



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

**(11) PI 0718711-4 B1**



**(22) Data do Depósito: 08/11/2007**

**(45) Data de Concessão: 24/09/2019**

---

**(54) Título:** MÉTODO DE CRIAR UMA CENA DE ÁUDIO, DISPOSITIVO DE COMPUTAÇÃO DE USUÁRIO E SISTEMA DISPOSTO PARA CRIAR UMA CENA DE ÁUDIO PARA UM MEIO VIRTUAL

**(51) Int.Cl.:** H04R 5/00.

**(30) Prioridade Unionista:** 08/11/2006 AU 2006906213.

**(73) Titular(es):** DOLBY LABORATORIES LICENSING CORPORATION.

**(72) Inventor(es):** PAUL BOUSTEAD; FARZAD SAFAEI.

**(86) Pedido PCT:** PCT AU2007001712 de 08/11/2007

**(87) Publicação PCT:** WO 2008/055305 de 15/05/2008

**(85) Data do Início da Fase Nacional:** 08/05/2009

**(57) Resumo:** MÉTODO DE CRIAR UMA CENA DE ÁUDIO, DISPOSITIVO DE COMPUTAÇÃO DE USUÁRIO E SISTEMA DISPOSTO PARA CRIAR UMA CENA DE ÁUDIO PARA UM MEIO VIRTUAL A presente invenção refere-se a uma cena de áudio que é criada para um avatar em um meio virtual de múltiplos avatares. Uma estrutura de ligação é criada entre os avatares. Uma cena de áudio é criada para cada avatar, com base nas associações de avatar com outros avatares.

Relatório Descritivo da Patente de Invenção para "**MÉTODO DE CRIAR UMA CENA DE ÁUDIO, DISPOSITIVO DE COMPUTAÇÃO DE USUÁRIO E SISTEMA DISPOSTO PARA CRIAR UMA CENA DE ÁUDIO PARA UM MEIO VIRTUAL**".

Campo da Invenção

[001] A presente invenção refere-se geralmente a um campo de comunicações de áudio imersivas e mais particularmente, porém por nenhum meio exclusivamente, a criação de uma cena de áudio imersiva em um meio ambiente de ponto a ponto.

Antecedente da Invenção

[002] Têm havido significativos avanços na criação de meios virtuais visualmente imersivos nos recentes anos. Estes avanços têm resultado na captação de ampla difusão de jogos que massivamente desempenham papel de múltiplo jogadores em que os participantes podem entrar em um meio virtual comum (tal como um campo de batalha) e são representados no meio virtual por um avatar, que está tipicamente na forma de um caractere animado.

[003] A captação difundida dos meios virtuais visualmente imersivos é devida, em parte, aos avanços significativos tanto na tecnologia de processamento de imagem que possibilita gráficos altamente detalhados e realísticos a serem gerados no meio virtual como no desenvolvimento de cartões de sons tridimensionais que empregam unidades de processamento de alta velocidade. Uma desvantagem principal com estes meios, todavia, é que os mecanismos de comunicação entre jogadores correntes são usualmente primitivos envolvendo trocas de palavra do texto ou comunicações de voz de "walkie-talkie". Engastar um ambiente de comunicação mais natural em que as vozes aparecem para vir dos avatares no mundo virtual correspondente ao jogador é complexo para implementar e caro para remessa. O áudio gerado por cada jogador que participa do meio virtual deve ser enviado a

cada um e a um outro jogador que está dentro da faixa da percepção auditiva. Para jogos massivamente de múltiplo jogadores, ambas as exigências da largura de banda a montante e a jusante para facilitar tal troca de áudio podem ser particularmente altas.

[004] Além do mais, o custo da CPU requerido para sintetizar todos os fluxos de áudio recebidos é alto, requerendo particularmente unidades de processamento poderosas a serem empregadas a fim de satisfazer as mínimas exigências de hardware para participar do meio virtual.

#### Definições

[005] A seguir, as definições são proporcionadas aos vários termos usados ao longo deste relatório descritivo:

[006] Cena de áudio – informação de áudio compreendendo sons combinados (por exemplo, vozes pertencentes a outros avatares e outra fonte de tempo real de som dentro do meio virtual) que são colocados de modo espacial e opcionalmente atenuados de acordo com uma distância entre uma fonte de recipiente do som. Uma cena de áudio pode também compreender efeitos de som que representam as características acústicas do meio.

#### Sumário da Invenção

[007] Em um primeiro aspecto da presente invenção, é provido um processo de criar uma cena de áudio para um avatar em um meio virtual incluindo uma pluralidade de avatares, o processo compreendendo as etapas de criar uma estrutura de ligação entre uma pluralidade de avatares; e sintetizar uma cena de áudio para cada avatar com base na sua associação com outros avatares.

[008] Vantajosamente, os aspectos da presente invenção proporcionam uma técnica de largura da banda menor para enviar cenas de áudio imersivo pela utilização de uma estrutura do tipo ponto a ponto. Enviar o serviço em tal maneira elimina a exigência de enviar tráfego

de tempo real através de um servidor central. Isto leva a significativa economia de custo quando grandes custos de parques de servidor e da largura da banda são comumente requeridos para enviar tal serviço. Ao invés, de acordo com os aspectos da presente invenção, o serviço pode ser enviado usando recursos da CPU e largura de banda dos pontos (avatars ligados) usando o serviço. Em adição, o envio de serviço pode facilmente ser escalado quando o número de avatares no meio virtual no meio ambiente cresce porque cada avatar novamente ligado adiciona recursos para suportar o serviço.

[0009] De acordo com um segundo aspecto, é provido um programa de computador compreendendo pelo menos uma instrução para controlar um computador para implementar um processo de acordo com o primeiro aspecto da invenção.

[0010] De acordo com um terceiro aspecto, é provido um meio legível de comutador proporcionando um programa de computador de acordo com o segundo aspecto da invenção.

[0011] De acordo com um quarto aspecto, é provido um dispositivo de computação de usuário disposto para realizar as etapas de processo de acordo com o primeiro aspecto da invenção.

[0012] De acordo com um quinto aspecto, é provido um sistema disposto a criar uma cena de áudio para um meio virtual, o sistema compreendendo: uma pluralidade de dispositivos de computação, cada dispositivo de computação sendo capaz de controlar pelo menos um avatar no meio virtual, em que cada dispositivo de computação é disposto para sintetizar uma cena de áudio de saída para pelo menos um avatar e comunica a cena de áudio de saída a pelo menos um outro dispositivo de computação.

#### Breve Descrição dos Desenhos

[0013] Não obstante quaisquer outras formas que podem cair dentro do escopo da presente invenção, uma concretização da presente

invenção será, a seguir, descrita, por meio do exemplo apenas, com referência aos desenhos anexos em que:

Figura 1 é um diagrama em bloco de um sistema disposto para realizar uma concretização da presente invenção;

Figura 2 mostra um exemplo de um layout de figuras avatares em um meio virtual;

Figura 3 mostra os ângulos e níveis de atenuação requeridos para sintetizar cada um dos fluxos de áudio recebidos por um avatar do meio virtual da Figura 2;

Figura 4 é um exemplo de uma conexão delta ligando os avatares ao meio virtual da Figura 2;

Figura 5 representa a aplicação de uma árvore de extensão mínima para determinar as ligações mais curtas entre os avatares da Figura 2; e

Figura 6 mostra uma borda sendo adicionada a árvore de extensão mínima, de acordo com uma concretização da presente invenção.

#### Descrição Detalhada das Concretizações Preferidas

[0014] Com referência à Figura 1, o sistema 100 inclui um servidor de meio virtual 102; servidor de controle 103; dispositivos de computação do usuário 104; e um sistema de comunicação 106.

[0015] A função primária do servidor de meio virtual 102 é manter a informação do estado para um meio virtual. Na presente concretização da invenção, o meio virtual é um campo de batalha de um jogo online de múltiplo jogadores e os avatares representam participantes (que são usuários dos dispositivos de computação do usuário 104) no meio virtual e estão na forma de soldados animados. A informação do estado, mantida pelo servidor de meio virtual 102 compreende, por exemplo, a posição dos avatares no meio virtual; isto é, o local dos soldados no campo da batalha.

[0016] Nota-se que a concretização não está restrita aos meios virtuais e avatares para jogos on-line de múltiplo jogadores. A concretização tem aplicação à uma faixa de meios virtuais incluindo, por exemplo, meios virtuais em um contexto de negócio (tais como reunião de equipes virtual) ou um contexto educacional (tal como uma aula virtual).

[0017] Para realizar as funções primárias de manutenção da informação do estado, o servidor de meio virtual 102 compreende hardware de computador incluindo uma placa-mãe, unidades de processamento central, memória de acesso aleatório, discos rígidos, hardware de rede e uma fonte de força. Em adição a hardware, o servidor de meio virtual 1-2 inclui um sistema de operação (tal como Linux, que pode ser obtido da Internet em website localizado em URL <http://www.redhat.com>) que reside no disco rígido e que coopera com o hardware para prover um meio no qual as aplicações de software podem ser executadas. Neste aspecto, o disco rígido do servidor de meio virtual 102 é carregado com uma aplicação do servidor de meio virtual (tal como motor Quake que pode ser obtido da Internet em um website localizado em URL <http://www.idsoftware.com>) para manter a informação do estado.

[0018] Um servidor de controle 103 é conectado ao servidor de meio virtual 102, via ligação de alta velocidade 105. O servidor de controle 103 incorpora o mesmo hardware que do servidor de meio virtual e é carregado com uma aplicação do servidor de controle que é disposta para interagir com o servidor de meio virtual 102 para obter informação que identifica os vários avatares presentes no meio virtual e o local dos avatares no meio virtual. Esta informação pode também incluir detalhes dos status dos avatares (por exemplo, ativos ou inativos) e detalhes de quaisquer barreiras de som dinâmicas. Usando algoritmos embutidos dentro da aplicação do servidor de controle, o ser-

vidor de controle 103 gera informação de sintetização de áudio imersivo que é comunicada a cada um dos dispositivos de computação do usuário. Como representado na Figura I, o servidor de controle 102 é também disposto para comunicar com os dispositivos de computação de usuário, via ligação 114.

[0019] Os dispositivos de computação de usuário 104 estão na forma de laptop ou desktop. Todavia, será facilmente apreciado que a concretização não está restrita aos dispositivos de comunicação de laptop ou desktop. É considerado que nas concretizações alternativas da presente invenção, os dispositivos de computação do usuário 104 (por exemplo, tal como dispositivos 04a-h, como mostrado) podiam ser dispositivos de comunicação sem fio portáteis tais como Nokia N-Gage e Playstation Portátil. Cada dispositivo de computação do usuário 104 compreende hardware de computador incluindo uma placa-mãe, unidade de processamento central, memória de acesso aleatório, um disco rígido ou dispositivo de armazenagem similar, fonte de força, monitor e uma entrada de informação do usuário (por exemplo, um teclado). Em adição ao hardware, o disco rígido de cada dispositivo de computação do usuário 104 é carregado com um sistema operativo capaz de interagir com o hardware do dispositivo de computação 104 para prover um meio em que as aplicações de software podem ser executadas. Neste aspecto, o disco rígido de cada dispositivo de computação do usuário 104 é carregado com uma aplicação do cliente de meio virtual e uma aplicação de cliente de comunicação de áudio imersivo.

[0020] A aplicação do cliente de meio virtual é disposta para enviar e receber a informação de estado para o meio virtual a e da aplicação do servidor de meio virtual carregada no servidor de meio virtual 102. A aplicação do cliente de comunicação de áudio imersivo é disposta para enviar e receber informação de áudio a e de outros clientes de

comunicação de áudio imersivo. Foi descrito previamente que cada dispositivo de computação do usuário 104 é carregado com um sistema operativo. A concretização pode ser facilmente disposta para operar quaisquer sistemas de operação diferentes carregados nos dispositivos de computação de usuário 104 incluindo, por exemplo, Microsoft Windows XP ou Linux (ambos seriam tipicamente usados quando os dispositivos de comutação 104 estivessem na forma de um computador desktop).

[0021] O sistema de comunicação 106 possibilita a aplicação do cliente de meio virtual de cada um dos dispositivos de computação de usuário 104 e a aplicação do servidor de meio virtual do sistema de meio virtual 102 para trocar dados (mais especificamente, a informação do estado) entre si. O sistema de comunicação 106 também possibilita a aplicação do cliente de comunicação de áudio imersivo de cada um dos dispositivos de computação de usuário 104 e servidor de controle 103 a trocar dados (mais especificamente, os detalhes de uma estrutura de ligação na forma de um gráfico de ponto a ponto) entre si.

[0022] Para suportar a troca de dados, o sistema de comunicação 106 inclui uma rede de dados 110 na forma da Internet para enviar e receber dados dos dispositivos de computação do usuário 104. A presente invenção não está restrita a ser usada com Internet e uma concretização alternativa da presente invenção pode ser empregada, por exemplo, 802.11 com base na rede sem fio ou similar. Para possibilitar o sistema de meio virtual 102 e os dispositivos de computação do usuário 104 para troca de dados via o sistema de comunicação 106, o servidor de meio virtual 102 é conectado à Internet 110 via uma ligação de comunicação 114 na forma de uma ligação de dados de alta velocidade.

[0023] A concretização da presente invenção primariamente diz

respeito às técnicas de ponto a ponto da largura da banda menor para prover o áudio imersivo aos usuários dos dispositivos de computação do usuário 104 que participam do meio virtual. Como tal, a seguinte descrição focaliza particularmente a funcionalidade da aplicação do servidor de controle carregada no servidor de controle 103 e da aplicação do cliente de comunicação de áudio imersivo carregada em cada um dos dispositivos de computação do usuário 104.

[0024] Como previamente discutido, a aplicação do servidor de controle residente no servidor de controle 103 é disposta para obter a informação relativa à localização dos avatares operando no meio virtual proveniente da aplicação do servidor de meio virtual. De acordo com a concretização descrita aqui, a informação da localização está na forma de coordenadas tridimensionais  $(x,y,z)$ . A aplicação do servidor de controle é adicionalmente configurada para obter detalhes de quaisquer barreiras de som dinâmicas (a seguir, informação da barreira) que podem efetuar a propagação de sons dentro do meio virtual. Uma vez que a informação de localização e informação de barreira têm sido obtidas, a aplicação do servidor de controle realiza as ações a seguir apresentadas:

calcular uma árvore ou uma série de árvores (se existirem inúmeros agrupamentos desiguais de avatares que não estão dentro da faixa auditiva entre si), com base na localização  $(x,y,z)$  de avatares no meio virtual. Estas árvores conectam avatares falantes aos avatares ouvintes. A aplicação do servidor de controle pode também levar em conta os fatores tais como largura da banda disponível e confiabilidade dos nós (isto é, correspondentes a cada um dos dispositivos de computação do usuário) para determinar como a árvore é construída. Por exemplo, se um nó na árvore não tiver uma largura de banda maior disponível, a aplicação pode restringir o número de outros nós na árvore que ela possa conectá-la. Se um nó não for confiável (por

exemplo, se o nó tiver uma história de grande retardo ou perda de pacote) então a aplicação de servidor de controle pode desejar tornar seguro que seja uma folha da árvore;

introduzir loops na árvore para reduzir a extensão do trajeto entre cada falante e cada ouvinte para formar um gráfico de ponto a ponto. A aplicação do servidor de controle não introduzirá loops que são mais curtos do que a extensão especificada para assegurar que não estão presentes nem a realimentação positiva nem ecos notáveis no fluxo de áudio eventualmente emitidos pelos dispositivos de computação do usuário; e

comunicar os seguintes detalhes associados com o gráfico de ponto a ponto aos dispositivos de computação do usuário 104 identificados no gráfico;

o conjunto de avatares que um dispositivo de comutação do usuário "A" deve criar conexões para o mesmo;

um valor de atenuação  $\alpha$  (alfa corresponde à extensão de uma ligação de gráfico) para aplicar aos fluxos de áudio que A enviará em cada uma de tais conexões; e

um ângulo  $\theta$  ( $\theta$  corresponde ao ângulo de uma ligação de gráfico) que os fluxos de áudio recebidos devem ser sintetizados.

[0025] Em relação à função da criação das estruturas de árvore (ver ponto 1 acima), qualquer estrutura de árvore apropriada podia ser empregada. Todavia, no caso da presente concretização, uma árvore de extensão mínima (MST) foi empregada. (É mínima com relação a uma métrica de custo associada com a extensão das ligações na árvore, que é proporcional à distância entre os avatares no meio virtual.)

[0026] Existem duas funções principais realizadas pela aplicação do cliente de áudio imersivo residindo em cada um dos dispositivos de computação do usuário 104. A primeira função é sintetizar os fluxos de áudio recebidos para criar uma cena de áudio imersivo para "playback"

para um usuário do dispositivo de computação do usuário 104. Este processo basicamente envolve o dispositivo de computação do usuário 104 recebendo um fluxo de áudio em cada uma das bordas do gráfico que é conectada a  $\{E_1, E_2 \dots E_n\}$  e sintetizando o fluxo de áudio em um ângulo  $\theta$  especificado pela informação de sintetizar (renderizar) áudio imersivo. A segunda função envolve o envio de um fluxo de áudio misto gerado pelo dispositivo de computação de usuário 104 a todos os outros dispositivos/avatars de computação do usuário conectados à borda. O fluxo de áudio misto contém uma mistura de ambos, um fluxo de áudio capturado pelo dispositivo de computador do usuário (isto é, a voz do cliente) e fluxos de áudio sendo recebidos em todas as outras bordas (excluindo, naturalmente a borda junto a qual o fluxo de áudio misto está para ser enviado). Antes do envio, o fluxo de áudio misto é atenuado usando o valor de atenuação  $\alpha$  para aquela borda particular, também especificada pela informação de sintetização de áudio imersivo.

[0027] Com referência às Figuras 1 a 6, um exemplo de um processo para sintetizar cenas de áudio imersivo por um dispositivo de computação do usuário será, a seguir, descrito.

[0028] Em uma primeira etapa, o servidor de meio virtual 102 passa informação do estado do meio virtual para a aplicação do servidor de cliente do servidor de cliente 103 para subsequente processamento. Dado um mundo virtual com  $N$  avatares ( $A_1, A_2, \dots A_N$ ) cada qual controlada por clientes ( $C_1, C_2, \dots C_N$ ). Amostras de áudio ( $V_i$ ) são geradas por cada um destes  $N$  clientes. Um gráfico é construído com  $F$  bordas ( $E_1, E_2, \dots E_F$ ). Cada uma destas bordas possui um ângulo associado  $\theta_j$  e fator de atenuação  $\alpha_j$  em que  $0 < j \leq F$ .

[0029] Etapa 1 - Colocar todos os avatares em um plano com suas coordenadas  $(x,y)$  correspondentes as suas coordenadas  $(x,y)$  no meio virtual como mostrado na Figura 2.

[0030] Etapa 2 - Criar uma conexão delta entre todos os nós do gráfico; se paredes existem, então removem quaisquer ligações entre os avatares que estão nos lados opostos das paredes.

[0031] Etapa 3 - Calcular uma árvore de extensão mínima (MST) para conectar todos os nós usando qualquer algoritmo apropriado – tal como algoritmo de Kruskal.

[0032] Etapa 4 - Remover todas as ligações no gráfico que não são parte da árvore de extensão mínima;

[0033] Etapa 5 - Embora loops possam ser adicionados sem exceder limites de transmissão nos nós e embora haja possíveis loops que excedam um limiar (por exemplo, dois tempos da faixa de audição), então:

(a) Ajustar extensão\_ da borda \_mínima= infinidade

(b) Para cada par de nós  $i, j$  que podem ouvir entre si.

(i) Calcular distância de loop mínima entre os dois nós como:  $\min\_loop_{ij} = SP_{ij} + VD_{ij}$ . Em que  $SP_{ij}$  é o trajeto mais curto na árvore entre  $i, j$  e  $VD_{ij}$  = a extensão de uma ligação direta entre  $i$  e  $j$ , se existir.

(ii) Se  $\min\_loop_{ij} >$  o limiar de loop (2x faixa de audição) e  $D_{ij} <$  extensão\_ da borda\_ mínima

(1) extensão\_ da borda\_ min =  $VD_{ij}$

(2) borda\_ min =  $\{i, j\}$

(c) Somar a borda referida como borda\_ min ao gráfico. Esta é a menor borda que introduz um loop aceitável.

[0034] A seguir, considerar 7 jogadores conectados ao meio virtual. Cada um dos jogadores está usando um PC diferente conectado ao Internet. Cada um dos jogadores está controlando um avatar separado – tais avatares são denominados A1 a A7. Os locais de  $x, y$  (espacial) destes avatares no mundo virtual são como mostrados na Tabela I.

Tabela 1: Local X,Y dos avatares

	X	Y
A1	228	191
A2	116	95
A3	205	150
A4	161	93
A5	216	235
A6	157	282
A7	127	221

[0035] Cada um dos avatares pode ser considerado sendo um nó. É admitido que cada um dos nós é capaz de enviar 4 fluxos e a faixa de audição de cada nó é 115.

[0036] Etapa 1 coloca os avatares em um plano com as mesmas coordenadas x,y quando elas ocupam o mundo virtual. Isto é mostrado na Figura 2.

[0037] Etapa 2 cria uma conexão de malha (Isto é, cada nó é conectado a cada um de outros nós) entre os nós. Esta etapa provê uma topologia de nó subjacente para uso pelo algoritmo de árvore de extensão mínima. A rede de malha para este exemplo é mostrada na Figura 4.

[0038] A distância entre cada um dos avatares é usada mais tarde no algoritmo e é calculada como o valor (matriz)  $VD_{ij}$ , como mostrado na Tabela 2.:

Tabela 2: Valores calculados de  $VD_{ij}$ 

	$VD_{2,1} = 148$	$VD_{3,1} = 47$	$VD_{4,1} = 119$
$VD_{1,2} = 148$		$VD_{3,2} = 105$	$VD_{4,2} = 45$
$VD_{1,3} = 47$	$VD_{2,3} = 105$		$VD_{4,3} = 72$
$VD_{1,4} = 119$	$VD_{2,4} = 45$	$VD_{3,4} = 72$	

$VD_{1,5} = 46$        $VD_{2,5} = 172$        $VD_{3,5} = 86$        $VD_{4,5} = 152$   
 $VD_{1,6} = 115$        $VD_{2,6} = 191$        $VD_{3,6} = 140$        $VD_{4,6} = 189$   
 $VD_{1,7} = 105$        $VD_{2,7} = 126$        $VD_{3,7} = 105$        $VD_{4,7} = 132$

$VD_{5,1} = 46$        $VD_{6,1} = 115$        $VD_{7,1} = 105$   
 $VD_{5,2} = 172$        $VD_{6,2} = 191$        $VD_{7,2} = 126$   
 $VD_{5,3} = 86$        $VD_{6,3} = 140$        $VD_{7,3} = 105$   
 $VD_{5,4} = 152$        $VD_{6,4} = 189$        $VD_{7,4} = 132$   
                           $VD_{6,5} = 75$        $VD_{7,5} = 90$   
 $VD_{5,6} = 75$                                        $VD_{7,6} = 68$   
 $VD_{5,7} = 90$        $VD_{6,7} = 68$

[0039] Na Etapa 3, o algoritmo de Kruskal foi usado para calcular uma árvore de extensão mínima. O algoritmo de Kruskal é um algoritmo comum para calcular árvores de extensão mínima e bem conhecido dos engenheiros de telecomunicação e expertos da teoria gráfica.

[0040] Na etapa 4, uma ligação que não é parte da árvore de extensão mínima é removida do gráfico. A saída deste estágio é mostrada na Figura 5.

[0041] O conjunto de nós que pode "ouvir" entre si é calculado na Etapa 5:

$H = [(3,1), (3,5), (3,7), (3,4), (3,2), (2,4), (1,5), (1,6), (1,7), (5,6), (7,6), (7,5)]$

e a distância de cada ligação de gráfico é, também, calculada como mostrado na Tabela 3:

**Tabela 3: Distância Entre Cada Ligação de Gráfico**

Ligação 2,4	45
Ligação 4,3	72
Ligação 3,1	47
Ligação 1,5	46

Ligação 5,6	75
Ligação 6,7	68

[0042] Por último, o trajeto mais curto para todos os elementos em H é calculado:

$$SP(3,1)=D(3,1)=47$$

$$SP(3,5)=D(3,1)+D(1,5)=47+46=93$$

$$SP(3,7)=D(3,1)+D(1,5)+D(5,6)+D(6,7)=47+46+75+68=236$$

$$SP(3,4)=D(3,4)=72$$

$$SP(3,2)=D(3,4)+D(4,2)=72+45=117$$

$$SP(2,4)=D(2,4)=45$$

$$SP(1,5)=D(1,5)=46$$

$$SP(1,6)=D(1,5)+D(5,6)=46+75=121$$

$$SP(1,7)=D(1,5)+D(5,6)+D(6,7)=46+75+68=189$$

$$SP(5,6)=D(5,6)=75$$

$$SP(7,6)=D(7,6)=68$$

$$SP(7,5)=D(7,6)+D(6,5)=68+75=143$$

$$\min\_loop_{ij} = Sp_{ij} + VF_{ij}$$

$$\min\_loop(3,1)=47+47=94$$

$$\min\_loop(3,5)=93+86=179$$

$$\min\_loop(3,7)=236+105=341$$

$$\min\_loop(3,4)=72+72=144$$

$$\min\_loop(3,2)=117+105=222$$

$$\min\_loop(2,4)=45+45=90$$

$$\min\_loop(1,5)=46+46=92$$

$$\min\_loop(1,6)=46+115=181$$

$$\min\_loop(1,7)=189+105=194$$

$$\min\_loop(5,6)=75+75=150$$

$$\min\_loop(7,6)=68+68=136$$

$$\min\_loop(7,5)=143+90=233$$

[0043] Então, ajustando extensão\_da borda\_ mín = infinidade, é

encontrado que o  $\text{loop\_min}(3,7) > 2 \times \text{faixa de audição}$  (230) e  $\text{VD}(3,7) < \text{extensão\_da borda\_min}$ ;

assim  $\text{extensão\_da borda\_mín} = \text{VD}(3,7) = 105$  e  $\text{borda\_min} = \{3,7\}$ .

[0044] Além do mais, nenhum outro  $\text{loop\_min}(i,j) > 2 \times \text{faixa\_de audição}$  existe.

[0045] A borda  $\{3,7\}$  é então adicionada ao gráfico como mostrado na Figura 6.

[0046] Etapa 5 é então repetida usando o gráfico atualizado (as ligações adicionadas pelas iterações prévias da Etapa 5) até não existirem loops deixados para adição.

[0047] A seguinte iteração da Etapa 5 constata que não existem mais loops aceitáveis para adicionar, assim o algoritmo para com apenas a adição de ligação 7,3.

[0048] Com referência à Figura 3, é mostrado um gráfico indicando os ângulos de sintetização e níveis de atenuação que são requeridos para sintetizar os fluxos de áudio de entrada da perspectiva do avatar 3. Como mostrado nas figuras precedentes, o avatar 3 receberá fluxos de áudio dos avatares 1, 7 e 4. A aplicação do cliente de meio virtual residente no dispositivo de computação do usuário que controla o avatar 3 sintetizará a cena de áudio como segue: o áudio chegando do avatar 1 será sintetizado no ângulo  $\theta_{31}$  para a esquerda do avatar 3. O áudio recebido do avatar 7 será sintetizado  $\theta_{37}$  graus para a esquerda do ouvinte; enquanto o áudio recebido do avatar 4 será sintetizado  $\theta_{34}$  para a direita do ouvinte, como mostrado na Figura 6.

[0049] Os fluxos de áudio que são recebidos dos avatares 1, 7 e 4 são uma mistura atenuada de todos os nós a montante. Por exemplo, o fluxo de áudio enviado do avatar 1 para o avatar 3 é uma mistura de todo o áudio recebido nos fluxos de entrada (exceto na ligação 3,1) misturados com a voz do avatar 1, que é gravada localmente pelo dis-

positivo de computação de usuário. Estas vozes são atenuadas de acordo com a extensão da ligação de áudio a ser recebido. Portanto, o fluxo de áudio que é enviado na ligação 3,1 é igual a: fluxo de áudio enviado pelo avatar 2 atenuada por  $\alpha_{12}$  misturado com voz localmente gerada do avatar 1.

[0050] Uma vez que o áudio proveniente de cada avatar é misturado com uma versão atenuada do áudio recebido em todas as ligações de entrada, o avatar 3 ouvirá eficazmente os avatares 1 e 5 que emanam da direção do avatar 1 com variações de distância devido aos níveis diferentes de atenuação. A voz do avatar 6 influenciará o áudio sendo ouvido pelo avatar 3 da direção de ambos os avatares 1 e 7. Todavia, se os níveis de atenuação estiverem ajustados corretamente, então o volume do avatar 6 será negligenciável. Deverá também ser notado que neste exemplo, se não existirem nenhum loop introduzido, então o avatar 3 não ouvirá o avatar 7 porque a atenuação no trajeto através de 5 e 1 teria sido muito grande.

[0051] Será facilmente apreciado que, mesmo embora a concretização da presente invenção tenha sido descrita no contexto de ser usado para suportar um meio virtual, a presente invenção não é restringida para ser usada com um meio virtual. Como tal, a presente invenção podia ser usada para prover um sistema de áudio imersivo para uso como, por exemplo, um sistema de conferenciar apenas áudio.

[0052] Embora a invenção tenha sido descrita com referência à presente concretização, será entendido por aqueles versados na técnica que alterações, mudanças e aperfeiçoamentos podem ser feitos e equivalentes podem ser substituídos para os seus elementos e suas etapas sem se distanciar do escopo da invenção. Em adição, muitas modificações podem ser feitas para adaptar a invenção a uma situação ou material particular aos ensinamentos da invenção sem distanciar-se do seu escopo central. Tais alterações, mudanças, modifica-

ções e aperfeiçoamentos, embora não descritos acima de modo expresso, são entrementes pretendidos e implicados para estarem dentro do escopo e espírito da invenção. Portanto, é pretendido que a invenção não seja limitada à concretização particular descrita aqui e incluirá todas as concretizações que caiam dentro do escopo das reivindicações independentes.

## REIVINDICAÇÕES

1. Método de criar uma cena de áudio para um avatar (AV 1 – AV 7) em um meio virtual incluindo uma pluralidade de avatares (AV 1 – AV 7), compreendendo as etapas de:

criar uma estrutura de ligação entre a pluralidade de avatares (AV 1 – AV 7), em que a estrutura de ligação é uma árvore de extensão mínima que conecta a pluralidade de avatares (AV 1 – AV 7); e

sintetizar uma cena de áudio para cada avatar (AV 1 – AV 7) com base na sua associação com outros avatares ligados (AV 1 – AV 7),

**caracterizado pelo fato de que** criar a estrutura de ligação compreende introduzir loops na árvore de extensão mínima, de modo que uma extensão mínima dos loops seja menor do que um valor predeterminado para evitar ecos nas cenas de áudio sintetizadas.

2. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a estrutura de ligação é operável para definir pelo menos um de ângulo de sintetização e fator de atenuação para aplicar aos fluxos de áudio nas ligações de entrada da estrutura de ligação.

3. Método, de acordo com a reivindicação 2, **caracterizado pelo fato de que** uma métrica de custo é calculada com base em uma distância entre dois avatares (AV 1 – AV 7) na estrutura de ligação.

4. Método, de acordo com a reivindicação 3, **caracterizado pelo fato de que** a métrica de custo define o fator de atenuação.

5. Método, de acordo com qualquer uma das reivindicações 2 a 4, **caracterizado pelo fato de que** um ângulo de uma ligação entre os dois avatares (AV 1 – AV 7) na estrutura de ligação define o ângulo de sintetização.

6. Método, de acordo com qualquer uma das reivindicações 1 a 5, **caracterizado pelo fato de que** o valor predeterminado é duas vezes a faixa de audição de um avatar (AV 1 – AV 7).

7. Método, de acordo com qualquer uma das reivindicações 1 a 6, **caracterizado pelo fato de que** ainda compreende a etapa de misturar a cena de áudio sintetizada com um fluxo de áudio gerado por pelo menos uma da pluralidade de avatares (AV 1 – AV 7).

8. Método, de acordo com a reivindicação 7, **caracterizado pelo fato de que** ainda compreende a etapa de enviar a cena de áudio mista para um outro avatar ligado (AV 1 – AV 7).

9. Dispositivo de computação de usuário (104), **caracterizado pelo fato de que** é disposto para realizar as etapas do método como definido em qualquer uma das reivindicações 1 a 8.

10. Sistema disposto para criar uma cena de áudio para um meio virtual de acordo com o método como definido em qualquer uma das reivindicações 1 a 8, **caracterizado pelo fato de que** compreende: uma pluralidade de dispositivos de computação (104), cada dispositivo de computação sendo capaz de controlar pelo menos um avatar (AV 1 – AV 7) no meio virtual, em que cada dispositivo de computação (104) é disposto para sintetizar uma cena de áudio de saída para o pelo menos um avatar (AV 1 – AV 7) e comunicar a cena de áudio de saída para pelo menos um outro dispositivo de computação (104).

11. Sistema, de acordo com a reivindicação 10, **caracterizado pelo fato de que** cada dispositivo de computação (104) é ainda disposto para receber uma cena de áudio de entrada a partir de pelo menos um outro dispositivo de computação (104) e sintetizar a cena de áudio de entrada com a cena de áudio de saída.

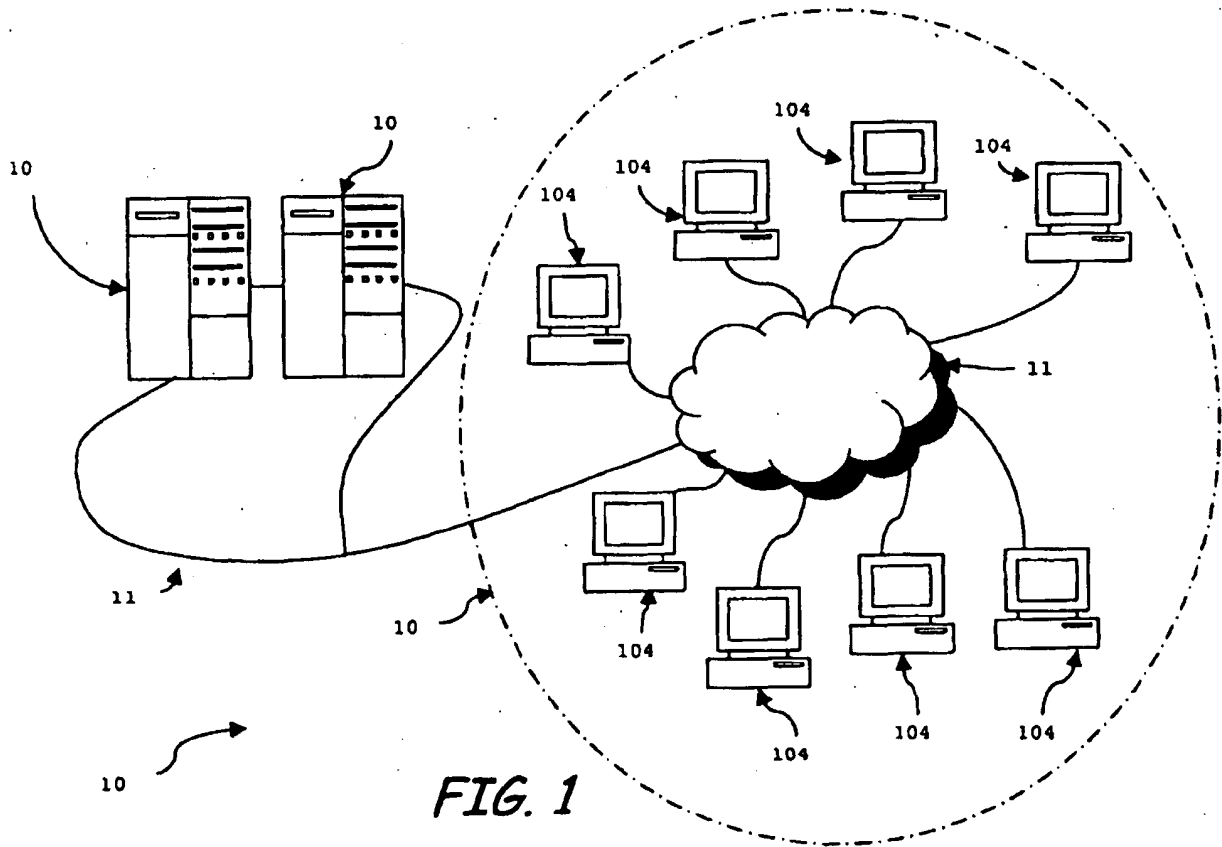
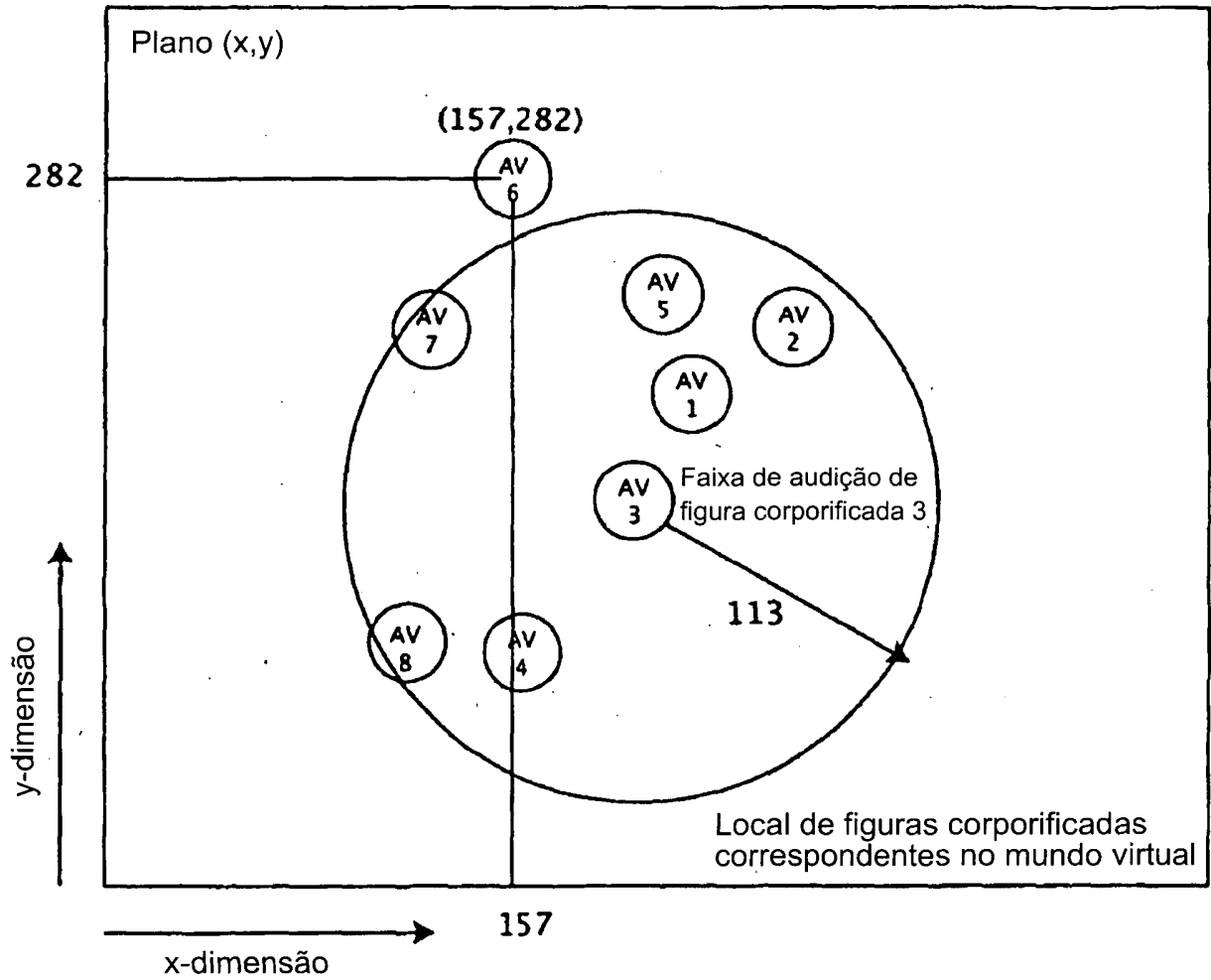
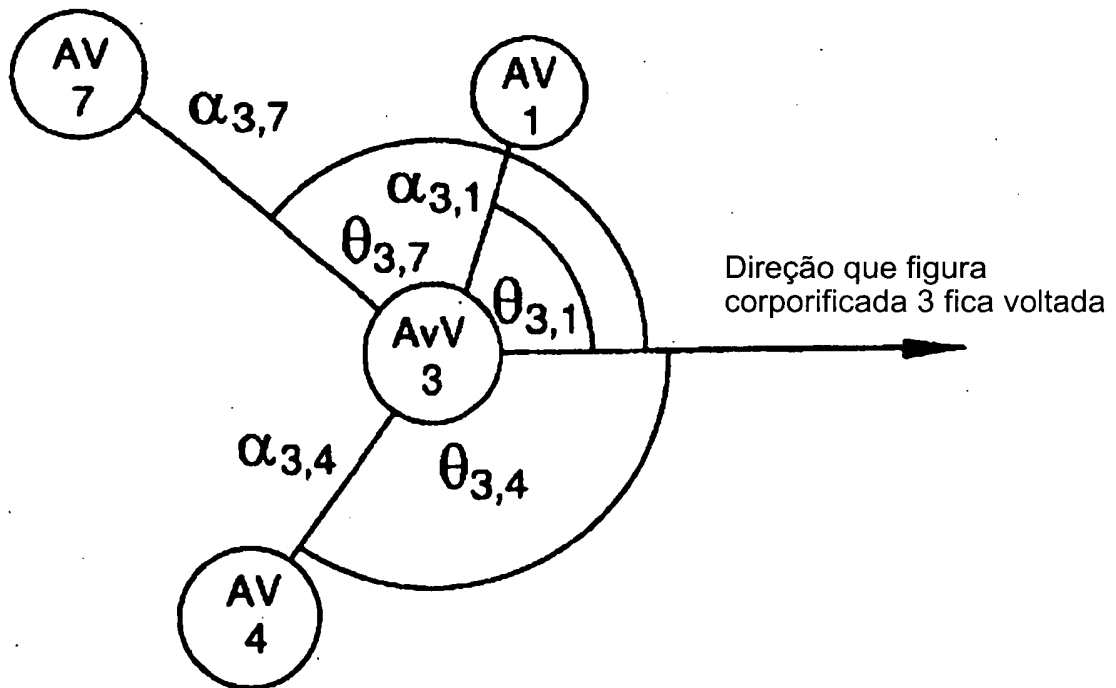
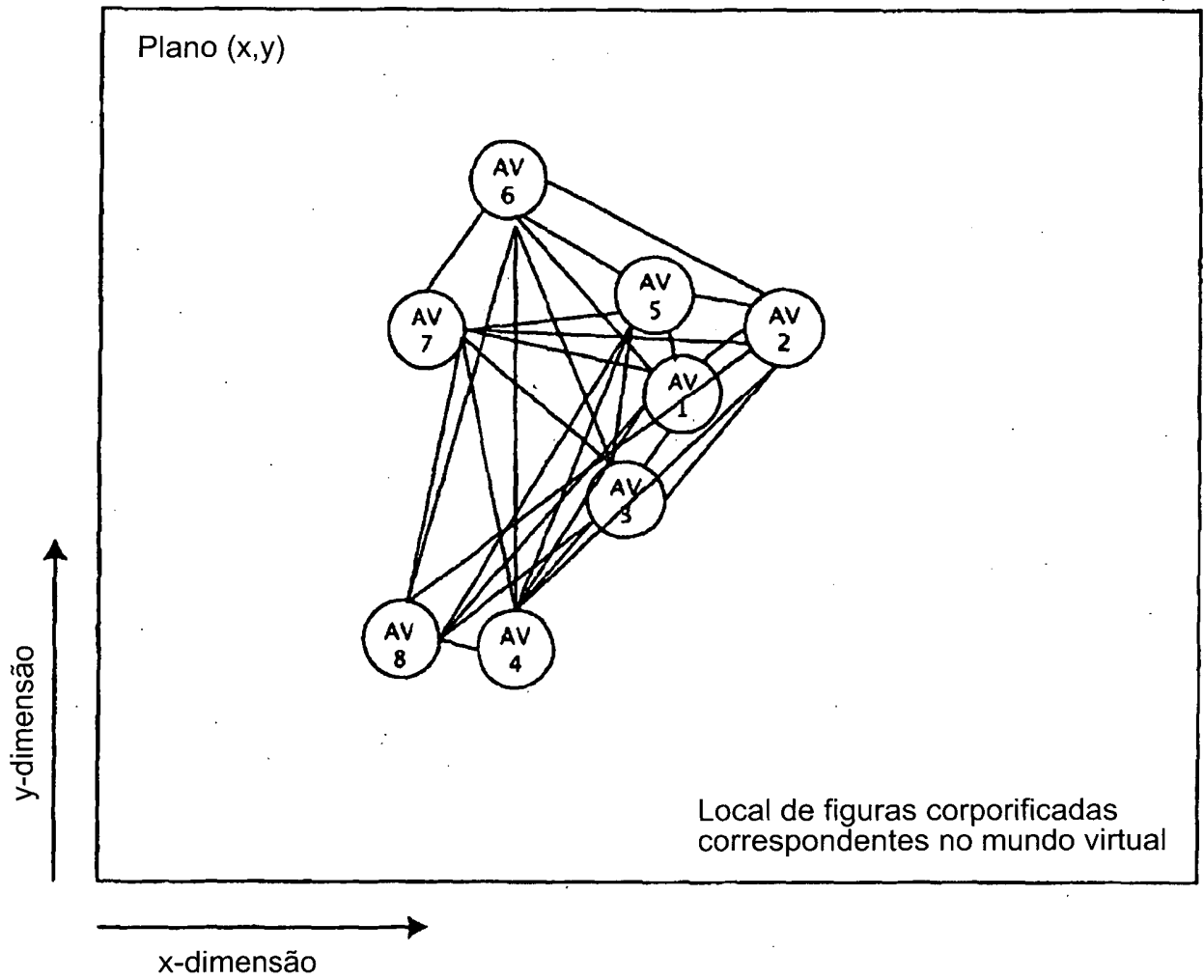


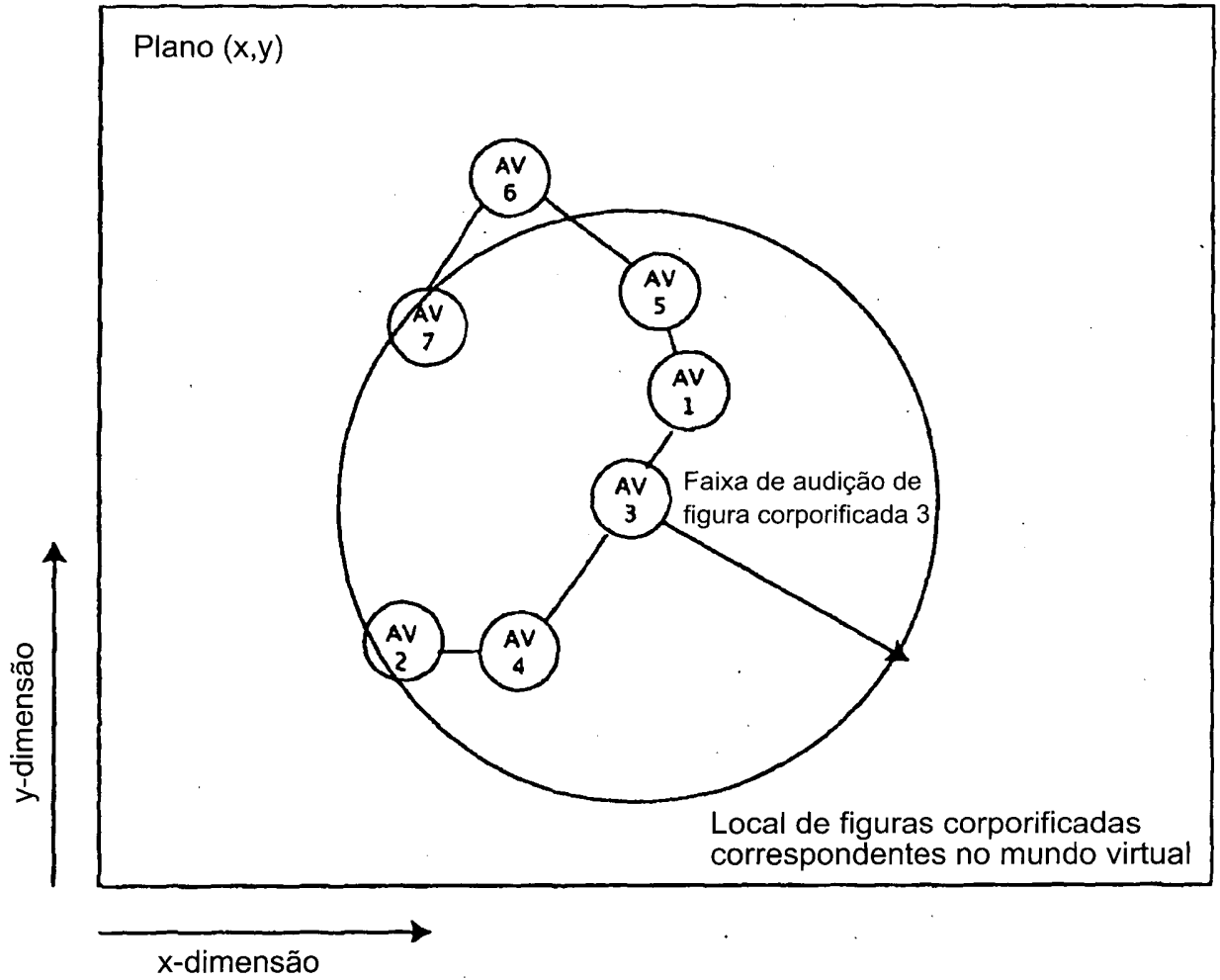
FIG. 1

**FIG. 2**

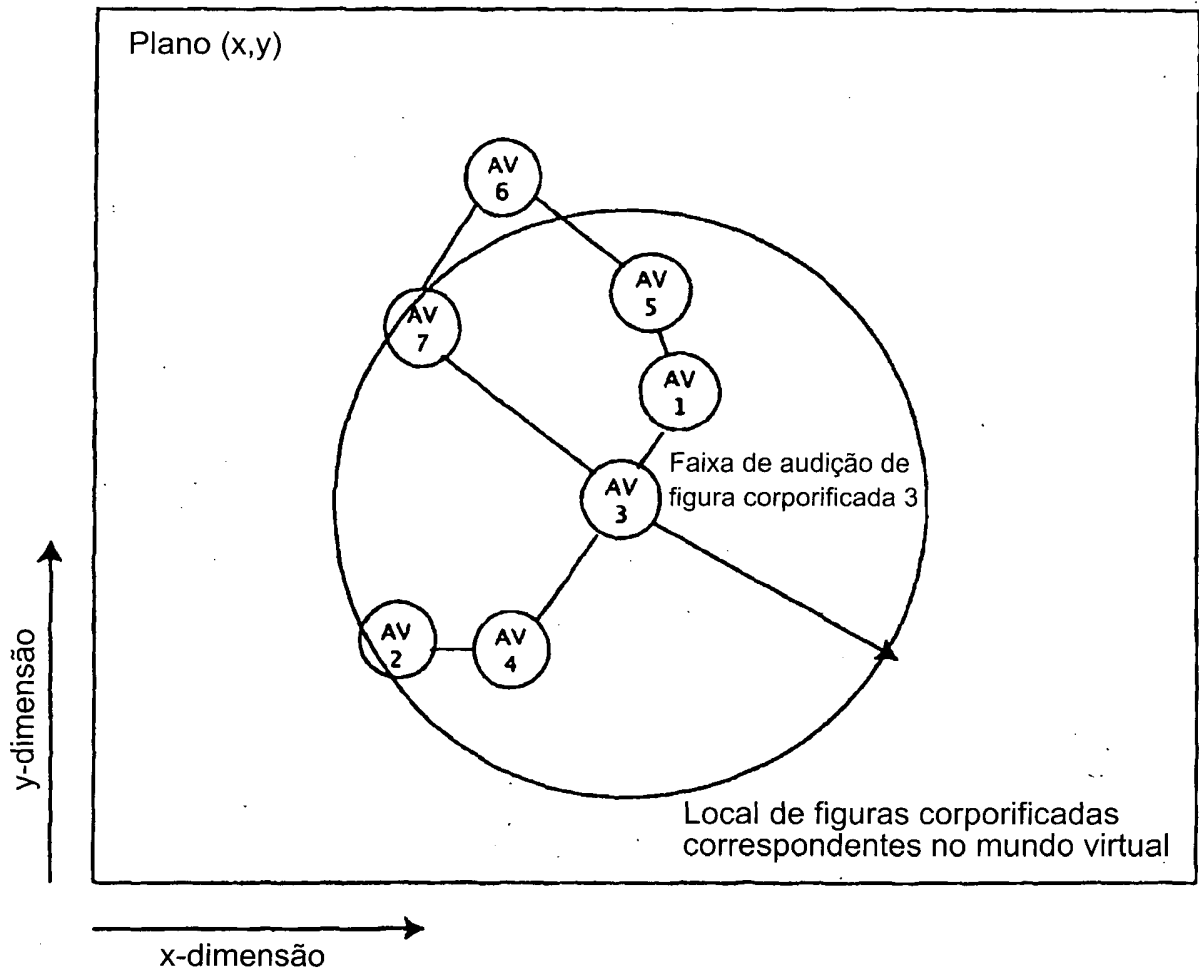
**FIG. 3**

## Gráfico de malha completo

**FIG. 4**



**FIG. 5**

**FIG. 6**