



INPI
INSTITUTO NACIONAL
DA PROPRIEDADE
INDUSTRIAL
Assinado
Digitalmente

REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DA INDÚSTRIA, COMÉRCIO EXTERIOR E SERVIÇOS
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

CARTA PATENTE Nº PI 0812029-3

O INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL concede a presente PATENTE DE INVENÇÃO, que outorga ao seu titular a propriedade da invenção caracterizada neste título, em todo o território nacional, garantindo os direitos dela decorrentes, previstos na legislação em vigor.

(21) Número do Depósito: PI 0812029-3

(22) Data do Depósito: 29/05/2008

(43) Data da Publicação do Pedido: 04/12/2008

(51) Classificação Internacional: G01L 19/00.

(30) Prioridade Unionista: GB 07 10211.4 de 29/05/2007.

(54) Título: MÉTODO DE RECUPERAR DADOS OCULTADOS, DISPOSITIVO DE TELECOMUNICAÇÕES, APARELHO DE OCULTAR DADOS, MÉTODO DE OCULTAR DADOS E CAIXA DE CONJUNTO SUPERIOR

(73) Titular: INTRASONICS S.A.R.L., Sociedade Luxemburguesa. Endereço: 12-14 Rue Léon Thyès, L-2636 Luxembourg, LUXEMBURGO(LU)

(72) Inventor: MICHAEL REYMOND REYNOLDS; PETER JOHN KELLY; JOHN RYE; IAN MICHAEL HOSKING.

Prazo de Validade: 10 (dez) anos contados a partir de 21/11/2018, observadas as condições legais

Expedida em: 21/11/2018

Assinado digitalmente por:

Alexandre Gomes Ciano

Diretor Substituto de Patentes, Programas de Computador e Topografias de Circuitos Integrados

Relatório Descritivo da Patente de Invenção para "MÉTODO DE RECUPERAR DADOS OCULTADOS, DISPOSITIVO DE TELECOMUNICAÇÕES, APARELHO DE OCULTAR DADOS, MÉTODO DE OCULTAR DADOS E CAIXA DE CONJUNTO SUPERIOR".

[001] A presente invenção refere-se a um sistema de comunicação. A invenção tem relevância particular, mas não exclusiva para sistemas de comunicações em que um aparelho de telefone como um telefone celular é provido com dados via um canal de dados acústicos.

[002] A WO02/45273 descreve um sistema de telefone celular em que dados ocultados podem ser transmitidos para um telefone celular dentro do áudio de um programa de televisão ou rádio. No contexto presente, os dados são ocultados no sentido de que eles são codificados no sentido de tentar esconder os dados no áudio de forma que não sejam obstrutivos para o usuário e sejam mascarados até certo ponto pelo áudio. Como aqueles versados na técnica observarão, o nível aceitável de audibilidade dos dados variará dependendo da aplicação e do usuário envolvido. Várias técnicas são descritas neste pedido antigo para codificar os dados dentro dos áudios, incluindo codificação de espectro espalhado, modulação de eco, codificação de faixa crítica, etc. No entanto, os inventores verificaram que o software de aplicação deve executar processamento significativo no sentido de ser capaz de recuperar os dados ocultados.

[003] Uma meta de uma modalidade, então, é para reduzir o requisito de processamento da aplicação de software.

[004] Em uma modalidade, é provido um método para recuperar dados ocultados de um sinal de áudio de entrada ou para identificar um sinal de áudio de entrada usando dispositivo de telecomunicações tendo um codificador de áudio para comprimir o sinal de áudio de entrada para

transmissão a uma rede de telecomunicações, sendo o método caracterizado por passar o sinal de áudio de entrada através da codificação - decodificação de áudio para gerar dados de áudio comprimidos e processar os dados de áudio comprimidos para recuperar os dados ocultados ou para identificar o sinal de áudio de entrada. Os inventores verificaram que passando a áudio de entrada através do codificador de áudio, a quantidade de processamento subsequente requerida para recuperar os dados ocultados ou para identificar o áudio de entrada pode ser significativamente reduzida. Em particular, este processamento pode ser executado sem ter que regenerar as amostras de áudio e então começar com as técnicas convencionais para recuperar os dados ocultados ou por identificar o sinal de áudio.

[005] Em uma modalidade, o codificador de áudio executa uma predição linear, LP, análise no áudio de entrada para gerar dados de LP representativos do áudio de entrada e em que a etapa de processamento processa os dados de LP para recuperar os dados ocultados ou para identificar o sinal de áudio de entrada. Preferencialmente, o codificador de áudio comprime os dados de LP para gerar os dados comprimidos de LP e a etapa de processamento inclui a etapa de regenerar os dados de LP dos dados de áudio comprimidos.

[006] Os dados de LP gerados pelo codificador podem incluir dados de filtro de LP, tais como coeficientes de filtro de LPC, polos de filtro ou frequências espectrais de linha e a etapa de processamento recupera os dados ocultados ou identifica o sinal de áudio usando este dados de filtro de LP.

[007] A etapa de processamento pode incluir a etapa de gerar uma resposta de impulso do filtro de síntese de LP ou a etapa de executar um algoritmo de Levinson-Durbin reverso nos dados de filtro de LP.

Quando gerando a resposta de impulso, sua autocorrelação é preferencialmente tomada de qual a presença ou ausência dos ecos pode ser identificada mais facilmente do que da resposta de impulso propriamente.

[008] Os dados de LP gerados pelo codificador de áudio podem incluir dados de excitação de LP (tais como índices de livro de código, posições de pulso de excitação, sinais de pulso, etc.) e a etapa de processamento pode recuperar os dados ocultos ou podem identificar o sinal de áudio usando estes dados de excitação de LP.

[009] Na maioria dos casos, os dados de LP incluirão tanto os dados de filtro de LP quanto os dados de excitação de LP e a etapa de processamento pode processar todos ou um subconjunto dos dados de áudio comprimidos correspondentes a um dos ditos dados de filtro de LP e ditos dados de excitação de LP para recuperar os dados ocultos.

[0010] Os dados podem ser ocultos dentro do sinal de áudio usando várias técnicas. No entanto, em uma modalidade preferida, os dados são ocultos no de áudio como um ou mais ecos do sinal de áudio. Os dados ocultos podem então ser recuperados por detectores de ecos. Cada símbolo dos dados a serem ocultos pode ser representado por uma combinação de ecos (ao mesmo tempo) ou como uma sequência de ecos dentro do sinal de áudio e a etapa de processamento pode incluir a etapa de identificar as combinações de ecos para recuperar os dados ocultos ou a etapa de monitorar a sequência de ecos no áudio para recuperar os dados ocultos.

[0011] Em uma modalidade, o codificador de áudio tem uma faixa de frequência operacional predefinida e os ecos são ocultos dentro do áudio dentro de uma parte predeterminada da faixa operacional, preferencialmente uma parte superior da faixa de frequência, e em que a etapa de processamento inclui uma etapa de filtrar para filtrar para fora frequências fora desta parte predeterminada. Por exemplo, onde o

codificador de áudio tem uma faixa operacional de 300 Hz a 3.4 kHz, o eco pode ser incluído somente na faixa entre 1 kHz e 3.4 kHz e mais preferencialmente entre 2 kHz e 3.4 kHz, já que isto pode reduzir os efeitos dos sinais de áudio cuja energia é tipicamente localizada dentro da parte mais baixa da largura de banda operacional. Em outra modalidade, o eco é incluído através da largura de banda operacional, mas a etapa de processamento ainda executa a filtragem, para reduzir os efeitos do áudio. Isto não é tão preferido já que parte do sinal de eco será perdida na filtragem também.

[0012] No sentido de ajudar a identificar a presença de ecos na saída de codificador de áudio, a etapa de processamento pode determinar um ou mais valores de autocorrelação, que ajuda a destacar os ecos. Filtragem intra-estrutura dos valores de autocorrelação pode ser também executada para reduzir os efeitos de componentes de áudio variando lentamente.

[0013] O codificador de áudio usado pode ser qualquer um de um grande número codificadores conhecidos como um codificador CELP, codificador AMR, codificador AMR de banda larga etc.

[0014] Em uma modalidade, a etapa de processamento pode determinar um gráfico de espectro da saída de dados de áudio comprimidos do codificador e então identificar recursos característicos (semelhantes a uma impressão digital) no gráfico de espectro. Estes recursos característicos identificam a entrada de áudio e podem ser usados para determinar informação de trilha para o áudio para produzir para o usuário ou que pode ser usado para sincronizar o dispositivo de telecomunicações para o sinal de áudio, por exemplo, produzir legendas relativas ao áudio.

[0015] Outra modalidade provê um dispositivo de telecomunicações, compreendendo: meios para receber sinais acústicos e para converter sinais acústicos recebidos em sinais de áudio

elétricos correspondentes, meios para testar os sinais de áudio elétricos para produzir amostras de áudio digital, meios de codificação de áudio para comprimir as amostras de áudio digital para gerar dados de áudio comprimidos para transmissão a uma rede de telecomunicações; e meios de processamento de dados, acoplados aos ditos meios de codificação de áudio, para recuperar dados ocultados fornecidos dentro do sinal acústico recebido ou para identificar o sinal acústico recebido.

[0016] Uma modalidade provê um aparelho para ocultar dados compreendendo: meios de codificação de áudio para receber e comprimir amostras de áudio digital representativas de um sinal de áudio para gerar dados de áudio comprimidos; meios para receber dados a serem ocultados dentro do sinal de áudio e para variar os dados de áudio comprimidos na dependência dos dados recebidos, para gerar dados de áudio comprimidos modificados; e meios para gerar amostras de áudio usando os dados de áudio comprimidos modificados, as amostras de áudio representando o sinal de áudio original e transportando os dados ocultados.

[0017] Outra modalidade fornece um método de dados ocultados em um sinal de áudio, o método compreendendo as etapas de adicionar um ou mais ecos ao áudio na dependência dos dados a serem ocultados no sinal de áudio e é caracterizado por filtrar o eco com filtro de passa alta antes de combiná-lo com o sinal de áudio. Os inventores verificaram que adicionando o eco somente em uma faixa de frequência mais alta do sinal de áudio, os ecos podem ser descobertos mais facilmente e reduz energia perdida em aplicar o eco através da faixa de áudio.

[0018] Estes e outros aspectos da invenção se tornarão claros da descrição detalhada seguinte de modalidades exemplificativas que são descritas com referência aos desenhos anexos, em que:

[0019] A figura 1 mostra esquematicamente um sistema de sinalizar para comunicar dados para um telefone celular via a parte de áudio de

um sinal de televisão;

[0020] A figura 2 é um diagrama em bloco esquemático ilustrando os componentes principais de um telefone celular incluindo aplicações de software para recuperar dados ocultos dentro de um sinal de áudio recebido;

[0021] A figura 3a é um diagrama em bloco esquemático ilustrando o processo executado por uma codificação - decodificação de áudio formando parte do telefone celular ilustrado na Figura 2;

[0022] A figura 3b ilustra um modelo de fonte-filtro subjacente à codificação de LP de sinais de áudio;

[0023] A figura 3c ilustra o modo no qual um filtro LPC inverso pode ser usado para gerar uma excitação ou sinal residual de um sinal de áudio de entrada;

[0024] A figura 4 é um diagrama em bloco esquemático ilustrando o processamento executado na saída codificação - decodificação de áudio para recuperar dados ocultos dentro do sinal de áudio;

[0025] A figura 5 é uma plotagem de autocorrelação da qual os dados ocultos podem ser determinados;

[0026] A figura 6 é um diagrama em bloco esquemático que ilustra um processo alternativo que pode ser executado para recuperar os dados ocultos;

[0027] A figura 7 é um diagrama em bloco esquemático que ilustra um modo alternativo adicional no qual os dados ocultos podem ser recuperados da saída codificação - decodificação de áudio;

[0028] A figura 8 é um diagrama em bloco esquemático que ilustra o modo no qual os dados ocultos podem ser recuperados da saída de parâmetros de excitação pela codificação - decodificação de áudio de áudio;

[0029] A figura 9 é uma saída de plotagem de autocorrelação pela seção de autocorrelação formando parte do circuito mostrada na Figura

8, da qual os dados ocultos podem ser identificados;

[0030] A figura 10 é diagrama em bloco esquemático que ilustra um refinamento para o circuito de processamento mostrada na Figura 4, no qual a resposta de impulso de um filtro de síntese de LPC é filtrado com filtro de passa alta para reduzir os efeitos de componentes de áudio de baixa frequência;

[0031] A figura 11 é um diagrama em bloco esquemático que ilustra um refinamento adicional do circuito de processamento mostrada na Figura 4 em que os coeficientes de LPC são filtrados com filtro de passa alta para remover coeficientes de ordem inferior relativo a componentes de áudio de frequência mais baixa;

[0032] A figura 12 ilustra um refinamento adicional da circuito de processamento mostrada na Figura 4 em que a plotagem de autocorrelação ilustrada na Figura 5 é filtrada por filtro de passa alta para remover auto correlações variando lentamente;

[0033] A figura 13 é um diagrama em bloco esquemático geral ilustrando um modo no qual os dados ocultos podem ser codificados dentro do áudio anterior à recepção pelo telefone celular;

[0034] A figura 14 é um diagrama em bloco geral que ilustra o modo no qual o telefone celular recupera os dados codificados usando o sistema ilustrado na Figura 13;

[0035] A figura 15 é um diagrama em bloco que ilustra um modo no qual os parâmetros gerados por um codificador de LPC podem ser modificados e recombinaados com um sinal residual para formar o áudio modificado para transmissão para o telefone celular; e

[0036] A figura 16 ilustra um modo alternativo no qual os parâmetros de excitação obtidos de um codificador de LPC são modificados e do qual um sinal residual é gerado para uso em sintetizar o áudio modificado com os coeficientes de LPC obtidos do codificador de LPC;
e

[0037] A figura 17 é um diagrama em bloco que ilustra o modo no qual a saída da codificação - decodificação de áudio pode ser processada para recuperar um gráfico de espectro para o áudio de entrada para usar na identificação ou caracterização do sinal de áudio de entrada.

Visão geral

[0038] A figura 1 ilustra uma primeira modalidade da invenção em que um sinal de dados $F(t)$, gerado por uma fonte de dados 1, é codificado dentro de uma trilha de áudio de uma fonte de áudio 3 por um codificador 5 para formar uma trilha de áudio modificado para um programa de televisão. Nesta modalidade, o sinal de dados $F(t)$ transporta sinais de gatilho para sincronizar a operação de uma aplicação de software executando em um telefone móvel 21 de um usuário com o programa de televisão. Como mostrado na Figura 1, a saída da trilha de áudio modificado pelo codificador 5 é então combinada com a correspondente trilha de vídeo, de uma fonte de vídeo 7, em um gerador de sinal 9 para formar um sinal de televisão transportando o programa de televisão. Nesta modalidade, a fonte de dados 1, a fonte de áudio 3, a fonte de vídeo 7 e o codificador 5 estão todos localizados em um estúdio de televisão e o sinal de televisão é distribuído por uma rede de distribuição 11 e, nesta modalidade, um sinal de frequência de rádio (FR) 13. O sinal de FR 13 é recebido por uma antena de televisão 15 que prove o sinal de televisão para uma televisão convencional 17. A televisão 17 tem uma tela (não mostrada) para mostrar a trilha de vídeo e um alto falante não mostrado para produzir a trilha de áudio modificado como um sinal acústico 19. Como mostrado, nesta modalidade, o telefone celular 21 detecta o sinal acústico 19 emitido pela televisão 17 usando um microfone 23 que converte o sinal acústico detectado em um sinal elétrico correspondente. O telefone celular 21 então decodifica o sinal elétrico

para recuperar o sinal de dados $F(t)$. O telefone celular 21 tem também componentes convencionais como um alto falante 25, uma antena 27 para comunicar com uma estação celular de base 35, uma tela 29, um teclado 31 para entrar números e letras e menu 33 para acessar opções de menu. Os dados recuperados do sinal de áudio podem ser usados para vários propósitos diferentes, como explicado na WO02/45273. Uma aplicação é para a sincronização de uma aplicação de software sendo executada no telefone celular 21 com o programa de televisão sendo mostrado na televisão 17. Por exemplo, pode haver um show de jogo de perguntas sendo mostrado na televisão 17 e o telefone celular 21 pode ser organizado para gerar e exibir perguntas relativas ao jogo de perguntas mostrado em sincronismo com o jogo de perguntas mostrado. As perguntas podem, por exemplo, ser pré-armazenadas no telefone celular 21 e produzidas quando um código de sincronização adequado é recuperado do sinal de dados $F(t)$. No fim do jogo de perguntas, a entrada de respostas pelo usuário no telefone celular 21 (via o teclado 31) pode então ser transmitida para um servidor remoto 41 via a estação de base de telefone celular 35 e a rede de telecomunicações 39. O servidor 41 pode então coletar as respostas recebidas de um grande número de usuários e ordená-las baseado no número de respostas corretas dadas e o tempo tomado para introduzir as respostas.

[0039] Esta informação de contagem de tempo poderia ser também determinada pelo telefone celular 21 e transmitida para o servidor remoto 41 junto com as respostas do usuário. Como aqueles versados na técnica observarão, o servidor 41 pode também processar a informação recebida dos diferentes usuários e coletar várias informações de perfil de usuário que ele pode armazenar na base de dados 43. Estas informações de perfil de usuário podem então ser usadas, por exemplo, para publicidade direcionada.

[0040] Depois que o servidor 41 identificou um ou mais usuários "vencedores", informação ou um prêmio pode ser enviado para aqueles usuários.

[0041] Por exemplo, uma mensagem pode ser enviada para eles através da rede de telecomunicações 39 junto com um cupom ou outro vale. Como mostrado pela linha tracejada 44 na Figura 1, o servidor 41 pode prover também a fonte de dados 1 com os dados a serem codificados dentro do áudio.

[0042] Como mencionado acima, os inventores imaginaram que o processamento requerido para ser executado pelo software que é executado no telefone celular 21 pode ser reduzido fazendo uso da codificação sendo executada pelo chip de codificação - decodificação de áudio dedicado. Em particular, os inventores verificaram que usando o processo de codificação inerente na codificação - decodificação de áudio como uma etapa inicial do processo de decodificação para recuperar os dados ocultos, reduz o processamento requerido pelo software para recuperar os dados ocultos.

Telefone celular

[0043] A figura 2 ilustra os componentes principais do telefone celular 21 usados nesta modalidade. Como mostrado, o telefone celular 21 inclui um microfone 23 para receber sinais acústicos e para convertê-los em sinais elétricos equivalentes. Estes sinais elétricos são então filtrados pelo filtro 51 para remover frequências não desejadas tipicamente fora da faixa de frequência de 300 Hz a 3.4 kHz (como definido no documento padrão EN300903, publicado por ETSI). O áudio filtrado é então digitalizado por um conversor analógico para digital 53, que testa o áudio filtrado em uma frequência de amostragem de 8 kHz, representando cada amostra tipicamente por um valor digital de 13 a 16 bits. O fluxo de áudio digitalizado (D(t)) é então introduzido na codificação - decodificação de áudio 55, o qual é uma codificação -

decodificação Adaptive MultiRate (AMR), cuja operação é descrita abaixo. A saída de áudio comprimido pela codificação - decodificação AMR 55 é então passada para uma unidade de processamento de RF 57 que modula o áudio comprimido sobre um ou mais sinais portadores de RF para transmissão para a estação de base 35 via a antena 27. Semelhantemente, os sinais de áudio comprimidos recebidos via a antena 27 são alimentados para a unidade de processamento de RF 57, que demodula os sinais de RF recebidos para recuperar os dados de áudio comprimidos do sinal portador de RF(s), que são passados para a codificação - decodificação AMR 55. A codificação - decodificação AMR 55 então decodifica os dados de áudio comprimidos para regenerar as amostras de áudio representadas desta forma, as quais são saídas para o alto falante 25 via o conversor digital para analógico 59 e o amplificador 61.

[0044] Como mostrado na Figura 2, a saída de dados de áudio comprimidos da codificação - decodificação AMR 55 (ou a unidade de processamento de RF 57) é também passada ao processador 63, que é controlado por softwares armazenado na memória 65. O software inclui software de sistema operacional 67 (para controlar a operação geral do telefone celular 21), um Browser 68 para acessar o software de internet e aplicação 69 para fornecer funcionalidade adicional ao telefone celular 21. Nesta modalidade, o software de aplicação 69 é configurado para fazer com que o telefone celular 21 a interaja com o programa de televisão na maneira discutida acima. Para fazer isto, o software de aplicação 69 é organizado para receber e processar a saída de dados de áudio comprimido da codificação - decodificação AMR 55 para recuperar os dados ocultos $F(t)$ que controlam o software de aplicação 69. Como será descrito em mais detalhe abaixo, o processamento dos dados de áudio comprimidos para recuperar os dados ocultos $F(t)$ pode ser executado sem ter que regenerar as

amostras de áudio digitalizado e ainda reduzindo o processamento que teria sido requerido pela aplicação de software 69 para recuperar os dados ocultados diretamente das amostras de áudio digitais.

[0045] A fim de recuperar os dados ocultados, o software de aplicação 69 é organizado para gerar e produzir dados (por exemplo, perguntas para o usuário) na tela 29 e para receber as respostas introduzidas pelo usuário via o teclado 31. A aplicação de software 69 então transmite as respostas do usuário para o servidor remoto 41 (identificado por um URL pré-armazenado, número E.164 ou o similar) junto com dados de contagem de tempo indicativos do tempo tomado pelo usuário para introduzir cada resposta (calculada pela aplicação de software 69 usando um temporizador interno (não mostrado)). A aplicação de software 69 pode exibir também informações de resultado recebidas de volta do servidor 41, indicativo de quão bem o usuário fez relativamente a outros usuários que fizeram parte no jogo de perguntas.

Codificação - decodificação AMR

[0046] Embora a codificação - decodificação AMR 55 seja bem conhecida e definida pelo corpo de padrões 3GPP (na documentação de Padrões TS 26.090 modalidade 3.1.0), uma descrição geral do processamento que ele executa, será dada agora com referência à figura 3 para que o leitor possa entender a descrição subsequente do processamento executado pelo software de aplicação 69.

[0047] A codificação - decodificação AMR 55 (codificador - decodificador Adaptive-Multi-Rate) converte áudio de dados amostrado de 8 kHz, na faixa de 300 Hz a 3.4 kHz em um fluxo de bits em um número de diferentes taxas de bits. A codificação - decodificação 55 é portanto altamente adaptada para situações onde as taxas de transmissão podem ser requeridas a variar. Sua taxa de bit de saída pode ser adaptada para igualar as condições de transmissão prevalentes, e por essa razão ela é um padrão 3G e atualmente usado

na maioria dos telefones celulares 21.

[0048] Embora a taxa de bit seja variável, os mesmos processos decodificação fundamental são empregados pela codificação - decodificação 55 em todas as taxas. Os processos de quantificação, a seleção de quais parâmetros são para ser transmitidos e a taxa de transmissão são variados para alcançar operação nas oito taxas de bit ou modos: 12.2, 10.2, 7.95, 7.40, 6.70, 5.90, 5.15 e 4.75 Kbits/s. Nesta modalidade o modo de taxa de 15 bit usado mais alto é (12.2 Kbits/s).

[0049] Existem quatro subsistemas de componentes principais na codificação - decodificação AMR 55 que são descritos abaixo. Eles são:

- Predição de Tom
- Análise de LPC
- Consulta de Livro de Código Fixado
- Livro de Código Adaptável

[0050] A codificação - decodificação AMR 55 os aplica naquela ordem, embora para propósitos presentes é mais fácil tratar predição de tom por último e como parte do processamento de livro de código adaptável. A codificação - decodificação AMR 55 é construída em torno de um sistema de codificação CELP (Predição Linear Excitada de Livro de Código). O sinal de áudio de entrada é dividido em 160 estruturas de amostra (f) e as estruturas são sujeitas à análise de predição linear para extrair um número pequeno de coeficientes por estrutura para codificar e transmitir. Estes coeficientes caracterizam o espectro a curto prazo do sinal dentro da estrutura. Além destes coeficientes, a codificação - decodificação AMR 55 computa também um LPC residual (também chamado da excitação) que é codificado usando os livros de código adaptáveis e fixados ajudados pelo preditor de tom. Estes subsistemas são descritos abaixo.

Análise de LPC

[0051] A análise de LPC é executada pela seção de análise de LPC

71 mostrada na Figura 3a. A LPC assume o modelo de fonte-filtro clássico de produção de fala (ilustrado na Figura 3b) no qual a fala é considerada como a saída de um filtro variando lentamente no tempo (filtro de síntese de LPC 72), excitado por pulsos regulares de glote para fala com voz, tais como em vogais, e ruído branco para fala sem voz, por exemplo /sh/, ou uma mistura dos dois para sons de misturados voz, tais como /z/ (representado pelo bloco de excitação 74). Embora baseado em um modelo de saída de fala, ele também prove um modelo válido para codificar todos os sons. O filtro de síntese 72 é assumido para ser todo-polo, isto é. ele tem somente ressonâncias. Esta suposição é a base do método de análise de LPC. Na notação de dados amostrados (plano-z) ele significa que a função de transferência é puramente um polinômio em z^{-1} no denominador da função de transferência, $H(z)$.

$$H(z) = \frac{1.0}{A(z)} = \frac{1.0}{\sum_{i=1}^p a_i z^{-i}} \quad (1)$$

[0052] A resposta da série de tempo S_n deste filtro para a excitação de entrada é então:

$$s_n = e_n + \sum_{i=1}^p a_i s_{n-i} \quad (2)$$

[0053] que diz que a saída s_n do sistema é a entrada, e_n mais uma soma linear ponderada das saídas p prévias. Isto é a base teórica de LPC. O limite p é a "ordem" de LPC que é normalmente fixada e na codificação - decodificação AMR 55 p é igual a dez. Na Codificação - decodificação AMR 55 (e outros sistemas baseados em LPC) a análise de predição linear é empregada para estimar os pesos ou coeficientes de filtro a_i , para cada estrutura do áudio de entrada. Uma vez estimados, eles são então convertidos em uma forma adequada para quantificação e transmissão.

[0054] Estimar eficazmente os coeficientes a_i requer aproximações e suposições a serem feitas. Todos os métodos de resolver os

coeficientes têm como meta minimizar a contribuição de e_n na equação (2) acima. A codificação - decodificação AMR 55 usa o método de autocorrelação, que significa resolver equações p lineares simultâneas; na forma de matriz:

$$\begin{pmatrix} R(0) & \cdots & R(p-1) \\ \vdots & \ddots & \vdots \\ R(p-1) & \cdots & R(0) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} R(1) \\ \vdots \\ R(p) \end{pmatrix} \quad (3)$$

[0055] ou em uma forma mais abreviada:

$$R_{ij} a_i = r_i \quad (4)$$

[0056] Os elementos, r_{ij} de R são os valores de autocorrelação para o sinal de áudio de entrada de atraso $|i - j|$. Como R é simétrica e todos os elementos de cada diagonal são iguais, ela está aberta a métodos recursivos rápidos para encontrar seu inverso. O algoritmo de Levinson-Durbin é usado no codificador de AMR 55.

Frequências Espectrais de Linha

[0057] Os coeficientes a_i são realmente difíceis para quantificar. Eles mudam bastante imprevisivelmente com o tempo e têm valores positivos e negativos sobre uma faixa indeterminada. A codificação - decodificação AMR 55 então usa uma seção de determinação de LSF 73 para converter estes coeficientes para frequências espectrais de linha antes da quantificação, que remove estas desvantagens e permite para a codificação eficiente dos coeficientes de LPC. Os coeficientes a_i são os pesos do filtro de síntese de todos-polos 72 e são os coeficientes de ordem p de um polinômio em z^{-1} , o qual pode ser fatorado para encontrar suas raízes. Estas raízes são as ressonâncias ou polos no filtro de síntese 72. Estes polos foram frequentemente quantificados para transmissão enquanto eles estão razoavelmente ordenados, têm valores médios e mudam mais premeditavelmente de estrutura para estrutura, que dá oportunidade para economia de bits, a qual a codificação a_i não faz. Frequências espectrais de linha (LSFs) são até melhores para isto do que os polos. É importante compreender que as

LSFs não são o mesmo que os polos do modelo todos-polos mas eles são relacionados. Sua derivação é envolvida, mas qualitativamente ela envolve escolher dois conjuntos de condições de limite em uma representação particular do filtro de síntese, uma condição de limite correspondente à quando a glote está perfeitamente aberta e a outra correspondente à quando a glote está perfeitamente fechada. Isto resulta em dois conjuntos de polos hipotéticos com largura de banda zero, isto é, ressonadores perfeitos.

[0058] As vantagens principais das LSFs são que:

[0059] as LSFs consistem em uma frequência somente, sua largura de banda é sempre zero (embora existam duas vezes tantas LSFs como existem polos)

[0060] as LSFs são teoricamente melhor ordenadas do que os polos)

[0061] As LSFs são deste modo sensíveis à codificação de taxa de bit muito baixa. Em particular, como mostrado na Figura 3a, a média (computada com antecedência e armazenada no armazenamento de dados 75) de cada LSF pode ser subtraída pela seção de subtração de média 77. Adicionalmente, já que o delta de LSF resultante não muda rapidamente com o tempo, um preditor 79 pode então ser usado para prever o valor de delta atual, que é subtraído do delta real pela seção de subtração de predição 81.

[0062] Os dados resultantes são então adicionalmente codificados por uma seção de quantificação vetorial (VQ) 83 a qual codifica dois valores de uma vez via um único índice, resultando em menos do que 1 bit por valor em alguns casos. A codificação - decodificação AMR 55 produz os valores de índice VQ deste modo obtidos para a estrutura atual como os dados de LPC codificado para transmissão para a estação de base 35.

Excitação de LP

[0063] Como mencionado acima, a codificação - decodificação AMR 55 codifica também a parte de excitação 74 do modelo ilustrado na Figura 3b. No sentido de fazer isto, a codificação - decodificação AMR 55 gera uma representação do sinal de excitação de forma que ela possa então codificá-lo. Como ilustrado na Figura 3c, ela faz isto gerando um filtro de LPC "inverso" 76 que pode gerar o sinal de excitação filtrando o sinal de áudio de entrada. O sinal de excitação obtido do filtro inverso 76 é às vezes também chamado de residual. Este filtro de LPC inverso 76 é realmente definido dos mesmos coeficientes a_i determinados acima, mas usando-os para definir um modelo todo-zero com a função de transferência:

$$A(z) = 1.0 + \sum_{i=1}^p a_i z^{-i} \quad (5)$$

[0064] Isto corresponde no domínio do tempo a um filtro:

$$e_n = s_n + \sum_{i=1}^p a_i s_{n-i} \quad (6)$$

[0065] O filtro de LPC inverso 76 definido por (6) consiste em zeros cancelando fora os polos no filtro de síntese todos-polos 72 definido por (2). Em teoria, se o sinal de áudio de entrada é filtrado usando o filtro inverso 76 e então o sinal de excitação gerado é filtrado pelo filtro de síntese 72, então chega-se de volta ao sinal de áudio de entrada (consequentemente o nome filtro LPC "inverso"). É importante notar que o sinal de áudio original não precisa ser fala para ocorrer uma reconstrução perfeita. Se a análise de LPC não fez um bom trabalho em representar o sinal de áudio de entrada, então existirá mais informação no residual.

[0066] É o trabalho da seção de livro de código fixa 87 e a seção de livro de código adaptável 89 da codificação - decodificação AMR 55 codificar o sinal de excitação. Um número relativamente grande de bits são usados na codificação - decodificação AMR 55 para codificar a excitação quando comparado ao número de bits usados para codificar

os LSFs: 206 fora de 244 por estrutura (84%) em modo de 12.2 Kbits/s e 72 fora de 95 (74%) em modo de 4.75 Kbits/s. É este uso de bits que permite a codificação - decodificação 20 AMR 55 codificar sinais de não-fala com algum efeito.

[0067] A excitação em fala com voz é caracterizada por uma série de cliques (pulsos) no tom de voz (mais ou menos 100 Hz a 130 Hz para um macho de adulto em fala normal, duas vezes aquelas para fêmeas e crianças). Em fala de não-voz ela é ruído branco (mais ou menos). Em fala misturada ele é uma mistura. Uma maneira de pensar sobre a excitação como o residual é imaginar que a análise de LPC retira as pancadas no espectro a curto prazo de áudio, deixando um residual com um espectro muito nivelado. Isto se aplica qualquer que seja a entrada de sinal.

[0068] Na codificação - decodificação AMR 55 o sinal de excitação é codificado como a combinação de um livro de código fixo e uma saída de livro de código adaptável. O livro de código adaptável não existe como qualquer coisa para consultar, mas é uma cópia das combinações prévias das saídas de livro de código combinadas alimentadas de volta no período predito pelo preditor de tom.

O Livro de Código Fixo

[0069] A seção de livro de código fixo 87 gera o sinal de excitação (e_i) para a estrutura atual usando a saída de coeficientes de LPC a_i da seção de análise de LPC 71 para a estrutura atual, fixar os pesos do filtro inverso 76 definido na equação (6) acima; e filtrando a estrutura atual do áudio de entrada com este filtro. A seção de livro de código fixa então identifica os pulsos ou padrões de livro de código fixo (armazenados no livro de código fixo 88) que melhor supre para novas coisas que acontecem no sinal de excitação, que modificará eficazmente a cópia retardada (atrasada) da excitação da estrutura prévia da seção de livro de código adaptável 89.

[0070] Cada estrutura é subdividida em quatro subestruturas cada uma das quais tem uma saída independentemente de livro de código fixo codificada. A excitação de livro de código fixo para uma subestrutura codifica a excitação como uma série de 5 trens intercalados de pares de pulsos de amplitude unitária. As posições possíveis para cada par de pulsos são mostradas na tabela abaixo para MR122 (o nome do modo de 12.2 kb /s do AMR). Como indicado acima esta codificação usa um número significativo de bits.

Trilha	Pulso	Posições
1	0, 5	0, 5, 10, 15, 20, 25, 30, 35
2	1, 6	1, 6, 11, 16, 21, 26, 31, 36
3	2, 7	2, 7, 12, 17, 22, 27, 32, 37
4	3, 8	3, 8, 13, 19, 23, 28, 33, 38
5	4, 9	4, 9, 14, 19, 24, 29, 34, 39

[0071] O sinal do primeiro pulso em cada trilha é também codificado; o sinal do segundo pulso é o mesmo que o primeiro a menos que ele caia antes na trilha quando ele é oposto. O ganho para a subestrutura é também codificado.

O Livro de Código Adaptável

[0072] O livro de código adaptável é uma cópia de tempo atrasada da parte prévia da excitação combinada e é importante na codificação de fala com voz. Devido ao fato da fala com voz ser regular, é possível codificar somente a diferença entre o período de tom atual e o prévio usando a saída de livro de código fixa. Quando adicionado a uma cópia salva do período de voz prévia, conseguiu-se a estimativa desta excitação da estrutura. O livro de código adaptável não é transmitido; o codificador e decodificador calculam o livro de código adaptável da saída combinada prévia e o atraso de tom atual.

Preditor de Tom

[0073] O propósito do preditor de tom (que forma parte da seção de

livro de código adaptável 89) é determinar o melhor atraso para usar para o livro de código adaptável. É um processo de dois estágios. O primeiro é uma única etapa, predição de tom de laço aberto que correlaciona a fala com amostras prévias para achar uma estimativa do período de voz se a fala é de voz ou a melhor taxa de repetição que minimiza uma medida de erro. Isto é seguido por uma predição de laço fechado repetida para conseguir o melhor atraso para a livro de código adaptável dentro de $1/6^o$ de uma amostra. Por essa razão a predição é parte do processo de livro de código adaptável no codificador. O cálculo é limitado pela abordagem dos dois estágios já que a segunda busca mais detalhada somente acontece sobre um pequeno número de amostras. A codificação - decodificação AMR 55 usa uma abordagem de análise por síntese, então seleciona o melhor atraso minimizando o erro quadrático médio entre as saídas e a fala de entrada para atrasos candidatos.

[0074] Portanto, para representar o sinal de excitação para a estrutura atual, os índices das saídas de codificação - decodificação AMR 55 os índices de livro de código fixo (um para cada subestrutura) determinados para a estrutura atual, o ganho de livro de código fixo, o atraso de livro de código adaptável e o ganho de livro de código adaptável. São estes dados e os dados codificados de LPC que são feitos disponíveis para o software de aplicação 69 sendo executado no telefone celular 21 e dos quais os dados ocultos têm que ser recuperados.

Ocultação e Recuperação de Dados

[0075] Existem vários modos em que os dados $F(t)$ podem ser ocultos dentro do sinal de áudio e o leitor é referido ao trabalho por Bender intitulado "Techniques For Data Hiding", Revista de Sistemas da IBM, Vol 35, no 384, 1996, para uma discussão detalhada de técnicas diferentes para dados em áudio. Na presente modalidade, os dados são

ocultados no áudio adicionando um eco para ao áudio, com o atraso de tempo do eco sendo variado para codificar os dados. Esta variação pode ser executada, por exemplo usando um esquema simples: nenhum eco correspondente a um binário zero e um eco corresponde a um binário um. Alternativamente, um binário pode ser representado pela adição de um eco em um primeiro atraso e um binário zero pode ser representado pela adição de um eco em um segundo atraso diferente. O sinal do eco pode ser também variado com os dados a serem ocultados. Em um esquema de codificação mais complexo um binário um pode ser representado por uma primeira combinação ou sequência de ecos (dois ou mais ecos ao mesmo tempo ou aplicados consecutivamente) e um binário zero pode ser representado por uma segunda combinação ou sequência diferente de ecos.

[0076] Nesta modalidade, ecos podem ser adicionados com atrasos de 20 0.75 ms e 1.00 ms e um binário um é representado adicionando um eco atenuado de 0.75 ms para uma primeira seção do áudio (tipicamente correspondente a várias estruturas de AMR) seguido por adicionar um eco atenuado de 1.00 ms em uma segunda seção do áudio; e um binário zero é representado adicionando um eco atenuado de 1.00 ms para uma primeira seção do áudio seguido de adicionar um eco atenuado de 0.75 ms em uma segunda seção do áudio. Portanto, no sentido de recuperar os dados ocultados, a aplicação de software tem que processar a saída codificada da codificação - decodificação AMR 55 para identificar as sequências de ecos recebidos no áudio e consequentemente os dados ocultados no áudio.

[0077] Tipicamente, ecos são identificados em sinais de áudio executando uma autocorrelação das amostras de áudio e identificando os cumes correspondentes a quaisquer ecos. No entanto, como mencionado acima, os dados ocultados são para ser recuperados da saída da codificação - decodificação AMR 55.

Recuperação de dados 1

[0078] A figura 4 ilustra um modo no qual os ecos podem ser detectados e os dados ocultos $F(t)$ recuperados pelo software de aplicação 69 da saída da codificação - decodificação AMR 55. Como mostrado, nesta técnica, o software de aplicação recupera os dados ocultos somente da saída de informação codificada de LPC pela seção VQ 83 mostrada na Figura 3. Como ilustrado na Figura 4, o primeiro processamento executado pelo software de aplicação 69 é executado pela seção VQ 91, que reverte a quantificação de vetor executada pela codificação - decodificação AMR 55. A saída da seção VQ 91, é então processada pela seção de adição de predição 93, que adiciona as predições delta LSF (determinadas pelo preditor 95) para as saídas da seção VQ 91. As médias LSF (obtidas do armazenamento de dados 97) são então adicionadas de retorno pela seção de adição de médias 99, para recuperar os LSFs para a estrutura atual. Os LSFs são então convertidos de volta para os coeficientes de LPC pela seção de conversão de LSF 101. Os assim determinados coeficientes \hat{a}_i não serão exatamente os mesmos que aqueles determinados pela seção de análise de LPC 71 na Figura 3, devido às aproximações e quantificação executada nas outras fases de processamento de AMR.

[0079] Como mostrado, nesta modalidade, os coeficientes de LPC determinado \hat{a}_i são usados para configurar um filtro de síntese de LPC 103 conforme a equação (2) acima. A resposta de impulso ($h(n)$) deste filtro de síntese 103 é então obtida aplicando um impulso (gerado pelo gerador de impulso 105) para o filtro 103 configurado deste modo. Os inventores verificaram que os ecos estão presentes dentro desta resposta de impulso ($h(n)$) e podem ser encontrados a partir de uma autocorrelação da resposta de impulso em torno dos atrasos correspondentes ao retardo do eco. Como mostrado, a seção de autocorrelação 107 executa estes cálculos de autocorrelação para os

atrasos identificados no armazenamento de dados 108. A figura 5 ilustra a autocorrelação obtida para todos os atrasos positivos. A plotagem identifica os atrasos como amostras desde o cume principal 108 à zero atraso. Então com uma taxa de amostragem 8 kHz, cada amostra corresponde a um atraso de 0,125 ms. Como mostrado, existe um cume inicial 108 à zero atraso, seguido por um cume 110 em um atraso de cerca de 1,00 ms (correspondente a 8 amostras desde a origem) – indicando que a estrutura atual tem um eco de 1,00 ms. Como aqueles versados na técnica observarão, não há necessidade de calcular a autocorrelação para todos os atrasos – só aqueles em torno dos atrasos correspondentes à onde os ecos são para serem encontrados (isto é, cerca de 0,75 ms e 1,00 ms).

[0080] Como mostrado na Figura 4, os valores de autocorrelação determinados pela seção de autocorrelação 107 são passados para uma seção de identificação de eco 109, que determina se existem quaisquer ecos na estrutura atual (por exemplo, limitando os valores de autocorrelação com um limite adequado para identificar quaisquer cumes nos atrasos relevantes). Os cumes identificados são então passados para a seção de recuperação de dados 111, que monitora a sequência de ecos identificados sobre estruturas vizinhas para detectar a presença de um binário um ou um binário zero dos dados ocultos $F(t)$. Deste modo, os dados ocultos são recuperados e podem então ser usados para controlar a operação do software de aplicação 69 na maneira descrita acima.

[0081] Os inventores verificaram que os requisitos computacionais para recuperar os dados ocultos deste modo são significativamente menos do que seriam requeridos recuperando os dados ocultos diretamente das amostras de áudio digitalizadas.

Recuperação de dados 2

[0082] Na modalidade descrita acima, foi determinada a

autocorrelação da resposta de impulso do filtro de síntese de LPC e da qual a presença dos ecos foi determinada para recuperar os dados ocultos. A figura 6 ilustra o processamento que pode ser executado de acordo com uma técnica alternativa para recuperar os dados ocultos. Como pode ser visto, comparando as Figuras 4 e 6, a diferença principal entre esta modalidade e a primeira modalidade é que os coeficientes de LPC regenerados \hat{a}_i para a estrutura atual são passados diretamente para a seção de autocorrelação 107, que calcula a autocorrelação da sequência de coeficientes de LPC. Esta modalidade é então uma simplificação da primeira modalidade. No entanto, os cumes na saída de autocorrelação nos atrasos de eco não são tão pronunciados quanto na primeira modalidade e então por essa razão esta modalidade mais simples não é preferida onde potência de processamento suficiente está disponível.

Recuperação de dados 3

[0083] A figura 7 ilustra o processamento que pode ser executado em uma terceira técnica para identificar a presença de ecos e a recuperação subsequente dos dados ocultos. Como pode ser visto comparando as Figuras 6 e 7, a diferença principal entre esta modalidade e a segunda modalidade é que os coeficientes de LPC regenerados \hat{a}_i para a estrutura atual são aplicados a uma seção de Levinson-Durbin reverso 114, que usa o algoritmo de Levinson-Durbin reverso para recomputar a matriz de autocorrelação R_{ij} da equação (3) sobre os coeficientes de LPC. Os valores determinados correspondem aos valores de autocorrelação do próprio sinal de áudio de entrada e incluirá, portanto, cumes em atrasos correspondentes ao atraso dos ecos ou de cada eco. A saída da seção de Levinson-Durbin 114 pode então ser processada como antes, para recuperar os dados ocultos. A desvantagem principal desta modalidade é que o algoritmo de Levinson-Durbin reverso é relativamente intensivo computacionalmente

e então onde existe potência de processamento limitado, esta modalidade não é preferida.

Recuperação de dados 4

[0084] Nas três modalidades acima, os dados ocultados são recupera dos processando a saída codificada de dados de filtro de LPC da codificação - decodificação AMR 55. A codificação - decodificação AMR 55 codificará os ecos nos dados de filtro de LPC desde que o atraso de eco é menor do que o comprimento do filtro de LPC. Como mencionado acima, o filtro de LPC tem uma ordem (p) de dez amostras. Com uma frequência de amostragem de 8 kHz, isto corresponde a um atraso máximo de 1.25 ms. Se um eco com um atraso mais longo for adicionado, então ele não poderá ser codificado nos coeficientes de LPC. No entanto, ele será codificado dentro do resíduo ou sinal de excitação. Para ilustrar isto, será descrita uma modalidade em que os binários um e zero são codificados no áudio usando ecos de 2 ms e 10 ms.

[0085] A figura 8 ilustra o processamento executado nesta modalidade pelo software de aplicação 69, para recuperar os dados ocultados. Como mostrado, nesta modalidade, o software de aplicação 69 recebe os dados codificados de excitação para cada estrutura como ela é produzida pela codificação - decodificação AMR 55. Os índices de livro de código fixo nos dados recebidos são usados, pela seção de livro de código fixo 121, para identificar os pulsos de excitação para a estrutura atual do livro de código fixo 123. Estes pulsos de excitação são então ampliados pelo ganho fixo correspondente, definido nos dados codificados recebidos da codificação - decodificação AMR 55. Os pulsos de excitação amplificados são então aplicados a uma máquina de adicionar 127, onde eles são adequadamente adicionados para ampliar e modalidades atrasadas de pulsos de excitação prévia obtidas passando a pulsos de excitação da estrutura prévia através do ganho

129 e um atraso de livro de código adaptável 131. O ganho de livro de código adaptável e atraso usados são definidos nos dados codificados recebidos da codificação - decodificação AMR 55. A saída da máquina de adicionar 127 é uma representação do sinal residual ou sinal de excitação para a estrutura atual. Como mostrado na Figura 8, esta representação de pulso (e_i) do sinal de excitação é então passado para uma seção de autocorrelação 107 que calcula sua autocorrelação para os diferentes atrasos definidos no armazenamento de dados de atrasos 108. A figura 9 ilustra a saída de autocorrelação da seção de autocorrelação 107 para todos os atrasos positivos, quando existe um eco de 2 ms no áudio recebido. Como mostrado, existe um cume principal 132 em um atraso zero e outro cume 134 em um atraso correspondente a 2 ms. Então, a saída da seção de autocorrelação 107 pode ser processada como antes pela seção de identificação de eco 109 e a seção de recuperação de dados 111 para recuperar os dados ocultos $F(t)$.

Refinamentos

[0086] Vários refinamentos para as modalidades descritas acima serão agora descritos com referência às Figuras 10, 11 e 12. Estes refinamentos foram feitos para aumentar a recuperação bem-sucedida dos dados ocultos e visar combater os efeitos de acústica de fala ou ambiente que pode mascarar a presença dos ecos. Estes refinamentos serão aplicados à primeira modalidade descrita acima, mas eles poderiam ser igualmente aplicados a outras modalidades.

[0087] Como pode ser visto comparando as Figuras 4 e 10, no primeiro refinamento, a resposta de impulso ($h(n)$) do filtro de síntese de LPC 103 para a estrutura atual é filtrada por um filtro de passa alta 151 para reduzir o efeito das frequências mais baixas na resposta de impulso. Os inventores verificaram que a informação de eco é tipicamente codificada na faixa de frequência mais alta da resposta de

impulso. Esta filtragem de passa alta então melhora a agudez dos cumes de autocorrelação para os ecos, tornando mais fácil para identificar sua presença. O filtro de passa alta 151 preferencialmente filtra para fora frequências abaixo de mais ou menos 2 kHz (correspondentes a uma frequência de um quarto da frequência de amostragem) embora algum ganho possa ainda ser feito filtrando para fora somente as frequências abaixo de mais ou menos 1 kHz. Como aqueles versados na técnica observarão, esta filtragem é uma filtragem "intra" estrutura (isto é, que filtra dentro somente da estrutura) que filtra para fora a parte de baixa frequência da resposta de impulso, embora pudesse ser também executada filtragem "entre" estruturas (por exemplo, para filtrar para fora recursos variando lentamente da resposta de impulso que acontece entre estruturas).

[0088] A figura 11 ilustra um modo alternativo de alcançar o mesmo resultado. Em particular, nesta modalidade, os coeficientes de LPC \hat{a}_i para a estrutura atual são passados através de um filtro de passa alta 153 antes de serem usados para configurar o filtro de síntese de LPC 103. Neste caso, o filtro de passa alta 153 remove os coeficientes correspondentes aos polos de frequência mais baixa do filtro de síntese 103. Isto é alcançado fatorando os coeficientes de LPC para identificar as frequências de polo e larguras de banda. Os polos em frequências abaixo do limite mais baixo são descartados e os polos restantes são usados para gerar um filtro de síntese somente de frequência mais alta 103. O processamento restante é como antes, e uma descrição adicional não será dada. Como aqueles versados na técnica observarão, esta filtragem é também uma filtragem "intra" estrutura, embora pudesse ser também executada filtragem "entre" estruturas.

[0089] A figura 12 ilustra um refinamento adicional que pode ser aplicado para aumentar a taxa de sucesso de recuperar os dados ocultos. Como mostrado, a diferença principal entre esta modalidade

e a modalidade mostrada na Figura 4 está na provisão de uma filtro de passa alta 155 para executar filtragem "entre" estruturas para filtrar para fora correlações variando lentamente (isto é, correlações que variam lentamente de estrutura para estrutura) na saída de autocorrelação que são tipicamente causadas pelo próprio áudio e a acústica ambiente em que o telefone celular 21 do usuário está localizado. Adicionalmente, ou em vez de filtrar para fora tais variações de estrutura, o filtro de passa alta 155 poderia executar filtragem "intra" estrutura para remover correlações de baixa frequência na saída de autocorrelação dentro de cada estrutura. Isto foi verificado estreitando os picos de correlação causados pelos ecos assim tornando-os mais fácil para identificar.

Esquema de Codificação Geral

[0090] Nas modalidades acima, dados foram ocultados dentro de um sinal de áudio adicionando ecos que têm diferentes atrasos. Como aqueles versados na técnica observarão, existem vários modos nos quais dados podem ser ocultados dentro do áudio e ainda serem passados através da codificação - decodificação AMR 55. Em termos gerais, os processos de ocultar e recuperar dados acima podem ser representados pelos diagramas de bloco geral mostrados nas Figuras 13 e 14 respectivamente. Como mostrado na Figura 13, o processo geral de ocultar dados pode ser considerado envolver uma operação de codificação semelhante 161 àquela executada pela codificação - decodificação AMR, gerar os parâmetros de AMR (que podem ser os parâmetros finais de saída de AMR ou parâmetros intermediários gerados no processamento de AMR). Um ou mais destes parâmetros são então variados 163 na dependência dos dados a serem ocultados dentro do áudio. Os parâmetros modificados são então decodificados 165 para gerar um sinal de áudio modificado que é transmitido como um sinal acústico e recebido pelo microfone 23 do telefone celular. Depois da filtragem e conversão analógica para digital, o codificador de áudio

167 então processa as amostras de áudio digitalizado na maneira descrita acima para gerar os parâmetros modificados. Os parâmetros modificados são então processados pela seção de processamento de parâmetro 169 para detectar a(s) modificação(s) que foram feitas aos parâmetros e então recuperar os dados ocultados.

[0091] No caso de adicionar ecos ao áudio para codificar os dados ocultados, isto pode facilmente ser feito na maneira descrita acima sem ter que executar o processamento de codificação detalhado no estúdio de televisão (ou onde quer que os dados são para serem ocultados dentro do áudio). Alternativamente, os ecos poderiam ser adicionados manipulando os parâmetros de saída ou parâmetros intermediários do processamento de codificação de AMR. Por exemplo, os ecos poderiam ser adicionados ao áudio adicionando uma constante a uma ou mais entradas da matriz de autocorrelação definida na equação (3) acima ou manipulando diretamente os valores de um ou mais dos coeficientes de LPC determinados da análise de LPC.

[0092] Os dados podem ser também ocultados por outros modos mais diretos de modular os parâmetros de codificação de áudio. Por exemplo, as frequências espectrais de linha geradas para o áudio podem ser modificadas (por, por exemplo, variando o bit menos significativo dos LSFs com os dados a serem ocultados), ou a frequência ou largura de banda dos polos dos quais os LSFs são determinados podem ser modificados conforme os dados a serem ocultados. Alternativamente ainda, os parâmetros de excitação podem ser modificados para transportar os dados ocultados. Por exemplo, a codificação - decodificação AMR 55 codifica o sinal de excitação usando livros de código fixos e adaptáveis que definem um trem de pulsos, com posições variáveis de pulso e sinais. Então, os dados poderiam ser ocultados variando o bit menos significativo das posições de pulso dentro de uma ou mais das trilhas ou subestruturas ou mudando o sinal

de trilhas ou subestruturas selecionadas.

[0093] Ao invés de aplicar ecos para ocultar os dados no áudio, a fase de um ou mais componentes de frequência do sinal de áudio pode ser variada na dependência dos dados a serem ocultados. A informação de fase do áudio é retida até certo ponto na posição dos pulsos codificados pelos livros de código fixos e adaptáveis. Então, esta codificação de fase pode ser detectada da saída da codificação - decodificação AMR 55 regenerando os pulsos de excitação dos livros de código e detectando as mudanças de fase 10 do(s) componente(s) de frequência relevante(s) com o tempo.

[0094] Como aqueles versados na técnica observarão, seria muito improvável que o sistema de estúdio usaria o modelo de codificador e decodificador AMR atual, já que a qualidade de áudio no estúdio de televisão será muito maior do que usado na codificação - decodificação AMR 55. Um sistema de estúdio pleno iria, então, dividir a faixa de áudio em uma faixa de AMR (entre 300 Hz e 3,4 kHz) e uma faixa de Não-AMR fora desta faixa. Iria então manipular a faixa de AMR como indicado acima, mas não reconstruiria o sinal da faixa de AMR usando o decodificador de AMR. Ao invés ele iria sintetizar o sinal de áudio da faixa de AMR do resíduo real de LPC obtido do sinal de áudio original e dados de LPC modificados, para produzir qualidade de áudio mais alta. Alternativamente, onde os parâmetros de excitação são modificados com os dados ocultados, um resíduo seria construído dos parâmetros modificados que iria então ser filtrado pelo filtro de síntese usando os coeficientes de LPC obtidos da análise de LPC. A faixa de AMR modificada iria então ser adicionada à faixa de Não-AMR para transmissão como parte do sinal de televisão. Este processo é ilustrado nas Figuras 15 e 16. Em particular, a Figura 15 ilustra o processamento que pode ser executado dentro do estúdio de televisão depois que o áudio original foi dividido na faixa de AMR e faixa de Não-AMR. Como

mostrado, a faixa de AMR de áudio é introduzida para um codificador de LPC 171 que executa a análise de LPC acima descrita para gerar os coeficientes de LPC a_i para a estrutura atual. Estes coeficientes são então passados para uma seção de variação de coeficiente 173 que varia um ou mais destes coeficientes na dependência dos dados a serem ocultados dentro do sinal de áudio. Os coeficientes de LPC modificados \hat{a}_i são então produzidos para configurar um filtro de síntese de LPC 175 conforme a equação (2) dada acima. Como mostrado na Figura 15, os coeficientes de LPC a_i gerados pelo codificador de LPC 171 são usados para configurar um filtro de LPC inverso 177 conforme a equação (6) acima. A estrutura de áudio da qual o conjunto atual de coeficientes de LPC são gerados é então passado através deste filtro de LPC inverso para gerar o LPC sinal residual (excitação) que é então aplicado ao filtro de síntese de LPC 175. Isto resulta na geração de um sinal de faixa de AMR de áudio modificado que é então combinado com o sinal de faixa de Não-AMR antes de ser combinado com a trilha de vídeo para distribuição.

[0095] A figura 16 ilustra o cenário alternativo onde os parâmetros de excitação são variados com os dados a serem ocultados. Em particular, como mostrado na Figura 16, a faixa de AMR de áudio é inicialmente processada por um codificador de LPC 171, que nesta modalidade gera e produz os dados de livro de código fixo e adaptável representando o sinal residual ou sinal de excitação. Estes dados de livro de código são então passados através de uma seção de variação 181, que varia os dados de livro de código no sentido de mudar a posição e/ou sinal de um ou mais pulsos representados pelos dados de livro de código fixo conforme os dados a serem ocultados dentro do sinal de áudio. Os dados de livro de código modificado são então produzidos para um gerador residual 183 que regenera um sinal residual correspondente que, quando processado pela codificação -

decodificação AMR 55 irá regenerar os dados de livro de código modificado fixo e adaptável. Isto pode ser alcançado, por exemplo, executando uma rotina iterativa para adaptar um começo residual até a codificação dela resulta na produção de dados de livro de código modificados pela seção de variação 181. Alternativamente, os dados de livro de código modificados podem ser usados para gerar os trens de pulsos que são diretamente usados como o sinal residual. Os intervalos entre os pulsos podem ser enchidos com ruído ou parte do sinal residual que pode ser gerado usando o filtro de LPC inverso e os coeficientes de LPC para a estrutura atual. Independentemente da técnica empregada, o assim gerado sinal residual é então passado para o filtro de síntese de LPC 175 o qual é configurado usando os coeficientes de LPC gerados pelo codificador de LPC 171. O filtro de síntese de LPC 175 então filtra o sinal residual aplicado para gerar a faixa de AMR de áudio modificado que é então combinada com a faixa de Não-AMR para regenerar o áudio para combinação com a trilha de vídeo.

Identificação de Áudio

[0096] Nas modalidades acima, dados foram ocultados dentro do áudio de um programa de televisão e estes dados foram recuperados por processamento adequado em um telefone celular. O processamento executado para recuperar os dados ocultados utiliza pelo menos parte do processamento que já é executado pela codificação - decodificação de áudio do telefone celular. Como mencionado acima, os inventores verificaram que isto reduz overhead computacional requerido para recuperar os dados ocultados. Vantagens semelhantes podem ser obtidas em outras aplicações onde não existe nenhum dado real ocultado dentro do áudio, mas em que, por exemplo, o áudio é para ser identificado de padrões acústicos (impressão digital) do próprio áudio. O modo no qual isto pode ser alcançado será agora descrito com referência a um sistema de identificação de música.

[0097] Atualmente, existem vários serviços de identificação de música, como o provido por Shazam. Estes serviços de identificação de música permitem aos usuários de telefones celulares 21 identificar uma música tocando atualmente, discando um número e tocando a música para o "handset". Os serviços então retornam texto do nome da trilha para o telefone. Tecnicamente, os sistemas operam pela colocação de uma chamada telefônica do telefone celular até um servidor remoto enquanto tocando a música para o telefone. O servidor remoto solta o telefonema depois de um período predeterminado, executa algumas comparações no som recebido contra padrões armazenados em uma base de dados para identificar a música e então envia uma mensagem de texto para o telefone com o título da trilha da música identificada.

[0098] Do material publicado dos inventores do sistema Shazam e outros, o processo geral usado para identificar trilhas é:

1. Converter o sinal de áudio bruto em um gráfico de espectro, que é normalmente alcançado calculando uma série de Transformadas de 5 Fourier Rápida (FFTs) sobrepostas.
2. Analisar o gráfico de espectro para determinar recursos característicos – estes são normalmente as posições de picos de energia, caracterizados pelo seu tempo e frequência.
3. Usar uma função de "hash" destes recursos e usar o resultado da função de "hash" para consultar uma base de dados para determinar um conjunto de entradas que podem comparar com o sinal de áudio.
4. Executar comparação padrão adicional contra estas comparações potenciais para determinar se o sinal de áudio é realmente uma comparação para qualquer daquelas identificadas da base de dados.

[0099] Convencionalmente, o gráfico de espectro para o áudio é determinado de uma série de Transformadas de Fourier Rápidas em

blocos sobrepostos de amostras de áudio digitalizadas para o sinal de áudio. Quando operando em uma rede de telefone móvel, o áudio de entrada estará comprimido pela codificação - decodificação AMR no telefone celular para transmissão sobre a interface de ar 37 para a rede de telefone móvel 35, onde o áudio comprimido é descomprimido para regenerar as amostras de áudio digitais. O servidor então executa a análise de Transformada de Fourier nas amostras de áudio digitais para gerar o gráfico de espectro para o sinal de áudio.

[00100] Os inventores têm imaginado que esta codificação e decodificação executadas pelo sistema de telefone móvel e então a análise de frequência subsequente executada pelo servidor de Shazam é perda de tempo e que um sistema semelhante pode ser implementado sem ter que decodificar de volta o áudio comprimido de amostras de áudio. Deste modo, o processamento de reconhecimento de trilha pode ser completamente executado dentro do telefone celular 21. O usuário, então, não deve colocar um telefonema para um servidor remoto para poder identificar a trilha que está sendo tocada. A O modo pelo qual isto é alcançado será agora descrito com referência à figura 17.

[00101] Em particular, a Figura 17 é um diagrama de bloco ilustrando o processamento executado por uma aplicação de software de reconhecimento de trilha (não mostrado) sendo executado no telefone celular 21. Como mostrado, nesta modalidade, a aplicação de software recebe os dados de LPC codificados de AMR e os dados de excitação de AMR codificado da codificação - decodificação AMR 55. Os dados de LPC codificados de AMR são então passados para a seção de VQ 91, seção de adição de predição 93, seção adição de média 99 e seção de conversão de LSF 101 como antes. O resultado deste processamento é os coeficientes de LPC regenerados \hat{a}_i . Os coeficientes de LPC para a estrutura atual são então passados para uma seção de FFT 201 que calcula sua Transformada de Fourier

Rápida. Semelhantemente, os dados de excitação codificados de AMR são decodificados pela seção de livro de código fixa 121, o ganho fixo 125, a adicionadora 127, o atraso de livro de código adaptável 121 e o ganho adaptável 129, para regenerar pulsos de excitação representando o residual para a estrutura de entrada. Estes pulsos decodificados são então introduzidos para a seção de FFT 203 para gerar a transformada de Fourier de pulsos de excitação. Como mostrado na Figura 17, as saídas das duas seções de FFT 201 e 203 são multiplicadas em conjunto junto pelo multiplicador 205 para gerar uma representação de frequência combinada para a estrutura atual. Esta saída de representação frequência combinada pelo multiplicador 205 deve corresponder aproximadamente à FFT das amostras de áudio digitais dentro da estrutura atual. Isto é por causa do modelo de fonte-filtro subjacente a análise de LPC executado pela codificação - decodificação AMR 55. Em particular, como descrito acima, a análise de LPC assume que a fala é gerada filtrando um sinal de excitação apropriado através de um filtro de síntese. Em outras palavras, o áudio é gerado convoluindo o sinal de excitação com a resposta de impulso do filtro de síntese, ou no domínio da frequência, multiplicando o espectro do sinal de excitação com o espectro do filtro de síntese de LPC.

[00102] Na modalidade presente, o espectro dos coeficientes de LPC é multiplicado com o espectro dos pulsos da excitação de livro de código. Estes são aproximações para o espectro do filtro de síntese de LPC e o espectro do sinal de excitação respectivamente. Então, a produção de espectro combinado do multiplicador 205 será uma aproximação do espectro do sinal de áudio digitalizado dentro da estrutura atual. Como mostrado na Figura 17, este espectro é então introduzido em uma seção geradora de gráfico de espectro 207 que gera um gráfico de espectro dos espectros recebidos para estruturas

adjacentes do sinal de áudio de entrada. O gráfico de espectro deste modo gerado é então passado para uma seção de comparar padrão 209 onde os recursos característicos do gráfico de espectro são usados para procurar padrões armazenados dentro de uma base de dados-padrão 211 para identificar a trilha de áudio sendo captada pelo microfone 23 do telefone celular. Como aqueles versados na técnica observarão, este padrão de comparação pode empregar técnicas de processamento semelhantes àqueles empregados no servidor do sistema Shazam, isto é., usando uma função "hash" primeiro para identificar uma parte da base de dados-padrão 211 para comparar com o gráfico de espectro de áudio. A saída de informação de trilha identificada pela seção de comparar padrão 209 é então produzida para exibição ao usuário na tela 29.

[00103] Os inventores verificaram que este processo exige significativamente menos computação do que converter os dados de áudio comprimidos de volta para as amostras de áudio digitalizadas e então tomando a Transformada de Fourier Rápida das amostras de áudio. Realmente, os inventores verificaram que este processamento exige menos processamento do que tomar as Transformadas de Fourier Rápida das amostras de áudio originais. Isto é porque, tomar a Transformada de Fourier Rápida dos coeficientes de LPC é relativamente simples já que existem somente dez coeficientes por estrutura e porque a Transformada de Fourier Rápida dos pulsos de excitação de livro de código é também relativamente direta já que os coeficientes de posição de pulso podem ser transformados no domínio de frequência simplesmente diferenciando as posições de pulso ou tendo eles pré-computados em uma tabela de consulta (já que existe um número limitado de posições de pulso definidas pela livro de código).

[00104] Como aqueles versados na técnica observarão, o gráfico de espectro resultante obtido desta maneira não é diretamente comparável

àquele derivado da FFT das amostras de áudio, devido às aproximações que são feitas. No entanto, o gráfico de espectro carrega informações adequadas e semelhantes ao gráfico de espectro convencional de forma que as mesmas ou semelhantes técnicas de comparar padrão podem ser usadas para o reconhecimento de áudio. Para melhores resultados, a informação-padrão armazenada na base de dados 211 é gerada preferencialmente de gráficos de espectro obtidos em uma maneira semelhante (isto é. da saída da codificação - decodificação AMR, em lugar de usar aquela gerada diretamente das amostras de áudio).

Modificações e Alternativas Adicionais

[00105] Várias modalidades foram descritas acima ilustrando o modo no qual uma codificação - decodificação de áudio em um telefone celular pode ser usada para reduzir o processamento subsequente executado por outras partes do telefone no sentido de recuperar informações ocultas ou para identificar um segmento de áudio de entrada. Como aqueles versados na técnica observarão, várias modificações e melhorias podem ser feitas para as modalidades acima e algumas destas modificações serão agora descritas.

[00106] Na modalidade de reconhecimento de áudio acima, toda a base de dados-padrão 211 foi armazenada dentro do telefone celular 21. Em uma modalidade alternativa, a seção de igualar padrões 209 pode ser organizada para gerar uma função de "hash" dos recursos característicos do gráfico de espectro gerado para o áudio e o resultado desta função "hash" pode então ser transmitida para um servidor remoto que baixa a informação-padrão apropriada para ser comparada com o gráfico de espectro de áudio. Deste modo a quantidade de dados que tem que ser armazenada dentro da base de dados-padrão 211 no telefone celular 21 pode ser mantida em um mínimo embora introduzindo somente uma demora relativamente pequena no processo

para recuperar padrões selecionados da base de dados distante.

[00107] Na modalidade de reconhecimento de áudio acima, as frequências espectrais de linha foram convertidas de volta para os coeficientes de LPC, que foram então transformados no domínio de frequência usando uma FFT. Em uma modalidade alternativa, o espectro para os dados de LPC pode ser diretamente determinado das frequências espectrais de linha ou dos polos derivados delas. Isto reduziria adicionalmente o processamento que é requerido para executar o reconhecimento de áudio. Nas modalidades anteriores descritas acima, os dados foram ocultados dentro do áudio e usados para sincronizar a operação do telefone a um programa de televisão sendo vista pelo usuário. Na última modalidade apenas descrita, não existe nenhum dado ocultado dentro do áudio e, ao invés, recursos característicos do áudio são identificados e usados para reconhecer o áudio. Como aqueles versados na técnica observarão, técnicas

[00108] 15 de reconhecimento de áudio semelhantes podem ser usadas nas modalidades de sincronização. Por exemplo, a aplicação de software que é executado no telefone pode sincronizar a si mesmo ao programa de televisão identificando partes predeterminadas dentro da trilha sonora de áudio. Este tipo de sincronização pode ser também usado para controlar a produção de sub20 títulos para o programa de televisão.

[00109] Nas modalidades anteriores descritas acima, os dados ocultados foram recuperados determinando valores de autocorrelação dos coeficientes de LPC ou a resposta de impulso do filtro de síntese. Este processamento de correlação não é essencial já que os dados ocultados podem ser encontrados monitorando os coeficientes ou a resposta de impulso diretamente. No entanto, o processamento de autocorrelação é preferido já que ele torna mais fácil para identificar os ecos.

[00110] Nos refinamentos descritos acima, foram usadas várias técnicas de filtragem de passa alta para filtrar para fora componentes de frequência baixa associadas ao áudio e à acústica ambiente. Em uma modalidade preferida, onde tal filtragem de passa alta é executada no telefone celular, o sinal de eco é preferencialmente somente adicionado (durante o processamento de ocultar) ao áudio na parte de frequência alta da faixa de AMR. Por exemplo, acima de 1 kHz e preferencialmente acima de 2 kHz somente. Isto pode ser alcançado, por exemplo, filtrando o sinal de áudio para remover os componentes de frequência mais baixa do AMR e então adicionando a saída filtrada para o áudio original com o atraso de tempo requerido. Isto é preferido já que reduz a energia no sinal de eco que será filtrado fora (e, portanto, perdido) pela filtragem de passa alta executada no telefone celular.

[00111] Nas modalidades acima, foi assumido que a codificação - decodificação de áudio usada pelo telefone celular é a codificação - decodificação AMR. No entanto, como aqueles versados na técnica observarão os princípios e conceitos descritos acima são também aplicáveis a outros tipos de codificação - decodificação de áudio e especialmente aqueles que se baseiam em uma análise de predição linear do áudio de entrada.

[00112] Nas modalidades acima, os vários processamentos da saída de dados de áudio comprimidos da codificação - decodificação de áudio foram executados por softwares sendo executado no telefone celular. Como aqueles versados na técnica observarão, alguns ou todos estes processamentos podem ser formados por circuitos de hardware dedicados, embora o software seja preferido devido a sua habilidade para ser adicionado ao telefone celular depois de fabricar e sua habilidade de ser atualizado uma vez carregado. O software para causar o telefone celular a operar na maneira acima pode ser provido como um sinal ou um portador tal como disco compacto ou outro meio portador.

[00113] Nas modalidades acima, o processamento foi executado dentro de um telefone celular. No entanto, como aqueles versados na técnica observarão, os benefícios se aplicarão para qualquer dispositivo de comunicação que tem uma codificação - decodificação de áudio embutida.

[00114] Nas modalidades anteriores descritas acima, dados foram ocultados dentro do áudio e usados para sincronizar a operação do telefone celular com o show de televisão sendo assistido pelo usuário. Como aqueles versados na técnica observarão, e como descrito na WO02/45273, existem vários outros usos para os dados ocultados. Por exemplo, os dados ocultados podem identificar um URL para uma localização distante ou podem identificar um código para ser enviado para um URL pré-armazenado para interpretação. Tais dados ocultados podem prover o usuário com informações adicionais sobre, por exemplo, o programa de televisão e/ou para prover ofertas especiais ou outra publicidade direcionada ao usuário.

[00115] Na modalidade acima, o programa de televisão foi transmitido ao usuário através de um vínculo de comunicação de RF 13. Como aqueles versados na técnica observarão, o programa de televisão pode ser distribuído para o usuário através de qualquer tecnologia de distribuição apropriada, como por TV a cabo, a Internet, TV de Satélite etc. Ele pode ser também obtido de um meio de armazenamento como um DVD e reproduzido por um aparelho de reprodução de DVD apropriado.

[00116] Nas modalidades acima, o telefone celular retirou o áudio de um programa de televisão. Como aqueles versados na técnica observarão, as técnicas acima podem ser também usadas onde o áudio é obtido de um rádio ou outro sistema de alto falante.

[00117] Nas modalidades acima, foi assumido que os dados foram ocultados dentro do áudio no terminal de estúdio de televisão do

sistema de televisão. Em uma modalidade alternativa, os dados podem ser ocultados dentro do áudio no terminal do usuário do sistema de televisão, por exemplo, por uma caixa de conjunto superior. A caixa de conjunto superior pode ser adaptada para ocultar os dados apropriados dentro do áudio antes de produzir o programa de televisão para o usuário.

[00118] Nas modalidades acima, a aplicação de software processou os dados de áudio comprimidos recebidos da codificação - decodificação AMR dentro do telefone celular 21. Em uma modalidade alternativa, a aplicação de software pode executar processo semelhante em dados de áudio comprimidos recebidos sobre a rede de telefone e providos para o processador 63 pela unidade de processamento de RF 57.

[00119] Nas modalidades acima, é assumido que a saída da codificação - decodificação de áudio não inclui os próprios coeficientes de LPC, mas outros parâmetros derivados deles, como as frequências espectrais de linha ou os polos de filtro do filtro de síntese de LPC. Como aqueles versados na técnica observarão, se a codificação - decodificação de áudio empregada no telefone celular 21 é tal que os coeficientes de LPC derivados por ela estão disponíveis para o processador 63 então o processamento inicial executado pelo software de aplicação para recuperar os coeficientes de LPC não é necessário e as aplicações de software podem trabalhar diretamente na produção de coeficientes de LPC pela codificação - decodificação de áudio. Isto reduzirá o processamento requerido adicional.

[00120] Como aqueles versados na técnica observarão, os valores precisos das taxas de bit, taxas de amostragem etc. descritas nas modalidades acima não são características essenciais da invenção e podem ser variadas sem se afastar da invenção.

REIVINDICAÇÕES

1. Método de recuperar dados ocultos de uma entrada de sinal de áudio ou de identificar uma entrada de sinal de áudio usando um dispositivo de telecomunicações (21) tendo um codificador de áudio (55) para comprimir um sinal de áudio de entrada para transmissão para uma rede de telecomunicações (39), o método sendo executado pelo dispositivo de telecomunicações (21), **caracterizado pelo fato de** passar o sinal de áudio de entrada através da codificação - decodificação de áudio para gerar dados comprimidos e processar os dados de áudio comprimidos para recuperar os dados ocultos ou para identificar o sinal de áudio de entrada.

2. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** o codificador de áudio (55) executa uma predição linear, LP, análise no áudio de entrada para gerar dados de LP representativos do áudio de entrada e em que a etapa de processamento processa os dados de LP para recuperar os dados ocultos ou para identificar o sinal de áudio de entrada.

3. Método de acordo com a reivindicação 2, **caracterizado pelo fato de que** o codificador de áudio (55) comprime os dados de LP para gerar os ditos dados comprimidos de LP e em que a dita etapa de processamento inclui a etapa de regenerar os dados de LP dos dados de áudio comprimidos.

4. Método, de acordo com a reivindicação 2 ou 3, **caracterizado pelo fato de que** os dados de LP compreendem dados de filtro de LP e a etapa de processamento recupera os dados ocultos ou identifica o sinal de áudio usando os dados de filtro de LP.

5. Método, de acordo com a reivindicação 4, **caracterizado pelo fato de que** a etapa de processamento inclui a etapa de gerar uma resposta de impulso de um filtro de síntese ou a etapa de executar um algoritmo de Levinson-Durbin inverso nos dados de filtro de LP.

6. Método, de acordo com a reivindicação 2, 3 ou 4, **caracterizado pelo fato de que** os dados de LP compreendem dados de excitação de LP e a etapa de processamento recupera os dados ocultados ou identifica o sinal de áudio usando os dados de excitação de LP.

7. Método, de acordo com a reivindicação 2 ou 3, **caracterizado pelo fato de que** os dados de LP compreendem dados de filtro de LP e dados de excitação de LP e em que a etapa de processamento processa um subconjunto dos dados de áudio comprimidos correspondentes para um dos ditos dados de filtro de LP e dos ditos dados de excitação de LP para recuperar os dados ocultados.

8. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado pelo fato de que** o sinal de áudio inclui dados ocultados definidos por um ou mais ecos do sinal de áudio e em que os processos de etapa de processo o comprimido de áudio para identificar a presença de ecos dentro do sinal de áudio para recuperar os dados ocultados.

9. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado pelo fato de que** cada símbolo de dados dos dados ocultados é representado por uma combinação de ecos ou uma sequência de ecos dentro do sinal de áudio e em que a etapa de processamento inclui a etapa de identificar as combinações de ecos para recuperar os dados ocultados ou a etapa de localizar uma sequência de ecos no áudio para recuperar os dados ocultados.

10. Método, de acordo com a reivindicação 8 ou 9, **caracterizado pelo fato de que** o codificador de áudio (55) tem uma faixa de frequência operacional predefinida e em que os ecos são ocultados dentro do áudio dentro de uma parte predeterminada da faixa operacional, preferencialmente uma parte superior da faixa de

frequência, e em que a etapa de processamento inclui uma etapa de filtrar para filtrar para fora as frequências fora da dita parte predeterminada.

11. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado pelo fato de que** a etapa de processamento determina um ou mais valores de autocorrelação para cada uma da sequência de estruturas de tempo do sinal de áudio e recupera os dados ocultos usando os valores de autocorrelação determinados.

12. Método, de acordo com a reivindicação 11, **caracterizado pelo fato de que** a etapa de processamento executa uma etapa de filtragem de passa alta dos valores de autocorrelação determinados para remover correlações variando lentamente.

13. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado pelo fato de que** a etapa de processamento recupera os dados ocultos ou identifica o áudio sem regenerar amostras de áudio digitalizadas dos dados de áudio comprimidos.

14. Dispositivo de telecomunicações (21) configurado para realizar o método como definido na reivindicação 1, o dispositivo de telecomunicações (21) compreendendo:

um microfone (23) para receber sinais acústicos e para converter os sinais acústicos recebidos em sinais de áudio elétricos correspondentes;

um conversor analógico digital (53) para amostrar os sinais de áudio elétricos para produzir amostras de áudio digitais;

um codificador de áudio (55) para comprimir as amostras de áudio digitais para gerar dados de áudio comprimidos para transmissão para uma rede de telecomunicações (39); e

caracterizado pelo fato de que

um processador de dados (115), acoplado ao dito codificador de áudio (55), para processamento dos dados de áudio comprimidos para recuperar dados ocultos transportados dentro do sinal acústico recebido ou para identificar o sinal acústico recebido.

15. Dispositivo, de acordo com a reivindicação 14, **caracterizado pelo fato de que** o codificador de áudio (55) é operável para executar uma predição linear, LP, análise no áudio de entrada para gerar dados de LP representativos do áudio de entrada e em que o processador de dados é operável para processar os dados de LP para recuperar os dados ocultos ou para identificar o sinal de áudio de entrada.

16. Dispositivo, de acordo com a reivindicação 15, **caracterizado pelo fato de que** o codificador de áudio (55) é operável para comprimir os dados de LP para gerar os ditos dados comprimidos de LP e em que o dito processador de dados é operável para regenerar os dados de LP dos dados de áudio comprimidos.

17. Dispositivo, de acordo com a reivindicação 15 ou 16, **caracterizado pelo fato de que** os dados de LP compreendem dados de filtro de LP e o processador de dados é operável para recuperar os dados ocultos ou para identificar o sinal de áudio usando os dados de filtro de LP.

18. Dispositivo, de acordo com a reivindicação 17, **caracterizado pelo fato de que** o processador de dados é operável para gerar uma resposta de impulso de um filtro de síntese ou para executar um algoritmo de Levinson-Durbin inverso nos dados de filtro de LP para recuperar os dados ocultos.

19. Dispositivo, de acordo com a reivindicação 15, 16 ou 17, **caracterizado pelo fato de que** os dados de LP compreendem dados de excitação de LP e o processador de dados é operável para recuperar os dados ocultos ou para identificar o sinal de áudio usando os dados

de excitação de LP.

20. Dispositivo, de acordo com a reivindicação 15 ou 16, **caracterizado pelo fato de que** os dados de LP compreendem dados de filtro de LP e dados de excitação de LP e em que o processador de dados é operável para processar um subconjunto dos dados de áudio comprimidos correspondentes a um dos ditos dados de filtro de LP e dos ditos dados de excitação de LP para recuperar os dados ocultados.

21. Dispositivo, de acordo com qualquer uma das reivindicações 14 a 20, **caracterizado pelo fato de que** o sinal de áudio inclui dados ocultados definidos por um ou mais ecos do sinal de áudio e em que o processador de dados é operável para processar os dados de áudio comprimidos para identificar a presença de ecos dentro do sinal de áudio para recuperar os dados ocultados.

22. Dispositivo, de acordo com qualquer uma das reivindicações 14 a 21, **caracterizado pelo fato de que** cada símbolo de dados dos dados ocultados, é representado por uma combinação de ecos ou uma sequência de ecos dentro do sinal de áudio e em que o processador de dados é operável para identificar as combinações de ecos para recuperar os dados ocultados ou para localizar uma sequência de ecos nos de áudio para recuperar os dados ocultados.

23. Dispositivo, de acordo com a reivindicação 21 ou 22, **caracterizado pelo fato de que** o codificador de áudio (55) tem uma faixa de frequência operacional predefinida e em que os ecos são ocultados dentro do áudio dentro de uma parte predeterminada da faixa operacional, preferencialmente uma parte superior da faixa de frequência, e em que o processador de dados é operável para filtrar para fora as frequências fora da dita parte predeterminada.

24. Dispositivo, de acordo com qualquer uma das reivindicações 14 a 23, **caracterizado pelo fato de que** o processador é operável para determinar um ou mais valores de autocorrelação para

cada uma de uma sequência de estruturas de tempo e é operável para recuperar os dados ocultos usando os valores de autocorrelação determinados.

25. Dispositivo, de acordo com a reivindicação 24, **caracterizado pelo fato de que** o processador de dados é operável para executar uma filtragem de passa alta dos valores de autocorrelação determinados para remover correlações variando lentamente.

26. Dispositivo, de acordo com qualquer uma das reivindicações 14 a 25, **caracterizado pelo fato de que** o processador de dados é operável para executar filtragem inter-estruturas e/ou intra-estrutura de passa alta quando recuperando os dados ocultos.

27. Dispositivo, de acordo com qualquer uma das reivindicações 14 a 26, **caracterizado pelo fato de que** o processador de dados é operável para recuperar os dados ocultos ou para identificar o áudio sem regenerar amostras de áudio digitalizadas dos dados de áudio comprimidos.

28. Aparelho de ocultar dados (5) compreendendo:

meios de codificação de áudio (161) para receber e comprimir amostras de áudio digital representativas um sinal de áudio para gerar dados de áudio comprimidos;

caracterizado pelo fato de que

meios (163) para receber dados para serem ocultos dentro do sinal de áudio e para variar os dados de áudio comprimidos na dependência dos dados recebidos, para gerar dados de áudio comprimidos modificados; e

meios (165) para gerar amostras de áudio usando os dados de áudio comprimidos modificados, as amostras de áudio representando o sinal de áudio original e transportando os dados ocultos.

29. Método de ocultar dados em um sinal de áudio para uso como a entrada de sinal de áudio método como definido na reivindicação 1, o método compreendendo as etapas de adicionar um ou mais ecos ao áudio na dependência dos dados a serem ocultados no sinal de áudio e **caracterizado pelo fato de** filtrar o eco com filtro de passa alta antes de combiná-lo com o sinal de áudio.

30. Caixa de conjunto superior compreendendo meios para receber um sinal de áudio (163), meios para ocultar dados no sinal de áudio recebido para uso como a entrada de sinal de áudio do método como definido na reivindicação 1, e meios para produzir o sinal de áudio com os dados ocultados a um usuário, **caracterizada pelo fato de que** a caixa de conjunto superior é operável para representar cada símbolo de dados dos dados a serem ocultados por uma combinação de ecos ou uma sequência de ecos dentro do sinal de áudio.

31. Caixa de conjunto superior, de acordo com a reivindicação 30, **caracterizada pelo fato de que** executa uma filtragem de passa alta de um ou mais dos ecos antes de adicionar aqueles ecos ao sinal de áudio.

32. Produto de instruções implementáveis em computador **caracterizada pelo fato de que compreende** instruções de computador implementáveis para causar um processador programável a executar as etapas de processamento como definidas em qualquer uma das reivindicações 1 a 13.

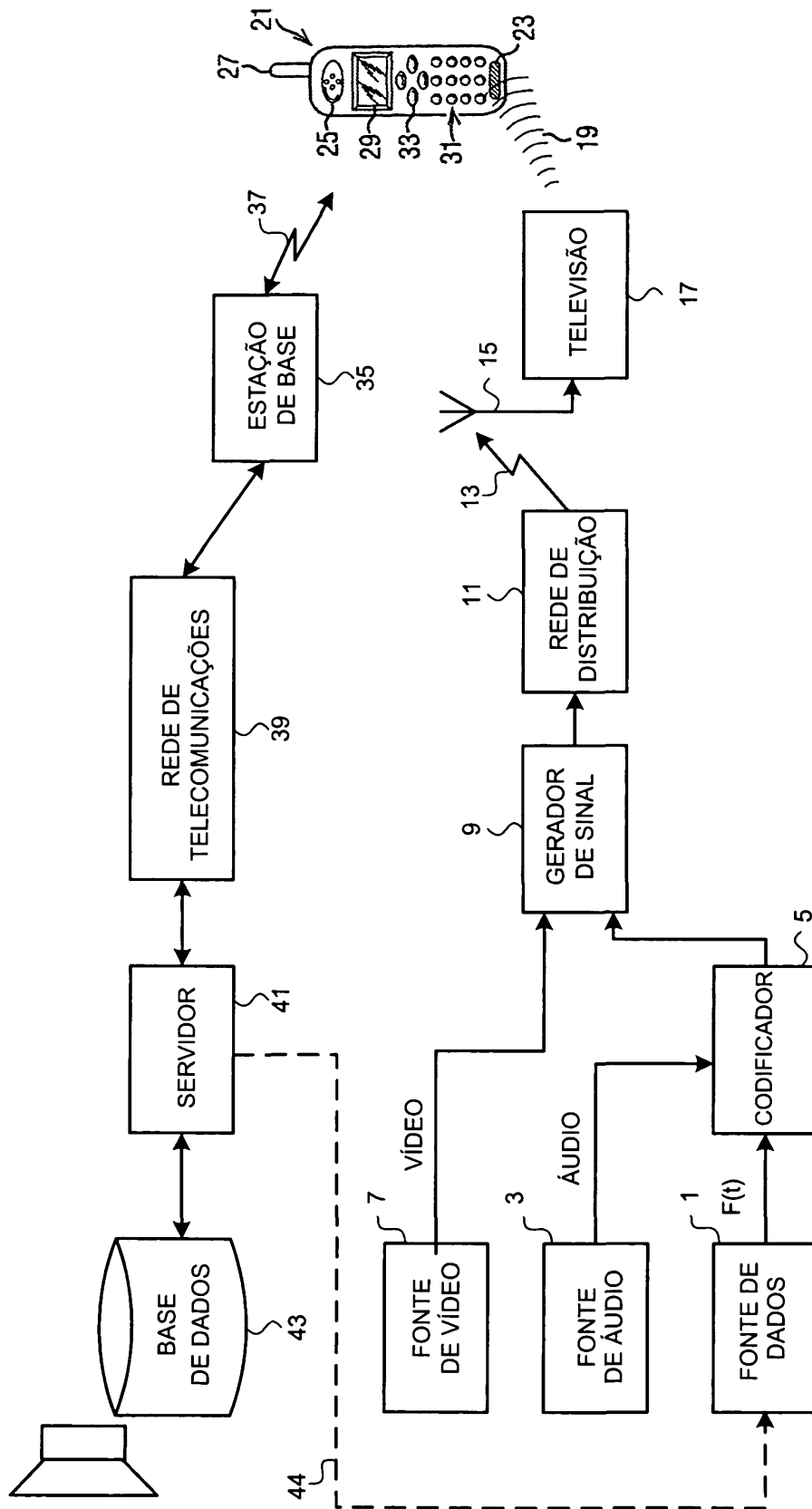


FIG. 1

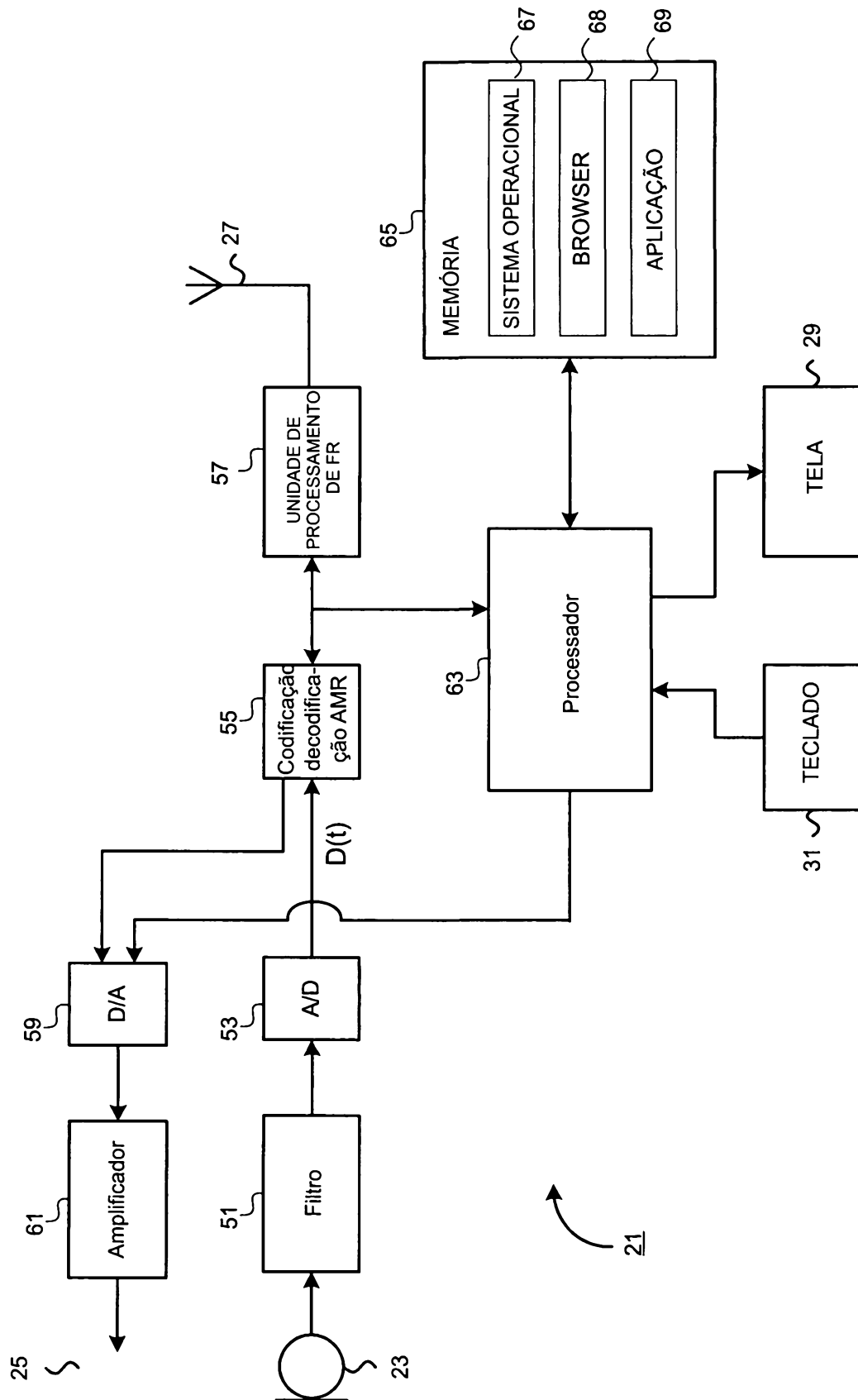


FIG. 2

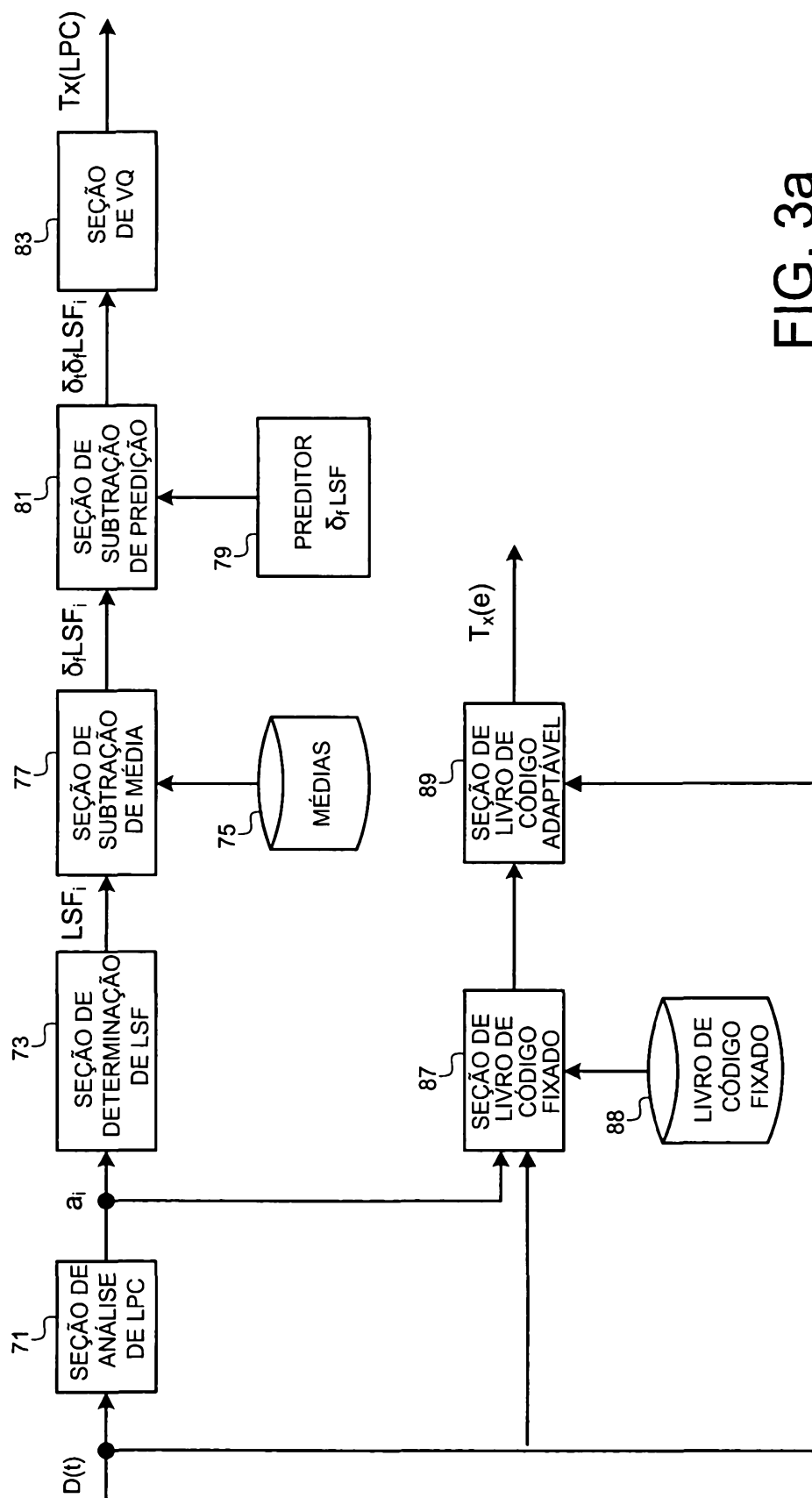


FIG. 3a

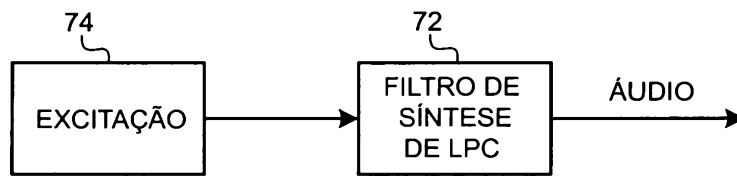


FIG. 3b

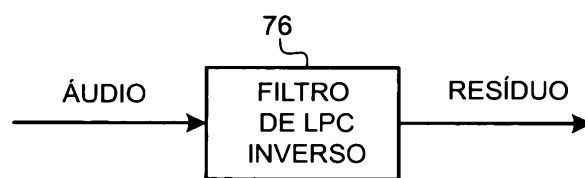


FIG. 3c

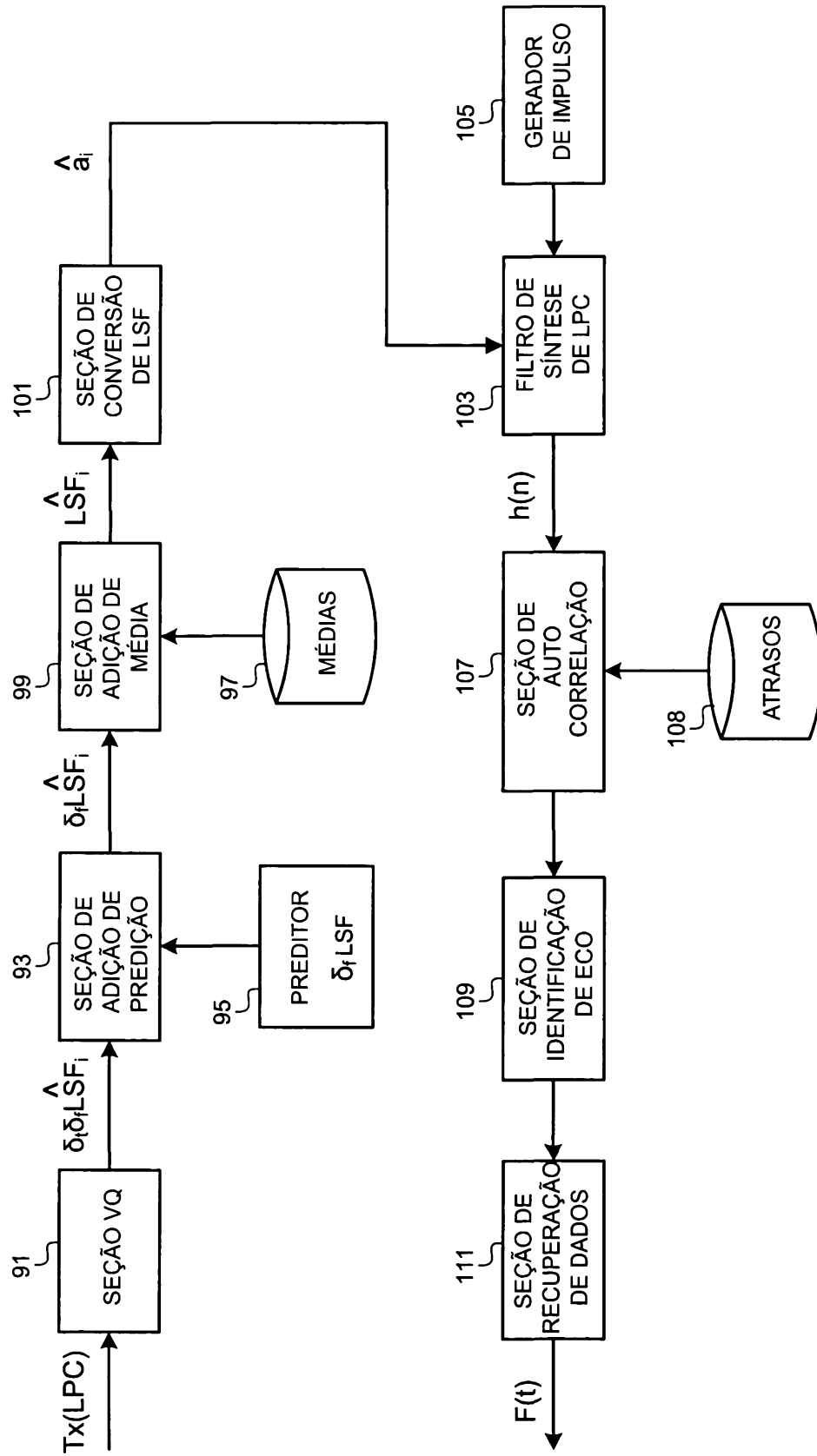
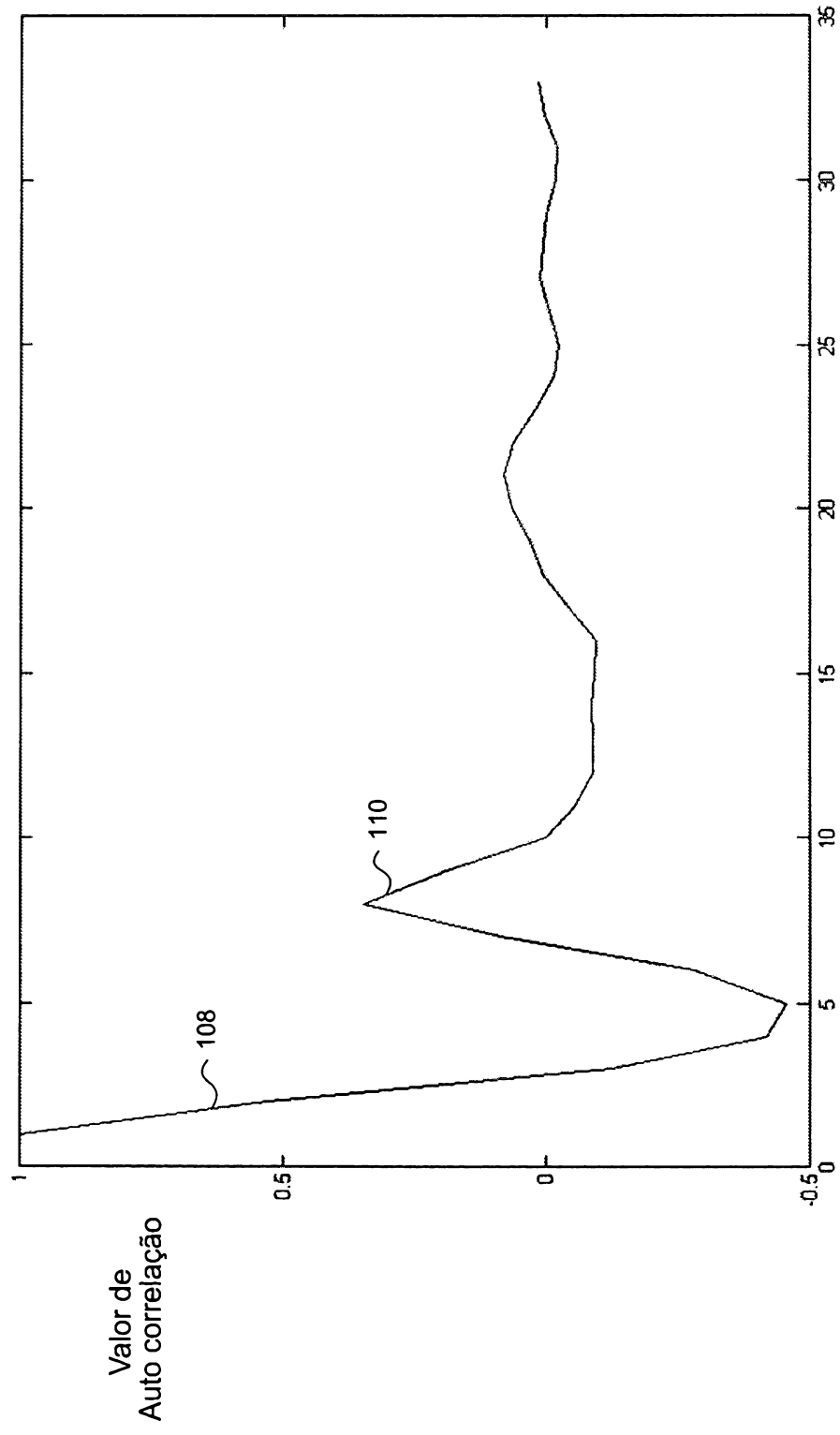


FIG. 4



Atraso nas Amostras

FIG. 5

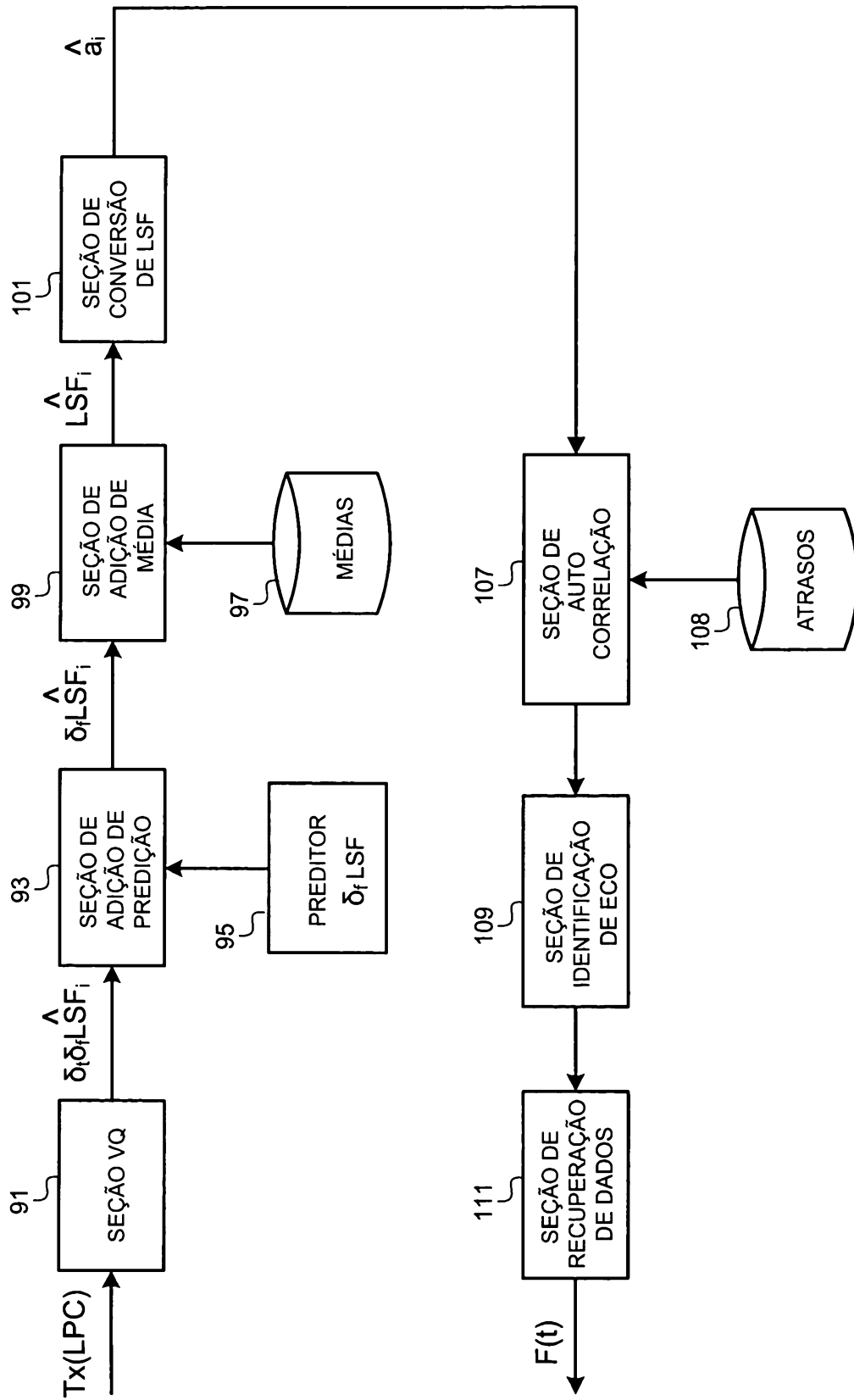


FIG. 6

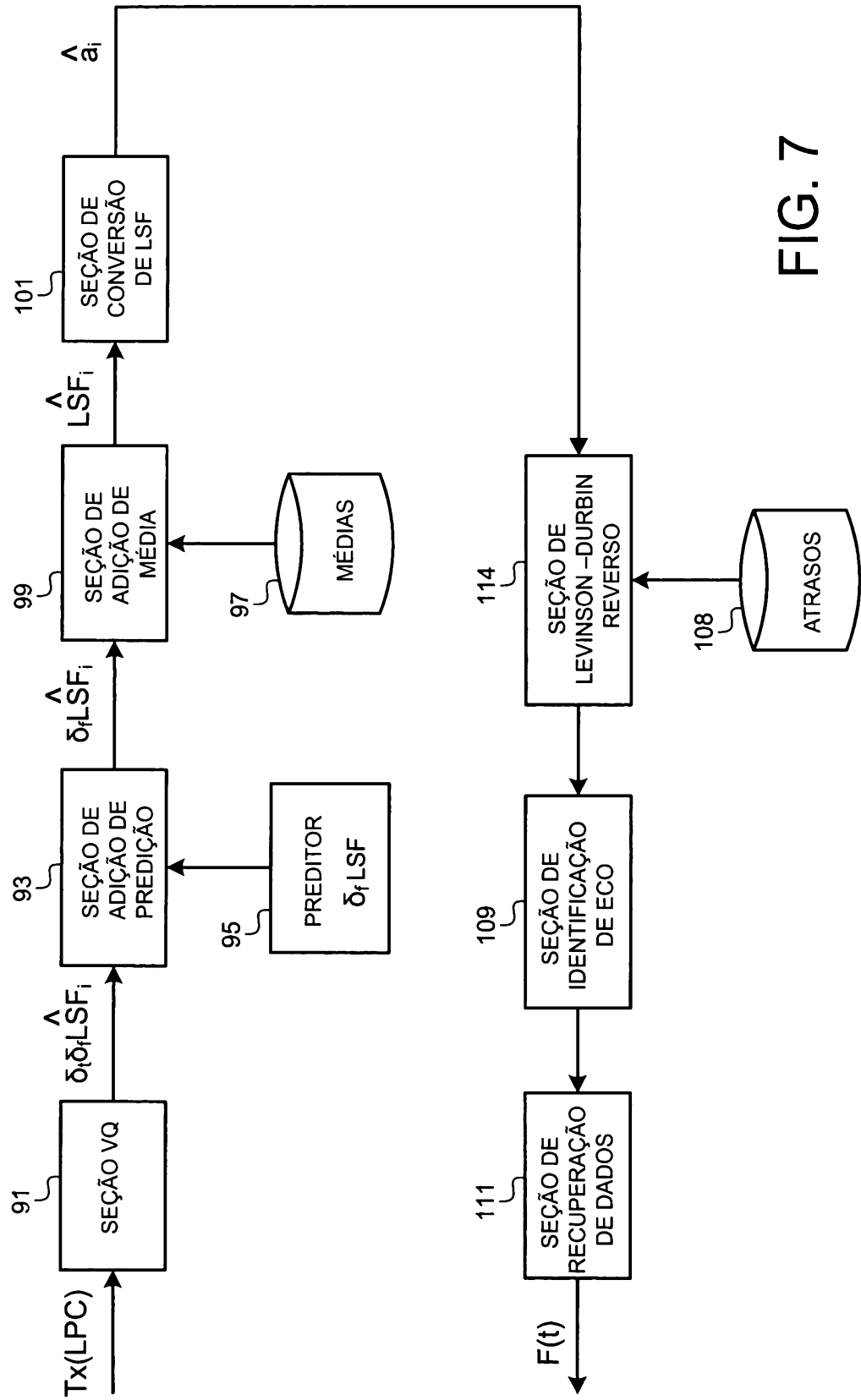


FIG. 7

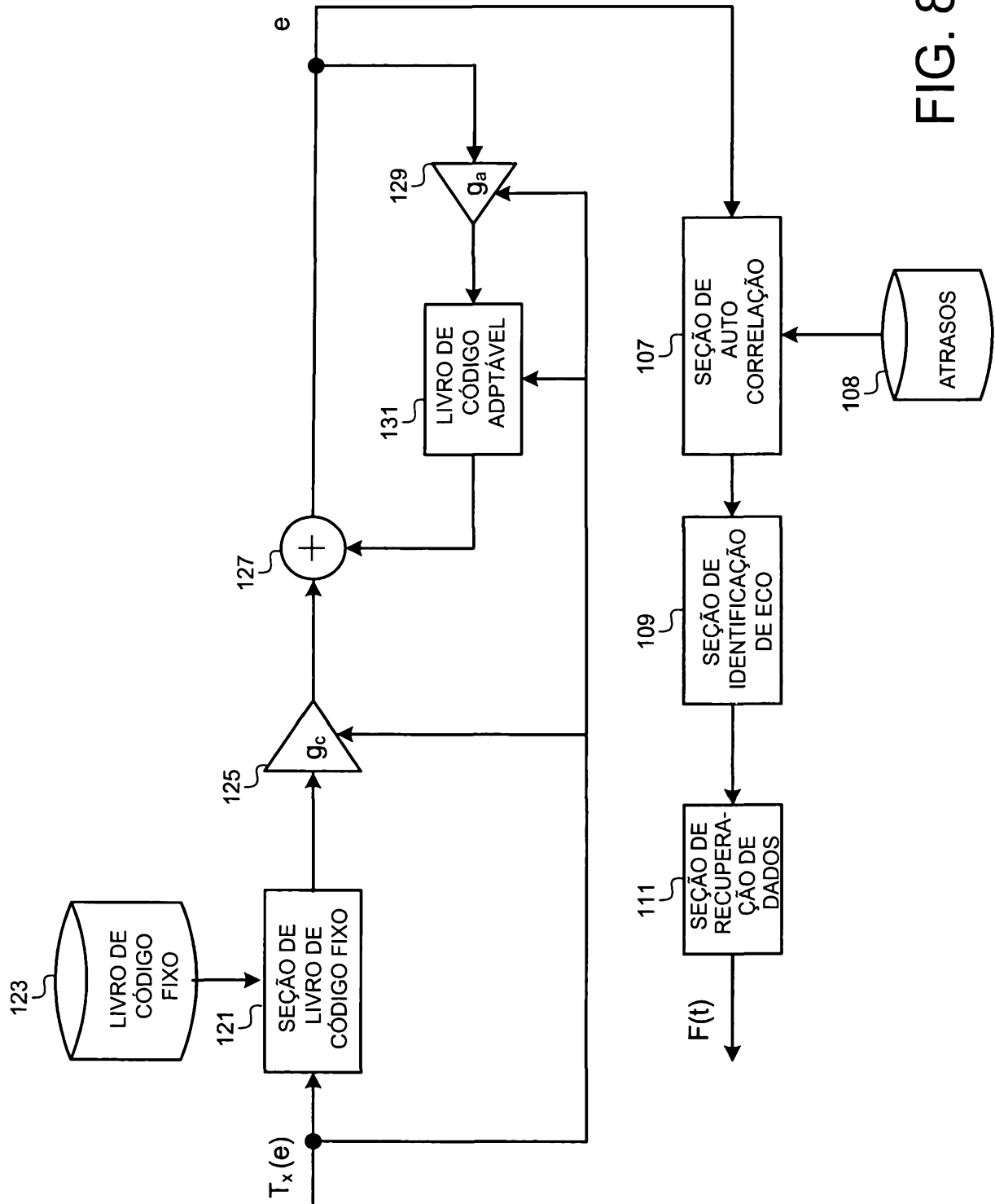


FIG. 8

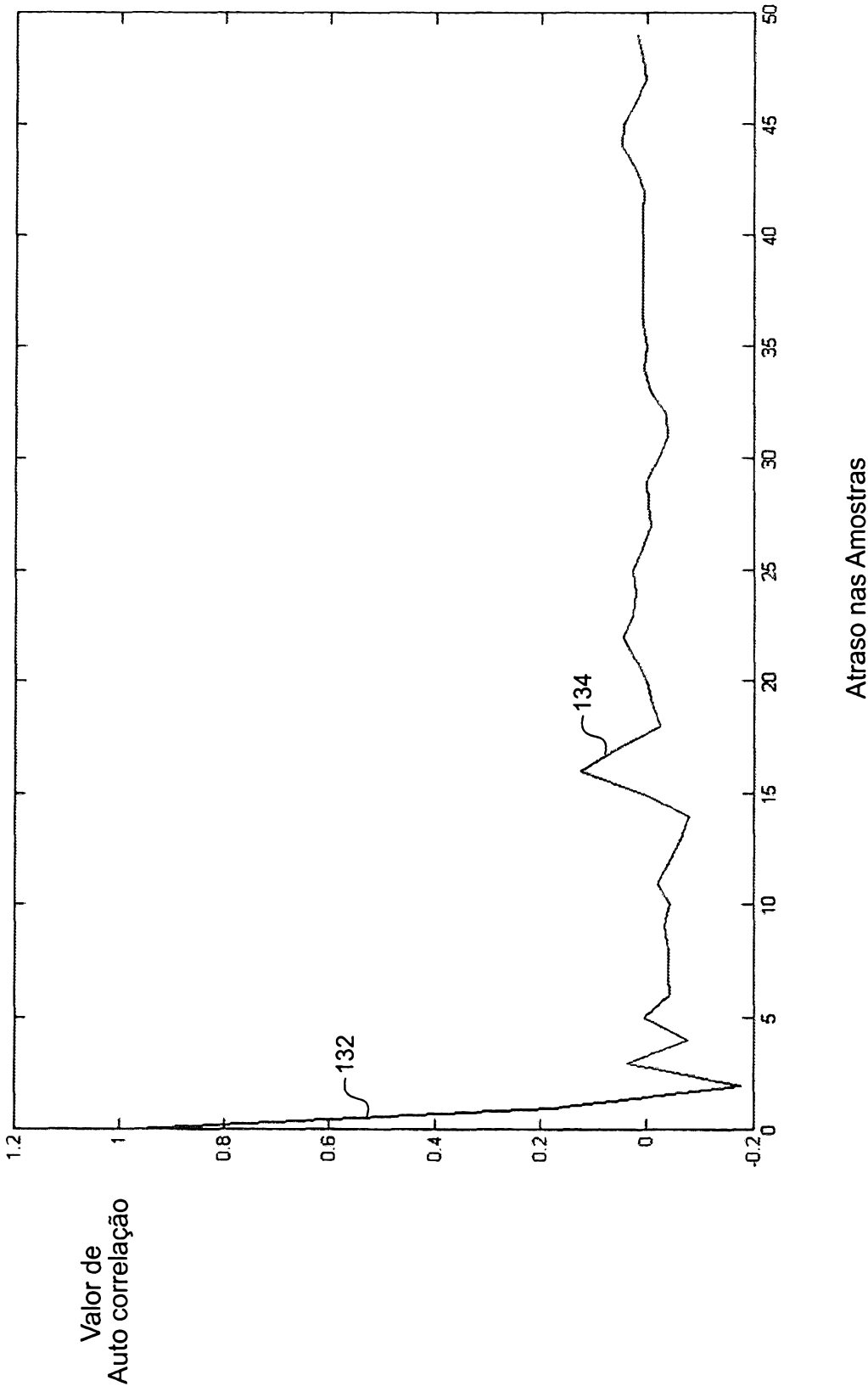


FIG. 9

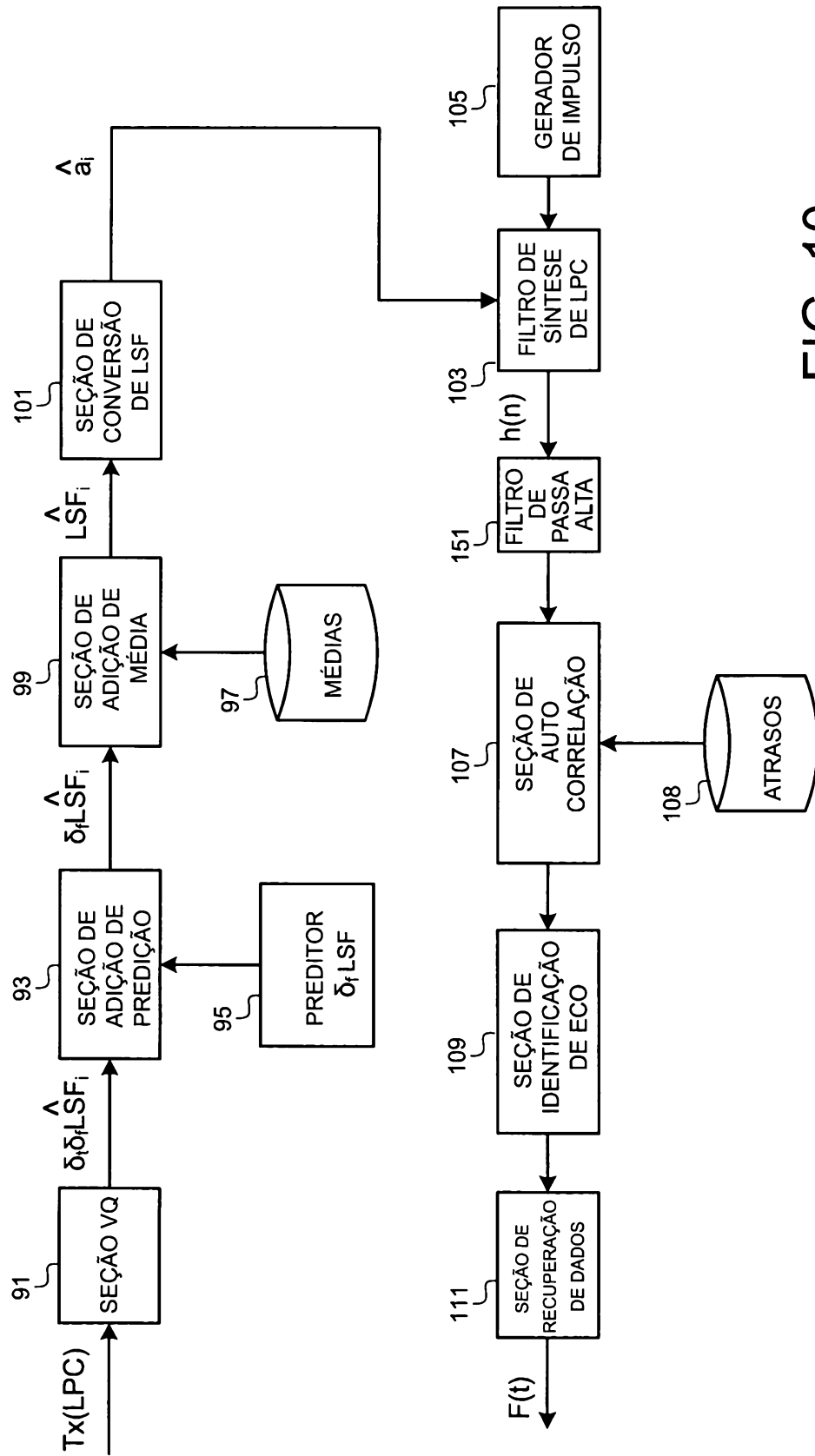


FIG. 10

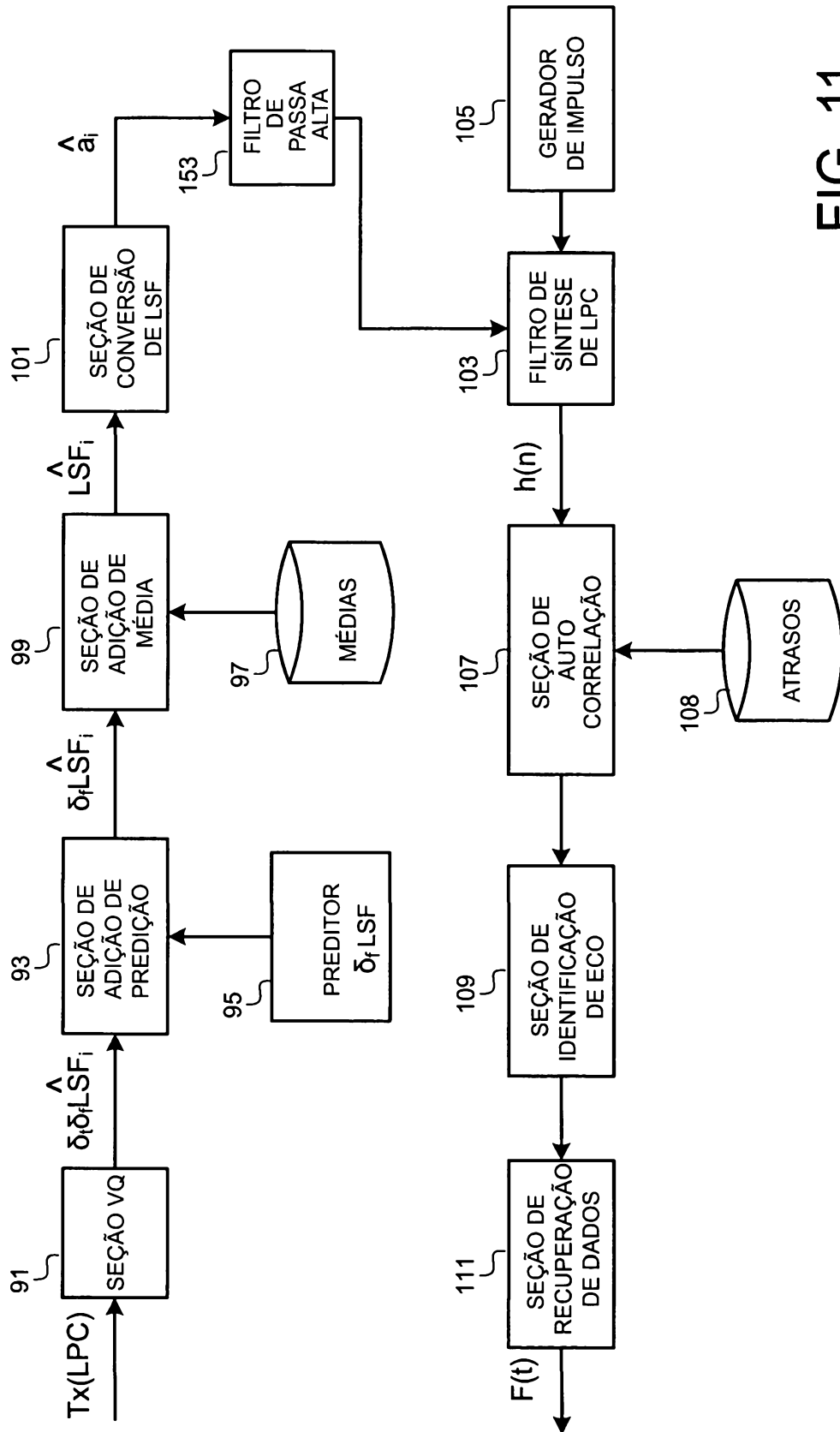


FIG. 11

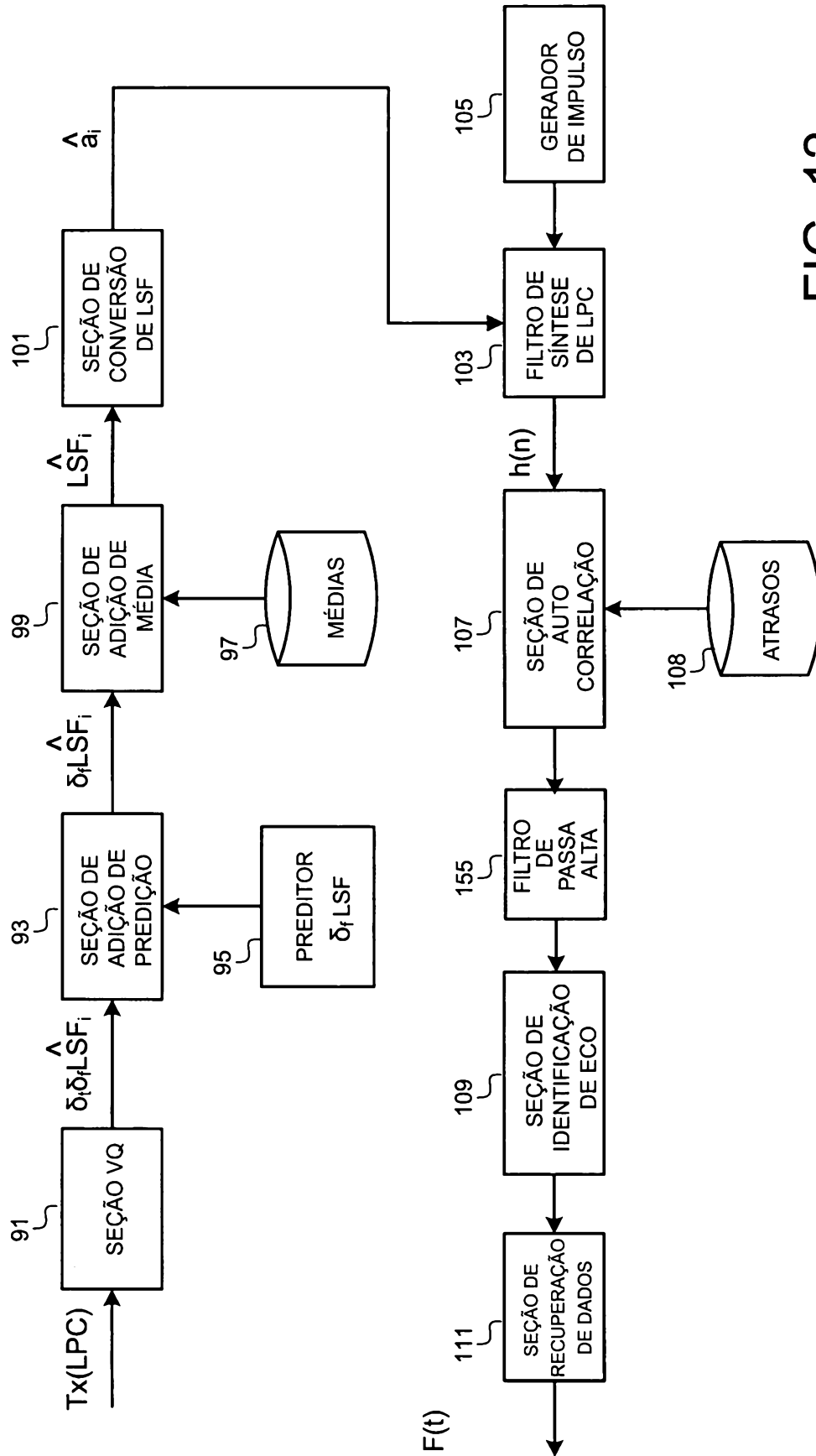


FIG. 12

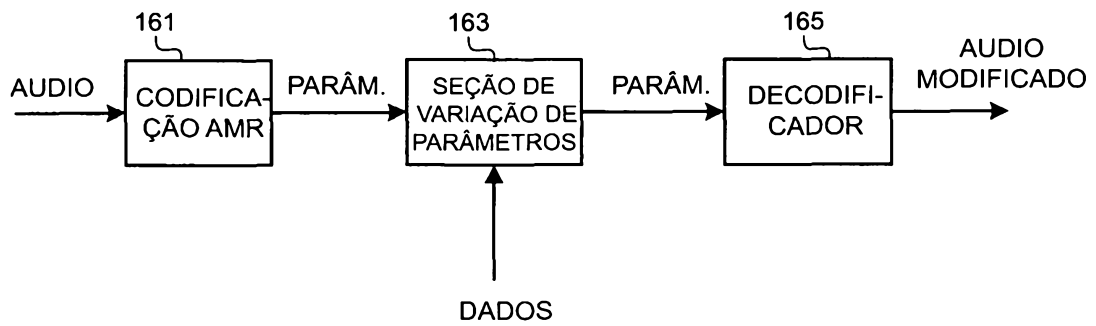


FIG. 13

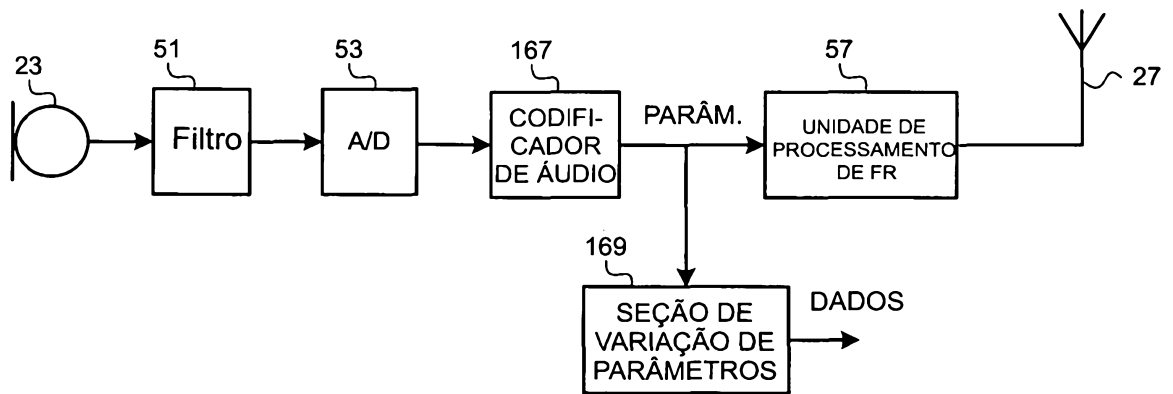


FIG. 14

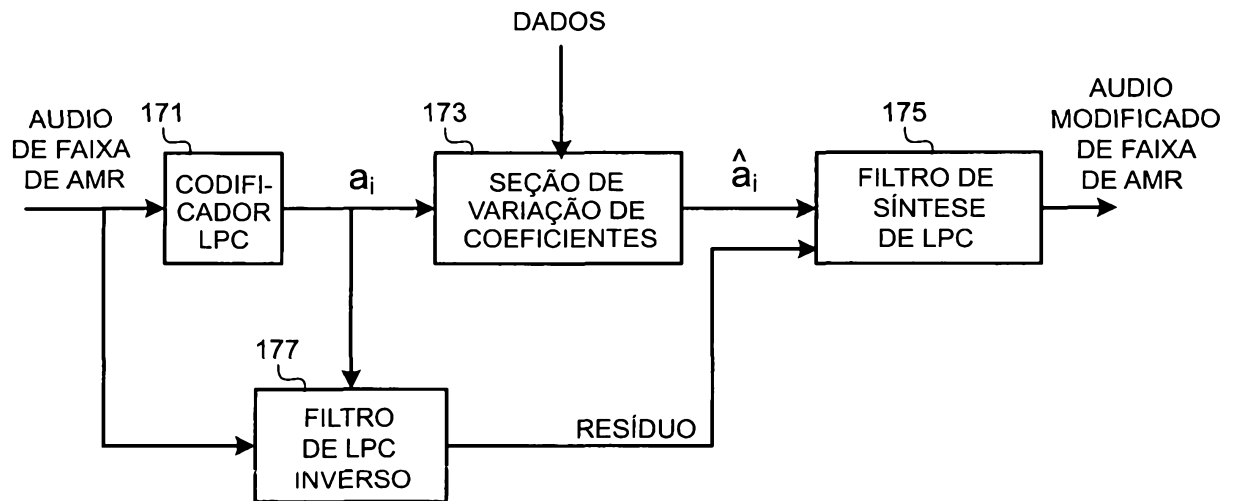


FIG. 15

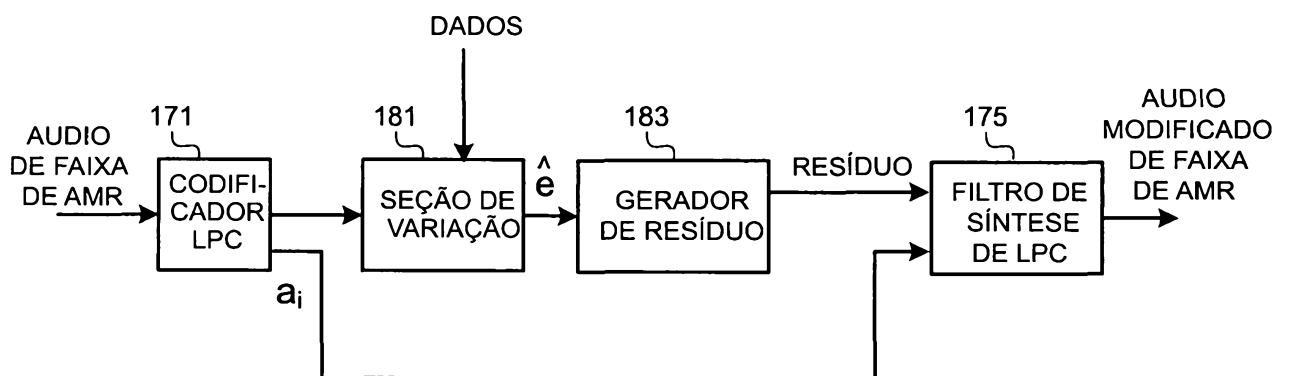


FIG. 16

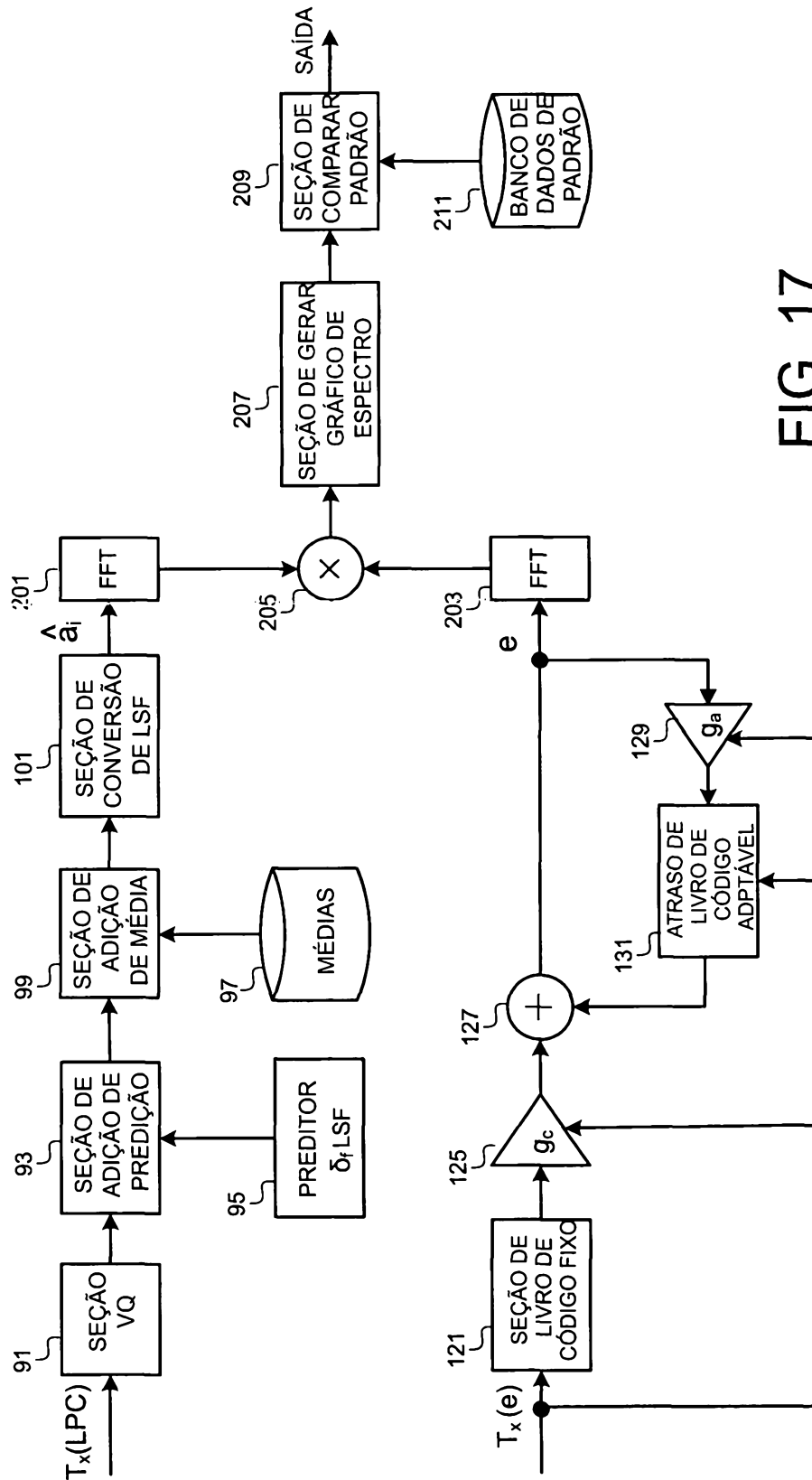


FIG. 17