



República Federativa do Brasil

Ministério do Desenvolvimento, Indústria,
Comércio e Serviços

Instituto Nacional da Propriedade Industrial



(11) BR 112015011758-9 B1

(22) Data do Depósito: 19/12/2013

(45) Data de Concessão: 18/04/2023

(54) Título: SISTEMA DE TRANSMISSOR PARA UM SISTEMA DE VIDEOCONFERÊNCIA, MÉTODO PARA OPERAR UM SISTEMA DE VIDEOCONFERÊNCIA E MEIO DE ARMAZENAMENTO DE COMPUTADOR

(51) Int.Cl.: H04N 7/14; H04N 7/15.

(52) CPC: H04N 7/147; H04N 7/15.

(30) Prioridade Unionista: 19/12/2012 US 13/719,314.

(73) Titular(es): MICROSOFT TECHNOLOGY LICENSING, LLC.

(72) Inventor(es): WILLIAM GEORGE VERTHEIN; SIMONE LEORIN.

(86) Pedido PCT: PCT US2013076671 de 19/12/2013

(87) Publicação PCT: WO 2014/100466 de 26/06/2014

(85) Data do Início da Fase Nacional: 21/05/2015

(57) Resumo: SISTEMA DE TRANSMISSOR PARA UM SISTEMA DE VIDEOCONFERÊNCIA, MÉTODO PARA OPERAR UM SISTEMA DE VIDEOCONFERÊNCIA E MEIO DE ARMAZENAMENTO DE COMPUTADOR. A presente invenção refere-se a um sistema de videoconferência que está configurado para selecionar um orador ativo enquanto evitando erroneamente selecionar um microfone ou câmera que está captando um sinal de áudio ou vídeo de um sinal remoto conectado. Uma determinação é feita se um sinal de áudio está acima de um nível limite. Se sim, então uma determinação é feita quanto a se um identificador está presente naquele sinal de áudio. Se sim, aquele sinal é ignorado. Se não, uma câmera é direcionada para a fonte de som identificada pelo sinal de áudio. Uma identificação é feita se um identificador está presente no sinal de vídeo daquela câmera. Se sim, a câmera é redirecionada. Se não, identificador(es) local(is) e/são inserido(s) no sinal de áudio e/ou no sinal de vídeo. O(s) sinal(is) identificado(s) e/são transmitido(s). Assim, o sistema ignorará o som ou vídeo que tem um identificador embutido de outro sistema de videoconferência.

Relatório Descritivo da Patente de Invenção para "**SISTEMA DE TRANSMISSOR PARA UM SISTEMA DE VIDEOCONFERÊNCIA, MÉTODO PARA OPERAR UM SISTEMA DE VIDEOCONFERÊNCIA E MEIO DE ARMAZENAMENTO DE COMPUTADOR**".

ANTECEDENTES

[0001] A videoconferência tornou-se difundida e muitos têm salas especialmente configuradas para as seções de videoconferência. Tais salas tipicamente contêm equipamento de videoconferência, tal como uma ou mais câmeras móveis e um ou mais microfones, os microfones tipicamente sendo colocados em localizações ao redor de uma mesa na sala para participantes. Uma Detecção de Orador Ativo (ASD) é frequentemente utilizada para selecionar uma câmera, ou para mover (panoramizar e/ou inclinar) uma câmera para mostrar a pessoa na sala que está falando e/ou para selecionar o microfone o qual estará ativo. Quando uma pessoa remota está falando, a sua imagem e/ou som vêm de um display de áudio-vídeo, tal como uma televisão (TV), monitor, ou outro tipo de display, na sala. Isto pode fazer com que a ASD erroneamente selecione a imagem na pessoa remota na TV que está falando ao invés de selecionar a última pessoa local que está ou estava falando.

[0002] Também, em seções de videoconferência de múltiplas localizações onde três ou mais localizações separadas estão em uma única seção de videoconferência, então, tipicamente, diversos painéis serão exibidos, um painel sendo maior do que os outros e mostrando a pessoa que está falando, e os outros painéis mostrando uma imagem de uma câmera nas outras localizações. Quando uma ASD errônea ocorre, como acima mencionado, o equipamento na sala onde uma pessoa está falando enviará um sinal para o equipamento nas outras localizações avisando que a pessoa na sua localização está falando e assim o display principal deve ser de uma câmera. Quando isto acontece o painel maior pode mudar de mostrar uma pessoa que está realmente falando

para mostrar uma imagem de uma tela de TV ou uma cadeira vazia. Assim, um problema com a ASD é que se o som do sistema de videoconferência for refletido ou for tão alto que este dispara a ASD então o som remoto pode ser retransmitido de volta para o sistema remoto e/ou fazer com que a câmera local focalize sobre uma cadeira vazia ou a tela de display mostrando a localização de videoconferência remota.

[0003] Uma técnica que foi utilizada para eliminar tal seleção de ASD errônea é localizar o traço de linha de escaneamento de imagem sobre a TV para determinar que o som está vindo de uma TV ao invés de uma pessoa local. As TVs de alta definição (HDTVs), no entanto, têm taxas de escaneamento progressivo altas (240 Hz ou melhor) e resoluções de imagem que são iguais às câmeras de modo que o traço de linha de escaneamento de imagem é de uso limitado quando uma HDTV está envolvida. Além disso, a ASD pode frequentemente ter problemas com o som ecoando ao redor de uma sala. Uma superfície refletiva de som, tal como uma janela ou um quadro coberto com vidro, pode refletir o som da TV em um modo que o som parece originar de uma pessoa local na mesa, mesmo se realmente não existir uma pessoa sentada naquela posição na mesa. Ainda, se uma gravação for feita da videoconferência, é dependente de um humano lembrar identificar precisamente a gravação com pelo menos, por exemplo, a data da videoconferência. Isto é frequentemente esquecido e feito mais tarde, algumas vezes com uma etiqueta errônea ou incompleta. É com relação a estas considerações e outras que a descrição feita aqui é apresentada.

SUMÁRIO

[0004] Tecnologias estão aqui descritas para um sistema de videoconferência que seleciona um orador ativo enquanto evitando erroneamente selecionar um microfone ou câmera que está captando um sinal de áudio ou vídeo de um sinal remoto conectado. Em uma implementação, um identificador é adicionado a um sinal de áudio e/ou vídeo que

sai. Se o microfone captar um som que contém o identificador do sistema remoto então o som é ignorado e a ASD não é implementada. Se o som não contiver o identificador remoto então o vídeo da câmera local é inspecionado. Se este contiver um identificador remoto então a ASD não é implementada. Se um identificador remoto não estiver presente em nenhum sinal então a ASD é implementada.

[0005] De acordo com uma modalidade aqui apresentada, um sistema de transmissor para um sistema de videoconferência tem um gerador de identificador para gerar pelo menos um de um identificador de áudio ou um identificador de vídeo, um combinador de sinal para pelo menos um de (i) combinar um sinal de áudio recebido com o identificador de áudio para produzir um sinal de áudio identificado ou (ii) combinar receber um sinal de vídeo recebido com o identificador de vídeo para produzir um sinal de vídeo identificado, e um transmissor para transmitir (i) o sinal de áudio identificado e o sinal de vídeo recebido, (ii) o sinal de áudio recebido e o sinal de vídeo identificado, ou (iii) o sinal de áudio identificado e o sinal de vídeo identificado. Um sistema de videoconferência remoto pode então utilizar os identificadores incorporados para distinguir os sons e imagens locais de sons e imagens remotas.

[0006] O método para operar um transmissor de um sistema de videoconferência inclui receber um sinal de áudio, receber um sinal de vídeo, gerando pelo menos um de um identificador de áudio ou um identificador de vídeo, pelo menos um de (i) combinar o sinal de áudio com o identificador de áudio para produzir um sinal de áudio identificado ou (ii) combinar o sinal de vídeo recebido com o identificador de vídeo para produzir um sinal de vídeo identificado, e transmitir (i) o sinal de áudio identificado e o sinal de vídeo, (ii) o sinal de áudio e o sinal de vídeo identificado, ou (iii) o sinal de áudio identificado e o sinal de vídeo identificado.

[0007] Um meio de armazenamento de computador tem instruções

executáveis por computador armazenadas neste. Estas instruções fazem com que o computador gere pelo menos um de um identificador de áudio ou um identificador de vídeo, pelo menos um de (i) combinar um sinal de áudio recebido com o identificador de áudio para produzir um sinal de áudio identificado ou (ii) combinar um sinal de vídeo recebido com o identificador de vídeo para produzir um sinal de vídeo identificado, e transmitir (i) o sinal de áudio identificado e o sinal de vídeo recebido, (ii) o sinal de áudio recebido e o sinal de vídeo identificado, ou (iii) o sinal de áudio identificado e o sinal de vídeo identificado.

[0008] Deve ser apreciado que o assunto acima descrito pode também ser implementado como um aparelho controlado por computador, um processo de computador, um sistema de computação, ou como um artigo de manufatura tal como um meio legível por computador. Estas e várias outras características serão aparentes da leitura da Descrição Detalhada seguinte e uma revisão dos desenhos associados.

[0009] Este Sumário está provido para introduzir uma seleção de conceitos em uma forma simplificada que estão adicionalmente abaixo descritos na Descrição Detalhada. Este sumário não pretende identificar as características chave ou características essenciais do assunto reivindicado, nem é pretendido que este Sumário seja utilizado para limitar o escopo do assunto reivindicado. Além disso, o assunto reivindicado não está limitado a implementações que resolvem qualquer ou todas as desvantagens notadas em qualquer parte desta descrição.

BREVE DESCRIÇÃO DOS DESENHOS

[0010] Figura 1 é uma configuração exemplar de um sistema de transmissor de um sistema de videoconferência.

[0011] Figura 2 é uma ilustração de um ambiente de sistema de videoconferência exemplar.

[0012] Figura 3 é um fluxograma que mostra uma detecção de identificador exemplar e uma técnica de controle de câmera e microfone.

[0013] Figura 4 é um fluxograma de uma técnica de gravação de informações exemplar.

[0014] Figura 5 é um diagrama de arquitetura de computador que mostra uma arquitetura de hardware e software de computador ilustrativa para um sistema de computação capaz de implementar aspectos das modalidades aqui apresentadas.

DESCRIÇÃO DETALHADA

[0015] A descrição detalhada seguinte está direcionada a tecnologias para videoconferência que podem corretamente selecionar um orador ativo enquanto evitando erroneamente selecionar um microfone ou câmera que está captando áudio ou vídeo de um sinal remoto conectado. Na descrição detalhada seguinte, referências são feitas aos desenhos acompanhantes que formam uma sua parte, e os quais estão mostrados por meio de ilustração modalidades ou exemplos específicos. Referindo agora aos desenhos, nos quais números iguais representam elementos iguais através de todas as diversas figuras, aspectos de um sistema de computação e uma metodologia para videoconferência serão descritos.

[0016] A Figura 1 é uma configuração exemplar de um sistema de transmissor 105 de um sistema de videoconferência 100. O sistema de transmissor 105 tem um sistema de seleção e controle de câmera e microfone 120, um gerador de identificador de vídeo 125, um combinador de sinal de vídeo 130 o qual provê um sinal de saída de vídeo 135, um gerador de identificador de áudio 140, e um combinador de sinal de áudio 145 o qual provê um sinal de saída de áudio 150. Os sinais de saída de vídeo e áudio podem ser difundidos ou transmitidos por um transmissor 155. O sistema de controle 120 pode também enviar sinais destinados para sistema remotos, avisando que este tem um orador ativo ao qual deve ser dado o maior painel se múltiplos painéis forem utilizados para exibir múltiplas localizações. O transmissor 155 pode usar qualquer

meio conveniente para enviar os sinais de saída de vídeo e áudio e quaisquer sinais de controle para um ou mais sistemas de receptor 160 em localizações remotas. Será apreciado que existe um sistema de transmissor 105 e um sistema de receptor 160 em cada localização, e que o sistema de transmissor 105 e o sistema de receptor 160 em uma localização podem estar combinados em um único dispositivo.

[0017] Uma ou mais câmeras 110 (110A-110N) e um ou mais microfones 115 (115A-115N) proveem sinais de vídeo e sinais de áudio, respectivamente, para o sistema de transmissor 105 e, mais especificamente, para o sistema de controle 120, o qual tem entradas para receber estes sinais. O sistema de seleção de controle de câmera e microfone 120 pode selecionar qual câmera 110 e qual microfone 115 serão utilizados para gerar a imagem e som locais, se mais do que um ou cada dispositivo for utilizado, pode controlar a panoramização, zoom, e/ou inclinação da câmera 110 selecionada se a câmera puder ser assim controlada, e pode gerar um sinal de controle ou outros para transmissão para os sistemas remotos.

[0018] O gerador de identificador de vídeo 125 e um gerador de identificador de áudio 140 geram identificadores de vídeo e áudio, respectivamente. Um combinador de sinal de vídeo 130 manipula ou modifica os pixels de vídeo no fluxo de vídeo para adicionar o identificador de vídeo e produzir um sinal de vídeo identificado 135. Um combinador de sinal de áudio 145 manipula ou modifica os bits no fluxo de áudio para produzir um sinal de áudio identificado 150. Isto pode ser considerado ser "identificar" um sinal ou adicionar um identificador a um sinal. Os geradores de identificador 125 e 140 podem estar incorporados em um único dispositivo, os combinadores de sinal 130, 145 podem estar incorporados em um único dispositivo, e um ou todos estes componentes podem estar incorporados como parte do sistema de controle 120.

[0019] Um fluxo de vídeo e/ou áudio é de preferência modificado

utilizando modos ou somente para níveis os quais são sutis e/ou indetectáveis por humanos, mas os quais podem ser detectados por análise algorítmica do fluxo de vídeo ou áudio. Um nível de distorção de menos do que um nível predeterminado pode ser imperceptível para um observador humano típico. Por exemplo, modificando o bit menos significativo em uma palavra de dados mesmo se o identificador estivesse em cada palavra geralmente não seria notável ou desagradável. Como outro exemplo, colocar um identificador de vídeo durante um intervalo de apagamento ou período de reconstituição em um quadro de vídeo, ou colchoar um identificador de vídeo no canto do fundo do display pode não ser notável ou desagradável. Mesmo colocando o identificador de vídeo como o bit mais significativo pode não ser notável ou desagradável se somente em um único pixel durante um quando.

[0020] O fluxo de vídeo e/ou áudio pode ser modificado, por exemplo, utilizando o bit ou bits menos significativos para transportar informações outras que o sinal de áudio ou vídeo inicial. Tal modificação pode ser feita em cada palavra de dados, palavras de dados alternadas, ou cada enésima palavra de dados, cada N milissegundos, antes ou após uma palavra ou bit de sincronização, etc.. Por exemplo, o(s) último(s) bit(s) da(s) palavra(s) de dado(s) apropriada(s) pode(m) sempre ter o mesmo valor, por exemplo, 0, 1,2, 3, etc., podem alternar entre valores, podem progredir através de valores, etc. Outras técnicas podem também ser utilizadas para identificar uma palavra de dados ou uma sua parte, como um identificador, ou como informações de identificação associadas com um identificador ou uma videoconferência. Como outro exemplo, uma palavra de dados inteira pode ser utilizada para este propósito. Por exemplo, se áudio for amostrado a uma taxa de 4000 amostras/segundo, então utilizando um número limitado destas palavras para transportar as informações de identificador não degradaria notavelmente a qualidade do áudio. Os sinais de vídeo proveem a oportunidade

de enviar ainda mais informações sem notavelmente degradar a qualidade do vídeo.

[0021] A Figura 2 é uma ilustração de um ambiente de sistema de videoconferência exemplar 200. Diversas pessoas 205 (205A-205C) estão reunidas ao redor de uma mesa 210, a qual tem sobre a mesma uma pluralidade de microfones 115 (115A-115E). Existe um display 215, o qual pode ser uma TV, que mostra uma pessoa remota 220. Também mostrado um alto-falante 225. Existe um sistema de transmissor 105 o qual está conectado nas câmeras e microfones, e um sistema de receptor 160 o qual está conectado no display e no orador. Como mencionado, o sistema de transmissor 105 e o sistema de receptor 160 podem estar, e tipicamente estão, incorporados em um único dispositivo e estão conectados por um meio de transmissão conveniente a um ou mais sistema de videoconferências remotos.

[0022] Quando uma pessoa local fala, tal como a pessoa 205B, o sistema de controle 120 detecta o sinal do microfone 115B, muda para o microfone 115B, muda para uma câmera 110B previamente apontada na direção da área da pessoa 115B, ou aponta uma câmera 110B na direção da área da pessoa 115B, e então transmite o sinal de áudio do microfone 115B e o sinal de vídeo da câmera 115B para a localização remota, possivelmente juntamente com um sinal indicando que a pessoa 205B deve ser proeminentemente exibida na tela remota. Apontar ou direcionar uma câmera, como aqui utilizado, é panoramizar, inclinar, e/ou fazer zoom coma câmera para conseguir uma imagem desejada de uma localização desejada.

[0023] Considere agora a situação em que um objeto ou superfície refletivo de som 230, tal como um espelho, quadro, ou janela está presente. O orador remoto 220 está falando e a voz do orador remoto 220 é transmitida para dentro da sala por um alto-falante 225. O som 235 do orador remoto 220 salta da superfície refletiva 230 e chega no microfone

115D. O sistema de controle 120 detecta a voz refletida 235 no microfone 115D e erroneamente determina que existe uma pessoa local no microfone 115D que está falando. O sistema de controle 120 muda para o microfone 115D e aponta uma câmera 110 na direção do espaço vazio próximo do microfone 115D. Assim, os sons e ecos refletidos podem causar problemas durante as seções de videoconferência. Isto pode ocorrer repetidamente até que a pessoa remota 220 para de falar ou alguém diminui o volume do alto-falante 225.

[0024] Para eliminar ou pelo menos reduzir tal ação ASD errônea, o sistema de transmissor 105 injeta um identificador(es) no sinal de áudio e/ou sinal de vídeo. O display 215 e o alto-falante 225 então reproduzirão estes identificadores nas suas saídas. Agora, considere novamente a situação em que o orador remoto 220 está falando e a voz do orador remoto 220 é transmitida para dentro da sala por um alto-falante 225. O som 235 do orador remoto 220 salta da superfície refletiva 230 e chega no microfone 115D. O sistema de controle 120 detecta a voz refletida 235 no microfone 115D mas também detecta o identificador na voz refletida 235. O sistema de controle 120 então determina que o som é do orador remoto, não um orador local, e portanto não executa nenhuma ação com relação à voz refletida.

[0025] Como outra proposta, quando a voz refletida 235 está presente no microfone 115D, o sistema de controle 120 pode ao invés, ou além disso, inspecionar a saída da câmera. Se o identificador de vídeo estiver presente então o sistema de controle 120 determina que o som é um som refletido, e portanto não executa nenhuma ação com relação à voz refletida.

[0026] Quando, no entanto, uma pessoa local 205B fala, o microfone 115B detecta a voz da pessoa local 205B, mas um identificador de áudio não está presente. O sistema de controle 120 então corretamente muda o microfone 115B e direciona a câmera 110 para a pessoa local

205B, e um identificador de vídeo não estará presente. Assim, o sistema de controle 120 corretamente determina que a pessoa 205B está falando e executa a ação apropriada. Será apreciado que algum som refletido 235 pode aparecer nos microfones 115B também. O volume do som refletido 235 será, no entanto, significativamente menor do que o volume da voz do orador local 205B, de modo que o identificador estará em um nível baixo demais para ser detectado pelo sistema de controle 120. Isto é, quando o som do microfone é digitalizado, o volume de identificador estará abaixo do(s) bit(s) menos significativo(s). O som refletido 235 pode também ser captado por outros microfones 115 também, mas o sistema de controle 120 rejeitará este microfones ou porque o seu volume é menor do que o volume no microfone 115B ou porque o identificador será prontamente detectável.

[0027] É possível, em algumas situações, que existirá uma câmera 240 no fundo da sala além das, ou invés das câmeras 110. Assuma que a pessoa remota 220 está falando e o som emitido pelo alto-falante 225 é recebido por um microfone 115A ou 115E. Um sistema convencional poderia erroneamente detectar este som recebido como um orador local e comutar para aquele microfone e direcionar a câmera 240 para a localização. Ao invés, com os identificadores aqui utilizados, o sistema de controle 120 detectará o identificador no sinal de áudio captado pelo microfone 115A ou 115E, determinará que a voz não é aquela de um orador local, e não mudará para o microfone 115A ou 115E. Também, o sistema de controle 120 pode apontar a câmera 240 na direção do display 215, detectar o identificador de vídeo que está sendo emitido pelo display 215, e então apontar a câmera 240 de volta para a sua direção original ou para uma direção padrão. Assim, os identificadores de áudio e vídeo melhoram a experiência de videoconferência reduzindo ou eliminando uma mudança errônea da câmera e/ou microfone causada pela voz do orador remoto.

[0028] Os identificadores podem também ser utilizados para a identificação da videoconferência, se desejado. Por exemplo, os identificadores podem conter informações referentes ao nome da companhia, hora, data, localização de sala, equipamento de transmissão utilizado tal como, mas não limitado a modelo, fabricante, número de série, versão de software, informações de marca registrada, informações de direitos autorais, informações de confidencialidade, informações de propriedade, protocolo ou padrão utilizado, etc. Todas estas informações não precisam ser transmitidas, nem todas as informações desejadas precisam ser transmitidas de uma vez, repetidamente, ou continuamente. Ao invés, os bits os quais identificam o identificador como tal precisam somente ser transmitidos frequentemente o bastante que o sistema de controle 120 possa reconhecer o identificador como tal. Assim, por exemplo, como acima mencionado, os bits os quais identificam o identificador poderiam somente ser transmitidos a cada N palavras de dados, as outras palavras de dados sendo utilizadas para a transmissão das informações acima mencionadas.

[0029] Além disso, as informações contidas no(s) identificador(es) não precisam ser obtidas da imagem apresentada pelo display 215 ou do som apresentado pelo alto-falante 225. Ao invés, e de preferência, estas informações são obtidas diretamente dos sinais de vídeo e/ou áudio recebidos pelo sistema de receptor 160.

[0030] A taxa de dados pode ser bastante lenta, mas, de preferência, a parte identificável do identificador é de preferência fornecida repetidamente em menos do que a metade da histerese do retardo de ASD. A parte identificável do identificador é ainda mais de preferência fornecida mais frequentemente de modo a acomodar os dados perdidos devido à interferência durante a transmissão ou ruído da sala. A velocidade de fornecimento das informações adicionais é menos sensível ao tempo e portanto podem ser transmitidas ao longo de um período de

tempo mais longo.

[0031] A Figura 3 é um fluxograma de uma técnica de controle de câmera e microfone e detecção de identificador exemplar 300. Após o início 305, uma determinação 310 é feita quanto a se qualquer sinal de áudio está acima de um nível limite. Se não, um retorno é feito para 310. Se sim, então uma determinação 315 é feita quanto a se um identificador está presente naquele sinal de áudio. Se sim, então aquele sinal de áudio é ignorado 317 e um retorno é feito para 310. Se não, então uma câmera é direcionada ou apontada 320 para a fonte de som identificada pelo sinal de áudio. Por exemplo, se o sinal de áudio for do microfone 115A então uma câmera 110 será apontada na direção da área servida pelo microfone 115A, ou uma câmera a qual foi previamente apontada na direção daquela área será selecionada.

[0032] Uma determinação é então feita 325 quanto a se um identificador está presente no sinal de vídeo daquela câmera. Se sim, então a câmera é redirecionada 330 para a sua posição anterior ou a câmera anterior é selecionada. Se não, então identificadores locais estão inseridos 335 no sinal de áudio e/ou sinal de vídeo. Os sinais identificados são então transmitidos. Um retorno é então feito para 310.

[0033] Assim se um microfone está captando som e existe um identificador de áudio embutido naquele som, ou se uma câmera está direcionada para a fonte daquele som estiver captando um identificador de vídeo embutido no sinal de vídeo, então o sistema ignorará aquele som e deixará os ajustes de microfone e câmera como estão. Se, no entanto, um identificador embutido não for detectado em qualquer sinal, então o microfone e/ou a câmera serão selecionados para transmissão daquele som e imagem para a videoconferência remota após inserção de um identificador local em pelo menos um daqueles sinais. Assim, um orador ativo é corretamente selecionado enquanto que sons remotos, refletidos são ignorados.

[0034] A Figura 4 é um fluxograma de uma técnica de gravação de informações exemplar 400. Após o início 405, uma determinação é feita 410 se a seção deve ser gravada. Se não então o procedimento termina 415. Se sim, então uma determinação é feita 420 quanto a se os identificadores estão presentes. Se nenhum identificador estiver presente então a seção é gravada 430. Se pelo menos um identificador estiver presente uma determinação é feita 425 se as informações estão presentes no(s) identificador(es). Se não, então a seção é gravada 430. Se sim, então a seção é gravada 435 com pelo menos algumas das informações. As informações a serem gravadas com a seção podem ser todas as informações incluídas no identificador ou podem ser somente uma porção pré-selecionada, tal como a data e hora.

[0035] Deve ser apreciado que as operações aqui descritas são implementadas (1) como uma sequência de atos implementados por computador ou módulos de programa que executam em um sistema de computação e/ou (2) como circuitos lógicos de máquina interconectados ou módulos de circuito dentro do sistema de computação. A implementação é um caso de escolha dependente do desempenho e outros requisitos do sistema de computação. Consequentemente, as operações lógicas aqui descritas são referidas variadamente como operações de estados, dispositivos estruturais, atos ou módulos. Estas operações, dispositivos estruturais, atos e módulos podem ser implementados em software, em firmware, em lógica digital de uso especial, e qualquer sua combinação. Deve também ser apreciado que mais ou menos operações podem ser executadas do que mostradas nas figuras e aqui descritas. Estas operações podem também ser executadas em uma ordem diferente do que aquela aqui descrita.

[0036] A Figura 5 mostra uma arquitetura de computador ilustrativa para um computador 500 capaz de executar os componentes de software aqui descritos para um sistema de videoconferência no modo

acima apresentado. A arquitetura de computador mostrada ilustra um desktop, laptop, ou computador servidor convencional e pode ser utilizada para executar qualquer aspecto dos componentes de software aqui apresentados descritos como executando no computador de cliente 104, nos computadores de servidor de interface inicial 106A-106N, ou nos computadores de servidor auxiliares 108A-108N. A arquitetura de computador mostrada inclui uma unidade de processamento central 502 ("CPU"), uma memória de sistema 508, que inclui uma memória de acesso randômico 514 ("RAM") e uma memória somente de leitura ("ROM") 516, e um barramento de sistema 504 que acopla a memória na CPU 502. Um sistema de entrada /saída básico que contém as rotinas básicas que ajudam a transferir as informações entre os elementos dentro do computador 500, tal como durante a partida, está armazenado na ROM 516. O computador 500 ainda inclui um dispositivo de armazenamento de massa 510 para armazenar um sistema de operação 518, programas de aplicação, e outros módulos de programa, os quais estão aqui descritos em maiores detalhes.

[0037] O dispositivo de armazenamento de massa 510 está conectado na CPU 502 através de um controlador de armazenamento de massa (não mostrado) conectado no barramento 504. O dispositivo de armazenamento de massa 510 e seu meio legível por computador associado proveem um armazenamento não volátil para o computador 500. Apesar da descrição de meio legível por computador aqui contida referir a um dispositivo de armazenamento de massa, tal como um disco rígido ou unidade de CD-ROM deve ser apreciado por aqueles versados na técnica que o meio legível por computador pode ser qualquer meio de armazenamento de computador disponível ou meio de comunicação que pode ser acessado pela arquitetura de computador 500.

[0038] Como exemplo, e não limitação o meio de armazenamento de computador pode incluir um meio volátil e não volátil, removível e não

removível implementado em qualquer método ou tecnologia para armazenamento de informações tal como instruções legíveis por estruturas de dados, módulos de programa ou outros dados. Por exemplo, o meio de armazenamento de computador inclui, mas não estão limitados a RAM, ROM, EPROM, EEPROM, memória instantânea ou outra tecnologia de memória de estado sólido, CD-ROM, discos versáteis digitais ("DVD"), HD-DVD, BLU-RAY ou outros armazenamentos ótico, cassetes magnéticos, fita magnética, armazenamento de disco magnético ou outros dispositivos de armazenamento magnético, ou qualquer outro meio o qual pode ser utilizado para armazenar as informações desejadas e o qual pode ser acessado pelo computador 500. Para propósitos das reivindicações, a frase "meio de armazenamento de computador", e suas variações, não inclui ondas ou sinais por si e/ou meio de comunicação.

[0039] O meio de comunicação inclui instruções legíveis por computador, estruturas de dados, módulos de programa ou outros dados em um sinal de dados modulados tal como uma onda portadora ou outro mecanismo de transporte e inclui qualquer meio de fornecimento. O termo "sinal de dados modulados" significa um sinal que tem uma ou mais de suas características mudadas ou ajustadas em uma maneira de modo a codificar as informações no sinal. Como exemplo, e não limitação, o meio de comunicação inclui um meio com fio tal como uma rede com fio ou uma conexão com fio direta, e um meio sem fio tal como um meio sem fio acústico, RF, infravermelho e outros. Combinações de qualquer um dos acima devem também estar incluídas dentro do escopo de meio legível por computador.

[0040] De acordo com várias modalidades, o computador 500 pode operar em um ambiente em rede utilizando conexões lógicas para computadores remotos através de uma rede tal como a rede 520. O computador 500 pode conectar com a rede 520 através de uma unidade de interface de rede 506 conectada no barramento 504. Deve ser apreciado

que a unidade de interface de rede 506 pode também ser utilizada para conectar a outros tipos de redes e sistemas de computador remotos. O computador 500 pode também incluir um controlador de entrada/saída 512 para receber e processar uma entrada de um número de outros dispositivos, incluindo um teclado, mouse, ou caneta eletrônica. Similarmente, um controlador de entrada/saída pode prover uma saída para uma tela de display, uma impressora, ou outro tipo de dispositivo de saída.

[0041] Como acima mencionado resumidamente, um número de módulos de programa e arquivos de dados pode estar armazenado no dispositivo de armazenamento de massa 510 e RAM 514 do computador 500, incluindo um sistema de operação 518 adequado para operar a operação de um desktop em rede, laptop, ou computador servidor. O dispositivo de armazenamento de massa 510 e RAM 514 podem também armazenar um ou mais módulos de programa os quais implementam as várias operações acima descritas. O dispositivo de armazenamento de massa 510 e a RAM 514 pode também armazenar outros tipos de módulos de programa.

[0042] Apesar do assunto aqui descrito ser apresentado no contexto geral de um ou mais módulos de programa que executam em conjunto a execução de um sistema de operação e programas de aplicação em um sistema de computador, aqueles versados na técnica reconhecerão que outras implementações podem ser executadas em combinação com outros tipos de módulos de programa. geralmente, os módulos de programa incluem rotinas, programas, componentes, estruturas de dados, e outros tipos de estruturas que executam tarefas específicas ou implementam tipos de dados abstratos específicos. Mais ainda aqueles versados na técnica apreciarão que o assunto aqui descrito pode ser praticado, se desejado, com outras configurações de sistema de computador incluindo dispositivos portáteis, sistemas de multiprocessador, eletrônica de

consumidor baseada em microprocessador ou programável, minicomputadores, computadores mainframe, e similares.

[0043] Com base no acima, deve ser apreciado que tecnologias para videoconferência estão aqui providas. Apesar do assunto aqui apresentado ter sido descrito em uma linguagem específica para características estruturais de computador, atos metodológicos e transformativos, maquinário de computação específico, e meio legível por computador, deve ser compreendido que a invenção definida nas reivindicações anexas não está necessariamente limitada às características, atos ou meios específicos aqui descritos. Ao invés as características, atos ou meios específicos estão descritos como formas exemplares para implementar as reivindicações.

[0044] O assunto acima descrito está provido como ilustração somente e não deve ser considerado como limitante. Várias modificações e mudanças podem ser feitas no assunto aqui descrito sem seguir as modalidades exemplares e aplicações ilustradas e descritas, e sem afastar do verdadeiro espírito e escopo da presente invenção no qual está apresentado nas reivindicações seguintes.

REIVINDICAÇÕES

1. Sistema de transmissor para um sistema de videoconferência, **caracterizado pelo fato de que** compreende:

um gerador de identificador para gerar um identificador de áudio;

um combinador para combinar um sinal de áudio com o identificador de áudio para produzir um sinal de áudio identificado; e

um transmissor para transmitir o sinal de áudio identificado e um sinal de vídeo correspondente; e

um sistema de controle operativo para:

determinar se o sinal de áudio está acima de um nível limite;

se o sinal de áudio foi determinado como estando acima do nível limite, então determinar se o sinal de áudio tem um identificador de áudio nele incorporado; e

se o sinal de áudio foi determinado como não tendo um identificador de áudio nele incorporado, então direcionar uma câmera em direção a uma fonte do sinal de áudio ou selecionar uma câmera que aponta em direção a uma fonte do sinal de áudio, em que a câmera produz o sinal de vídeo correspondente.

2. Sistema de transmissor, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende um sistema de controle para receber uma pluralidade de sinais de áudio, e para selecionar um dos sinais de áudio recebidos como o sinal de áudio a ser combinado com o identificador de áudio, em que o sistema de controle seleciona um dos sinais de áudio recebidos que não contém um identificador de áudio a partir de outro sistema de videoconferência.

3. Sistema de transmissor, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a distorção do sinal de áudio causada pelo identificador de áudio estar abaixo de um nível predeterminado.

4. Sistema de transmissor, de acordo com a reivindicação 1,

caracterizado pelo fato de que o sistema de controle ser ainda operativo para:

determinar se o sinal de vídeo correspondente tem um identificador de vídeo nele incorporado; e

se o sinal de vídeo foi determinado como tendo um identificador de vídeo nele incorporado, então redirecionar a câmera para uma fonte anterior de um sinal de áudio ou selecionar uma câmera previamente selecionada.

5. Sistema de transmissor, de acordo com a reivindicação 4, **caracterizado pelo fato de que** o sistema de controle ser ainda operativo para:

se o sinal de vídeo correspondente foi determinado como não tendo um identificador de vídeo nele incorporado, então instruir o combinador a combinar o sinal de áudio com o identificador de áudio para produzir o sinal de áudio identificado e instruir o transmissor a transmitir o sinal de áudio identificado e o sinal de vídeo correspondente.

6. Sistema de transmissor, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende:

um gerador de identificador de vídeo;

um combinador de vídeo para combinar o sinal de vídeo correspondente com o identificador de vídeo para produzir um sinal de vídeo identificado; e

em que o transmissor também transmite o sinal de vídeo identificado.

7. Método para operar um sistema de videoconferência, o método **caracterizado pelo fato de que** compreende as etapas de:

receber um sinal de áudio;

receber um sinal de vídeo correspondente;

gerar um identificador de áudio;

determinar se o sinal de áudio está acima de um nível limite;

se o sinal de áudio foi determinado como estando acima do nível limite, então determinar se o sinal de áudio tem um identificador de áudio nele incorporado;

se o sinal de áudio foi determinado como não tendo um identificador de áudio nele incorporado, então direcionar uma câmera em direção a uma fonte do sinal de áudio ou selecionar uma câmera que aponta em direção a uma fonte do sinal de áudio, em que a câmera produz o sinal de vídeo correspondente;

combinar o sinal de áudio com o identificador de áudio para produzir um sinal de áudio identificado; e

transmitir o sinal de áudio identificado e o sinal de vídeo correspondente.

8. Método, de acordo com a reivindicação 7, **caracterizado pelo fato de** existir uma pluralidade de sinais de áudio recebidos e ainda compreender selecionar um sinal de áudio para combinar o qual não contenha um identificador de áudio de outro sistema de videoconferência.

9. Método, de acordo com a reivindicação 7, **caracterizado pelo fato de que** a distorção do sinal de áudio causada pelo identificador de áudio estar abaixo de um nível predeterminado.

10. Método, de acordo com a reivindicação 7, **caracterizado pelo fato de que** ainda compreende, antes de combinar o sinal de áudio com o identificador de áudio:

determinar se o sinal de vídeo correspondente tem um identificador de vídeo nele incorporado; e

se o sinal de vídeo foi determinado como tendo um identificador de vídeo nele incorporado, então redirecionar a câmera para uma fonte anterior de um sinal de áudio ou selecionar uma câmera previamente selecionada.

11. Método, de acordo com a reivindicação 10,

caracterizado pelo fato de que ainda compreende, antes de combinar o sinal de áudio com o identificador de áudio:

se o sinal de vídeo correspondente foi determinado como não tendo um identificador de vídeo nele incorporado, então realizar a combinação e transmissão.

12. Método, de acordo com a reivindicação 11, **caracterizado pelo fato de que** ainda compreende:

antes da transmissão, gerar um identificador de vídeo e combinar o sinal de vídeo correspondente com o identificador de vídeo para produzir um sinal de vídeo identificado; e

em que transmitir o sinal de vídeo correspondente compreende transmitir o sinal de vídeo identificado.

13. Meio de armazenamento de computador, **caracterizado pelo fato de que** possui um método para:

determinar se um sinal de áudio recebido está acima de um nível limite;

se o sinal de áudio recebido foi determinado como estando acima do nível limite, então determinar se o sinal de áudio recebido tem um identificador de áudio nele incorporado;

se o sinal de áudio recebido foi determinado como não tendo um identificador de áudio nele incorporado, então direcionar uma câmera em direção a uma fonte do sinal de áudio recebido ou selecionar uma câmera que aponta em direção a uma fonte do sinal de áudio recebido, em que a câmera produz um sinal de vídeo correspondente;

gerar um o identificador de áudio;

combinar o sinal de áudio recebido com o identificador de áudio para produzir um sinal de áudio identificado; e

transmitir o sinal de áudio identificado e o sinal de vídeo correspondente.

14. Meio de armazenamento de computador, de acordo com

a reivindicação 13, **caracterizado pelo fato de que** o meio de armazenamento de computador ainda compreende um método armazenado nele que, quando executado por um computador, faz o computador detectar identificadores de áudio de um sistema de videoconferência remoto em uma pluralidade de sinais de áudio recebidos, e selecionar, como o sinal de áudio a ser combinado com o identificador de áudio, um dos sinais de áudio recebidos o qual não contenha um identificador de áudio de outro sistema de videoconferência.

15. Meio de armazenamento de computador, de acordo com a reivindicação 13, **caracterizado pelo fato de que** o meio de armazenamento de computador ainda compreende um método armazenado nele que, quando executado por um computador, faz ainda o computador, antes de combinar o sinal de áudio com o identificador de áudio,

determinar se o sinal de vídeo correspondente tem um identificador de vídeo nele incorporado; e

se o sinal de vídeo correspondente foi determinado como tendo um identificador de vídeo nele incorporado, então redirecionar a câmera para uma fonte anterior de um sinal de áudio ou selecionar uma câmera previamente selecionada.

16. Meio de armazenamento de computador, de acordo com a reivindicação 15, **caracterizado pelo fato de que** o meio de armazenamento de computador ainda compreende um método armazenado nele que, quando executado por um computador, faz ainda o computador, antes de combinar o sinal de áudio com o identificador de áudio,

se o sinal de vídeo correspondente foi determinado como não tendo um identificador de vídeo nele incorporado, então realizar a combinação e transmissão.

17. Meio de armazenamento de computador, de acordo com a reivindicação 13, **caracterizado pelo fato de que** o meio de armazenamento de computador ainda compreende um método armazenado

nele que, quando executado por um computador, faz ainda o computador:

antes da transmissão, gerar um identificador de vídeo e combinar o sinal de vídeo correspondente com o identificador de vídeo para produzir um sinal de vídeo identificado; e

em que transmitir o sinal de vídeo correspondente compreende transmitir o sinal de vídeo identificado.

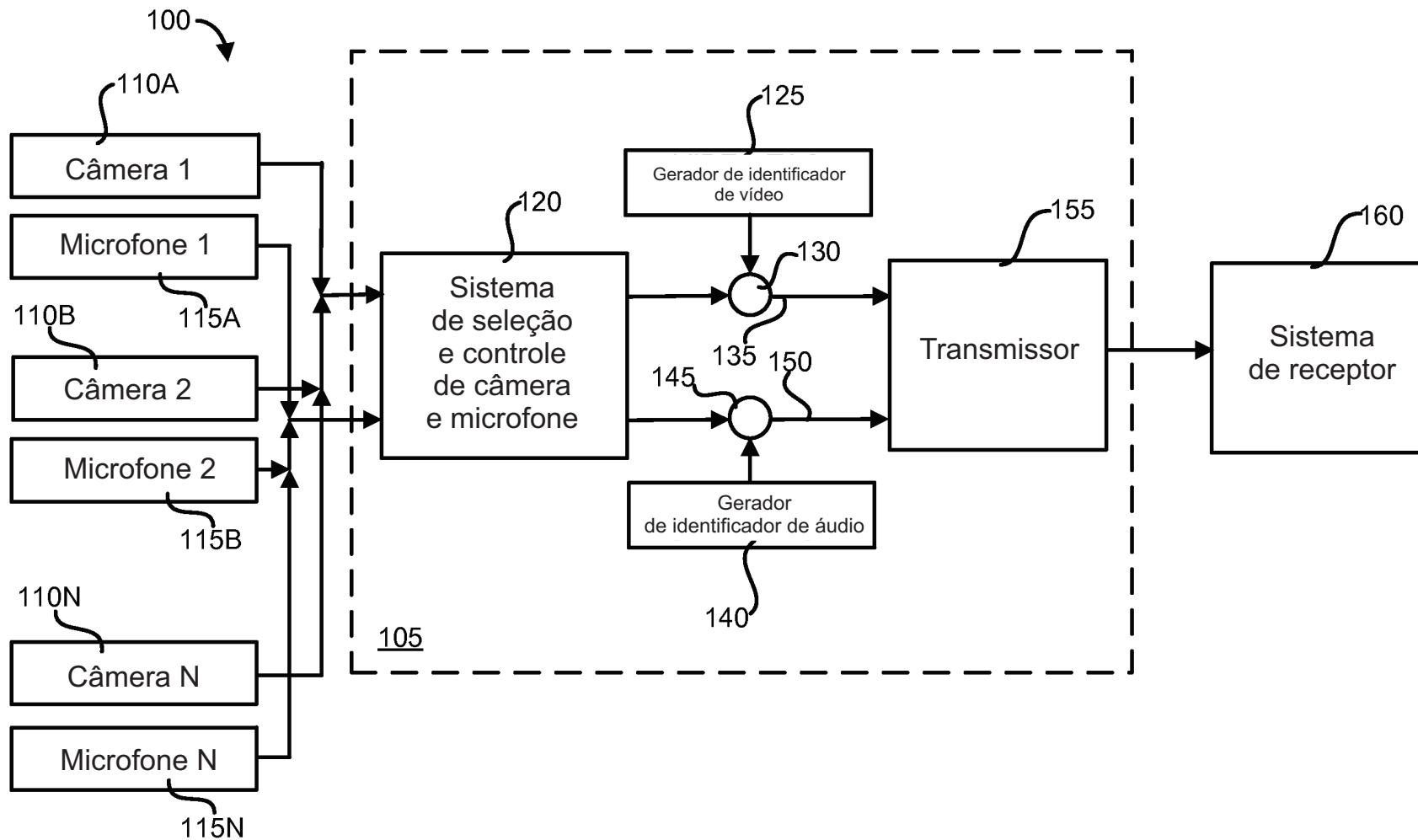


FIG. 1

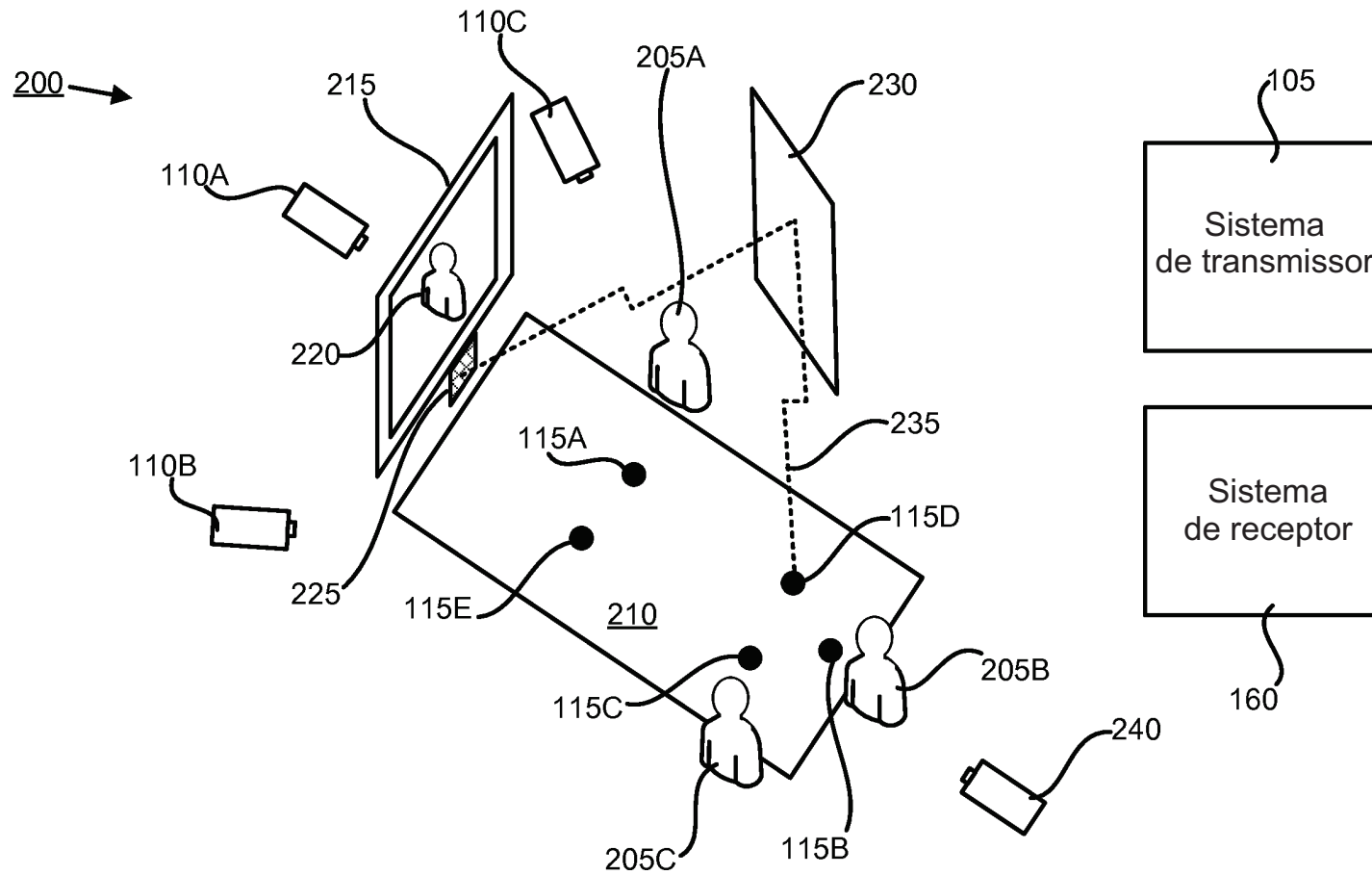


FIG. 2

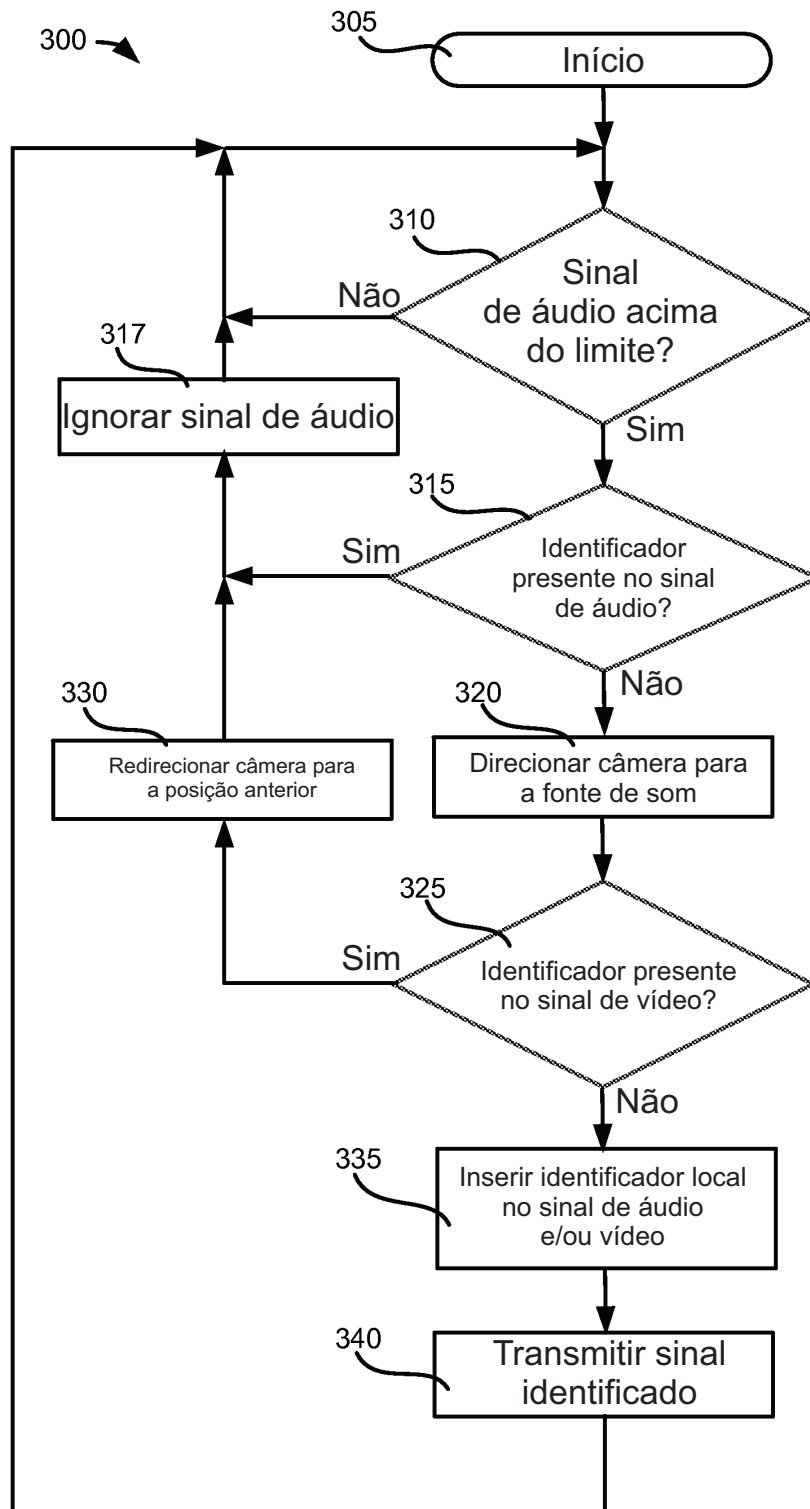
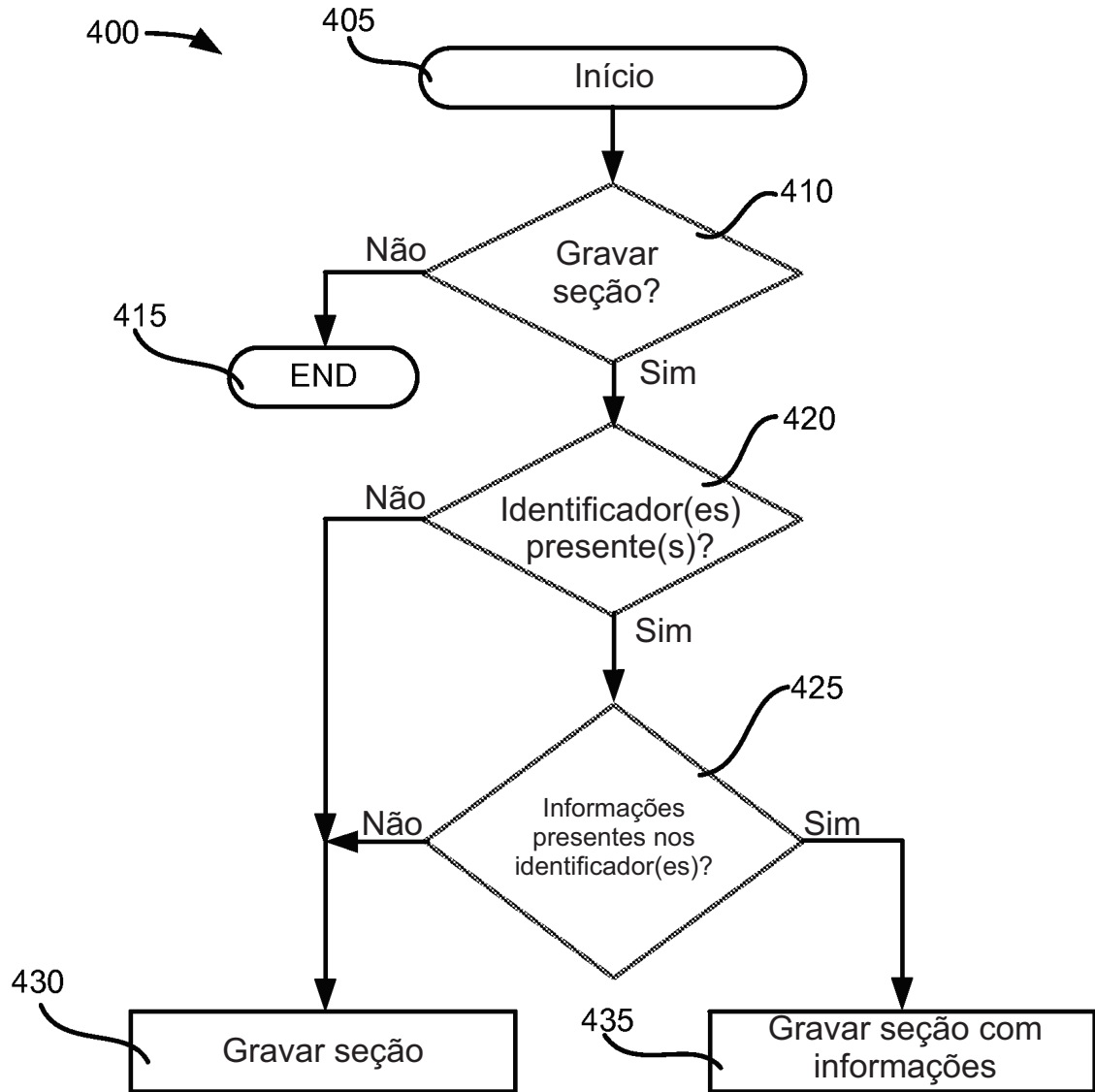


FIG. 3

**FIG. 4**

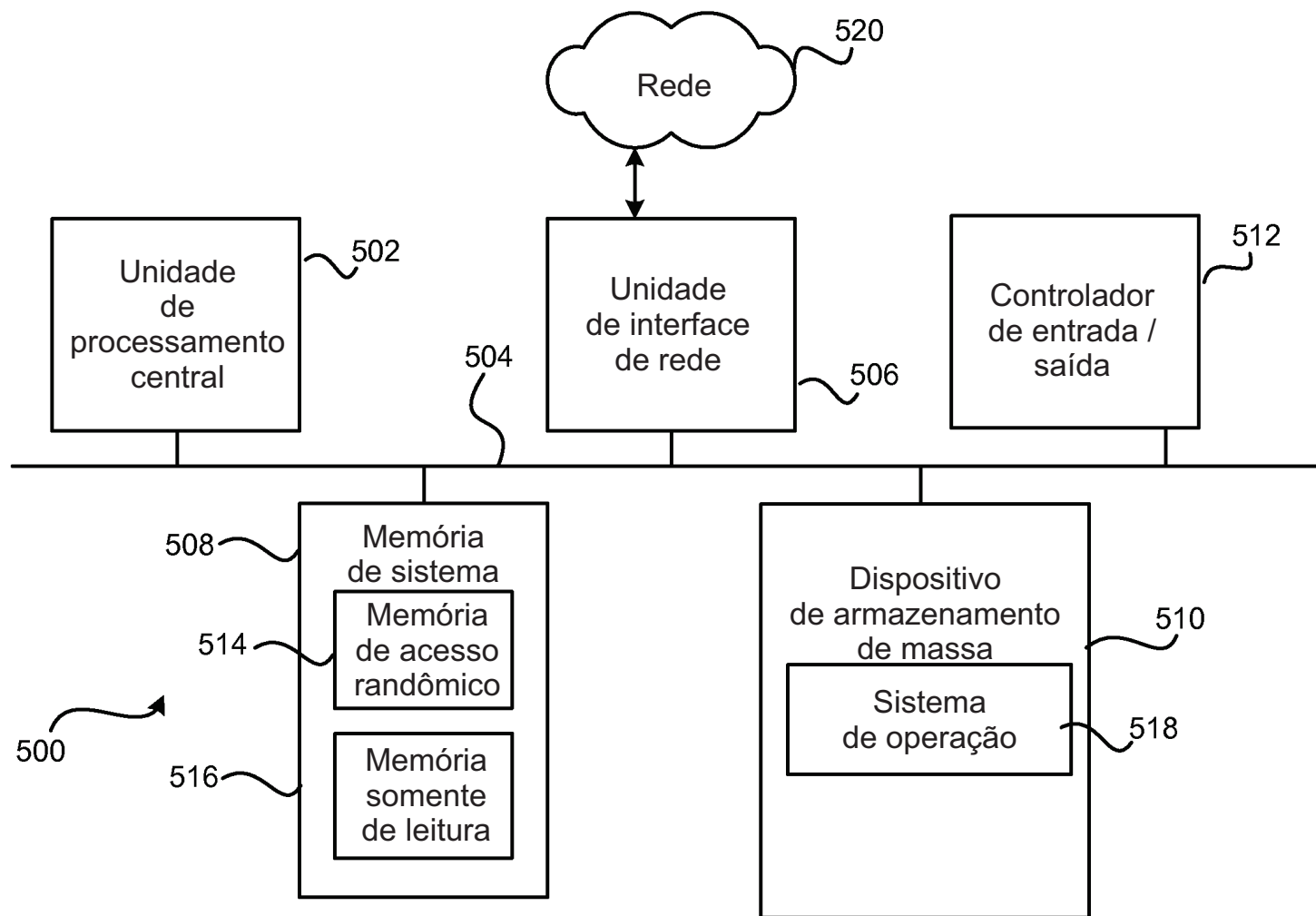


FIG. 5