



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

(11) PI 0912307-5 B1



(22) Data do Depósito: 29/04/2009

(45) Data de Concessão: 24/11/2020

(54) Título: MÉTODO DE COMUNICAÇÃO POR VÍDEO TRIDIMENSIONAL

(51) Int.Cl.: H04N 7/14; H04N 7/15; H04N 13/10; H04N 13/239; H04N 13/359; (...).

(52) CPC: H04N 7/147; H04N 7/152; H04N 13/10; H04N 13/239; H04N 13/359; (...).

(30) Prioridade Unionista: 30/05/2008 CN 2008 10098355.3.

(73) Titular(es): HUAWEI DEVICE (SHENZHEN) CO., LTD..

(72) Inventor(es): YUAN LIU.

(86) Pedido PCT: PCT CN2009071546 de 29/04/2009

(87) Publicação PCT: WO 2009/143735 de 03/12/2009

(85) Data do Início da Fase Nacional: 29/11/2010

(57) Resumo: MÉTODO, APARELHO E SISTEMA PARA COMUNICAÇÃO POR 3D. A presente invenção refere-se a tecnologias de comunicação de vídeo, e mostra um método, um aparelho e um sistema para uma comunicação de vídeo 3D. Considerando-se que a técnica anterior permite apenas uma comunicação de vídeo 3D de par a par e é incapaz de se adequar às exigências de comunicação, o método de comunicação de vídeo 3D provido aqui inclui: o recebimento de uma informação de vídeo de pelo menos duas partes de sessão, em que a informação de vídeo de pelo menos uma parte de sessão é uma informação de vídeo 3D; e o processamento da informação de vídeo recebida a partir de pelo menos duas partes de sessão de acordo com a informação de estado de pelo menos duas partes de sessão e as capacidades de exibição de um terminal. A solução técnica segundo a presente invenção é aplicável a redes de comunicação de vídeo.

Relatório Descritivo da Patente de Invenção para "**MÉTODO DE COMUNICAÇÃO POR VÍDEO TRIDIMENSIONAL**".

CAMPO DA INVENÇÃO

[0001] A presente invenção se refere a tecnologias de comunicação de vídeo e, em particular, a um método, um aparelho e um sistema para uma comunicação de vídeo tridimensional (3D).

ANTECEDENTES

[0002] A imagem de vídeo tradicional é um meio de informação bidimensional (2D), o qual apresenta apenas o conteúdo da cena, mas ignora a informação de profundidade, tal como a distância e a localização do objeto. Portanto, a imagem de vídeo 2D tradicional é incompleta. As pessoas estão acostumadas a observar coisas de forma binocular e, portanto, mais informação além de uma ilustração 2D é requerida para que as pessoas obtenham a informação espacial necessária. A tecnologia de vídeo 3D provê ilustrações as quais estão em conformidade com os princípios visuais 3D e portam uma informação de profundidade, reconstrói o objeto físico de forma autêntica, apresenta a cena com a sensação de profundidade, perspectiva e realidade, e é uma tendência da tecnologia de vídeo. O desenvolvimento da tecnologia de vídeo 3D começou nos anos 1980, mas a tecnologia não foi aplicada massivamente devido à imaturidade técnica, alto custo do dispositivo de exibição e falta de padronização. Nos últimos anos, com o aperfeiçoamento das tecnologias relevantes, a queda do custo do dispositivo e o lançamento de normas relevantes, a tecnologia de vídeo 3D é colocada em uso prático, e a comunicação de vídeo 3D se torna uma questão palpitante no campo de pesquisa de vídeo 3D.

[0003] A técnica anterior mostra uma estação de trabalho de videoconferência 3D. A estação de trabalho inclui uma câmera de vídeo dupla para a captura de sinais de vídeo e um aparelho de cálculo de

imagem de profundidade para a criação de uma imagem de profundidade. Os sinais de vídeo e a informação de profundidade são transmitidos como dados de vídeo 3D. Além disso, a estação de trabalho inclui um dispositivo de exibição 3D para a exibição de imagens 3D. O aparelho de cálculo de imagem de profundidade pode estimar a localização de um usuário, e é configurado para exercer controle para a geração de imagens 3D.

[0004] A técnica anterior também mostra um terminal de comunicação de imagem 3D para a exibição de uma informação de imagem 3D. O terminal inclui: um módulo de entrada de imagem 3D com múltiplas câmeras para a captura de imagens 3D; um módulo de exibição de imagem 3D, configurado para a exibição de uma informação de imagem 3D; e um módulo de comunicação, configurado para a transmissão de pelo menos uma informação de imagem 3D obtida pelo módulo de entrada de imagem 3D. O módulo de exibição de imagem 3D é composto por um dispositivo de exibição de paralaxe horizontal/vertical que integra tipos de formação de imagem. As câmeras são dispostas pelo menos a partir do topo para baixo e a partir da esquerda para a direita próximas do dispositivo de exibição de imagem 3D.

[0005] No processo de desenvolvimento da presente invenção, o inventor descobre pelo menos estes defeitos na técnica anterior: a comunicação de vídeo 3D é capaz de uma comunicação apenas de par a par e exibir a imagem de vídeo 3D de uma única entidade, e não é capaz de realizar uma comunicação de vídeo 3D de entidade múltipla. Portanto, com o rápido desenvolvimento da tecnologia de comunicação, a implementação da comunicação de vídeo de entidade múltipla 3D ou uma comunicação de vídeo de entidade múltipla híbrida de 2D/3D se torna uma exigência urgente.

SUMÁRIO

[0006] As modalidades da presente invenção provêm um método, um aparelho e um sistema para uma comunicação de vídeo 3D para a implementação de uma comunicação de vídeo de entidade múltipla 3D ou uma comunicação de vídeo de entidade múltipla híbrida de 2D/3D.

[0007] A solução técnica segundo a presente invenção é implementada da forma a seguir:

[0008] Um método de comunicação de vídeo 3D provido em uma modalidade da presente invenção inclui:

o recebimento de uma informação de vídeo de pelo menos duas entidades de sessão, onde a informação de vídeo de pelo menos uma entidade de sessão é uma informação de vídeo 3D; e

o processamento da informação de vídeo recebida a partir de pelo menos duas entidades de sessão e capacidades de exibição de um terminal.

[0009] Um sistema de comunicação de vídeo 3D provido em uma modalidade da presente invenção inclui:

pelo menos duas entidades de sessão, configuradas para o envio de uma informação de vídeo para um terminal de comunicação por vídeo, onde a informação de vídeo de pelo menos uma entidade de sessão é uma informação de vídeo 3D; e

um terminal de comunicação de vídeo, configurado para: receber a informação de vídeo de pelo menos entidades de sessão, e o processamento da informação de vídeo recebida a partir de pelo menos duas entidades de sessão de acordo com a informação de estado de pelo menos duas entidades de sessão e as capacidades de exibição do terminal.

[00010] Um sistema de comunicação de vídeo 3D provido em uma modalidade da presente invenção inclui:

uma unidade de controle de ponto múltiplo (MCU), configurada para: receber uma informação de vídeo de pelo menos duas

entidades de sessão, onde a informação de vídeo de pelo menos uma entidade de sessão é uma informação de vídeo 3D; processar a informação de vídeo recebida a partir de pelo menos duas entidades de sessão de acordo com a informação de estado de pelo menos duas entidades de sessão e as capacidades de exibição de um terminal, e obter a informação de vídeo processada; e

um terminal de comunicação de vídeo, configurada para a obtenção e a exibição da informação de vídeo processada.

[00011] Um aparelho de comunicação de vídeo 3D provido em uma modalidade da presente invenção inclui:

um módulo de recepção, configurado para obter uma informação de vídeo e uma informação de estado de pelo menos duas entidades de sessão, bem como capacidades de exibição de um terminal, onde a informação de vídeo de pelo menos uma entidade de sessão é uma informação de vídeo 3D; e

um módulo de processamento, configurado para processamento da informação de vídeo recebida a partir de pelo menos duas entidades de sessão de acordo com a informação de estado de pelo menos duas entidades de sessão e as capacidades de exibição do terminal.

[00012] Através do método, do aparelho e do sistema para comunicação de vídeo 3D providos em uma modalidade da presente invenção, um terminal capaz de exibir as imagens 3D recebe uma informação de vídeo de pelo menos duas entidades de sessão, e exibe a informação de vídeo recebida a partir de cada entidade de sessão de uma forma 3D ou 2D, assim se implementando uma comunicação de vídeo 3D de entidade múltipla ou uma comunicação de vídeo de entidade múltipla híbrida de 2D/3D. Portanto, o usuário pode realizar uma comunicação de vídeo 3D com múltiplas entidades de sessão simultaneamente, e obter uma melhor experiência da comunicação de

vídeo 3D.

BREVE DESCRIÇÃO DOS DESENHOS

[00013] A FIG. 1 mostra os princípios básicos de um vídeo 3D em uma modalidade da presente invenção;

a FIG. 2 mostra uma imagem de olho esquerdo e o mapa de paralaxe correspondente em uma modalidade da presente invenção;

a FIG. 3 mostra uma estrutura de rede de um sistema de comunicação de vídeo 3D em uma modalidade da presente invenção;

a FIG. 4 mostra um primeiro modo de exibição de comunicação de vídeo 3D em uma modalidade da presente invenção;

a FIG. 5 é um fluxograma de um método para um terminal para sintetização de uma informação de vídeo de todas as entidades de sessão em uma modalidade da presente invenção;

a FIG. 6 mostra o segundo modo de exibição de comunicação de vídeo 3D em uma modalidade da presente invenção;

a FIG. 7 é um fluxograma do terceiro modo de exibição de comunicação de vídeo 3D em uma modalidade da presente invenção;

a FIG. 8 mostra uma conferencia de ponto múltiplo que emprega uma MCU em uma modalidade da presente invenção;

a FIG. 9 é um fluxograma de um método para uso de uma MCU para a implementação de uma comunicação de vídeo 3D em uma modalidade da presente invenção;

a FIG. 10 mostra uma estrutura de um sistema de comunicação de vídeo 3D em uma modalidade da presente invenção; e

a FIG. 11 mostra uma estrutura de um outro sistema de comunicação de vídeo 3D em uma modalidade da presente invenção.

DESCRIÇÃO DETALHADA DAS MODALIDADES

[00014] De modo a se tornarem mais claros a solução técnica, os objetivos e os méritos da presente invenção, o que vem a seguir descreve a presente invenção em maiores detalhes com referência aos

desenhos associados e às modalidades preferidas.

[00015] A comunicação de vídeo 3D simula a visão humana, e usa câmeras para a captura de uma imagem de olho esquerdo e uma imagem de olho direito. Como resultado, quando uma pessoa assiste a um vídeo 3D, o olho esquerdo vê a imagem de olho esquerdo e o olho direito vê a imagem de olho direito. Portanto, a pessoa vê imagens estereoscópicas e percebe a profundidade da cena. As imagens de vídeo 3D são categorizadas em imagens de olho esquerdo e imagens de olho direito, e existe uma paralaxe entre os dois tipos de imagem.

[00016] Conforme mostrado na FIG. 1, um sistema de câmera dupla paralela é aplicado em uma modalidade da presente invenção. Na parte esquerda da FIG. 1, A e B são dois pontos em uma cena real, O_L e O_R são os centros óticos da câmera esquerda e da câmera direita, respectivamente, e a distância entre O_L e O_R é C. Neste caso, um sistema de câmera paralela é aplicado, especificamente, a linha que conecta O_L e O_R é paralela ao plano de projeção e ao terreno. As distâncias a partir dos centros óticos das duas câmeras até o plano de posição fechada são ambos f, especificamente, os comprimentos focais das câmeras são ambos f. A'_L e A'_R são projeções de A no plano de projeção com respeito à câmera esquerda e à câmera direita, respectivamente; B'_L e B'_R são projeções de B no plano de projeção com respeito à câmera direita, respectivamente. Conforme mostrado na parte direita da FIG. 1, suponha que exista uma câmera cujo centro ótico O_M caia no ponto médio da linha entre O_L e O_R , a projeção B no plano de projeção com respeito à câmera O_M será B_M . Neste caso, apontando O_M como a origem das coordenadas, as coordenadas espaciais de B serão (x, y, z), pela observação da projeção de B sobre o plano de projeção com respeito à câmera O_M da perspectiva do plano Y-Z, a relação geométrica mostrada na parte direita da FIG. 1 pode ser obtida. Na parte direita da FIG. 1, suponha que a linha vertical de O_M até o plano

de projeção intercepte o plano em F, a distância entre O_M e F será f, especificamente, $O_M F = f$; supondo que a linha vertical a partir de B até $O_M F$ intercepte $O_M F$ em E, $BE = y$ e $O_M E = z$. Devido ao fato de o triângulo $O_M E B$ ser similar ao triângulo $O_M F B_M$, as fórmulas a seguir são inferidas:

$$\frac{B_M F}{BE} = \frac{O_M F}{O_M E} \Rightarrow \frac{B_M F}{y} = \frac{f}{z} \Rightarrow B_M F = \frac{yf}{z} \quad (1)$$

[00017] De acordo com a fórmula (1) e a FIG. 1, o valor de comprimento de $B_M F$ é a coordenada de eixo Y do ponto de projeção B_M no plano de projeção. Da mesma forma, a coordenada de eixo X de B_M no plano de projeção pode ser deduzida. Se as coordenadas de eixo X e de eixo Y de B_M no plano de projeção forem representadas por u e v respectivamente, a fórmula a seguir se aplicará:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{xf}{z} \\ \frac{yf}{z} \end{bmatrix} \quad (2)$$

[00018] A câmera O_L se desvia da câmera O_M $-c/2$ no eixo X, e a câmera O_R se desvia da câmera O_M no eixo X $+c/2$. Portanto, para um ponto X (x, y, z), as fórmulas a seguir refletem a coordenada de eixo X u_L e a coordenada de eixo Y v_L da posição fechada de X no plano de projeção com respeito à câmera O_L , e refletem a coordenada de eixo X u_R e a coordenada de eixo Y v_R da posição fechada de X no plano de projeção com respeito à câmera O_R :

$$\begin{bmatrix} u_L \\ v_L \end{bmatrix} = \begin{bmatrix} \frac{(x + \frac{c}{2})f}{z} \\ \frac{yf}{z} \end{bmatrix} \quad \begin{bmatrix} u_R \\ v_R \end{bmatrix} = \begin{bmatrix} \frac{(x - \frac{c}{2})f}{z} \\ \frac{yf}{z} \end{bmatrix} \quad (3)$$

[00019] Devido ao sistema de câmera paralela, a distância entre um ponto em uma imagem e sua contraparte na outra imagem na direção vertical (eixo Y) pode ser julgada como 0, e a distância entre o ponto e

sua contraparte apenas existe na direção horizontal, conforme expresso abaixo:

$$\begin{cases} u_L - u_R = \frac{cf}{z} \\ v_L - v_R \approx 0 \end{cases} \quad (4)$$

[00020] Conforme pode ser visto, uma paralaxe na direção horizontal existe entre a imagem de olho esquerdo e a imagem de olho direito. A paralaxe está relacionada à profundidade do ponto de formação de imagem na cena. Portanto, a paralaxe pode ser convertida na profundidade e vice-versa.

[00021] Se comparado com o vídeo 2D, o vídeo 3D envolve mais dados, e os dados criados massivamente são difíceis de armazenar e transmitir. Portanto, os dados redundantes nos dados de vídeo 3D precisam ser eliminados tanto quanto possível. Atualmente, a redundância de dados dos dados de vídeo 3D no domínio de espaço e no domínio de tempo pode ser eliminada através de uma predição intraquadro e de uma predição interquadro, e redundância de dados de domínio de espaço dos dados de vídeo 3D precisa ser eliminada entre a imagem de olho esquerdo e a imagem de olho direito. Portanto, uma tecnologia de compensação de estimativa de paralaxe/profundidade para eliminação de redundância de domínio de espaço entre a imagem de olho esquerdo e a imagem de olho direito torna-se uma tecnologia chave na codificação de vídeo 3D. De modo similar à tecnologia de compensação de estimativa de movimento, a tecnologia de compensação de estimativa de paralaxe/profundidade usa a relevância entre duas ou mais imagens para a eliminação da redundância de dados de domínio de espaço entre as imagens. A imagem de paralaxe/profundidade é obtida, geralmente, através de uma cobertura estereoscópica para a imagem de olho esquerdo e a imagem de olho direito.

[00022] Conforme mostrado na FIG. 2, a imagem no lado esquerdo

é a imagem de olho esquerdo, e a imagem no lado direito é um mapa de paralaxe (o que usa uma informação de luminância para expressão da informação de disparidade de cada ponto de pixel na imagem). A FIG. 2 revela que a paralaxe da área de imagem mais próxima do observador é maior, e a luminância dessa área de imagem é mais alta; a paralaxe na área de imagem mais distante do observador é menor, e a luminância dessa área de imagem é mais baixa. Portanto, de acordo com a relação entre a paralaxe e a profundidade, um mapa de paralaxe pode ser facilmente convertido em um mapa de profundidade. Uma imagem de olho direito pode ser derivada a partir de uma imagem de olho esquerdo de acordo com a informação de paralaxe/profundidade (o método de conversão básico é: um pixel na imagem de olho esquerdo é adicionado a um desvio de paralaxe para a obtenção do pixel correspondente na imagem de olho direito). Através da tecnologia precedente, a transmissão da informação de dados na imagem de olho direito é evitada ou reduzida, e a eficiência de codificação é melhorada. Mais ainda, com a informação de profundidade na cena filmada, a imagem de um ponto de vista em que uma câmera virtual é colocada pode ser gerada de acordo com um modelo de formação de imagem de câmera, e o usuário pode ver a imagem da cena de um ponto de vista em que nenhuma câmera está colocada. Portanto, quando o usuário assiste a uma imagem estereoscópica através de um visor estereoscópico automático, a paralaxe nunca muda, devido à mudança da localização e a imagem estereoscópica nunca muda nos olhos do observador, e, assim, o usuário pode ver as imagens 3D apropriadas em localizações diferentes.

[00023] Conforme mostrado na FIG. 3, nesta modalidade, o sistema de comunicação de vídeo pode ser estabelecido em uma rede de domínio de pacote comutado, tal como uma rede de área local (LAN), uma rede digital de serviço integrado (ISDN) de banda estreita e uma

ISDN de banda larga ou E1, e é composto por gateways, uma MCU, terminais de comunicação de vídeo 2D, terminais de comunicação de vídeo 3D e outros dispositivos. O gateway permite uma comunicação bidirecional em tempo real com um terminal na rede de domínio de pacote comutado ou um outro gateway. A MCU é um nó em uma rede, o qual permite uma comunicação de vídeo de entidade múltipla (por exemplo, uma conferência de ponto múltiplo) para três ou mais terminais e os gateways. A MCU também conecta dois terminais para a realização de uma comunicação de vídeo de ponto a ponto, a qual então é estendida para uma comunicação de vídeo de entidade múltipla. A MCU pode incluir um controlador de ponto múltiplo (MC) e um processador de ponto múltiplo (MP). O MC provê a função de controle para a comunicação de vídeo de entidade múltipla, e pode realizar uma negociação de capacidade com os terminais e controlar os recursos de conferência. Sob o controle do MC, o MP realiza um processamento centralizado (tal como mistura e comutação) para as transmissões contínuas de áudio, vídeo e/ou dados na comunicação de vídeo de entidade múltipla. O terminal de comunicação de vídeo 2D é um videofone, um terminal de videoconferência ou um terminal de comunicação de vídeo de computador pessoal. O terminal de comunicação de vídeo 3D é capaz de processar e exibir imagens 3D.

[00024] Na estrutura de rede precedente, antes de um terminal iniciar uma sessão de comunicação de vídeo, o terminal negocia as capacidades com o par através do terminal ou do MC. Se ambas as entidades forem sistemas de comunicação de vídeo 3D, ambas as entidades poderão ver as imagens de vídeo 3D em tempo real. Se uma entidade for um terminal de vídeo 2D, o qual não atende às condições da comunicação de vídeo 3D, a comunicação de vídeo 2D será realizada entre ambas as entidades.

[00025] O terminal de comunicação de vídeo 3D em uma modalidade

da presente invenção pode receber a informação de vídeo 3D enviada por múltiplas entidades de sessão. O formato de informação de vídeo pode ser: imagem de olho esquerdo + imagem de olho direito, imagem de olho esquerdo + informação de profundidade/paralaxe, ou imagem de olho esquerdo + informação de profundidade/paralaxe + informação residual de imagem de olho direito. A informação de profundidade/paralaxe é usada porque:

a informação de profundidade/paralaxe pode ser usada para a realização de uma codificação e uma decodificação efetivas das imagens de vídeo 3D; e

a informação de profundidade/paralaxe pode ser usada para a reconstrução de outros pontos de vista virtuais de forma efetiva (especificamente, pontos de vista em que nenhuma câmera física é colocada).

[00026] Em uma modalidade da presente invenção, o terminal de comunicação de vídeo 3D pode sintetizar a informação de vídeo de todas as entidades de sessão.

[00027] Em um sistema de terminal de comunicação de vídeo 3D em uma modalidade da presente invenção, o terminal de comunicação de vídeo pode sintetizar a informação de vídeo de múltiplas entidades de sessão de acordo com o modo de exibição selecionado pelo usuário ou o modo de exibição pré-definido. Há três modos de exibição:

[00028] Modo 1: o terminal exibe múltiplas entidades de sessão simultaneamente em modos diferentes, especificamente, exibe a entidade de sessão atualmente interagindo com o terminal em um modo de exibição 3D e exibe outras entidades de sessão em um modo de exibição 2D.

[00029] Conforme mostrado na FIG. 4, de acordo com o modo 1, a cena da entidade de sessão atualmente interagindo com o terminal é exibida de uma forma 3D (a janela de entidade de sessão no lado

esquerdo superior da tela), e as outras entidades de sessão atuais (às quais o tem presta menos atenção) são exibidas em um modo 2D. No modo 1, o usuário do terminal percebe a cena como se o usuário estivesse no foco.

[00030] Nesta modalidade, o terminal precisa detectar todas as entidades de sessão para confirmar a entidade de sessão atual. A entidade atual pode ser uma entidade de sessão atualmente falando com o usuário do terminal, ou uma entidade de sessão especificada pelo usuário de terminal. Além disso, quando a entidade de sessão atual do usuário de terminal muda, é necessário trocar o modo de exibição das entidades de sessão em tempo real, especificamente, o terminal precisa determinar o modo de exibição de todas as entidades de sessão uma a uma. Por brevidade de descrição, esta modalidade toma uma entidade de sessão do terminal como um exemplo para a descrição de como o terminal confirma o modo de exibição da entidade de sessão.

[00031] Conforme mostrado na FIG. 5, o método para o terminal sintetizar uma informação de vídeo de todas as entidades de sessão nesta modalidade inclui as etapas a seguir:

[00032] Etapa 501: detectar a informação de estado da entidade de sessão cujo modo de exibição precisa ser determinado, e detectar as capacidades de exibição do terminal.

[00033] Nesta etapa, a informação de estado da entidade de sessão pode incluir uma informação indicando se a entidade de sessão está falando com o terminal, e as câmaras de exibição do terminal incluem uma informação indicando se o terminal suporta o modo de exibição 3D.

[00034] Etapa 502: julgar se a entidade de sessão é a entidade de sessão atual, e se o terminal suporta o modo de exibição 3D. Se esse for o caso, realizar a etapa 503; caso contrário, realizar a etapa 505.

[00035] Nesta etapa, os critérios para julgar a entidade de sessão podem ser pré-regulados de acordo com as condições reais. Por

exemplo, o terminal pode tomar a entidade de sessão atualmente falando com o terminal como a entidade de sessão atual de acordo com a informação de estado, ou especificar uma entidade de sessão como a entidade de sessão atual. Se o usuário de terminal especificar uma entidade de sessão como a entidade de sessão atual, a ação de detecção da informação de estado na etapa 501 pode ser omitida.

[00036] Além disso, nesta etapa, é necessário julgar se o terminal suporta o modo de exibição 3D de acordo com as capacidades de exibição do terminal.

[00037] Etapa 503: sintetizar a informação de vídeo 3D a ser exibida.

[00038] Nesta etapa, pelo uso da informação de vídeo 3D (por exemplo, imagem de olho esquerdo + imagem de profundidade/paralaxe) transmitida a partir da entidade de sessão, imagens de ponto de vista múltiplo (por exemplo, imagem de olho esquerdo e imagem de olho direito) adequadas para a posição de assistência do usuário de terminal são obtidas, ou pelo menos uma imagem de ponto de vista e um mapa de profundidade/paralaxe é obtido, e a apresentação de imagem 3D é realizada para a sintetização da informação de visão 3D para exibição.

[00039] Etapa 504: exibição da informação de vídeo 3D na janela de entidade de sessão precedente. O processo termina.

[00040] Etapa 505: sintetizar a informação de vídeo 2D.

[00041] Nesta etapa, a entidade de sessão precedente é exibida através de imagens de vídeo 2D. Portanto, é necessário apenas selecionar uma imagem de vídeo 2D (por exemplo, a imagem de olho esquerdo ou a imagem de olho direito) ou sintetizar uma imagem de vídeo 2D (por exemplo, uma imagem de ponto de vista médio sintetizada pelo uso da imagem de olho esquerdo e da imagem de olho direito) para apresentação, e a informação de vídeo 2D é gerada.

[00042] Etapa 506: exibição da informação de vídeo 2D na janela de

entidade de sessão precedente. O processo termina.

[00043] O terminal realiza o processamento precedente para cada entidade de sessão, confirma o modo de exibição de cada entidade de sessão, de modo que apenas a entidade de sessão atual seja exibida no modo 3D, e as outras entidades de sessão não atuais são exibidas no modo 2D.

[00044] Além disso, nesta modalidade, o usuário de terminal pode selecionar uma entidade de sessão diretamente, e trocar o modo de exibição de entidade múltipla para o modo de exibição de sessão única (por exemplo, um modo de exibição de tela cheia). Como resultado, a entidade de sessão selecionada é exibida no modo 3D ou 2D automaticamente, e as outras entidades de sessão não são exibidas. Neste caso, a entidade de sessão selecionada pode ser a entidade de sessão atual ou não.

[00045] Nesta modalidade, o modo 1 pode suportar dois modos de exibição:

(1) Um modo de exibição que envolve uma exibição estereoscópica única

[00046] Neste modo de exibição, uma exibição estereoscópica única exibe múltiplas entidades de sessão através de múltiplas janelas. Neste modo de exibição, cada janela corresponde de forma única a uma entidade de sessão. A entidade de sessão atual pode ser exibida através de uma janela de vídeo 3D (especificamente, uma janela para exibição de imagens de vídeo 3D), e as outras entidades de sessão podem ser exibidas através de uma janela de vídeo 2D (especificamente, uma janela para exibição de imagens de vídeo 2D). Atualmente, muitas exibições estereoscópicas suportam uma troca flexível entre o modo de exibição 2D e o modo de exibição 3D, por exemplo, a exibição estereoscópica baseada em sequência de quadro, especificamente, a exibição estereoscópica emprega o armazenamento

temporário de imagem de imagem de olho esquerdo e imagem de olho direito para implementação da exibição estereoscópica. Além disso, a imagem estereoscópica da biblioteca gráfica aberta (OpenGL) também suporta uma exibição das imagens de vídeo 3D em uma janela e uma exibição de imagens de vídeo 2D em outras janelas.

(2) Um modo de exibição que envolve múltiplas exibições

[00047] No modo de exibição, múltiplas exibições são usadas para a exibição de múltiplas entidades de sessão, e cada entidade de sessão corresponde de forma única a uma exibição. Em um modo de exibição como esse, a exibição correspondente à entidade de sessão atual exibe as imagens de vídeo 3D, e as exibições correspondentes às outras imagens de vídeo 2D de exibição de entidades de sessão.

[00048] Modo 2: o terminal exibe múltiplas entidades de sessão simultaneamente no modo de exibição de entidade múltipla, e o usuário de terminal seleciona o modo de exibição de cada entidade de sessão.

[00049] Conforme mostrado na FIG. 6, no modo 2, o usuário de terminal pode selecionar o modo de exibição de cada entidade de sessão de acordo com as condições de aplicação reais. Isto é, o usuário de terminal pode especificar o modo de exibição de cada entidade de sessão, e gravar o modo de exibição especificado de cada entidade de sessão na informação de estado da entidade de sessão. Por exemplo, o usuário de terminal especifica o modo de exibição das entidades para as sessões 1 e 4 como sendo o modo de exibição 3D, e especifica o modo de exibição das pares para as sessões 2 e 3 como sendo o modo de exibição 2D. Neste caso, o terminal pode sintetizar a informação de vídeo de todas as entidades de sessão de acordo com o modo de exibição especificado na informação de estado de cada entidade de sessão. Isto é, o terminal pode sintetizar uma ou mais transmissões contínuas de vídeo simultaneamente para a obtenção e a exibição das imagens de vídeo 3D de uma ou mais entidades de sessão. Nesta

modalidade, o terminal sintetiza a informação de vídeo de todas as entidades de sessão da forma similar às etapas 501 a 506 no modo 1 e difere apenas pelo fato de, no modo 2, as condições de julgamento na etapa 502 são mudadas para: julgar se o modo de exibição da entidade de sessão precedente é um modo de exibição 3D, ou mudadas para: julgar se a entidade de sessão precedente é uma entidade de sessão especificada.

[00050] Além disso, o usuário de terminal pode selecionar uma entidade de sessão e, então, trocar o modo de exibição de entidade múltipla para o modo de exibição de sessão única para exibição da entidade de sessão selecionada (especificamente, a exibição da entidade de sessão selecionada sozinha, sem exibição de outras entidades de sessão). O modo de exibição de sessão única pode ser um modo de exibição 3D ou 2D. Como o modo 1, o modo 2 suporta o modo de exibição que envolve uma exibição estereoscópica única e o modo de exibição que envolve múltiplas exibições. Além disso, quando da sintetização das imagens de vídeo 3D, o modo 2 controla as paralaxes das imagens 3D de cada entidade de sessão, assim se evitando uma interferência entre as imagens 3D exibidas e melhorando a experiência do usuário.

[00051] Modo 3: o terminal exibe múltiplas entidades de sessão simultaneamente no modo de exibição de entidade múltipla. Todas as entidades de sessão são exibidas no modo de exibição 2D, uma entidade de sessão pode ser selecionada para mudança do modo de exibição de entidade múltipla para o modo de exibição de sessão única (por exemplo, o modo de tela cheia), e a entidade de sessão selecionada é exibida no modo de exibição 3D.

[00052] Conforme mostrado na FIG. 7, no modo 3, para os dispositivos que não suportam simultaneamente a exibição das imagens 2D e das imagens 3D, mas suportam um modo de exibição de sessão

única (tais como óculos estereoscópicos), as ilustrações de entidade múltipla são exibidas no modo de exibição 2D. Após o usuário de terminal selecionar uma entidade de sessão (por exemplo, a entidade de sessão 1), a informação indicando que a entidade de sessão serve como uma entidade de sessão selecionada é gravada na informação de estado da entidade de sessão selecionada, e o terminal pode trocar a entidade de sessão selecionada para o modo de exibição de sessão única de tela cheia, de acordo com a informação de estado da entidade de sessão selecionada no modo de exibição 3D. Além disso, no modo 3, quando a entidade de sessão selecionada é exibida no modo de exibição de sessão única, o terminal precisa sintetizar a informação de vídeo 3D para exibição de acordo com a informação de vídeo da entidade de sessão selecionada. Quando se muda de volta para a exibição de ilustração múltipla, o terminal precisa sintetizar a informação de vídeo 2D para exibição de acordo com a informação de vídeo da entidade de sessão selecionada. O método do terminal sintetizando a informação de vídeo de todas as entidades de sessão é o mesmo que o método de sintetização no modo 2, e não é repetido aqui adicionalmente.

[00053] Através do método para comunicação de vídeo 3D provido nesta modalidade, o terminal capaz de exibir imagens 3D recebe uma informação de vídeo de pelo menos duas entidades de sessão, e exibe a informação de vídeo recebida a partir de cada entidade de sessão em uma forma 3D ou 2D, assim implementando uma comunicação de vídeo 3D de entidade múltipla ou uma comunicação de vídeo de entidade múltipla híbrida 2D/3D. Portanto, o usuário pode realizar uma comunicação de vídeo 3D com múltiplas entidades de sessão simultaneamente, e obter uma experiência melhor na comunicação de vídeo 3D.

[00054] Em uma modalidade da presente invenção, uma MCU pode

synthetizar a informação de vídeo de todas as entidades de sessão.

[00055] Nos três modos de exibição nas modalidades precedentes, o terminal de comunicação de vídeo 3D synthetiza as ilustrações de vídeo de múltiplas entidades de sessão, e, então, processa o modo de exibição das entidades de sessão. Na modalidade a seguir, uma MCU capaz de processar sessões de vídeo 3D e synthetizar imagens 3D é usada para a synthetização de ilustrações de vídeo de entidade múltipla e, então, enviar as transmissões contínuas de vídeo synthetizado para o terminal de comunicação de vídeo para exibição.

[00056] Conforme mostrado na FIG. 8, nesta modalidade, a rede de comunicação inclui os terminais A, B e C capazes de comunicação de par a par, e inclui os terminais D, E e F incapazes de comunicação de par a par. Portanto, a MCU capaz de processar sessões de vídeo 3D e a synthetização de imagens 3D pode ser usada para a implementação da conferência de ponto múltiplo, por exemplo, o terminal A realiza uma conferência de ponto múltiplo com os terminais D, E e F.

[00057] Conforme mostrado na FIG. 9, o método para a MCU implementar a comunicação de vídeo 3D nesta modalidade inclui as etapas a seguir:

[00058] Etapa 901: obtenção das capacidades de exibição do terminal de comunicação e do modo de exibição de cada entidade de sessão.

[00059] Nesta etapa, a MCU precisa obter as capacidades de exibição de cada terminal de comunicação através de um protocolo de comunicação de vídeo, e obter o modo de exibição de cada entidade de sessão de acordo com a informação de estado de cada entidade de sessão. As capacidades de exibição do terminal de comunicação incluem: se o terminal de comunicação é capaz de capturar vídeos 3D, se o terminal de comunicação suporta o modo de exibição 3D, e se o modo de exibição de todas as entidades de sessão é o modo de

exibição de entidade múltipla. A informação de estado de entidade de sessão pode incluir uma informação indicando se a entidade de sessão é a entidade de sessão atual, uma entidade de sessão especificada ou uma entidade de sessão selecionada. O modo de exibição de cada entidade de sessão pode ser obtido de acordo com essa informação.

[00060] Etapa 902: julgar se o terminal suporta o modo de exibição 3D. Se o terminal suportar o modo de exibição 3D, realizar a etapa 904; caso contrário, realizar a etapa 903.

[00061] Etapa 903: sintetizar as transmissões contínuas de vídeo 2D das múltiplas entidades de sessão, e realizar a etapa 907.

[00062] Nesta etapa, devido ao fato de o terminal não suportar o modo de exibição 3D, a MCU sintetiza a informação de vídeo de múltiplas entidades de sessão a serem exibidas pelo terminal atual em transmissões contínuas de vídeo 2D de múltiplas entidades de sessão.

[00063] Etapa 904: julgar se o modo de exibição de todas as entidades de sessão é o modo de exibição 3D. Se o modo de exibição de todas as entidades de sessão for o modo de exibição 3D, realizar a etapa 905; caso contrário, realizar a etapa 906.

[00064] Etapa 905: sintetizar as transmissões contínuas de vídeo 3D das múltiplas entidades de sessão, e realizar a etapa 907.

[00065] Nesta etapa, a MCU sintetiza as transmissões contínuas de vídeo 3D de múltiplas entidades de sessão de acordo com a informação de vídeo de todas as entidades de sessão. Por exemplo, as imagens de olho esquerdo de todas as entidades de sessão são sintetizadas em um quadro de imagem de olho esquerdo, as imagens de profundidade/paralaxe de todas as entidades de sessão são sintetizadas em um quadro de imagens de profundidade/paralaxe, e o quadro da imagem de olho esquerdo e o quadro da imagem de profundidade/paralaxe são sintetizados na transmissão contínua de vídeo 3D de múltiplas entidades de sessão.

[00066] Etapa 906: sintetizar as transmissões contínuas de vídeo 2D e 3D de múltiplas entidades de sessão.

[00067] Nesta etapa, a MCU sintetiza as transmissões contínuas de vídeo 2D e 3D de múltiplas entidades de sessão de acordo com as capacidades de exibição e o modo de exibição do terminal, bem como o modo de exibição de cada entidade de sessão. Por exemplo, a MCU sintetiza as imagens de olho esquerdo de todas as entidades de sessão em um quadro de imagem de olho esquerdo, sintetiza as imagens de profundidade/paralaxe da entidade de sessão selecionada (por exemplo, a entidade de sessão atual a qual precisa ser exibida no modo de exibição 3D, ou a entidade de sessão especificada pelo usuário de terminal, ou a entidade de sessão selecionada pelo usuário de terminal) em um quadro de imagem de profundidade/paralaxe, sintetiza o quadro da imagem de olho esquerdo e o quadro de imagem de profundidade/disparidade nas transmissões contínuas de vídeo 3D de múltiplas entidades de sessão, e sintetiza a informação de vídeo de todas as entidades de sessão não selecionadas (por exemplo, a informação de vídeo 2D ou a imagem de olho esquerdo) em transmissões contínuas de vídeo 2D de múltiplas entidades de sessão.

[00068] Etapa 907: enviar as transmissões contínuas de vídeo sintetizadas para o terminal, respectivamente.

[00069] Após a etapa 907, o terminal sintetiza as transmissões contínuas de vídeo de todas as entidades de sessão para exibição de acordo com as transmissões contínuas de vídeo sintetizadas recebidas a partir da MCU.

[00070] Especificamente, o terminal sintetiza as transmissões contínuas de vídeo de todas as entidades de sessão para exibição de acordo com as transmissões contínuas de vídeo sintetizadas recebidas. Por exemplo, quando as transmissões contínuas de vídeo sintetizadas recebidas pelo terminal são transmissões contínuas de vídeo 2D

sintetizadas, o terminal divide as transmissões contínuas de vídeo 2D sintetizadas e sintetiza a informação de vídeo 2D de cada entidade de sessão para exibição; quando as transmissões contínuas de vídeo sintetizadas recebidas pelo terminal são transmissões contínuas de vídeo 2D/3D sintetizadas, o terminal divide as transmissões contínuas de vídeo 2D/3D sintetizadas e sintetiza a informação de vídeo 2D de entidades de sessão diferentes para exibição, ou sintetiza a informação de vídeo 3D para ser exibida no modo 3D (por exemplo, combinando uma imagem de olho esquerdo nas transmissões contínuas de vídeo sintetizadas com o mapa de profundidade/paralaxe correspondente à imagem de olho esquerdo em uma imagem de olho direito correspondente à imagem de olho esquerdo e formando a informação de vídeo 3D) e exibe a informação; quando as transmissões contínuas de vídeo sintetizadas recebidas pelo terminal são transmissões contínuas de vídeo 3D sintetizadas, o terminal divide as transmissões contínuas de vídeo 3D sintetizadas e sintetiza a informação de vídeo 3D de cada entidade de sessão para exibição (por exemplo, combinando uma imagem de olho esquerdo na informação de imagem sintetizada com o mapa de profundidade/paralaxe correspondente à imagem de olho esquerdo em uma imagem de olho direito correspondente à imagem de olho esquerdo e formando a informação de vídeo 3D) e exibe a informação.

[00071] Através do método para comunicação de vídeo 3D provido nas modalidades da presente invenção, o terminal capaz de exibir as imagens 3D recebe uma informação de vídeo de pelo menos duas entidades de sessão, e exibe a informação de vídeo recebida a partir de cada entidade de sessão de uma forma 3D ou 2D, assim implementando a comunicação de vídeo 3D de entidade múltipla ou uma comunicação de vídeo de entidade múltipla híbrida de 2D/3D. Portanto, o usuário pode realizar uma comunicação de vídeo 3D com

múltiplas entidades de sessão simultaneamente, e obter uma melhor experiência na comunicação de vídeo 3D.

[00072] É compreensível para aqueles versados na técnica que todas ou parte das etapas das modalidades de método precedentes podem ser implementadas por um hardware instruído por um programa. O programa pode ser armazenado em um meio de armazenamento que pode ser lido em computador. Quando é executado, o programa realiza as etapas das modalidades de método precedentes. O meio de armazenamento pode ser qualquer meio adequando para o armazenamento de códigos de programa, por exemplo, uma memória apenas de leitura (ROM), uma memória de acesso randômico (RAM), um disco magnético ou um disco compacto.

[00073] Conforme mostrado na FIG. 10, o sistema de comunicação de vídeo 3D em uma modalidade da presente invenção inclui pelo menos dois terminais de comunicação de vídeo, onde pelo menos um terminal de comunicação de vídeo é uma informação de vídeo capaz de exibir uma informação de vídeo 3D. A FIG. 10 mostra dois terminais de comunicação de vídeo capazes de exibirem uma informação de vídeo 3D.

[00074] O sistema em uma modalidade da presente invenção inclui:
uma MCU, configurada para: receber uma informação de vídeo de pelo menos duas entidades de sessão, onde a informação de vídeo de pelo menos uma entidade de sessão é uma informação de vídeo 3D; processar a informação de vídeo recebida a partir de pelo menos duas entidades de sessão de acordo com a informação de estado de pelo menos duas entidades de sessão e as capacidades de exibição de um terminal, e obter a informação de vídeo processada; e
um terminal de comunicação de vídeo, configurada para a obtenção e a exibição da informação de vídeo processada.

[00075] Neste sistema, todos os terminais de comunicação de vídeo

podem realizar comunicações de vídeo de entidade múltipla, e o terminal de comunicação de vídeo capaz de exibir uma informação de vídeo 3D pode receber uma informação de vídeo a partir de múltiplas entidades de sessão e exibir a informação de vídeo recebida de todas as entidades de sessão simultaneamente no modo 3D ou 2D.

[00076] A MCU no sistema é configurada para sintetizar a informação de vídeo de todas as entidades de sessão, e implementar uma comunicação de vídeo de entidade múltipla entre pelo menos dois terminais.

[00077] Conforme mostrado na FIG. 10, a MCU ainda inclui:

um módulo de recepção, configurado para obter uma informação de vídeo e uma informação de estado de pelo menos duas entidades de sessão, bem como capacidades de exibição de um terminal, onde a informação de vídeo de pelo menos uma entidade de sessão é uma informação de vídeo 3D; e

um módulo de processamento, configurado para processamento da informação de vídeo recebida a partir de pelo menos duas entidades de sessão de acordo com a informação de estado de pelo menos duas entidades de sessão e as capacidades de exibição do terminal.

[00078] Conforme mostrado na FIG. 10, o módulo de processamento ainda pode incluir:

uma unidade de detecção, configurada para a detecção de um modo de exibição da entidade de sessão de acordo com a informação de estado obtida da entidade de sessão; e

uma unidade de sintetização, configurada para: a sintetização da informação de vídeo 3D da entidade de sessão de acordo com a informação de vídeo da entidade de sessão, se a unidade de detecção determinar que o modo de exibição da entidade de sessão é um modo de exibição 3D e as capacidades de exibição do terminal suportarem uma

exibição 3D; e a sintetização de uma informação de vídeo bidimensional (2D) da entidade de sessão de acordo com a informação de vídeo da entidade de sessão, se a unidade de detecção determinar que o modo de exibição da entidade de sessão é um modo de exibição 2D.

[00079] Opcionalmente, a unidade de detecção é configurada para a detecção do modo de exibição da entidade de sessão de acordo com a informação de estado obtida da entidade de sessão; a unidade de sintetização é configurada para: a sintetização das transmissões contínuas de vídeo 3D da entidade de sessão de acordo com a informação de vídeo da entidade de sessão, quando a unidade de detecção determinar que o modo de exibição da entidade de sessão é o modo de exibição 3D e as capacidades de exibição do terminal suportam a exibição 3D; e sintetizar as transmissões contínuas de vídeo 2D da entidade de sessão de acordo com a informação de vídeo da entidade de sessão, quando a unidade de detecção determinar que o modo de exibição da entidade de sessão é um modo de exibição 2D.

[00080] Além disso, em uma outra modalidade da presente invenção, o sistema de comunicação de vídeo 3D precedente não inclui a MCU precedente. Ao invés disso, o terminal de comunicação de vídeo 3D capaz de exibir uma informação de vídeo 3D sintetiza a informação de vídeo de todas as entidades de sessão.

[00081] Conforme mostrado na FIG. 11, o sistema de comunicação de vídeo 3D em uma modalidade da presente invenção inclui:

pelo menos duas entidades de sessão, configuradas para o envio de uma informação de vídeo para um terminal de comunicação por vídeo, onde a informação de vídeo de pelo menos uma entidade de sessão é uma informação de vídeo 3D; e

o terminal de comunicação de vídeo, configurado para: receber a informação de vídeo de pelo menos entidades de sessão, e o processo da informação de vídeo recebida a partir de pelo menos duas

entidades de sessão de acordo com a informação de estado de pelo menos duas entidades de sessão e as capacidades de exibição do terminal.

[00082] O terminal de comunicação de vídeo 3D na FIG. 11 ainda pode incluir:

um módulo de recepção, configurado para obter uma informação de vídeo e uma informação de estado de pelo menos duas entidades de sessão, bem como capacidades de exibição de um terminal, onde a informação de vídeo de pelo menos uma entidade de sessão é uma informação de vídeo 3D; e

um módulo de processamento, configurado para processamento da informação de vídeo recebida a partir de pelo menos duas entidades de sessão de acordo com a informação de estado de pelo menos duas entidades de sessão e as capacidades de exibição do terminal.

[00083] O módulo de processamento ainda pode incluir:

uma unidade de detecção, configurada para a detecção de um modo de exibição da entidade de sessão de acordo com a informação de estado obtida da entidade de sessão; e

uma unidade de sintetização, configurada para: a sintetização da informação de vídeo 3D da entidade de sessão de acordo com a informação de vídeo da entidade de sessão, se a unidade de detecção determinar que o modo de exibição da entidade de sessão é um modo de exibição 3D e as capacidades de exibição do terminal suportarem uma exibição 3D; e a sintetização de uma informação de vídeo 2D da entidade de sessão de acordo com a informação de vídeo da entidade de sessão, se a unidade de detecção determinar que o modo de exibição da entidade de sessão é um modo de exibição 2D.

[00084] Opcionalmente, conforme mostrado na FIG. 11, o módulo de processamento ainda pode incluir:

uma unidade de determinação, configurada para a determinação de uma entidade de sessão selecionada de acordo com uma instrução de seleção de um usuário; e

uma unidade de deslocamento, configurada para o deslocamento de um modo de exibição de entidade múltipla para um modo de exibição de sessão única e exibição da entidade de sessão selecionada determinada pela unidade de determinação no modo de exibição de sessão única, onde a entidade de sessão selecionada é exibida no modo de exibição 3D ou 2D.

[00085] Opcionalmente, a unidade de detecção é configurada para a detecção do modo de exibição da entidade de sessão de acordo com a informação de estado obtida da entidade de sessão e detectar se a entidade de sessão é a entidade de sessão selecionada; e a unidade de sintetização é configurada para: a sintetização da informação de vídeo 3D para exibição de acordo com a informação de vídeo da entidade de sessão selecionada, se a unidade de detecção determinar que a entidade de sessão é a entidade de sessão selecionada, o modo de exibição da entidade de sessão é um modo de exibição 3D, e as capacidades de exibição do terminal suportarem um modo de exibição de sessão única; e a sintetização da informação de vídeo 2D a ser exibida para um usuário de acordo com a informação de vídeo de cada entidade de sessão, se a unidade de detecção determinar que a entidade de sessão não é a entidade de sessão selecionada ou o modo de exibição da entidade de sessão é um modo de exibição 2D. Em conclusão, através do método, do aparelho e do sistema providos nas modalidades da presente invenção, é conveniente implementar uma comunicação de vídeo 3D de entidade múltipla ou uma comunicação de vídeo de entidade múltipla híbrida de 2D/3D, e o usuário pode realizar uma comunicação de vídeo 3D com múltiplas entidades de sessão simultaneamente e desfrutar melhor da experiência de comunicação de vídeo 3D.

REIVINDICAÇÕES

1. Método de comunicação por vídeo tridimensional (3D) para uso em um sistema de comunicação de vídeo de entidade múltipla, **caracterizado pelo fato de que** o sistema de comunicação de vídeo de entidade múltipla compreende uma unidade de controle de ponto múltiplo (MCU), pelo menos duas entidades de sessão, e um terminal correspondendo a uma entidade de sessão nas duas entidades de sessão, uma informação de estado de cada entidade sessão indica que o modo de exibição de cada entidade de sessão é modo de exibição 3D ou que os modos de exibição das entidades de sessão não são todos o modo de exibição 3D, as capacidades de exibição do terminal compreendem modo de exibição 2D ou modo de exibição 3D, e o método compreende as etapas de:

receber, pelo MCU, informação de vídeo das pelo menos duas entidades de sessão, em que a informação de vídeo da pelo menos uma entidade de sessão é uma informação de vídeo 3D; e

processar, pelo MCU, a informação de vídeo recebida a partir das pelo menos duas entidades de sessão de acordo com a informação de estado das pelo menos duas entidades de sessão e capacidades de exibição de um terminal,

em que a etapa de processar a informação de vídeo recebida a partir das pelo menos duas entidades de sessão de acordo com a informação de estado das pelo menos duas entidades de sessão e as capacidades de exibição do terminal compreende:

sintetizar (903), pelo MCU, transmissões contínuas de vídeo bidimensionais (2D) de cada entidade de sessão de acordo com a informação de vídeo de cada entidade de sessão se as capacidades de exibição do terminal não suportam exibição 3D;

sintetizar (905), pelo MCU, transmissões contínuas de vídeo 3D de cada entidade de sessão de acordo com a informação de vídeo

de cada entidade de sessão se as capacidades de exibição do terminal suportam a exibição 3D e a informação de estado de cada entidade de sessão indica que um modo de exibição de cada entidade de sessão é um modo de exibição 3D;

sintetizar (906), pelo MCU, as transmissões contínuas de vídeo 2D e 3D de cada entidade de sessão de acordo com a informação de vídeo de cada entidade de sessão e o modo de exibição de cada entidade de sessão se as capacidades de exibição do terminal suportam a exibição 3D e a informação de estado de cada entidade de sessão indica que os modos de exibição das entidades de sessão não são todas do modo de exibição 3D; e

enviar (907) as transmissões contínuas de vídeo sintetizadas para o terminal.

2. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a etapa de sintetizar (905) as transmissões contínuas de vídeo 3D de cada entidade de sessão de acordo com a informação de vídeo de cada entidade de sessão compreende:

sintetizar imagens de ponto de vista único de um mesmo tipo na informação de vídeo de todas as entidades de sessão em um quadro de uma imagem de ponto de vista único, sintetizar imagens de profundidade/paralaxe na informação de vídeo de todas as entidades de sessão em um quadro de imagem de profundidade/paralaxe, e sintetizar o quadro de uma imagem de ponto de vista único e o quadro de imagem de profundidade/paralaxe nas transmissões contínuas de vídeo 3D.

3. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a etapa de sintetizar (906) as transmissões contínuas de vídeo 2D e 3D de cada entidade de sessão de acordo com a informação de vídeo de cada entidade de sessão e o modo de exibição de cada entidade de sessão compreende:

sintetizar imagens de ponto de vista única de um mesmo tipo na informação de vídeo de todas as entidades de sessão em um quadro de imagem de ponto de vista único, sintetizar imagens de profundidade/paralaxe das entidades de sessão exibidas em um modo 3D em um quadro de imagem de profundidade/paralaxe, e sintetizar o quadro de imagem de ponto de vista único e o quadro de imagem de profundidade/paralaxe nas transmissões contínuas de vídeo 3D; e

sintetizar a informação de vídeo de todas as entidades de sessão exibidas em um modo 2D para as transmissões contínuas de vídeo de 2D.

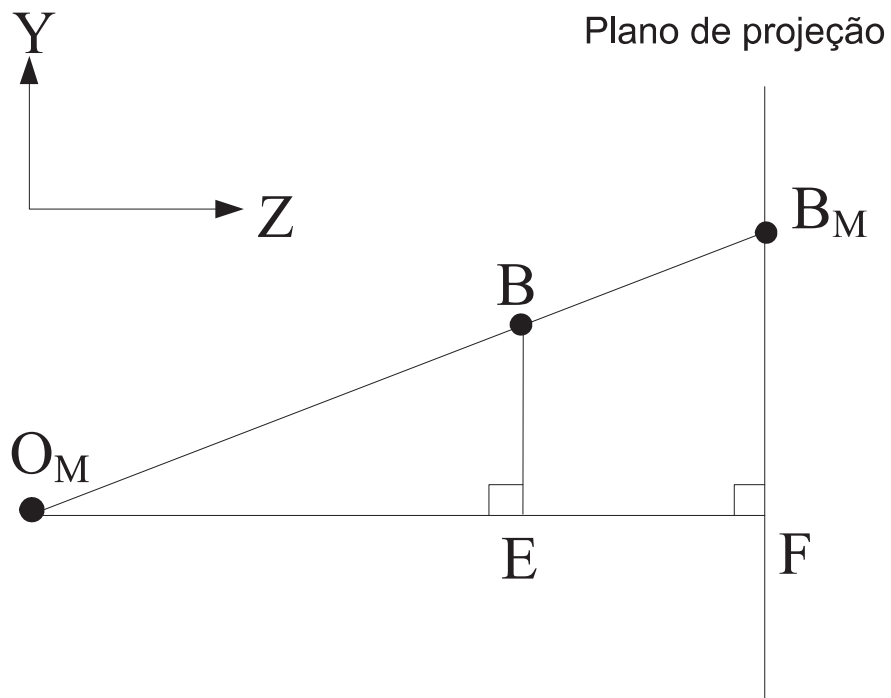
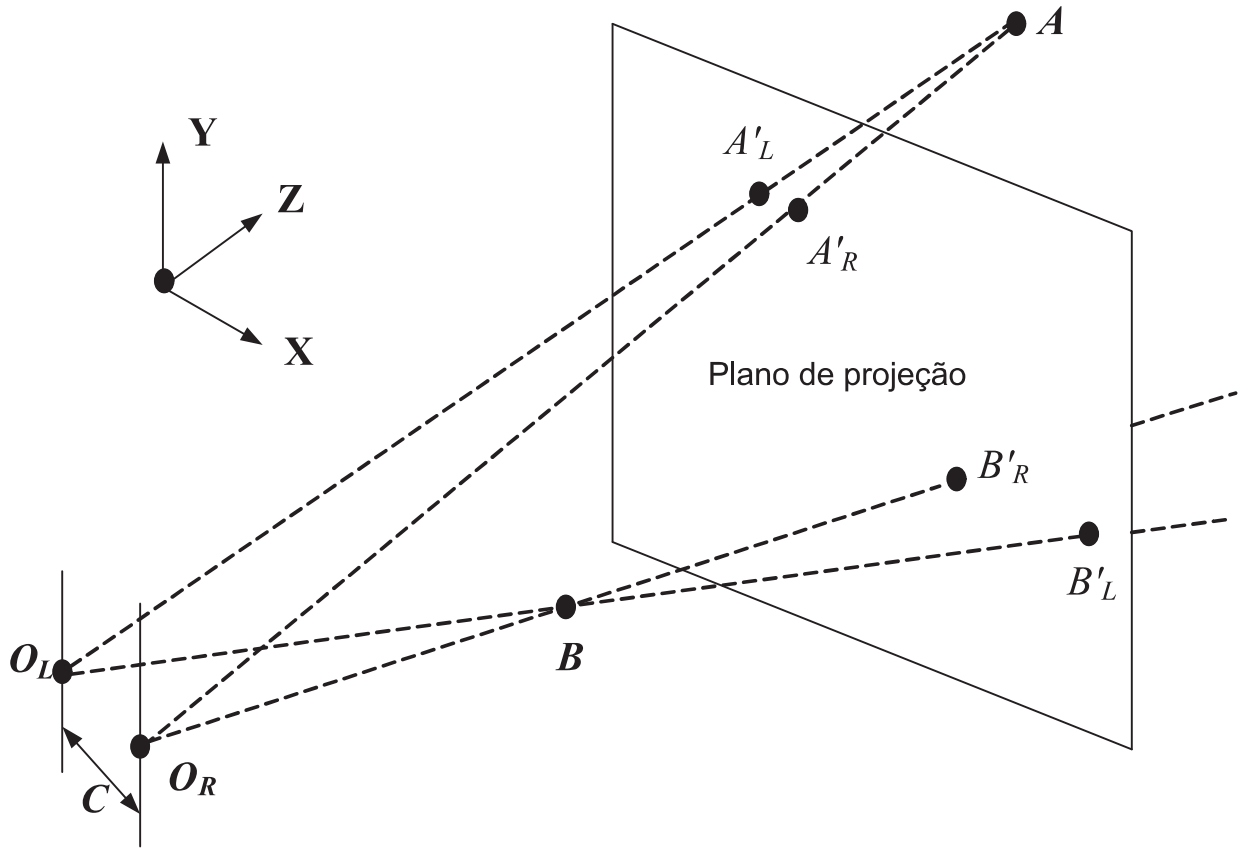


FIG. 1



FIG. 2

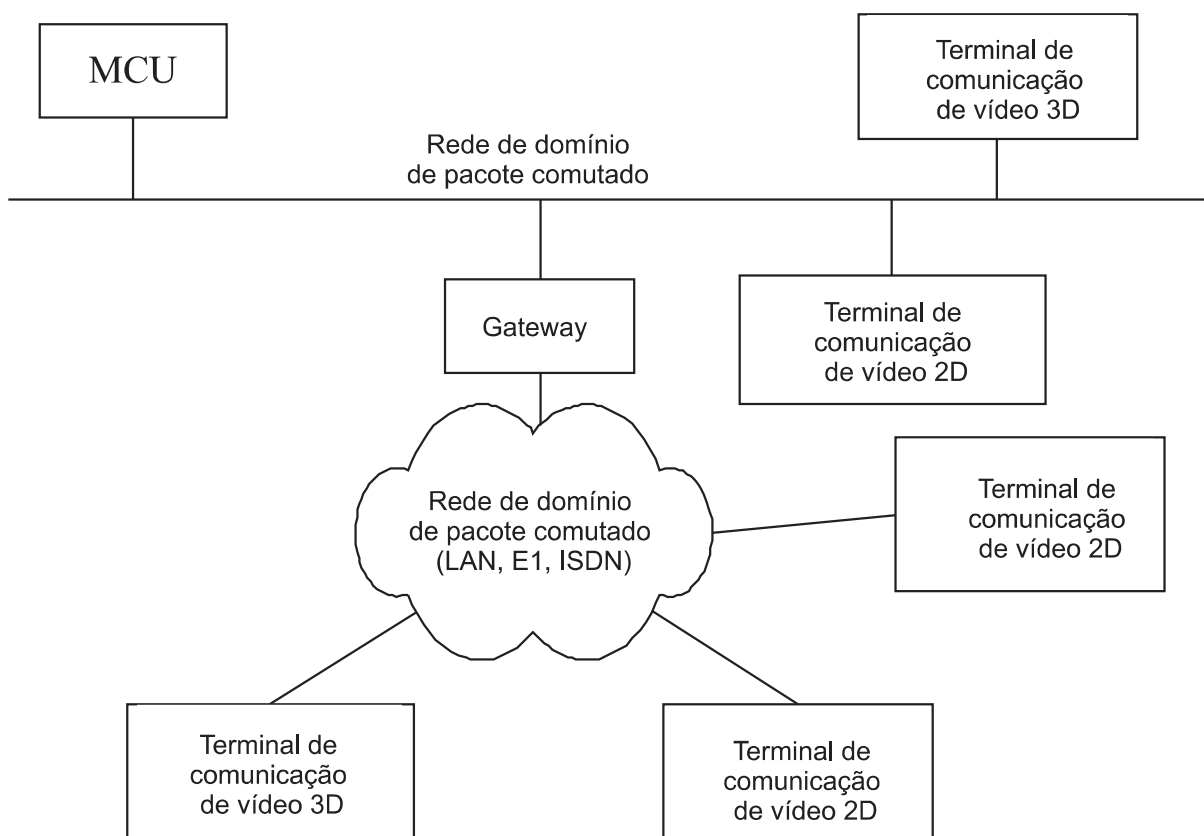


FIG. 3

Imagem de vídeo 3D (entidade de sessão atual)	Imagem de vídeo 2D
Imagem de vídeo 2D	Imagem de vídeo 2D

FIG. 4

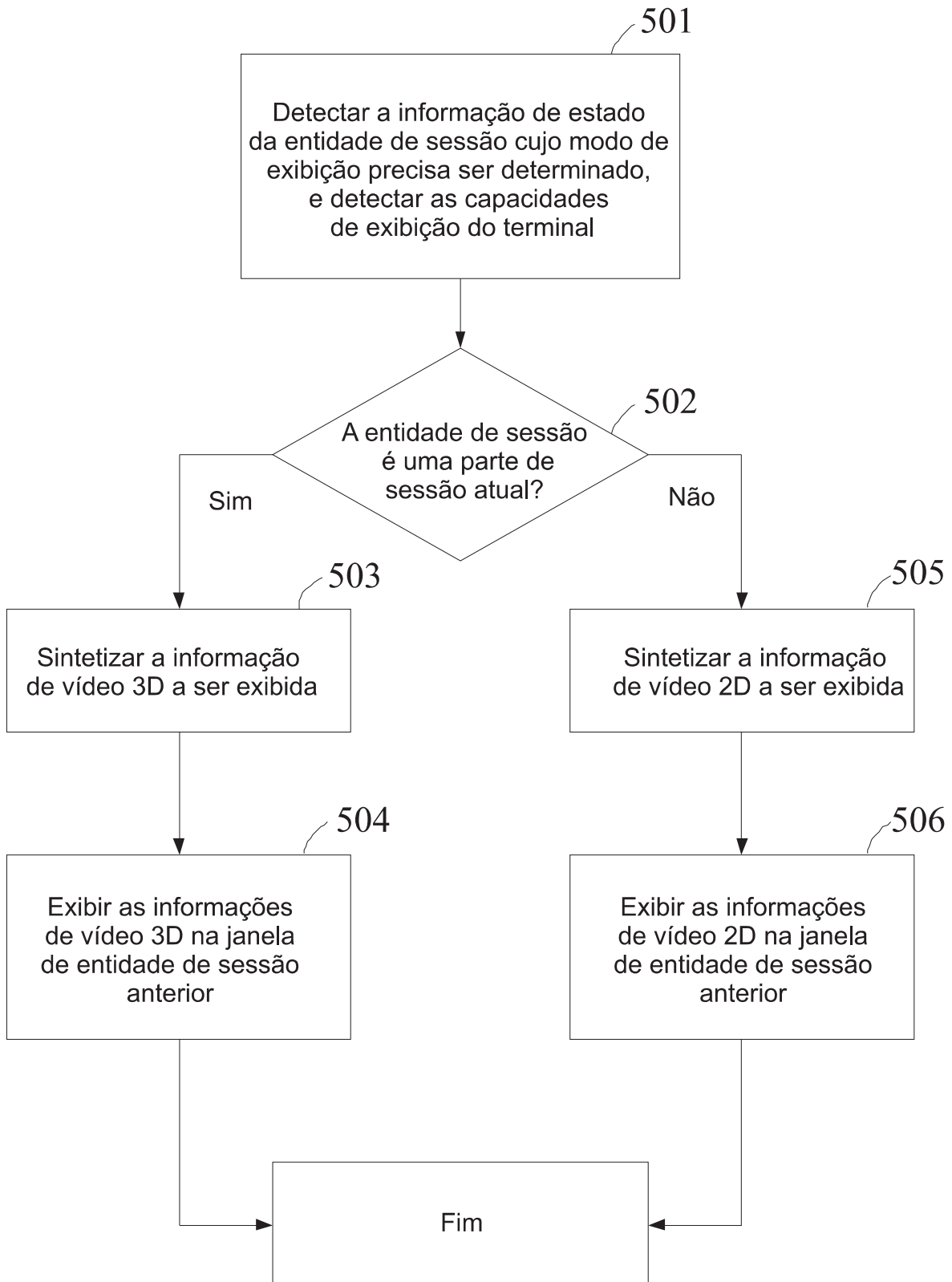


FIG. 5

Imagem de vídeo 3D (entidade de sessão 1)	Imagem de vídeo 2D (entidade de sessão 2)
Imagem de vídeo 2D (entidade de sessão 3)	Imagem de vídeo 3D (entidade de sessão 4)

FIG. 6

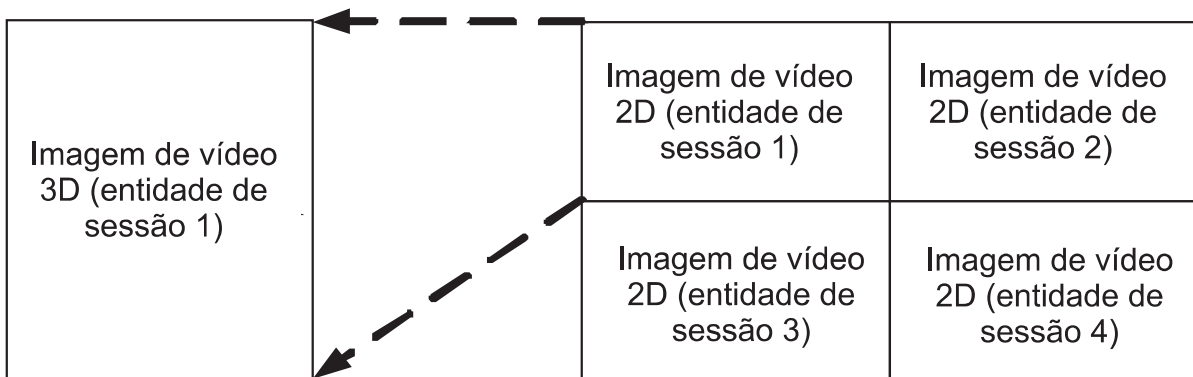


FIG. 7

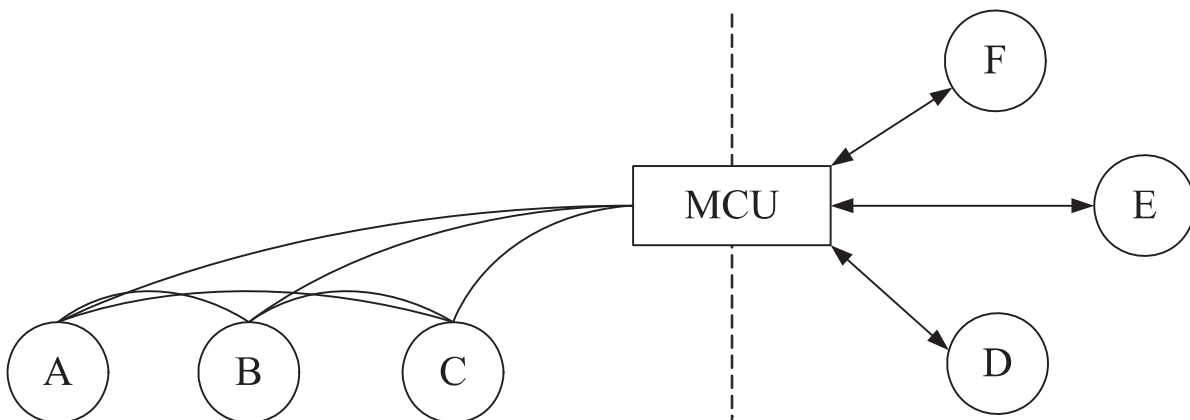


FIG. 8

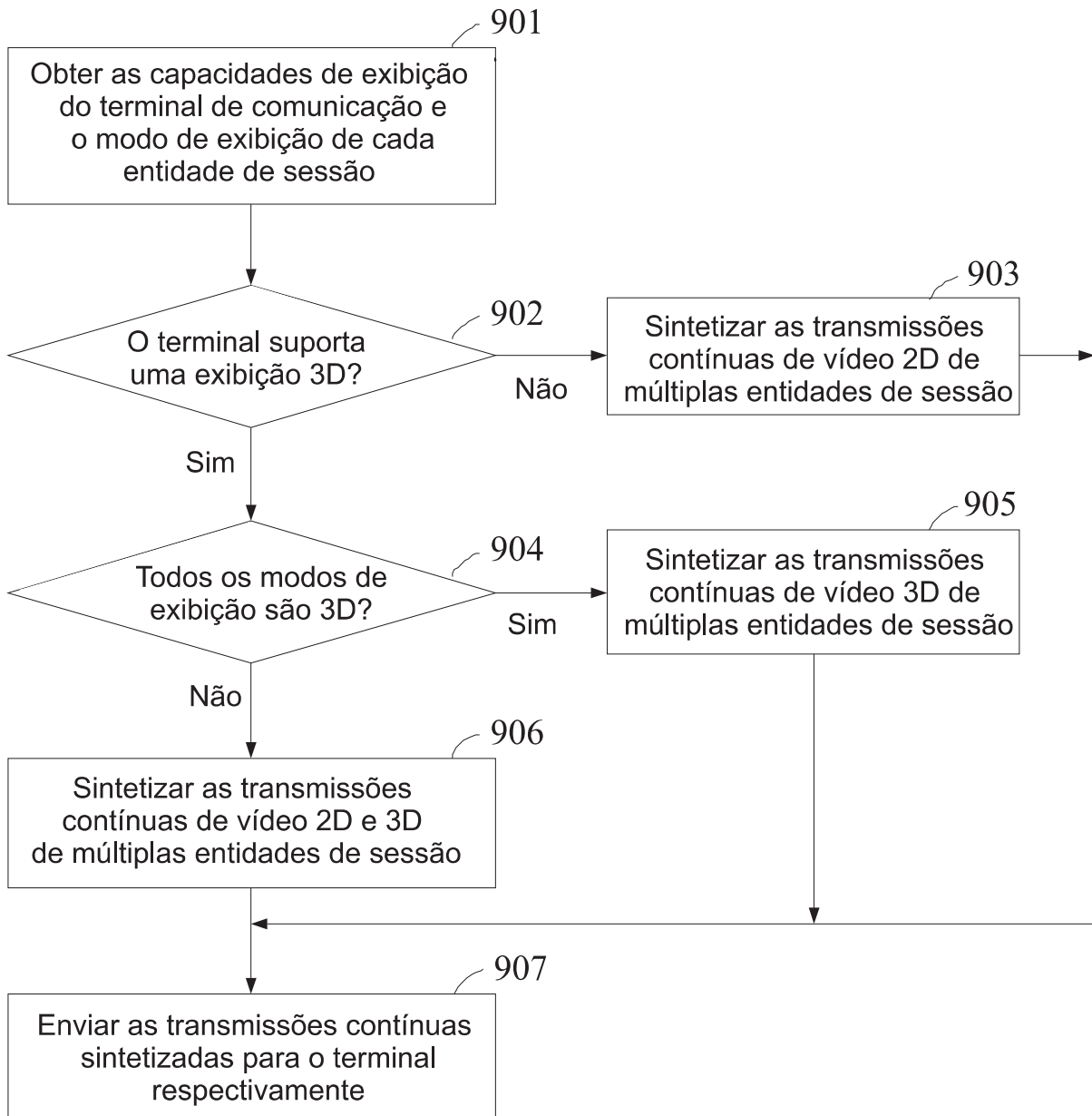


FIG. 9

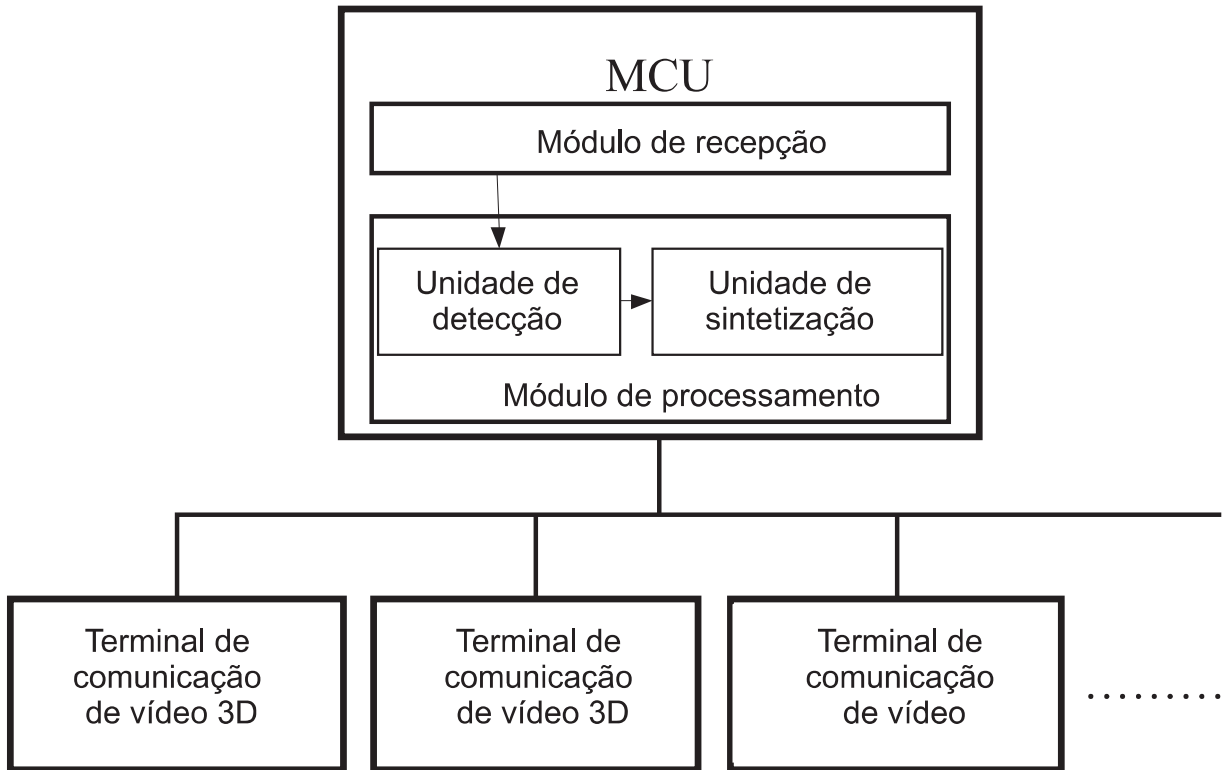


FIG. 10

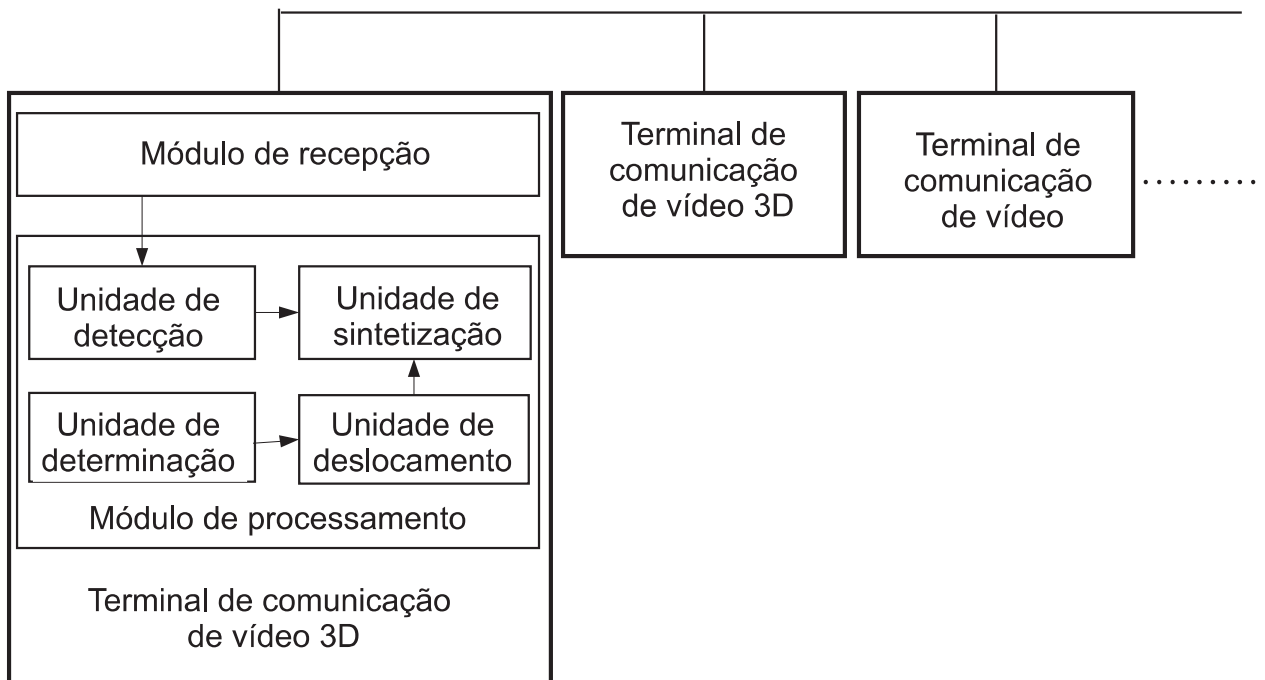


FIG. 11