível relação de *timing*, com referência à Figura 9. A Figura 9 ilustra uma representação gráfica de um processamento de blocos diferentes do sinal de áudio de entrada original 110. A primeira representação gráfica 910 descreve uma evolução temporal do sinal de áudio de entrada original 110, onde uma abscissa 912 designa o tempo. O sinal de áudio de entrada 110 compreende uma parte de sinal transiente 920, um comprimento que pode ser variável. Como uma referência de *timing*, os intervalos de processamento ou blocos de processamento 922a, 922b, 922c do processador de sinal 140, são ilustrados na representação gráfica 910. Conforme é possível observar, a duração da parte de sinal

transiente 920 pode ser menor que a duração temporal dos intervalos de processamento 922a, 922b, 922c. Entretanto, em alguns casos, a duração temporal da parte de sinal transiente pode, inclusive, ser mais longa que a duração temporal dos intervalos de processamento, ou estender-se mais além de apenas em um intervalo de processamento. Em alguns casos, os intervalos de processamento 922a, 922b, 922c podem, também, ser de sobreposição de tempo.

[000133] Uma representação gráfica 930 representa um sinal de áudio com redução de transiente 132, que pode ser obtida pela substituição de transiente realizada pelo repositório de sinal transiente 130. Conforme é possível observar, a parte de sinal transiente 920 foi substituída por uma parte do sinal de reposição.

[000134] Uma representação gráfica 950 descreve o sinal de áudio processado 142, que pode ser obtido, por exemplo, com o uso de processamento em bloco do sinal de áudio com redução de transiente 132. O processamento pode, por exemplo, ser realizado com o uso de um vocoder de fase e com *downsampling*. Neste processamento, os blocos podem, opcionalmente, ser janelados, sendo os blocos opcionalmente sobrepostos.

[000135] Uma representação gráfica adicional 970 representa o sinal de áudio processado 120, no qual o transiente (ou sua versão modificada) foi reinserido pelo reinsersor de sinal transiente 150.

[000136] É importante observar que a parte de sinal transiente 920 poderia apresentar um impacto em todo o bloco 1'', se a parte de sinal transiente 920 foi considerada em um

processamento em bloco, uma vez que a energia transiente poderia, tipicamente, ser distribuída sobre todo o bloco, em um referido processamento em bloco. Assim, se a parte de sinal transiente foi considerada no processamento em bloco, a energia geral do bloco poderia, possivelmente, ser falseada pela energia transiente. Ademais, o transiente poderia, tipicamente, ser distribuído (ou seja, ampliado), se o transiente tiver sido afetado pelo processamento em bloco. Por outro lado, o processamento separado do transiente permite limitar o impacto do transiente em um intervalo de tempo de 1'' do sinal de áudio processado 120, que está associado ao transiente. Assim, é possível evitar uma distribuição da parte de sinal transiente no sentido de um bloco completo do processamento em bloco do sinal, no processador de sinal 140. Pelo contrário, a duração da parte de sinal transiente, no sinal de áudio processado 120, pode ser determinada pelo processamento transiente realizado pelo processador de transiente 160. Alternativamente, é possível inserir a parte de sinal transiente 920 no sinal de áudio processado 142, em sua duração original, se desejado. Assim, é possível evitar uma distribuição indesejada da energia transiente no processador de sinal 140.

[000137] Distribuição de tempo do sinal de áudio

[000138] Conforme é possível observar, a partir da descrição acima, o conceito da invenção para manipular um sinal de áudio, que compreende um evento transiente, pode ser utilizado em várias aplicações diferentes. Por exemplo, o referido conceito pode ser aplicado em qualquer processamento de sinal de áudio, no qual os transientes poderiam ser degradados pelo processamento de sinal e nos quais ele é, não obstante, desejável para manter os

transientes. Por exemplo, vários tipos de processamento de sinal de áudio não linear poderiam resultar em resultados gravemente degradados, na presença de transientes. Ademais, alguns tipos de filtragem temporal poderiam ser significativamente afetados pela presença de transientes. Ademais, qualquer processamento em bloco de um sinal de áudio poderia, tipicamente, ser degradado pela presença de transientes, uma vez que a energia dos transientes poderia ser indistinta sobre todo o bloco de processamento, resultando assim em artefatos audíveis.

[000139] Não obstante, a extensão de tempo dos sinais de áudio pode ser considerada como sendo uma aplicação particularmente importante do conceito presente para manipular um sinal de áudio que compreende um evento transiente. Por esta razão, os detalhes referentes a esta aplicação serão descritos a seguir.

[000140] A seguir, serão descritas algumas desvantagens dos conceitos convencionais para a extensão de sinais de áudio, para permitir uma compreensão das vantagens do conceito da invenção. A extensão de tempo dos sinais de áudio, por um vocoder de fase, compreende partes de sinal transiente “indistintas” pela dispersão, uma vez que a denominada coerência vertical (no sentido de uma relação de fase específica entre componentes de diferentes bandas de frequência) do sinal está comprometida. Os métodos que trabalham com os denominados métodos *overlap-add* (OLA) podem gerar pré-ecos de interrupção e ecos tardios do evento sonoro transiente. Estes problemas podem, de fato, ser atingidos por uma extensão de tempo mais acentuada, no ambiente de transientes. Entretanto, se ocorrer uma transposição, o fator de transposição

não mais será constante no ambiente dos transientes, ou seja, o tom dos constituintes de sinal sobreposto (possivelmente tonal) será alterado e será percebido como interrompido.

[000141] Se os transientes forem cortados e se a lacuna resultante for estendida, uma lacuna bastante grande deverá ser preenchida em seguida. Se os transientes seguem mutuamente de modo estreito, as lacunas grandes podem, possivelmente, ser sobrepostas.

[000142] A seguir, será descrito um novo método para transformação de sinais. O método ora apresentado resolve os problemas acima mencionados.

[000143] De acordo com um aspecto deste método, uma seção janelada, que contém o transiente, é interpolada ou extrapolada a partir do sinal a ser manipulado (por exemplo, o sinal de áudio de entrada original 110). Se a aplicação for essencial ao tempo, ou seja, se o atraso deve ser evitado, é possível, preferencialmente, optar pela extrapolação. Se o futuro é conhecido como uma denominada visão adiante, e se o atraso não desempenha uma parte muito importante, a interpolação será preferida.

[000144] Em algumas configurações, o método pode ser, essencialmente, compatível com as seguintes etapas e será ilustrado nas Figuras 10 e 11.

1. Reconhecimento do transiente;
2. Determinação do comprimento do transiente;
3. O transiente é salvo;
4. Extrapolação e/ou interpolação;
5. Aplicação do método atual, por exemplo, vocoder de fase;

6. Reinserção do transiente salvo; e
7. Possivelmente (opcional) re-ilustragem (para modificação da taxa de ilustragem).

[000145] Quando esta sequência é realizada, a duração de tempo do transiente é encurtada no *downsampling*. Se isso não for desejado, o transiente pode ser modulado para assim permanecer na banda de frequência desejada, antes de ser reinserido depois do chaveamento de mudança (etapas 6 e 7 intercambiáveis).

[000146] A seguir, serão descritos alguns detalhes com referência à Figura 10. A Figura 10 ilustra uma representação gráfica de sinais diferentes, que podem aparecer em uma configuração do equipamento 100, de acordo com a Figura 1. A representação da Figura 10 está desenhada em sua totalidade com 1000. Uma representação do sinal 1010 descreve uma evolução temporal do sinal de áudio de entrada original 110. Como é possível observar, o sinal de áudio de entrada 110 compreende uma parte de sinal transiente 1012, uma largura variável (ou duração), a qual pode ser determinada pelo detector de transiente 130a em um modo adaptado ao sinal. A parte de sinal transiente 1012 pode ser eliminada pelo repositor de sinal transiente 130 e pode ser substituída por uma parte do sinal de reposição. Adequadamente, o sinal de áudio com redução de transiente 132 pode ser obtido, que é ilustrado em uma representação de sinal 1020. Uma parte do sinal de reposição é ilustrada no número de referência 1022, substituindo a parte de sinal transiente 1012. O sinal de áudio com redução de transiente 132 pode ser processado em um formato em bloco, onde diferentes janelas de processamento (que determina a granularidade do processamento em bloco, e que também são

denominadas “grãos”) são ilustradas em uma representação de sinal 1030. Por exemplo, para cada bloco (ou “grão”), é possível obter um conjunto de coeficientes espectrais, para assim formar uma representação de domínio tempo-frequência do sinal de áudio com redução de transiente 132. Um processamento com vocoder de fase pode ser aplicado na representação de domínio tempo-frequência do sinal de áudio com redução de transiente 132, assim é obtido um sinal de duração aumentada. Para esta finalidade, é possível obter coeficientes de domínio tempo-frequência interpolados. Os coeficientes de domínio tempo-frequência podem ser utilizados para construir um sinal de domínio de tempo, cuja duração temporal é estendida quando comparada ao sinal de áudio de entrada original, ao passo que é mantido o tom, ou seja, o número de períodos de sinal é aumentado. O sinal obtido pela operação do vocoder de fase é ilustrado na representação de sinal 1040. Conforme é possível observar a partir da representação gráfica 1040, uma denominada “área transiente de corte”, na qual uma parte do sinal de reposição foi inserida para substituir a parte de sinal transiente, é mudada em relação a uma posição temporal da parte de sinal transiente no sinal de áudio de entrada original 110 (quando considerado em referência ao início do sinal de áudio de entrada).

[000147] Subsequentemente, a parte do sinal transiente, que foi previamente substituído, é reinserida, por exemplo, pelo reinsersor de sinal transiente 150. Por exemplo, a parte de sinal transiente, descrita pelo sinal transiente 152, pode ser submetida ao desvanecimento cruzado na versão processada 142 do sinal de áudio com redução de transiente. Um resultado da reinserção de transiente é ilustrado em uma representação gráfica 1050.

[000148] Em um *downsampling* subsequente, é possível reduzir uma duração temporal do sinal de áudio processado 120. O *downsampling* pode, por exemplo, ser realizado por um condicionador de sinal 170. O *downsampling* pode, por exemplo, compreender uma alteração de escala de tempo. Alternativamente, é possível reduzir vários pontos de amostra. Como consequência, uma duração temporal do sinal de *downsampling* é reduzida, quando comparada a um sinal fornecido pelo vocoder de fase. Ao mesmo tempo, vários períodos podem ser mantidos pelo *downsampling*, quando comparado ao sinal fornecido pelo vocoder de fase. Adequadamente, o tom do sinal submetido ao *downsampling*, que é ilustrado em uma representação de sinal 1050, pode ser aumentado, quando comparado ao sinal fornecido pelo vocoder de fase (ilustrado na representação do sinal 1040).

[000149] A Figura 11 ilustra outra representação de sinal que representa sinais que aparecem em outra configuração do equipamento 100, da Figura 1. O processamento é semelhante ao processamento explicado com referência à Figura 10, para que assim sejam ora descritas somente diferentes na ordem do processamento, e para que assim as representações de sinal idêntico e características de sinal serão desenhadas com numerais idênticos de referência, nas Figuras 10 e 11.

[000150] No processamento de sinal representado, na representação de sinal 1100, o *downsampling* é realizado antes da reinserção de sinal transiente. Assim, a representação de sinal 1150 ilustra o sinal submetido ao *downsampling*, sem uma parte de sinal transiente inserida. Entretanto, a parte de sinal transiente é comutada na frequência, com o uso de uma operação de comutação

de frequência transiente 1160, que pode ser realizada pelo processador transiente 160. O sinal transiente com frequência comutada (frequência comutada em relação à parte de sinal transiente substituída pelo repositor de sinal transiente 130) pode ser reinserido no sinal de áudio processado submetido ao *downsampling* 142 pelo reinsersor de sinal transiente 150. O resultado da reinserção transiente é ilustrado em uma representação de sinal 1170.

[000151] Ajuste da parte de sinal transiente

[000152] A seguir, será descrito como o sinal transiente 152 pode ser combinado ao sinal de áudio processado 142, com o uso do insersor de sinal transiente 150. Por exemplo, o insersor de sinal transiente 150 pode ser configurado para cortar uma área transiente do sinal de áudio processado 142, na qual o sinal transiente da área transiente 152 deve ser inserido. É possível considerar aqui que as partes limitantes do sinal transiente 152 podem, temporariamente, ser sobrepostas por partes limitantes da área transiente de corte. Nesta parte limitante de sobreposição, pode ocorrer um desvanecimento cruzado entre o sinal de áudio processado 142 e o sinal transiente 152. O sinal transiente 152 também pode ser comutado pelo tempo, em relação ao sinal de áudio processado 142, para que assim a forma de onda das partes limitantes da área transiente abrangida seja trazida em um bom acordo com a forma de onda das partes limitantes do sinal transiente 152.

[000153] O ajuste adequado pode ser realizado pelo cálculo da correlação cruzada máxima dos limites do recesso resultante, com limites da parte transiente (onde o recesso pode ser causado

pelo corte da área transiente, a partir do sinal de áudio processado 142). Desta forma, a qualidade subjetiva de áudio do transiente não é mais comprometida pela dispersão e pelos efeitos de eco.

[000154] A determinação precisa da posição do transiente, para fins de seleção de um corte adequado, pode ser realizada, por exemplo, com o uso de um centro flutuante de cálculo de gravidade da energia sobre um período de tempo adequado.

[000155] O ajuste ideal do transiente, em conformidade com a correlação cruzada máxima, pode exigir uma compensação discreta no tempo sobre sua posição original. Entretanto, devido à existência de um pré-mascaramento temporal e, em particular, de efeitos pós-mascaramento, a posição do transiente reinserido não deve ser exatamente compatível com a posição original. Devido ao período mais longo de ação do pós-mascaramento, uma comutação do transiente na direção positiva de tempo é favorecida neste contexto. Pela inserção da parte de sinal original, uma alteração na taxa de ilustragem leva a uma alteração no timbre ou no tom. Contudo, isso é particularmente mascarado pelo transiente, por meio dos mecanismos de mascaramento psicoacústicos.

[000156] Processamento de transiente

[000157] Se o transiente deve ser menos tonal, antes de sua reinserção após o corte, por exemplo, uma vez que ele é simplesmente adicionado ao sinal processado, a parte transiente janelada respectiva deverá ser processada de modo adequado. Neste contexto, é possível realizar a filtragem inversa (LPC).

[000158] Uma abordagem alternativa será brevemente descrita a seguir:

1. Determinar a Transformada de Fourier de Tempo Curto (STFT) (por exemplo, da parte de sinal transiente descrita pela informação transiente 134) para obter um espectro;
2. Determinar o Cepstrum (por exemplo, do espectro da parte de sinal transiente);
3. Filtrar por passa-altas o cepstrum (os primeiros coeficientes são estabelecidos em 0), para obter uma filtragem por passa-altas do espectro;
4. Dividir o espectro (por exemplo, da parte de sinal transiente) pelo espectro filtrado (por exemplo, da parte de sinal transiente), para obter um espectro suavizado; e
5. Inverter a transformada (por exemplo, do espectro suavizado) ao domínio de tempo (por exemplo, para obter o sinal transiente processado 152).

[000159] O sinal resultante exibe (no mínimo, de modo aproximado) o mesmo envelope espectral como sinal de saída, mas tem partes tonais perdidas.

[000160] Método

[000161] Configuração, de acordo com a invenção, compreende um método para manipular um sinal de áudio que compreende um evento transiente. A Figura 12 ilustra um fluxograma do referido método 1200.

[000162] O método 1200 compreende uma etapa 1210 de substituição da parte de sinal transiente, que compreende o evento transiente do sinal de áudio, com uma parte do sinal de reposição adaptada às características de energia do sinal de uma ou mais partes de sinal não transiente do sinal de áudio ou para uma característica de energia de sinal da parte de sinal transiente,

para obter um sinal de áudio com redução de transiente.

[000163] O método 1200 compreende ainda uma etapa 1220 de processamento do sinal de áudio com redução de transiente para obter uma versão processada do sinal de áudio com redução de transiente.

[000164] O método 1200 compreende ainda uma etapa 1230 de combinação da versão processada do sinal de áudio com redução de transiente com uma representação de sinal transiente, em uma forma original ou processada, um teor transiente da parte de sinal transiente.

[000165] O método 1200 pode ser complementado por qualquer característica ou funcionalidade ora descrita, também em relação ao equipamento da invenção acima, ou seja, embora alguns aspectos sejam descritos no contexto de um aparelho, está claro que estes aspectos também representam uma descrição do método respectivo, onde um bloco ou dispositivo se refere a uma etapa do método ou uma característica de uma etapa do método. De modo análogo, os aspectos descritos no contexto de uma etapa de método também representam uma descrição de um bloco respectivo, item ou característica de um equipamento respectivo.

[000166] Programa de computador

[000167] De acordo com determinados requisitos de implementação, as configurações da invenção podem ser implementadas no hardware ou no software. A implementação pode ser realizada com o uso de uma mídia de armazenamento digital, por exemplo, um disco flexível, DVD, Blue-Ray, CD, CD-ROM, PROM, EPROM, EEPROM ou memória flash, com sinais de controle passíveis de leitura eletrônica nele armazenados, capaz de atuar em um

sistema de informática programável, para que assim o referido método seja executado. Portanto, a mídia de armazenamento digital deve ser passível de leitura em computador.

[000168] Algumas configurações, de acordo com a invenção, compreendem uma mídia de transporte de dados, com sinais de controle passíveis de leitura eletrônica, que são capazes de atuar em um sistema de informática programável, para que assim seja executado um dos métodos ora descritos.

[000169] Geralmente, as configurações da presente invenção podem ser implementadas como um produto de programa de computador com um código de programa, sendo o código de programa operacional para realizar um dos métodos, quando o produto de programa de computador seja executado em um computador. O código de programa pode, por exemplo, ser armazenado em um dispositivo transportador passível de leitura em máquina.

[000170] Outras configurações compreendem o programa de computador para realizar um dos métodos ora descritos, armazenado em um dispositivo transportador passível de leitura em máquina, ou seja, uma configuração do método da invenção é, portanto, um programa de computador que tem um código de programa para realizar um dos métodos ora descritos, quando o programa de computador é executado em um computador.

[000171] Uma configuração adicional dos métodos da invenção é, portanto, um dispositivo transportador de dados (ou uma mídia de armazenamento digital ou uma mídia passível de leitura em computador) compreendendo, o programa de computador nele gravado para realizar um dos métodos ora descritos.

[000172] Uma configuração adicional dos métodos da

invenção é, portanto, um fluxo de dados ou uma sequência de sinais que representa o programa de computador para realizar um dos métodos ora descritos. O fluxo de dados ou sequência de sinais pode, por exemplo, ser configurado para transferência via conexão de comunicação de dados, por exemplo, via Internet.

[000173] Uma configuração adicional compreende um meio de processamento, por exemplo, um computador, ou um dispositivo lógico programável, configurado ou adaptado para realizar um dos métodos ora descritos.

[000174] Uma configuração adicional compreende um computador com um programa de computador instalado para realizar um dos métodos ora descritos.

[000175] Em algumas configurações, é possível utilizar um dispositivo lógico programável (por exemplo, um arranjo de portas programável em campo - FPGA) para realizar algumas ou todas as funcionalidades dos métodos ora descritos. Em algumas configurações, o arranjo de portas programável em campo - FPGA pode atuar com um microprocessador para realizar um dos métodos ora descritos. Geralmente, os métodos são preferencialmente realizados por qualquer equipamento de hardware.

[000176] Conclusão

[000177] Para resumir o acima exposto, as configurações, de acordo com a presente invenção, compreendem um método novo de tratar eventos sonoros, que não são ou não podem ser processados por meio da rotina atual de processamento (por exemplo, com o uso do processador de sinal). Em algumas configurações, o método da invenção consiste, essencialmente, na extrapolação ou interpolação da parte de sinal contendo os eventos sonoros que devem ser

separadamente processados. Após o processamento, as partes transientes separadamente tratadas são novamente adicionadas. Este processamento não está limitado à extensão de tempo e frequência, mas, geralmente, pode ser empregado no processamento de sinal, quando o processamento atual do sinal é deletério à parte de sinal transiente (ou se negativamente afetado por partes de sinal transiente).

[000178] A seguir, algumas vantagens do novo método são descritas, que podem ser obtidas em algumas configurações. Com o novo método, os artefatos (como dispersão, pré-eco ou ecos tardios), que podem surgir durante o processamento do transiente com o uso de métodos de extensão de tempo e de transposição, são efetivamente apresentados. O comprometimento potencial da qualidade das partes de sinal sobrepostas (possivelmente, tonais) é evitado.

[000179] As configurações, de acordo com a invenção, podem ser aplicadas em diferentes campos de aplicação. O método é, por exemplo, adequado para qualquer aplicação de áudio onde as velocidades de reprodução dos sinais de áudio, ou de seus tons, devam ser alteradas.

[000180] Para resumir o acima apresentado, foi descrito um meio e um método para um tratamento separado de eventos sonoros em sinais de áudio para evitar artefatos.

[000181] Configuração 2

[000182] Outra configuração da invenção será descrita a seguir, tomando-se como referência as Figuras 13 a 16.

[000183] Primeiramente, serão discutidos os detalhes referentes a uma detecção de transiente. De modo subsequente, o

tratamento do transiente será explicado com referência às Figuras 13 e 14. Os resultados do tratamento do transiente serão discutidos com referência à Figura 15. Os aperfeiçoamentos adicionais do tratamento do transiente serão explicados com referência à Figura 16. Ademais, será fornecida uma avaliação de desempenho da configuração e serão efetuadas algumas conclusões.

[000184] Configuração 2 - Detecção de transiente

[000185] Para implementar o conceito inventado, é importante detectar a presença de transientes para permitir uma substituição dos transientes e para um tratamento separado dos transientes.

[000186] Além da aplicação de extensão de tempo de um lado, uma ampla variedade de métodos de processamento de sinal exige conhecimento sobre um teor transiente do sinal de áudio. Os exemplos proeminentes são decisões de comprimento de bloco (B. Edler, "Coding of audio signals with over-lapping block transform and adaptive window functions (in German)," *Frequenz*, vol. 43, no. 9, pp. 252-256, Sept. 1989) ou codificação separada de sinais transientes e estacionários (Oliver Niemeyer and Bernd Edler, "Detection and extraction of transients for audio coding," in *AES 120th Convention*, Paris, France, 2006) em codificadores de áudio de transformada, modificação de componentes transientes (M. M. Goodwin and C. Avendano, "Frequency-domain algorithms for audio signal enhancement based on transient modification," *Journal of the Audio Engineering Society*, vol. 54, pp. 827-840, 2006.) e segmentação de sinal de áudio (P. Brossier, J.P. Bello, and M.D. Plumbley, "Real-time temporal segmentation of note objects in music signals," in

ICMC, Miami, USA, 2004). Tão numerosas quanto suas aplicações são as abordagens para detectar transientes. Mais comumente, a detecção é realizada pelo processamento de uma função de detecção (J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M.B. Sandler, "A tutorial on onset detection in music signals," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 5, pp. 1035–1047, Sept. 2005), ou seja, uma função com máximos locais que coincidem com a ocorrência de transientes. Vários métodos propostos derivam da referida função de detecção por meio da investigação da magnitude (ponderada) ou do envelope de energia dos sinais de sub-banda, o sinal de banda larga, sua função derivativa ou diferença relativa (consulte, por exemplo, as referências(A. Klapuri, "Sound onset detection by applying psychoacoustic knowledge," in ICASSP, 1999) e (P. Masri and A. Bateman, "Improved modelling of attack transients in music analysis-resynthesis," in ICMC, 1996).)

[000187] Outros métodos calculam o desvio entre a fase medida e a prevista (consulte, por exemplo, C. Duxbury, M. Davies, and M. Sandler, "Separation of transient information in musical audio using multiresolution analysis techniques," in *DAFX*, 2001), uma avaliação combinada da fase e das magnitudes de sinais de sub-banda (consulte, por exemplo, C. Duxbury, M. Sandler, and M. Davies, "A hybrid approach to musical note onset detection, " in *DAFX*, 2002) ou o erro produzido pelo preditor linear adaptativo (consulte, por exemplo, W-C. Lee and C-C. J. Kuo, "Musical onset detection based on adaptive linear prediction," in *ICME*, 2006). Pela seleção de pico, a presença de um transiente e sua localização no tempo é derivada como uma

decisão binária ou uma função de detecção contínua é aplicada para controlar o comportamento da unidade de modificação (consulte, por exemplo, a referência M. M. Goodwin and C. Avendano, "Frequency-domain algorithms for audio signal enhancement based on transient modification," *Journal of the Audio Engineering Society.*, vol. 54, pp. 827-840, 2006).

[000188] Com uma decisão binária, atribuições equivocadas, devido a classificações errôneas na detecção da etapa, podem provocar comprometimentos graves em algumas aplicações. Para o presente algoritmo, um falso negativo (ou seja, ausência de um transiente) pode ser pior que um falso positivo (ou seja, detecção de transiente não existente). O primeiro poderia levar a um componente transiente indistinto, ao passo que o último somente resultaria em uma interpolação supérflua se a interpolação for adequadamente realizada.

[000189] Os valores absolutos ponderados resumidos dos blocos de transformada de Fourier de tempo curto são utilizados para a detecção de áreas transientes. Esta função ilustra aumentos acentuados durante os transientes de ataque e também é capaz de indicar a queda de sinais percussivos e reverberação associada. A seleção de pico na função de detecção suavizada foi considerada com o uso de um limite adaptativo, com base em um cálculo de percentil conforme descrito, por exemplo, na referência J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M.B. Sandler, "A tutorial on onset detection in music signals," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 5, pp. 1035-1047, Sept. 2005.

[000190] Para resumir o acima apresentado, diferentes

conceitos para detecção transiente são conhecidos na técnica e podem ser aplicados em um equipamento inventado. Por exemplo, o conceito acima descrito para a detecção de um transiente pode ser utilizado no detector transiente 130a do repositório de sinal transiente 130.

[000191] Configuração 2 - Tratamento de transiente

[000192] A seguir, será descrito o tratamento de um transiente tomando-se como referência as Figuras 13 e 14. A Figura 13 ilustra uma representação gráfica de uma eliminação e interpolação de transiente. A Figura 14 ilustra uma representação gráfica de uma extensão de tempo e de uma reinserção de transiente. Assim, as representações esquemáticas nas Figuras 13 e 14 ilustram a sequência das etapas de processamento do algoritmo apresentado.

[000193] A primeira linha 1310 da Figura 3 ilustra o sinal original (ou seja, o sinal de áudio 110) que contém um evento transiente 1312. Em resposta à (ou por meio da) detecção deste transiente 1312, uma área transiente (por exemplo, que se estende a partir de uma posição inicial de área transiente 1314 para uma posição final de área transiente 1316) é definida (por exemplo, pelo detector de transiente 130a) que é, subsequentemente, subtraído do sinal, ou seja, primeiramente, o transiente é detectado e janelado. Em segundo lugar, ele é subtraído a partir do sinal. Um sinal, no qual o transiente é subtraído, é ilustrado na Ref. [B20]. O transiente em si é armazenado para uso posterior. Até esta etapa, o algoritmo é idêntico aquele descrito na Ref. [B8], apesar do fato de que a janela de corte ora utilizada é retangular (linha espessa

pontilhada). Para armazenamento do transiente, um intervalo de guarda de alguns milissegundos é precedido e anexado, bem como a janela é estreitada (linha fina contínua) para definir áreas de desvanecimento cruzado para uma reinserção suave do transiente armazenado nos sinais sem transiente com eliminação de tempo.

[000194] Subsequentemente, a característica mais importante do algoritmo da invenção, de acordo com a presente configuração - a interpolação para preencher a lacuna - é aplicada, ou seja, finalmente, a lacuna resultante é preenchida por meio da interpolação. Um resultado da interpolação pode ser observado em uma linha inferior da Figura 13, na Referência Nº 1330. Uma vez que o sinal é, tipicamente, quase-estacionário depois da interpolação, ele pode, agora, ser estendido sem a introdução de artefatos desagradáveis. Um resultado da referida extensão é ilustrado na primeira linha da Figura 14, na Referência Nº 1410. A região transiente na posição transposta é identificada e preparada para inserção do transiente janelado anteriormente armazenado. Portanto, a janela cônica (que foi aplicada para extração e/ou armazenamento do transiente e que é ilustrada por uma fina linha contínua na representação gráfica, na Referência Nº 1310) é invertida e aplicada ao sinal, para permitir que o transiente seja novamente adicionado. Um resultado deste processo é ilustrado na Referência Nº 1420. Finalmente, o transiente armazenado é adicionado ao sinal estendido, conforme é possível observar na representação gráfica, na Referência Nº 1430.

[000195] Para resumir o acima apresentado, a eliminação

e a interpolação de transiente da lacuna, que é provocada pela eliminação do transiente, são ilustradas na Figura 13. Primeiramente, o transiente é detectado e janelado. Em segundo lugar, ele é subtraído do sinal. Finalmente, a lacuna resultante é preenchida pela interpolação. A Figura 14 ilustra a extensão de tempo e a reinserção de transiente, que segue a eliminação e a interpolação de transiente. Primeiramente, o sinal quase-estacionário é estendido, por exemplo, com o uso de um vocoder ora descrito. Subsequentemente, a posição para o transiente no sinal com tempo estendido é preparada pela multiplicação com a janela invertida que foi utilizada para armazenar o transiente na Figura 14. Finalmente, o transiente é novamente adicionado ao sinal, ou seja, finalmente o transiente armazenado é adicionado ao sinal estendido.

[000196] Configuração 2 - Resultados do tratamento de transiente

[000197] A seguir, serão discutidos alguns resultados do tratamento da invenção de transiente tomando-se como referência a Figura 15. A Figura 15 ilustra uma representação gráfica das etapas do tratamento da invenção de transiente na aplicação com extensão de tempo, com vocoder de fase. Uma primeira linha contém o sinal sem extensão e uma segunda linha contém as portas estendidas. Devem ser observadas diferentes expectativas de tempo utilizadas nas representações gráficas na primeira linha e na segunda linha.

[000198] A Figura 15 ilustra os resultados das etapas algorítmicas diferentes, com base de castanholas misturadas a um diapasão.

[000199] Na Figura 15a é descrito um gráfico de forma de onda do sinal de entrada original, com uma indicação das áreas transientes detectadas. A Figura 15b ilustra as áreas transientes de corte que são interpoladas (em uma etapa posterior) para resultar no sinal estacionário sem transiente, apresentado na Figura 15c. A Figura 15d contém as áreas transientes, que incluem os intervalos de guarda do desvanecimento cruzado, ao passo que a Figura 15e ilustra o sinal interpolado (e tipicamente, o tempo estendido) que é amortecido com a janela de desvanecimento cruzado inverso, nas posições transientes excluídas de tempo. Concluindo, a Figura 15f ilustra a saída final do algoritmo de extensão de tempo.

[000200] Assim, a Figura 15a representa o sinal de áudio 110. A Figura 15e representa o sinal de áudio com redução de transiente 132. A Figura 15d representa o sinal transiente 152. A Figura 15f representa o sinal de áudio processado 120.

[000201] Configuração 2 - aperfeiçoamentos do tratamento de transiente

[000202] Foi descoberto que diferentes conceitos referentes à interpolação das áreas transientes de corte podem ser importantes em alguns casos. Por exemplo, a interpolação sobre uma área transiente pode ser difícil se o sinal anterior ao transiente difere consideravelmente do sinal depois do transiente. Neste caso, dificilmente é possível prever o envolvimento do sinal durante o evento transiente, em alguns casos. A Figura 16 ilustra a referida situação, simplificada pelo uso da avaliação possível de apenas uma das respectivamente duas parciais, pelo modo do exemplo. O

algoritmo (por exemplo, o algoritmo para realizar a interpolação para preencher a lacuna) deve decidir para um envolvimento do tom (do sinal interpolado para preencher a lacuna). O mesmo se aplica a sinais de banda larga mais complexos. Uma solução possível para superar o problema reside na previsão de avançar e retroceder com desvanecimento cruzado mutuamente. Assim, a previsão de avançar e retroceder com desvanecimento cruzado mutuamente pode ser aplicada ao processar o sinal interpolado para preencher a lacuna.

[000203] Este problema está ilustrado na Figura 16 e é apresentada uma solução, de acordo com um aspecto da invenção. A Figura 16 ilustra que a interpolação do transiente (ou seja, interpolação da lacuna provocada pela eliminação do transiente) é difícil, se o sinal for acentuadamente alterado durante o transiente. Há modos infinitos de contornos de tom durante a variação de interpolação (ou seja, interpolação da lacuna provocada pela eliminação do transiente). A Figura 16a ilustra uma representação gráfica de um sinal que contém um evento transiente na forma de uma representação de tempo-frequência. Uma variação transiente, ou seja, um intervalo que foi identificado como um intervalo de tempo transiente, é desenhado com 1610. A Figura 16 b ilustra uma representação gráfica de diferentes possibilidades para obter uma parte temporal do sinal de áudio de entrada, durante o qual um transiente foi detectado e eliminado. Conforme é possível observar, se houver primeiramente um tom que precede temporariamente o intervalo de tempo 1620, durante o qual o transiente é eliminado do sinal de áudio de entrada, e em segundo lugar, um tom temporariamente

depois do intervalo de tempo 1620, é necessário determinar uma evolução de tom para preencher a lacuna, que é deixada pela eliminação do intervalo de tempo transiente 1620. Conforme é possível observar, por exemplo, é possível que o tom precedente ao intervalo de tempo 1620 seja extrapolado adiante (na direção de tempo), para obter o tom durante o intervalo de tempo 1620 (consulte a linha tracejada 1620). Alternativamente, é possível extrapolar em retrocesso (na direção temporal) um tom, que está presente depois do intervalo de tempo 1620, para o intervalo de tempo 1620 (consulte a linha pontilhada 1632). Alternativamente, é possível interpolar, durante o intervalo de tempo 1620, entre um tom que está presente antes do intervalo de tempo 1620 e um tom que está presente depois do intervalo de tempo 1620 (consulte a linha pontilhada 1634). Naturalmente, são possíveis esquemas diferentes para obter uma evolução de tom durante o intervalo de tempo 1620 (lacuna causada pela eliminação de transiente).

[000204] Um impacto do sinal de áudio processado finalmente obtido, depois da reinserção de sinal, é ilustrado na Figura 16c. Conforme é possível observar, a parte de sinal transiente reinserido (que reflete um teor transiente original ou processado da parte de sinal transiente) pode ser, temporariamente, mais curto que o sinal de áudio processado (por exemplo, tempo estendido) 142, que foi processado sem o teor transiente. Assim a escolha do conceito para preencher a lacuna provocada pela eliminação do transiente no sinal de áudio 132 pode, de fato, apresentar um impacto audível no sinal de áudio processado 120, inclusive depois da reinserção de

transiente, por exemplo, se a parte transiente reinserida (descrita pelo sinal transiente 152) é mais curta que o resultado processado do preenchimento de lacuna no sinal de áudio processado 142. A referência é feita ao intervalo de tempo 140 que precede o transiente reinserido e um intervalo de tempo 142 depois do transiente reinserido.

[000205] Para resumir o acima exposto, é ilustrado com referência à Figura 16, que a interpolação da área transiente exige certa consideração se o sinal for acentuadamente alterado durante o transiente. Há modos infinitos de contornos de tom durante a variação de interpolação. A Figura 16a ilustra um sinal que contém um evento transiente. A Figura 16b ilustra diferentes possibilidades para interpolações de variação transiente, que estão indicadas por linhas pontilhadas. A Figura 16c ilustra um sinal estendido. Conforme as regiões interpoladas estendidas se estendem além das partes transientes, o sinal interpolado é audível e pode levar a artefatos perceptíveis.

[000206] Configuração 2 - Avaliação de desempenho

[000207] Para obter certo discernimento para o desempenho perceptual do método proposto, é realizada uma audição informal. Os sinais selecionados incluem itens com características de sinal transiente e estacionário, para avaliar o benefício do novo esquema para sinais transientes, enquanto, ao mesmo tempo, é garantido que os sinais estacionários não são degradados.

[000208] Este teste informal revelou um benefício significativo para a combinação acima mencionada do diapasão e

das castanholas, em comparação ao algoritmo de extensão de tempo do software do estado da técnica. O resultado ilustrou uma preferência em algoritmos de extensão de tempo com base no VF sobre WSOLA, quando o foco é levado aos sinais transientes.

[000209] Os sinais estendidos do mundo real, com o novo método, também foram, às vezes, preferidos sobre outros métodos.

[000210] Conclusão

[000211] Para resumir o acima apresentado, um novo esquema de tratamento de transiente foi descrito, que pode ser vantajosamente utilizado para os algoritmos de extensão de tempo. Alterar a velocidade ou o tom dos sinais de áudio, sem afetar os demais respectivos, é frequentemente utilizado para a produção de música e para reprodução criativa, como a remixagem. Isso também é utilizado para outros fins, como extensão de largura de banda e aperfeiçoamento da velocidade. Embora os sinais estacionários possam ser estendidos sem prejudicar a qualidade, os transientes frequentemente não são bem mantidos depois da extensão, ao serem utilizados os algoritmos convencionais. A presente invenção demonstra uma abordagem para o tratamento do transiente em algoritmos de extensão de tempo. As regiões transientes são substituídas por sinais estacionários. Os transientes eliminados resultantes são salvos e reinseridos ao sinal de áudio estacionário com tempo ampliado, depois da extensão de tempo.

[000212] Um desafio é lançado pela tarefa para estender uma combinação de um sinal bastante tonal, como um diapasão e um sinal percussivo, como as castanholas.

[000213] Embora alguns métodos convencionais aproximadamente preservem o envelope de um sinal, na versão de tempo estendido, bem como suas características espectrais, e seja esperado um evento percussivo com tempo ampliado para reduzir lentamente o original, a presente invenção segue a hipótese oposta que o escalamento de tempo dos sinais musicais, o objetivo é preservar o envelope dos eventos transientes. Portanto, algumas configurações, de acordo com a invenção, somente estendem o componente mantido para atingir um efeito que se parece com o mesmo instrumento tocado, em uma disposição diferente (consulte, por exemplo, a Ref. [B3]). Para atingir isso, os componentes de sinal transiente e estacionário são separadamente tratados, de acordo com a invenção.

[000214] As configurações, de acordo com a invenção, têm como base um conceito que foi descrito na publicação [B8], na qual foi demonstrado como os transientes podem ser preservados na extensão de tempo e frequência, com o vocoder de fase. Na referida abordagem, os transientes são cortados do sinal antes que ele seja estendido. A eliminação da parte transiente resulta em lacunas no sinal, que são estendidas pelo processo do vocoder de fase. Depois da extensão, os transientes são novamente adicionados ao sinal, com um *surrounding* que se ajusta às lacunas estendidas. Contudo, foi descoberto que a solução compreende algumas vantagens para vários sinais. Entretanto, também foi descoberto que, pelo corte dos transientes, surgem novos artefatos, conforme as lacunas introduzem novas partes não estacionárias ao sinal, em particular, em limites das lacunas introduzidas. É possível

observar as referidas partes não estacionárias, por exemplo, na Figura 15b.

[000215] As configurações do método da invenção ora descrito apresentam a vantagem sobre as técnicas descritas, por exemplo, em publicações [B3], [B6], [B7] que permitem extensão de tempo sem a necessidade de alterar o fator de extensão no *surrounding* de um transiente. O método da invenção apresenta pontos em comum com os métodos descritos, por exemplo, nas referências [B8] e [B5]. O esquema da invenção divide o sinal em uma parte transiente e um sinal quase-estacionário sem transiente. Por outro lado, no método descrito em [B8], as lacunas, que surgem do corte de transientes, são substituídas pelos sinais estacionários. Um método de interpolação é utilizado para calcular uma continuação dos sinais próximos ao período de lacuna em toda a lacuna. A parte quase-estacionária resultante é bem ajustada aos algoritmos de extensão de tempo. Devido ao fato de que este sinal agora (ou seja, depois da interpolação ou da extrapolação) não mais inclua transientes ou lacunas, os artefatos dos transientes estendidos e das lacunas estendidas podem ser evitados. Depois da realização da extensão, os transientes substituem as partes do sinal interpolado. A técnica reside na detecção da correção dos transientes e na interpolação perceptivelmente correta da parte estacionária. Entretanto, à parte da interpolação, outras técnicas de preenchimento podem ser utilizadas, conforme descrito acima.

[000216] Para melhor resumir o acima apresentado, em algumas configurações descritas acima, o objetivo foi estender

uma combinação de um tonal estrito e um sinal transiente, como um diapasão acrescido de castanholas, sem qualquer artefato perceptível. Foi mostrado que a presente invenção provê um avanço significativo no sentido de seu objetivo. Um dos aspectos importantes da presente invenção reside na identificação correta de um evento transiente, especialmente em seu início exato, e mais difícil, em sua queda e sua reverberação associada. Uma vez que a queda e a reverberação de um evento transiente são sobrepostas às partes estacionárias do sinal, estas partes devem ser meticulosamente tratadas para evitar flutuações perceptíveis depois da nova adição de partes estendidas do sinal.

[000217] Alguns ouvintes tendem a preferir versões nas quais a reverberação é estendida com as partes de sinal mantidas. Esta preferência se opõe ao objetivo atual de considerar um transiente e sons associados, como uma entidade. Portanto, em alguns casos, é necessário um maior discernimento na preferência dos ouvintes.

[000218] Entretanto, a ideia e a abordagem do princípio, de acordo com a presente invenção, comprovaram seu valor e aplicação em um caso especial. Não obstante, espera-se que a variação das aplicações da presente invenção possa, inclusive, ser estendida. Devido a sua estrutura, o algoritmo da invenção pode ser facilmente adaptado para ser utilizado para uma manipulação da parte transiente, por exemplo, pela alteração de seu nível comparado às partes de sinal estacionário.

[000219] Uma aplicação adicional possível do método da invenção poderia se atenuar ou obter arbitrariamente

transientes para repetição. Isso poderia ser explorado para alteração da sonoridade de eventos transientes como bateria, ou inclusive, para sua eliminação completa, uma vez que a separação do sinal em transiente e parte estacionária é inerente ao algoritmo.

[000220] As configurações acima descritas são meramente ilustrativas para os princípios da presente invenção. É entendido que as modificações e variações das organizações e dos detalhes ora descritos serão evidentes a outras pessoas capacitadas na técnica. Portanto, a intenção é limitada somente pelo escopo das reivindicações independentes da patente e não por detalhes específicos apresentado pelo modo de descrição e explicação das configurações ora apresentadas.

[000221] Referências

[000222] [A1] J.L. Flanagan and R.M. Golden, "The Bell System Technical Journal, November 1966", pages 1394 to 1509;

[000223] [A2] United States Patent 6,549,884, Laroche, J. & Dolson, M.: "Phase-vocoder pitch-shifting";

[000224] [A3] Jean Laroche and Mark Dolson, "New Phase-Vocoder Techniques for Pitch-Shifting, Harmonizing and Other Exotic Effects", by Proc.

[000225] [A4] Zölzer, U: "DAFX: Digital Audio Effects", Wiley & Sons, Edition: 1 (26 February 2002), pages 201-298;

[000226] [A5] Laroche L., Dolson M.: "Improved phase vocoder timescale modification of audio", IEEE Trans. Speech and Audio Processing, vol. 7, no. 3, pp. 323-332;

[000227] [A6] Emmanuel Ravelli, Mark Sandler and Juan P. Bello: "Fast implementation for non-linear time-scaling of stereo

- audio", Proc. of the 8th Int. Conference on Digital Audio Effects
 [000228] (DAFx'05), Madrid, Spain, September 20-22, 2005;
- [000229] [A7] Duxbury, C., M. Davies, and M. Sandler
 (2001, December): "Separation of transient information in musical
 audio using multiresolution analysis techniques". In: Proceedings
 of the COST G-6 Conference on Digital Audio Effects (DAFX-01),
 Limerick, Ireland;
- [000230] [A8] Röbel A.: "A NEW APPROACH TO TRANSIENT
 PROCESSING IN THE PHASE VOCODER", Proc. Of the 6th Int. Conference
 on Digital Audio Effects (DAFx-03), London, UK, September 8-11,
 2003.
- [000231] [B1] T. Karrer, E. Lee, and J. Borchers,
 "Phavorit: A phase vocoder for real-time interactive time-
 stretching," in *Proceedings of the ICMC 2006 International
 Computer Music Conference*, New Orleans, USA, November 2006, pp.
 708-715.
- [000232] [B2] T. F. Quatieri, R. B. Dunn, R. J. McAulay,
 and T. E. Hanna, "Time-scale modifications of complex acoustic
 signals in noise," Technical report, Massachusetts Institute of
 Technology, February 1994.
- [000233] [B3] C. Duxbury, M. Davies, and M. B. Sandler,
 "Improved time-scaling of musical audio using phase locking at
 transients," in *112th AES Convention*, Munich, 2002, Audio
 Engineering Society.
- [000234] [B4] S. Levine and Julius O. Smith III, "A
 sines+transients+noise audio representation for data
 compression and time/pitchscale modifications," 1998.
- [000235] [B5] T. S. Verma and T. H. Y. Meng, "Time scale

modification using a sines+transients+noise signal model," in *DAFX98*, Barcelona, Spain, 1998.

[000236] [B6] A. Röbel, "A new approach to transient processing in the phase vocoder," in *6th Conference on Digital Audio Effects (DAFx-03)*, London, 2003, pp. 344-349.

[000237] [B7] A. Röbel, "Transient detection and preservation in the phase vocoder," in *Int. Computer Music Conference (ICMC 03)*, Singapore, 2003, pp. 247-250.

[000238] [B8] F. Nagel, S. Disch, and N. Rettelbach, "A phase vocoder driven bandwidth extension method with novel transient handling for audio codecs," in *126th AES Convention*, Munich, 2009.

[000239] [B9] M. Dolson, "The phase vocoder: A tutorial," *Computer Music Journal*, vol. 10, no. 4, pp. 14-27, 1986.

[000240] [B10] B. Edler, "Coding of audio signals with over-lapping block transform and adaptive window functions (in german)," *Frequenz*, vol. 43, no. 9, pp. 252-256, Sept. 1989.

[000241] [B11] Oliver Niemeyer and Bernd Edler, "Detection and extraction of transients for audio coding," in *AES 120th Convention*, Paris, France, 2006.

[000242] [B12] M. M. Goodwin and C. Avendano, "Frequency-domain algorithms for audio signal enhancement based on transient modification," *Journal of the Audio Engineering Society*, vol. 54, pp. 827-840, 2006.

[000243] [B13] P. Brossier, J.P. Bello, and M.D. Plumbley, "Real-time temporal segmentation of note objects in music signals," in *ICMC*, Miami, USA, 2004.

- [000244] [B14] J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M.B. Sandler, "A tutorial on onset detection in music signals," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 5, pp. 1035-1047, Sept. 2005.
- [000245] [B15] A. Klapuri, "Sound onset detection by applying psychoacoustic knowledge," in ICASSP, 1999.
- [000246] [B16] P. Masri and A. Bateman, "Improved modelling of attack transients in music analysis-resynthesis," in ICMC, 1996.
- [000247] [B17] C. Duxbury, M. Davies, and M. Sandler, "Separation of transient information in musical audio using multiresolution analysis techniques," in DAFX, 2001.
- [000248] [B18] C. Duxbury, M. Sandler, and M. Davies, "A hybrid approach to musical note onset detection," " in DAFX, 2002.
- [000249] [B19] W-C. Lee and C-C. J. Kuo, "Musical onset detection based on adaptive linear prediction," in ICME, 2006.
- [000250] [Edler] O. Niemeyer and B. Edler, "Detection and extraction of transients for audio coding", presented at the AES 120th Convention, Paris, France, 2006;
- [000251] [Bello] J.P. Bello et al., "A Tutorial on Onset Detection in Music Signals", IEEE Transactions on Speech and Audio Processing, Vol. 13, No. 5, September 2005;
- [000252] [Goodwin] M. Goodwin, C. Avendano, "Enhancement of Audio Signals Using Transient Detection and Modification", presented at the AES 117th Convention, USA, October 2004;
- [000253] [Walther] Walther et al., "Using Transient Suppression in Blind Multi-channel Upmix Algorithms", presented at

the AES 122th Convention, Austria, May 2007;

[000254] [Maher] R.C. Maher, "A Method for Extrapolation of Missing Digital Audio Data", JAES, Vol. 42, No. 5, May 1994;

[000255] [Daudet] L. Daudet, "A review on techniques for the extraction of transients in musical signals", book series: Lecture Notes in Computer Science, Springer Berlin/Heidelberg, Volume 3902/2006, Book: Computer Music Modeling and Retrieval, pp. 219-232.

REIVINDICAÇÕES

1. EQUIPAMENTO PARA MANIPULAR UM SINAL DE ÁUDIO COMPREENDENDO UM EVENTO TRANSIENTE, sendo o dito Equipamento (100) para a manipulação de um sinal de áudio (110), caracterizado por compreender um evento transiente, o equipamento (100) compreendendo:

um repositor de sinal transiente (130) configurado para substituir uma parte do sinal transiente, compreendendo o evento transiente, do sinal de áudio por uma parte do sinal de reposição adaptado às características da energia do sinal de uma ou mais partes do sinal não transiente do sinal de áudio, ou para uma característica da energia do sinal da parte do sinal transiente, para obter um sinal de áudio com redução de transiente (132);

um processador de sinais (140) configurado para processar o sinal de áudio com redução de transiente (132), para obter uma versão processada (142) do sinal de áudio com redução de transiente; e

um reinsersor de sinal transiente (150) configurado para combinar a versão processada (142) do sinal de áudio com redução de transiente (132) com um sinal transiente (152) representando, em uma forma original ou processada, um teor transiente da parte do sinal transiente;

em que o repositor de sinal transiente (130) é configurado para extrapolar valores de amplitude de uma ou mais partes do sinal que precedem a parte do sinal transiente, para obter valores de amplitude da parte do sinal de reposição, e

em que o repositor de sinal transiente (130) é

configurado para extrapolar valores de fase de uma ou mais partes do sinal que precedem a parte do sinal transiente para obter valores de fase da parte do sinal de reposição.

2. Equipamento (100) para a manipulação de um sinal de áudio (110), caracterizado por compreender um evento transiente, o equipamento (100) compreendendo:

um repositor de sinal transiente (130) configurado para substituir uma parte do sinal transiente, compreendendo o evento transiente, do sinal de áudio por uma parte do sinal de reposição adaptado às características da energia do sinal de uma ou mais partes do sinal não transiente do sinal de áudio, ou a uma característica da energia do sinal da parte do sinal transiente, para obter um sinal de áudio com redução de transiente (132);

um processador de sinais (140) configurado para processar o sinal de áudio com redução de transiente (132), para obter uma versão processada (142) do sinal de áudio com redução de transiente; e

um reinsersor de sinal transiente (150) configurado para combinar a versão processada (142) do sinal de áudio com redução de transiente (132) com um sinal transiente (152) representando, em uma forma original ou processada, um teor transiente da parte do sinal transiente;

em que o repositor de sinal transiente (130) é configurado para interpolar entre um valor de amplitude de uma parte do sinal que precede uma parte do sinal transiente e um valor de amplitude de uma parte do sinal que se segue à parte do sinal transiente, para obter um ou mais valores de amplitude da

parte do sinal de reposição, e

em que o repositore de sinal transiente (130) é configurado para interpolar entre um valor de fase de uma parte do sinal que precede a parte do sinal transiente e um valor de fase de uma parte do sinal que se segue à parte do sinal transiente, para obter um ou mais valores de fase da parte do sinal de reposição.

3. Equipamento (100) para a manipulação de um sinal de áudio (110), caracterizado por compreender um evento transiente, o equipamento (100) compreendendo:

um repositore de sinal transiente (130) configurado para substituir uma parte do sinal transiente, compreendendo o evento transiente, do sinal de áudio por uma parte do sinal de reposição adaptado às características da energia do sinal de uma ou mais partes do sinal não transiente do sinal de áudio, ou de uma característica da energia do sinal da parte do sinal transiente, para obter um sinal de áudio com redução de transiente (132);

um processador de sinais (140) configurado para processar o sinal de áudio com redução de transiente (132), para obter uma versão processada (142) do sinal de áudio com redução de transiente; e

um reinsersore de sinal transiente (150) configurado para combinar uma versão processada (142) do sinal de áudio com redução de transiente (132) com um sinal transiente (152) representando, em uma forma original ou processada, um teor transiente da parte do sinal transiente;

em que o repositore de sinal transiente (130) é

configurado para extrapolar, em um domínio tempo-frequência, coeficientes no domínio tempo-frequência de valores complexos associados a uma parte do sinal não transiente do sinal de áudio (110) que precede a parte do sinal transiente, para obter coeficientes no domínio tempo-frequência da parte do sinal de reposição, ou

em que o repositor de sinal transiente (130) é configurado para interpolar, em um domínio tempo-frequência, entre coeficientes no domínio tempo-frequência de valores complexos associados a uma parte do sinal não transiente do sinal de áudio (110) que precedem a parte do sinal transiente, e coeficientes no domínio tempo-frequência de valores complexos associados a uma parte do sinal não transiente do sinal de áudio que se segue à parte do sinal transiente, para obter coeficientes no domínio tempo-frequência da parte do sinal de reposição.

4. Equipamento (100), de acordo com uma das reivindicações de 1 a 3, caracterizado por o repositor de sinal transiente (130) ser configurado para prover a parte do sinal de reposição de maneira que a parte do sinal de reposição represente um sinal de tempo tendo uma evolução temporal suavizada quando comparada à parte do sinal transiente, de maneira que um desvio entre uma energia da parte do sinal de reposição e uma energia de uma parte do sinal não transiente do sinal de áudio (110) que precede a parte do sinal transiente ou que se segue à parte do sinal transiente seja menor que um valor limite predeterminado.

5. Equipamento (100), de acordo com uma das reivindicações de 1 a 4, caracterizado por o repositor de sinal transiente (130) ser configurado para aplicar um ruído ponderado

para obter os valores de amplitude da parte do sinal de reposição, ou

para aplicar um ruído ponderado para obter os valores de fase das partes do sinal de reposição.

6. Equipamento (100), de acordo com uma das reivindicações de 1 a 4, caracterizado por o repositório de sinal transiente (130) ser configurado para combinar componentes não transientes da parte do sinal transiente com os valores extrapolados ou interpolados, para obter a parte do sinal de reposição.

7. Equipamento (100), de acordo com uma das reivindicações de 1 a 6, caracterizado por o repositório de sinal transiente (130) ser configurado para obter partes do sinal de reposição de comprimentos variáveis na dependência de um comprimento da presente parte do sinal transiente.

8. Equipamento (100), de acordo com uma das reivindicações de 1 a 7, caracterizado por o processador de sinais (140) ser configurado para processar o sinal de áudio com redução de transiente (132) de maneira que uma dada parte do sinal temporal da versão processada (142) do sinal de áudio com redução de transiente seja dependente de uma pluralidade de partes do sinal temporal mudadas temporalmente do sinal de áudio com redução de transiente (132).

9. Equipamento (100), de acordo com uma das reivindicações de 1 a 8, caracterizado por o processador de sinais (140) ser configurado para realizar um processamento com base em bloco de tempo do sinal de áudio com redução de transiente 132, para obter a versão processada (142) do sinal de áudio com redução

de transiente; e

em que o repositor de sinal transiente 130 é configurado para ajustar a duração da parte do sinal transiente para ser substituída pela parte do sinal de reposição com uma resolução temporal que é melhor que a duração de um bloco de tempo, ou para substituir uma parte do sinal transiente tendo uma duração temporal menor que a duração do bloco de tempo com uma parte do sinal de reposição tendo uma duração temporal menor que a duração do bloco de tempo.

10. Equipamento (100), de acordo com uma das reivindicações de 1 a 9, caracterizado por o processador de sinais (140) ser configurado para processar o sinal de áudio com redução de transiente (132) de maneira dependente da frequência, de maneira que o processamento introduza mudanças de fase dependentes da frequência degradantes de transiente no sinal de áudio com redução de transiente (132).

11. Equipamento (100), de acordo com uma das reivindicações de 1 a 10, caracterizado por o repositor de sinal transiente (130) compreender um detector de transientes (130a), em que o detector de transientes (130a) é configurado para prover um limite de detecção com variação de tempo para a detecção do transiente no sinal de áudio (110) de maneira que o limite de detecção siga um envelope do sinal de áudio em uma constante de tempo com ajuste suave, e

em que o detector de transientes é configurado para mudar a constante de tempo suave em resposta à detecção de um transiente e/ou na dependência de uma evolução temporal do sinal de áudio.

12. Equipamento (100), de acordo com uma das reivindicações de 1 a 11, caracterizado por o equipamento (100) compreender um processador transiente (160) configurado para receber informações transientes (134) e para obter, com base nas informações transientes (134), um sinal transiente processado (152) no qual são reduzidos os componentes tonais, e

em que o reinsensor de sinal transiente (150) é configurado para combinar a versão processada (142) do sinal de áudio com redução de transiente (132) com o sinal transiente processado (152) provido pelo processador transiente (160).

13. Equipamento (100), de acordo com uma das reivindicações de 1 a 12, caracterizado por o repositor de sinal transiente (130) compreende um detector de transientes (130a, 130c) configurado para detectar uma parte do sinal transiente do sinal de áudio (110) com base no monitoramento do sinal de áudio (110), ou com base em informações auxiliares que acompanham o sinal de áudio, e para determinar um comprimento da parte do sinal transiente;

em que o repositor de sinal transiente (130) é configurado para levar em conta o comprimento da parte do sinal transiente determinado pelo detector de transientes (130a, 130c);

em que o repositor de sinal transiente (130) é configurado para extrapolar, em um domínio tempo-frequência, coeficientes no domínio tempo-frequência de valores complexos associados a uma parte do sinal não transiente do sinal de áudio (110) que precedem a parte do sinal transiente, para obter coeficientes no domínio tempo-frequência da parte do sinal de reposição, ou

em que o repositório de sinal transiente (130) é configurado para interpolar, em um domínio tempo-frequência, entre coeficientes no domínio tempo-frequência de valores complexos associados a uma parte do sinal não transiente do sinal de áudio (110) que precedem a parte do sinal transiente, e coeficientes no domínio tempo-frequência de valores complexos associados a uma parte do sinal não transiente do sinal de áudio que se segue à parte do sinal transiente, para obter coeficientes no domínio tempo-frequência da parte do sinal de reposição;

em que o processador de sinais (140) é configurado para realizar um processamento de sinal de áudio com degradação de transiente por distensão de tempo ou compressão de tempo, de maneira que o sinal processado (142) provido pelo processador de sinais (140) compreenda uma duração maior ou menor que a duração do sinal não processado (132) recebido pelo processador de sinal de áudio; e

em que o equipamento (100) é configurado para adaptar um escalonamento de tempo ou razão de amostragem do sinal obtido pelo reinsersor de sinal transiente (150) de maneira que pelo menos componentes não transientes do sinal obtido pelo reinsersor de sinal transiente (150) sejam transpostos por frequência quando comparados à entrada do sinal de áudio (110) no repositório de sinal transiente (130).

14. Equipamento (100), de acordo com uma das reivindicações de 1 a 13, caracterizado por um reinsersor de sinal transiente (150) é configurado para fazer o desvanecimento cruzado da versão processada (142) do sinal de áudio com redução de transiente (132) com um sinal transiente (152) representando, em

uma forma original ou processada, um teor transiente da parte do sinal transiente.

15. Método (1200) para a manipulação de um sinal de áudio compreendendo um evento transiente, caracterizado por o método compreender:

substituir (1210) uma parte do sinal transiente, compreendendo o evento transiente, do sinal de áudio por uma parte do sinal de reposição adaptado às características da energia do sinal de uma ou mais partes do sinal não transiente do sinal de áudio, ou às características da energia do sinal da parte do sinal transiente, para obter um sinal de áudio com redução de transiente;

processar (1220) o sinal de áudio com redução de transiente, para obter uma versão processada do sinal de áudio com redução de transiente; e

combinar (1230) a versão processada do sinal de áudio com redução de transiente com um sinal transiente representando, em uma forma original ou processada, um teor transiente da parte do sinal transiente;

em que os valores de amplitude de uma ou mais partes do sinal que precedem a parte do sinal transiente são extrapolados para obter valores de amplitude da parte do sinal de reposição, e em que valores de fase de uma ou mais partes do sinal que precedem a parte do sinal transiente são extrapolados para obter valores de fase da parte do sinal de reposição; ou

em que é feita uma interpolação entre um valor de amplitude de uma parte do sinal que precede a parte do sinal transiente e um valor de amplitude de uma parte do sinal que se

segue à parte do sinal transiente, para obter um ou mais valores de amplitude da parte do sinal de reposição, e

em que é feita uma interpolação entre um valor de fase de uma parte do sinal que precede a parte do sinal transiente e um valor de fase de uma parte do sinal que segue um ou mais valores de fase da parte do sinal de reposição; ou

em que coeficientes no domínio tempo-frequência de valores complexos associados a uma parte do sinal não transiente do sinal de áudio que precedem a parte do sinal transiente são extrapolados em um domínio tempo-frequência, para obter coeficientes no domínio tempo-frequência da parte do sinal de reposição; ou

em que é feita uma interpolação, em um domínio tempo-frequência, entre coeficientes no domínio tempo-frequência de valores complexos associados a uma parte do sinal não transiente do sinal de áudio que precedem a parte do sinal transiente, e coeficientes no domínio tempo-frequência de valores complexos associados a uma parte do sinal não transiente do sinal de áudio que se segue à parte do sinal transiente, para obter coeficientes no domínio tempo-frequência da parte do sinal de reposição.

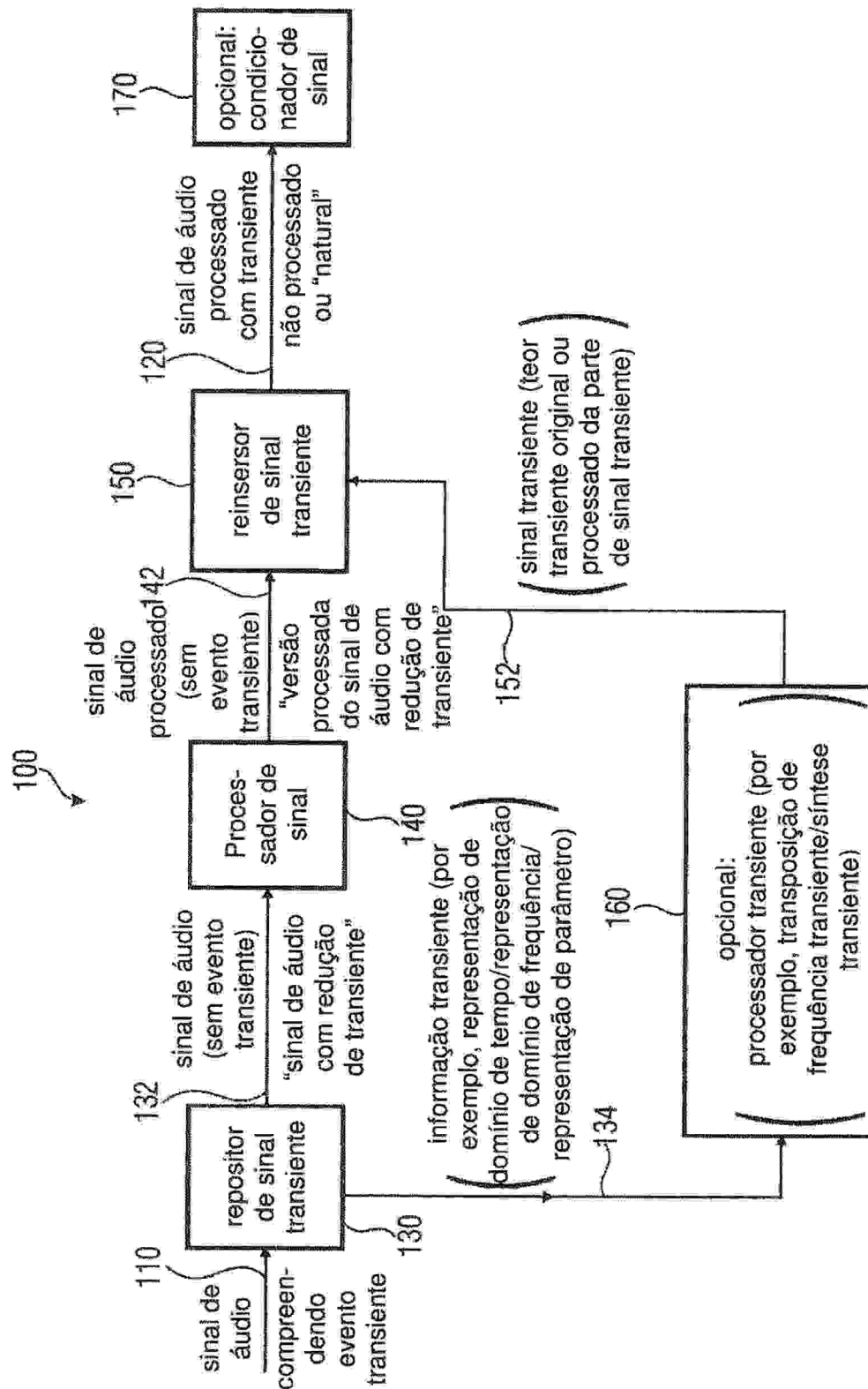


FIG 1

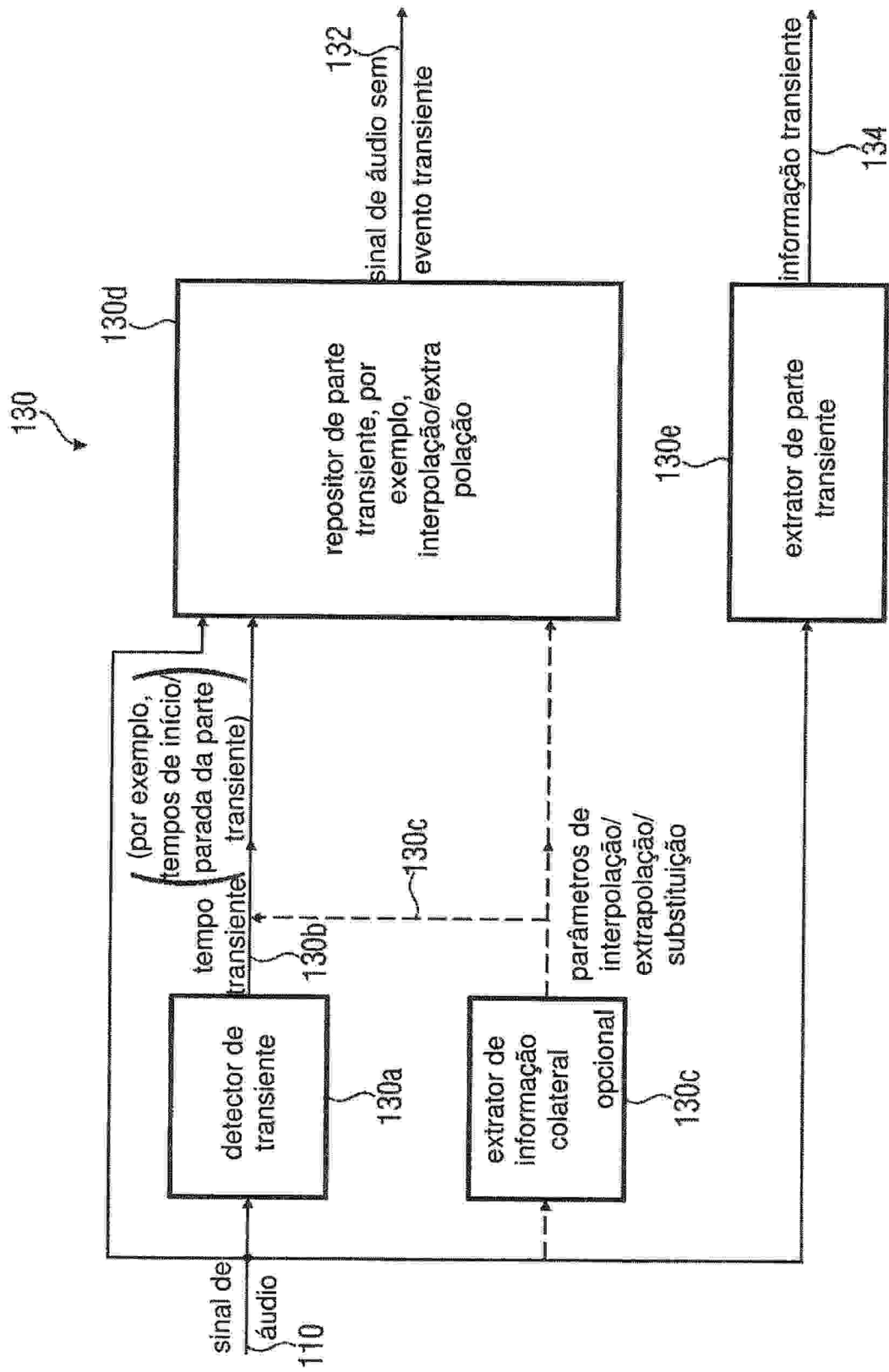


FIG 2

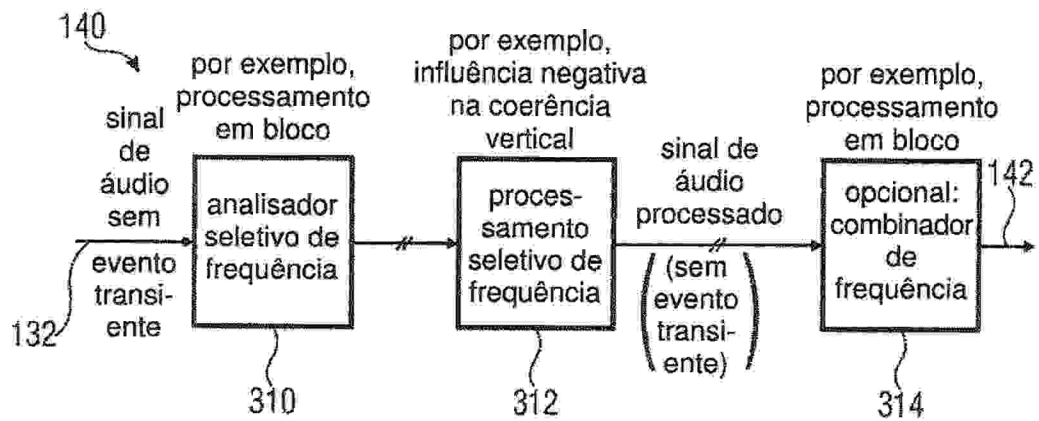


FIG 3A

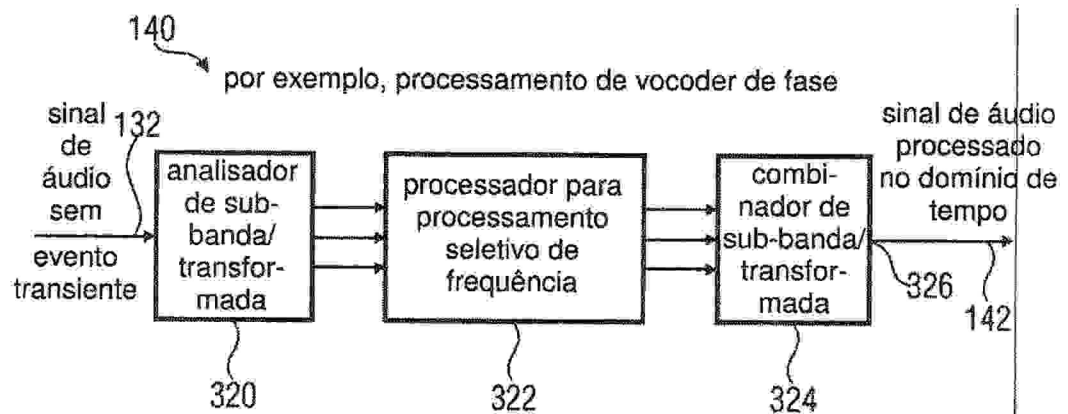


FIG 3B

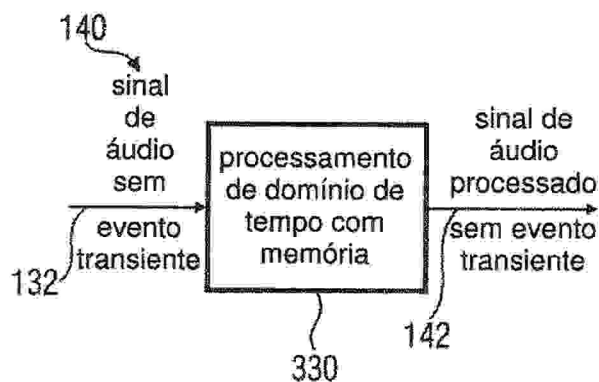


FIG 3C

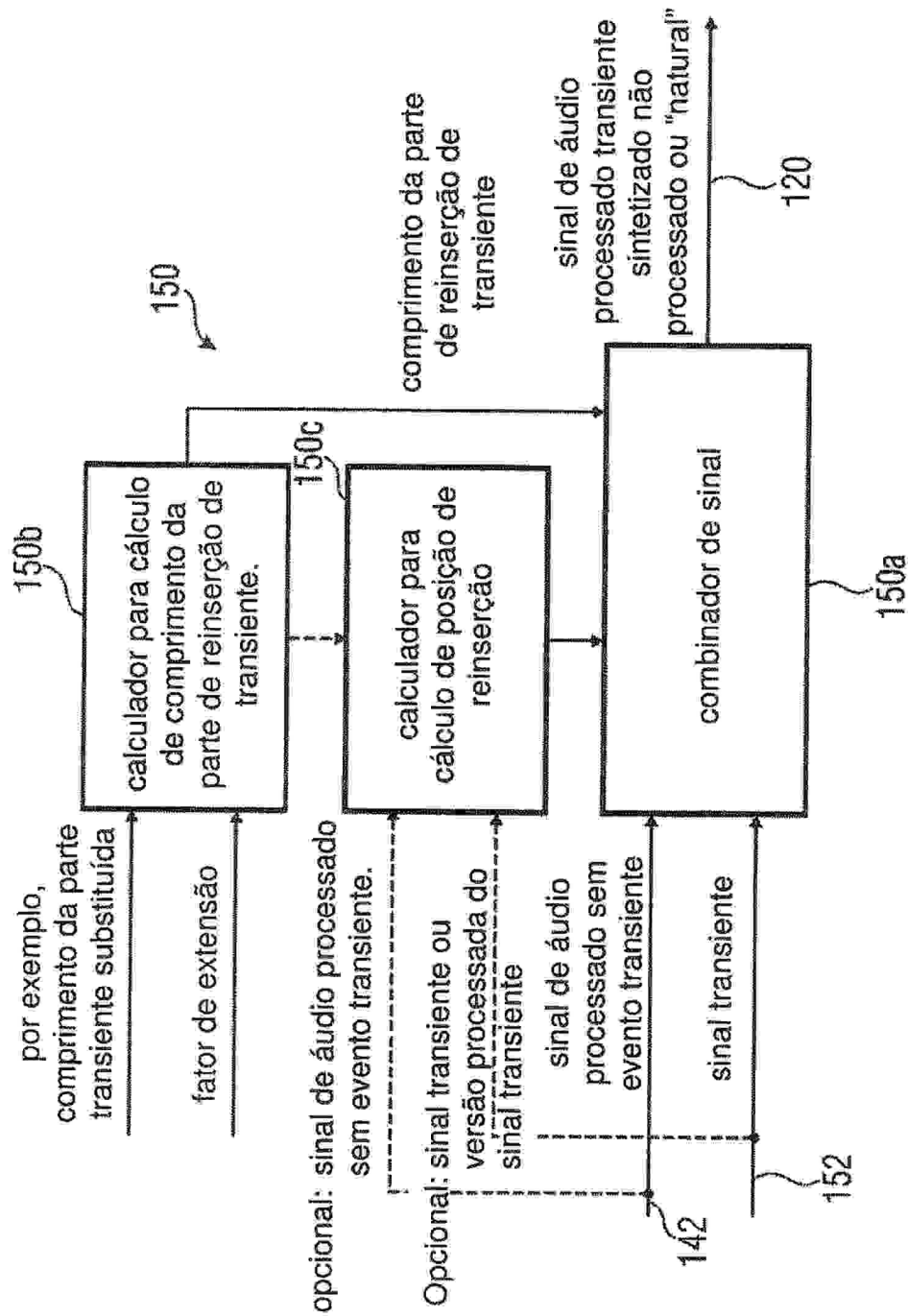


FIG 4

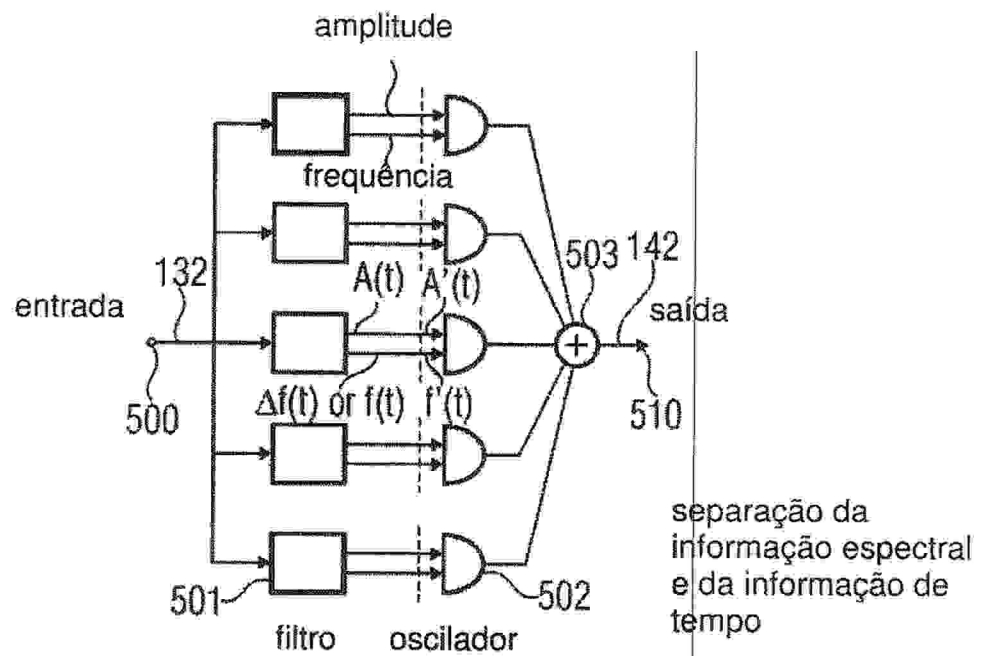


FIG 5A
(implementação de banco de filtros)

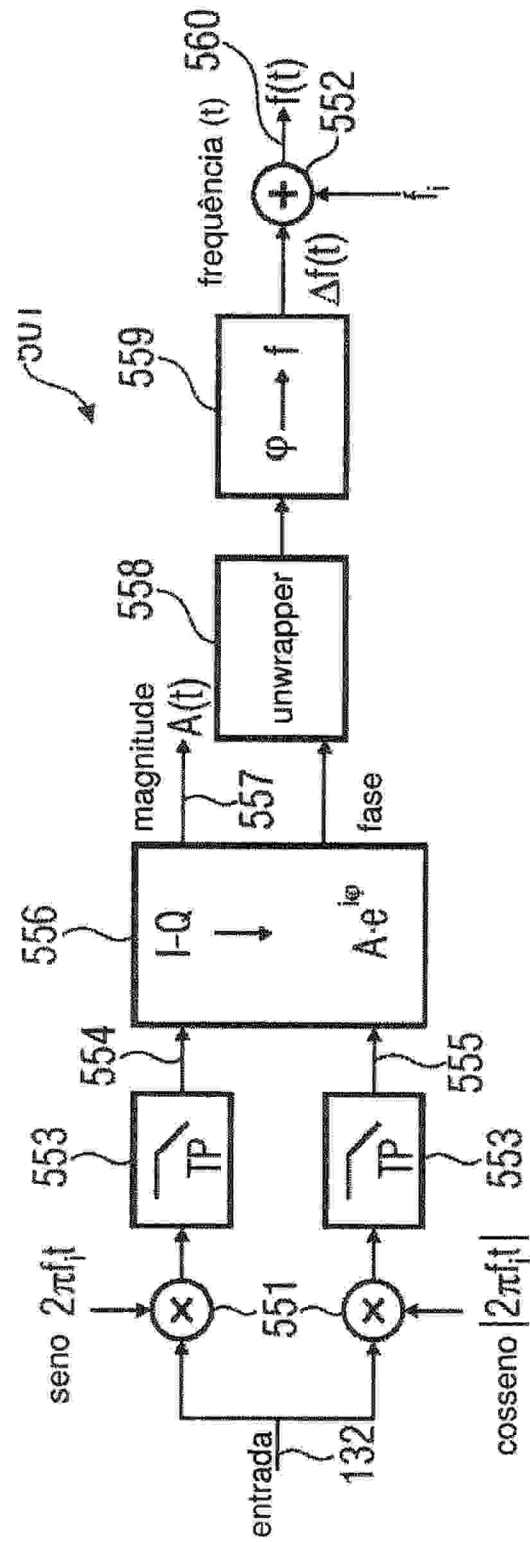


FIG 5B

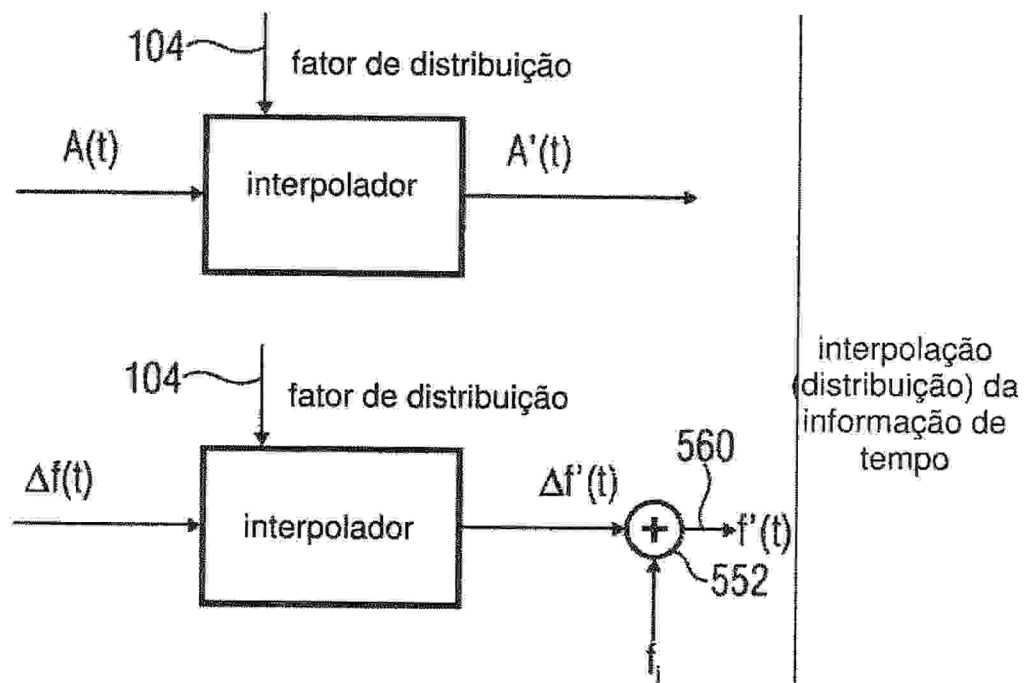


FIG 5C

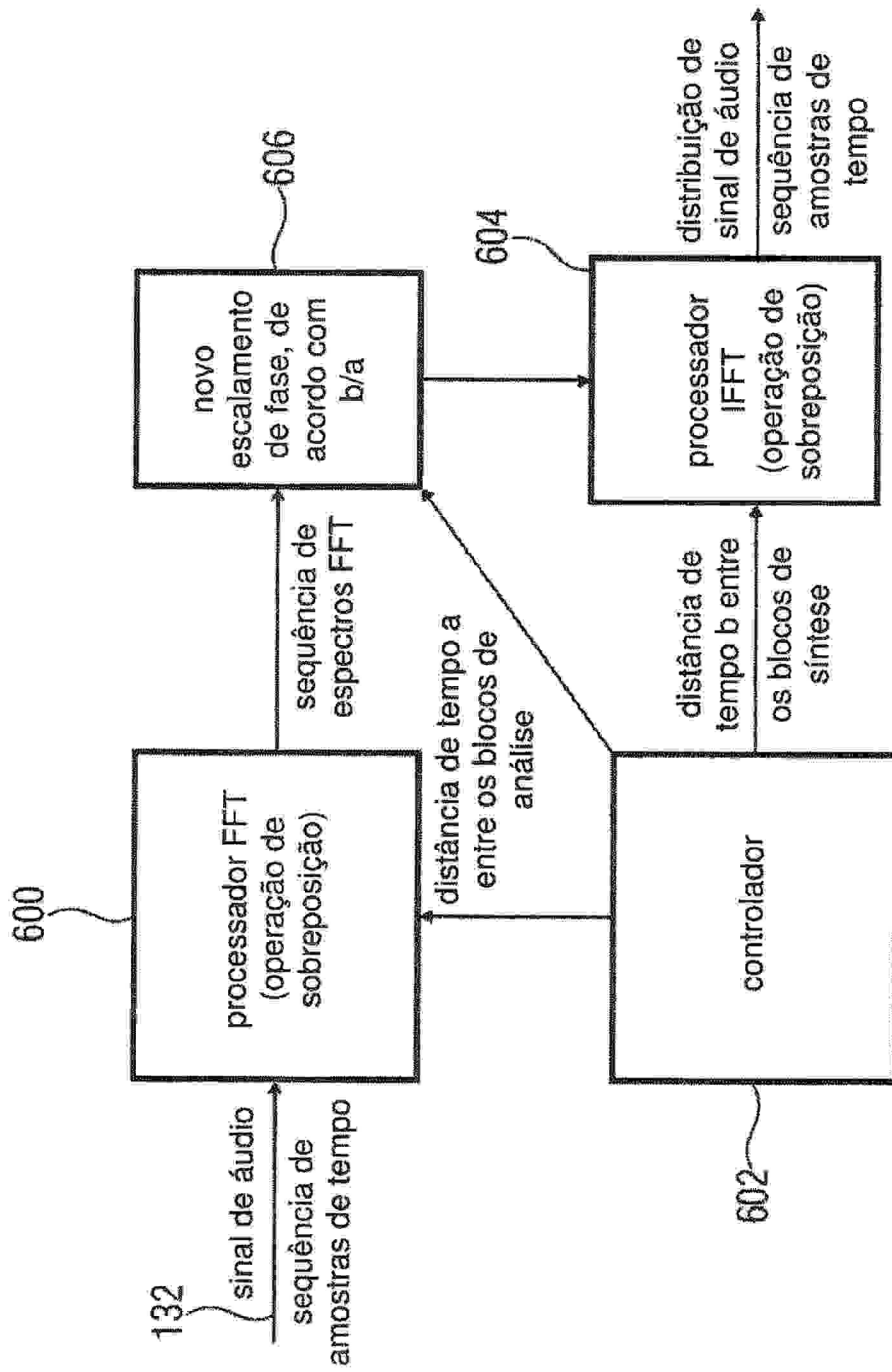
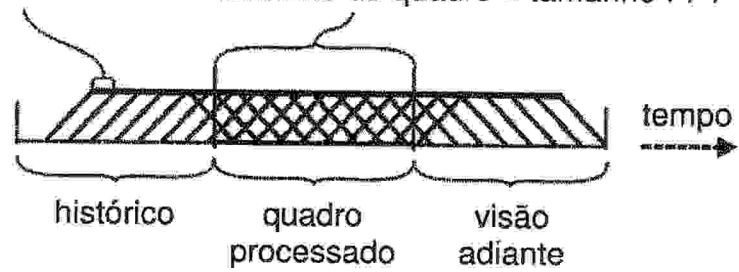


FIG 6
(implementação de transformada)

9/17

tamanho da grade de
tempo de análise =
FFT tamanho/8

tamanho de quadro = tamanho FFT



tamanho da grade de
tempo de síntese =
2.FFT tamanho/8

2.FFT tamanho/8

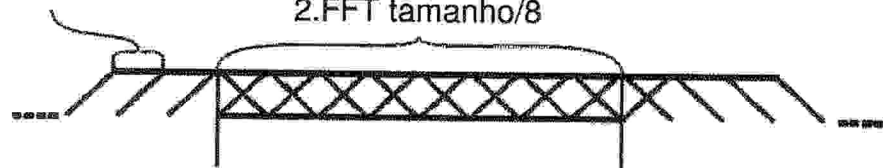


FIG 7

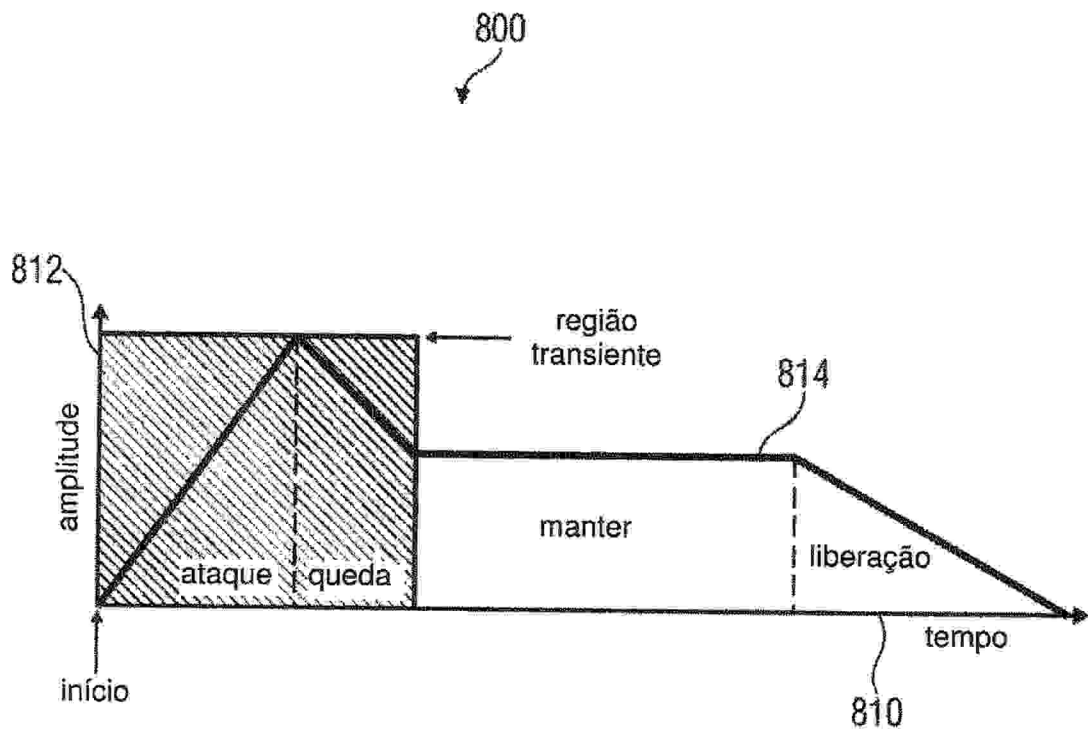
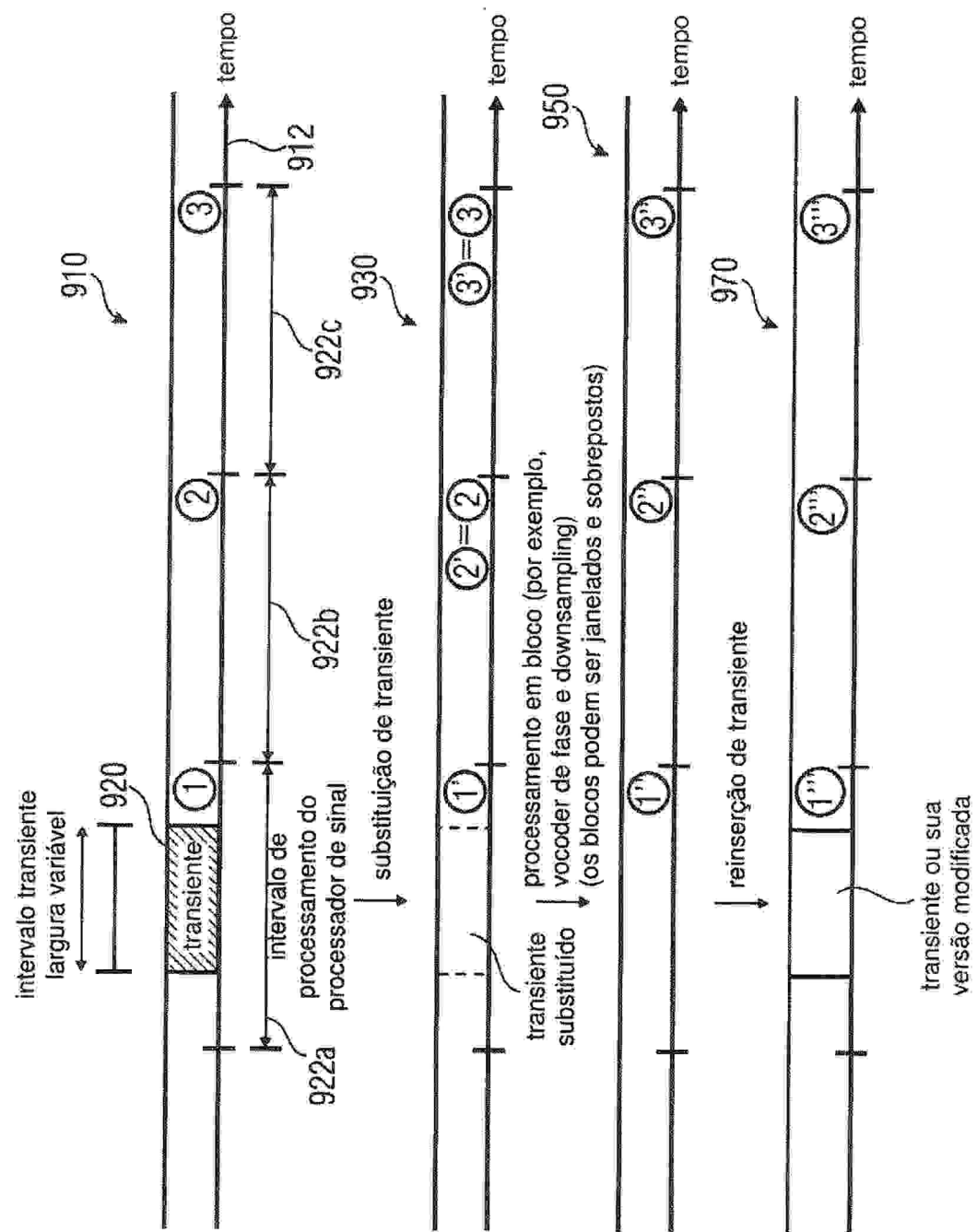


FIG 8

FIG 9



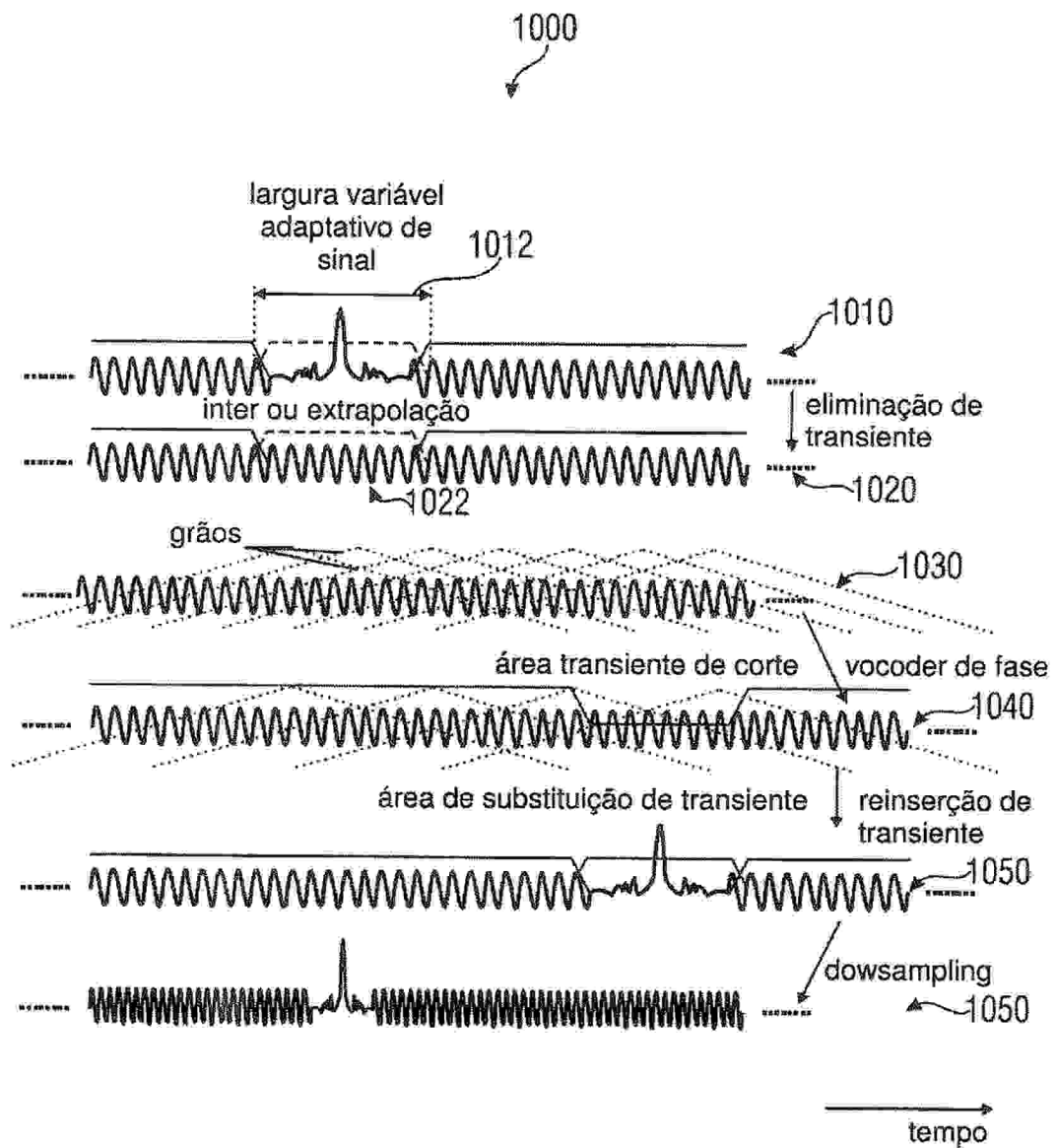


FIG 10

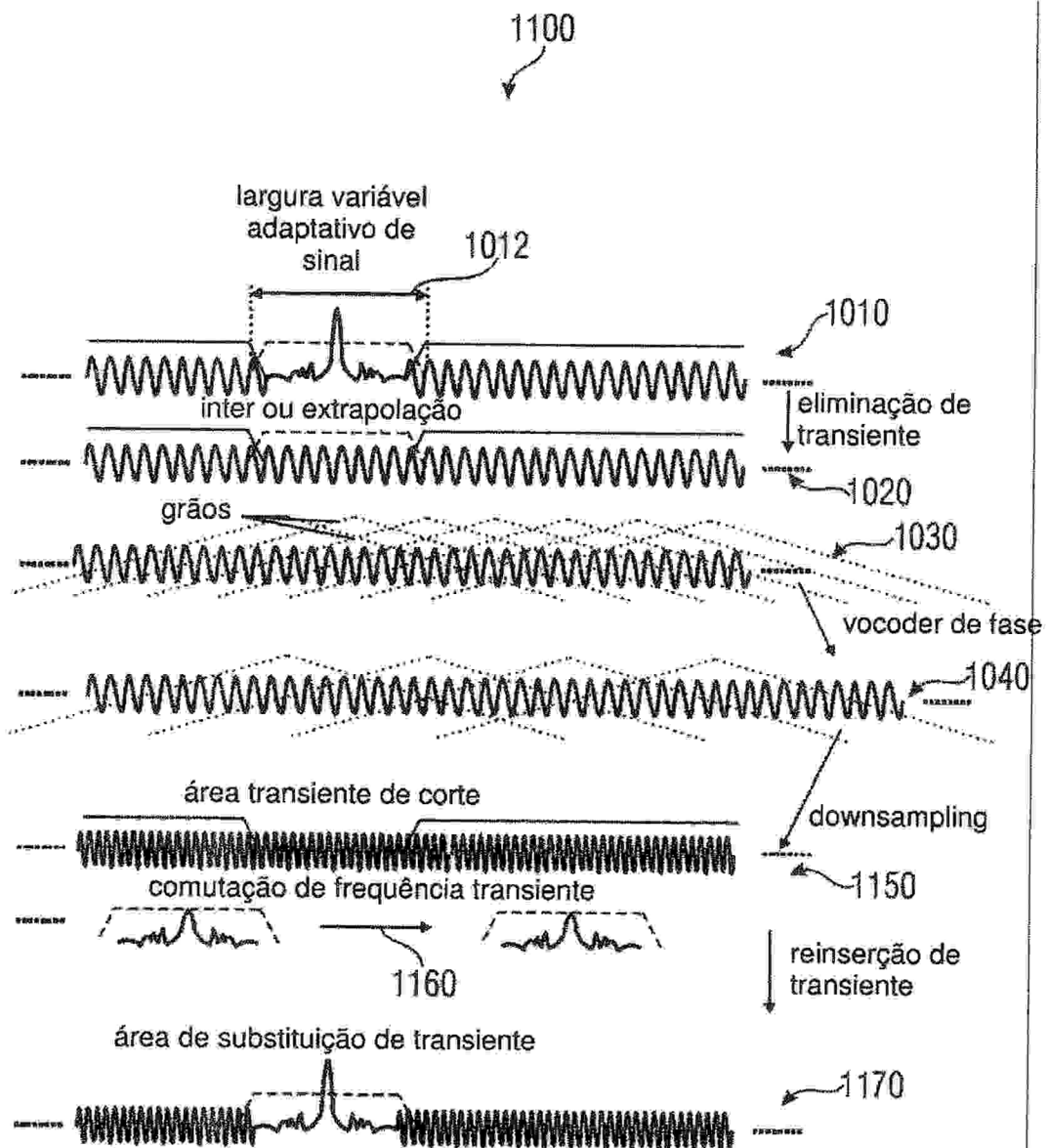


FIG 11

1200

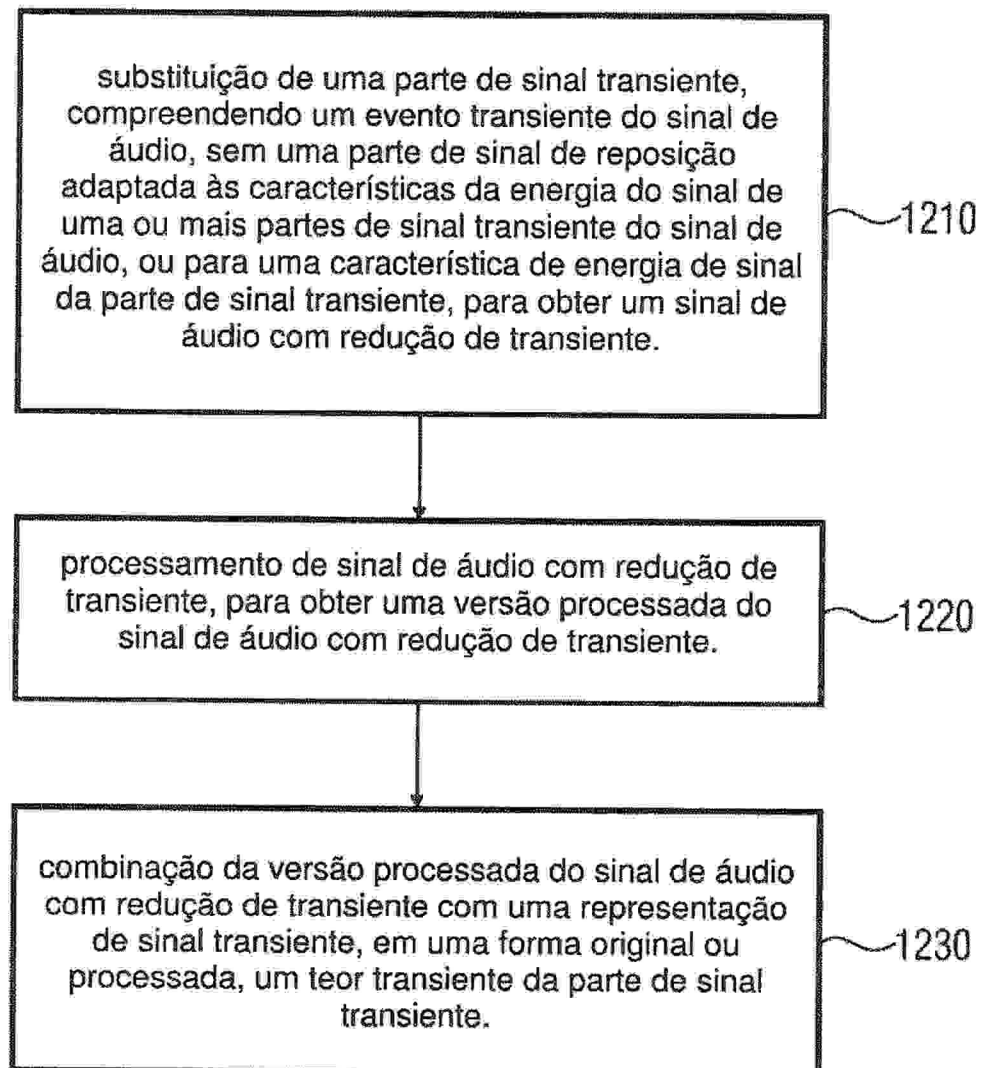


FIG 12

FIG 13

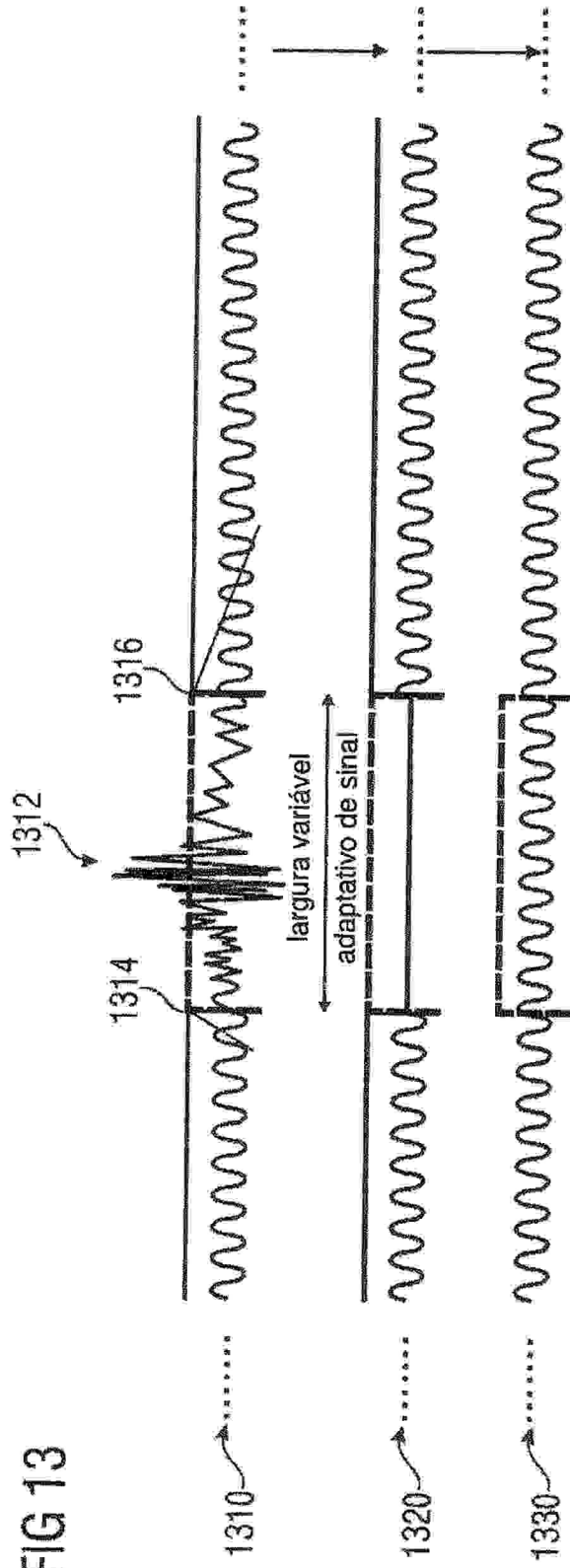


FIG 14

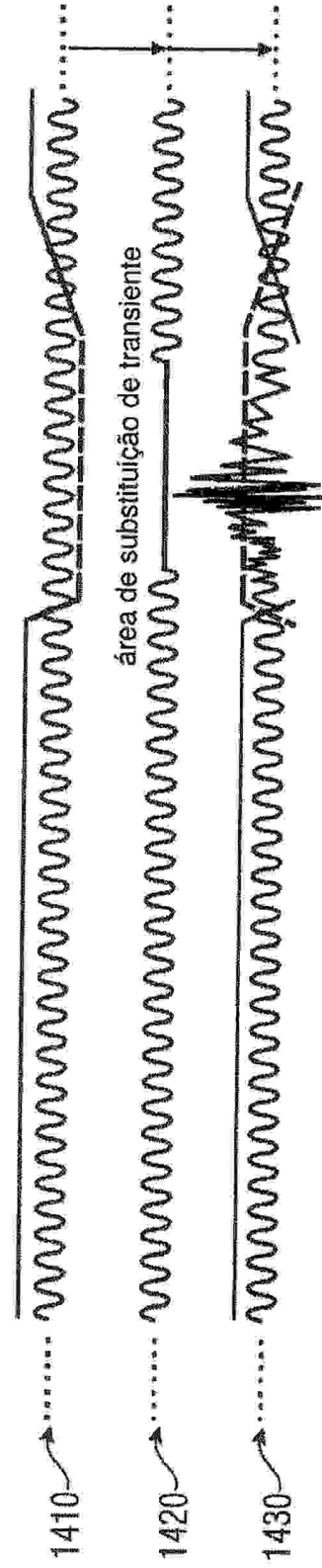
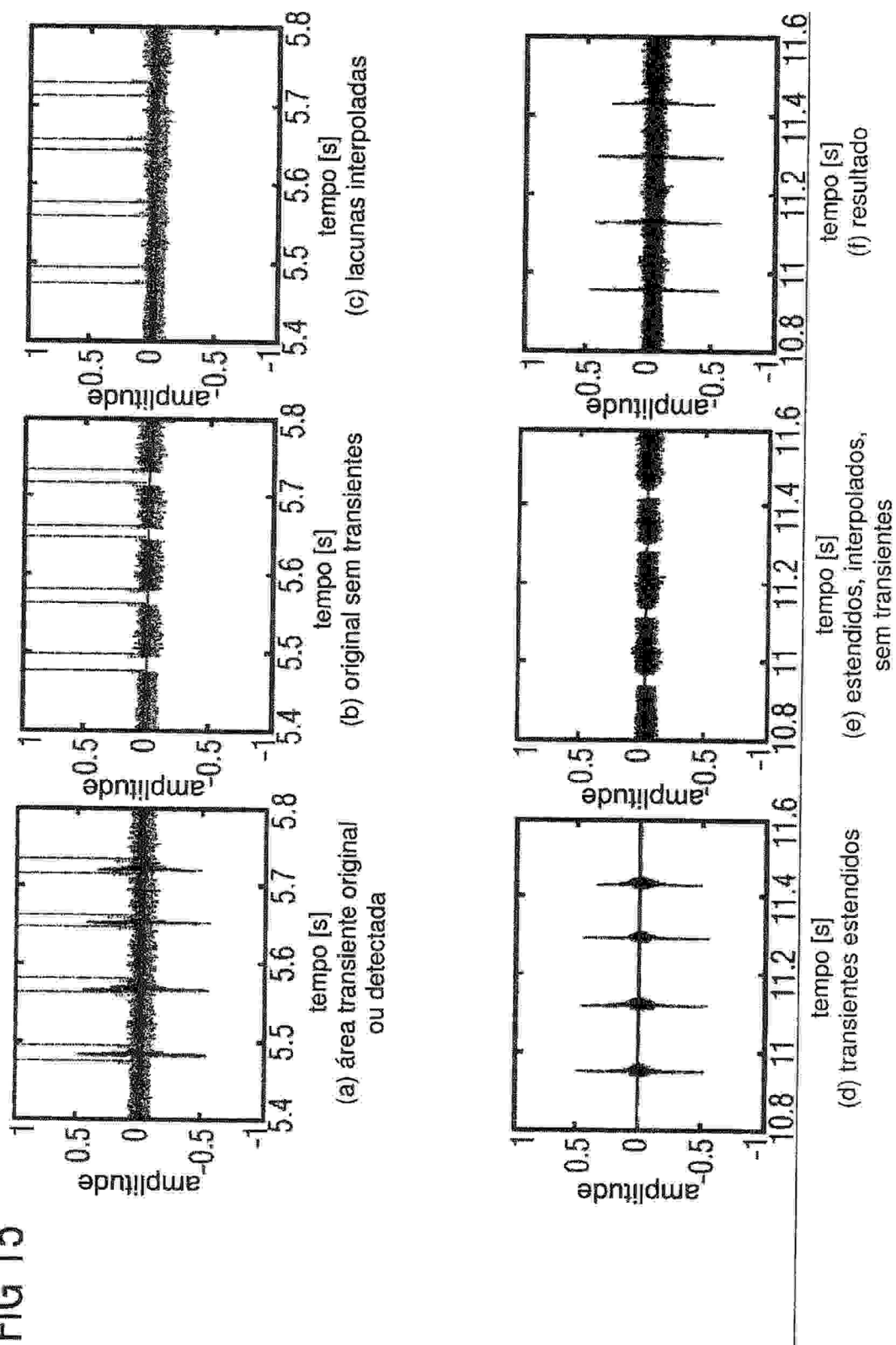


FIG 15



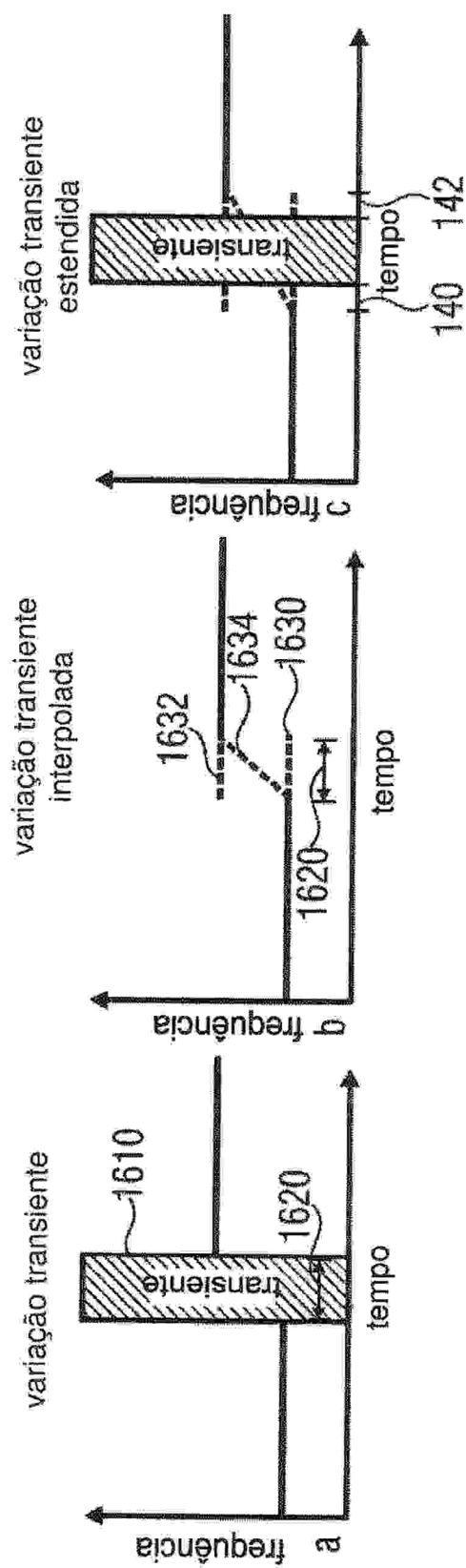


FIG 16