

2

Tecnologias para interações virtuais

As interações virtuais ocorrem através de ambientes virtuais, que muitas vezes são ambientes gerados por computação gráfica, onde pessoas fisicamente distantes compartilham esse ambiente, conectadas via rede, interagindo entre si e/ou com objetos deste mundo virtual.

Estes mundos virtuais no entanto, não estão limitados apenas a ambientes gráficos 3D e também usam diversas mídias tais como vídeo, som, voz, imagens. O ponto principal destes ambientes virtuais está em combinar tecnologia com interação/comunicação de humanos.

As propriedades dos mundos virtuais variam de acordo com a aplicação, que dentre as principais pode-se citar: entretenimento, educação à distância, chat 3D (em ambiente gráfico 3D), videochat (usando imagem de câmera), eventos sociais e culturais como conferências, comércio eletrônico em *virtual shopping malls* e *showrooms*, suporte remoto a usuários, treinamento militar e industrial, projetos colaborativos. O desenvolvimento desses mundos virtuais envolve as tecnologias de realidade virtual, sistemas com simulação de vida, as comunidades virtuais, animação, agentes inteligentes, comportamento e sistemas distribuídos, por exemplo.

A presença independente de múltiplos usuários diferencia mundos virtuais dos sistemas padrões de realidade virtual ou sistemas padrões de jogos. A possibilidade de compartilhar objetos nos ambientes virtuais diferencia-os das salas de chat convencionais e a interação em tempo real diferencia da comunicação por e-mail ou web (54).

No passado, “surfar” na internet, mandar e-mail, ou *chat* usando apenas texto eram as únicas experiências possíveis em *Cyberspace*. Com o desenvolvimento de mundos virtuais, essa conotação vem mudando e procurando proporcionar cada vez mais a integração das pessoas com os ambientes virtuais e entre as pessoas, através desses ambientes.

Mundos virtuais compartilhados são uma nova mídia de comunicação. Estes são espaços tridimensionais ou bidimensionais onde cada pessoa tem sua representação virtual, chamada *avatar*, nos casos de aplicações gráficas. Os avatares podem mover-se, conversar via texto ou voz e interagir em um

contexto comum de informações e aplicações.

Os chats normalmente usam a ferramenta texto, o que ocasiona muitas limitações apesar de permitir que a pessoa analise o que está escrevendo antes de mandar. É crescente o número de ambientes que permitem o uso de voz, tornando a comunicação mais natural e espontânea. Os *videochats* podem ser considerados mundos virtuais pela sua aplicação e usam a própria imagem de vídeo da pessoa para representá-la, em tempo real, adicionada à transmissão de texto, imagens instantâneas do usuário, e voz. As plataformas de mundos virtuais 3D na internet estão começando a olhar para a inclusão da imagem de webcam do usuário integrada ao mundo gráfico 3D. Existem projetos nesta direção, como o Invite-EU e o Moove Roomancer que serão apresentados na Seção 2.2.

Nosso processo natural de comunicação inclui expressões faciais e posturas corporais como formas de expressar sentimentos, satisfação, concordância ou discordância. Incluir toda esta naturalidade como parte da comunicação em interações virtuais torna-se um assunto bastante complexo. Existem muitos softwares desenvolvidos e em desenvolvimento, assim como muitas pesquisas em andamento buscando atingir este objetivo.

2.1 Tecnologias

A seção apresenta as tecnologias usadas para promover interações virtuais, com ênfase na visualização, discutindo suas características e desenvolvimento. Alguns aspectos apresentados, no entanto, serão mais aplicáveis aos ambientes virtuais gráficos.

2.1.1 Principais elementos que habitam os mundos virtuais (4)

Ambientes virtuais não podem ser apenas um conjunto de imagens ou modelos 3D sem vida, pois estes objetivam reproduzir o mundo real. Quem habita os mundos virtuais?

Avatar é a representação virtual de cada pessoa participante do mundo virtual que será visualizada pelos outros participantes e que navegará no ambiente virtual. O avatar pode ter diferentes formas: humanos, animais, cartoons, formas geométricas, imagem real obtida de uma câmera, etc., podendo portanto ser 3D ou 2D, de acordo com o ambiente. Esse ambiente é também chamado de *cyberspace*.

Bot é um programa que se comporta como um usuário em um ambiente de chat ou executa algum tipo de serviço útil aos usuários. Esse termo é derivado

da palavra *robot*, e de certa forma os bots são os robots do mundo virtual. Um bot é como um avatar, exceto pelo fato de que não tem um humano “habitando” esse bot. É apenas uma peça automática de software (13).

Biota é uma espécie de vida virtual que popula mundos virtuais e pode se assemelhar a plantas, animais ou outras formas de vida. Eles possuem vida independente no sistema e podem ou não ser interativos. Biotas mutam, se reproduzem, crescem e morrem de acordo com regras no mundo virtual. Programas genéticos, redes neurais e teoria da complexidade são importantes tecnologias usadas no desenvolvimento de biotas (9).

Agentes são programas autônomos projetados para atender necessidades dos usuários tais como fazer busca de informações, por exemplo. Os agentes se assemelham aos bots na função desempenhada mas não necessariamente têm uma representação física como o bot. Os agentes também não têm a finalidade de auto-existir, como os biotas. Existe uma discussão em torno de quando um programa é simplesmente um programa ou um agente, e isso segue uma gradação relacionada com automaticidade, orientação para uma tarefa, flexibilidade, comunicação, etc.

2.1.2

Componentes de um sistema de mundos virtuais (4)

Para obter as propriedades necessárias a um mundo virtual é preciso garantir aos usuários uma visão consistente do estado do sistema. Mundos virtuais são sistemas dinâmicos por ser um processamento em tempo real, com usuários alterando freqüentemente o estado do sistema. O estado dinâmico compartilhado consiste nesta mudança de informação que as múltiplas máquinas devem manter sobre o sistema, tais como: o que e quem está participando no momento, sua localização e suas ações, e dar a todos os participantes a mesma informação e visualização simultânea.

A precisão deste estado compartilhado é fundamental para criar ambientes realísticos. Sem isso cada usuário poderá agir independentemente e não haverá mais o senso de compartilhamento de lugar e tempo entre os usuários.

Pode-se dividir esse estado dinâmico compartilhado em dois:

- Estado Local - é o que ocorre localmente em cada usuário, ou seja, as alterações que cada participante gera localmente em sua máquina e manda para o sistema. Essa ação do usuário está diretamente ligada à interação, que é a forma que o usuário tem para se integrar ao sistema e expressar as suas ações.
- Estado Global - é tratado pelo sistema que atualiza todos os valores recebidos dos participantes e distribui para os mesmos, para garantir a

consistência. Na construção de um mundo virtual é fundamental considerar uma forma de manter uma visão consistente do estado dinâmico para os participantes.

2.1.3

Visualização síncrona e assíncrona (39)

Analisando no sentido amplo da perspectiva de visualização que o sistema dá a todos os usuários do ambiente virtual, pode-se definir dois tipos de visualização para ambientes distribuídos.

A primeira é uma visão síncrona do ambiente, i. e., todos os usuários compartilham a visualização da mesma cena. Por exemplo, um simulador de vôo distribuído onde as imagens são coordenadas para dar a sensação de que todos os usuários compartilham a mesma cabine de piloto, com uma mesma visualização. Sincronismo requer ao mesmo tempo alta confiabilidade e baixa latência. No entanto, estes ambientes que precisam de visualização sincronizada, por razões práticas, muitas vezes ficam restritos a LANs (*Local Area Network*).

A CAVE usa uma abordagem similar para sincronizar cada quadro da imagem projetada; estes quadros são renderizados em diferentes máquinas para melhorar a performance do *rendering* (processo de geração de imagens gráficas). Visão síncrona é também importante para sistemas de CAD onde as pessoas precisam ter a mesma visão de um projeto de engenharia, por exemplo. A área de jogos também é uma outra aplicação que requer visualização síncrona do ambiente. Visualização síncrona é importante para pequenos mundos em que são necessárias manipulações colaborativas de objetos e aplicações sobre LANs.

A segunda forma é a visualização assíncrona. Neste paradigma, múltiplos usuários têm controle individual sobre quando e o que eles podem ver no ambiente virtual, concorrentemente. Participantes podem estar fisicamente separados em uma rede local ou a grandes distâncias. Ambientes virtuais de grande porte usam visualização assíncrona também pelo alto custo de sincronização em redes de longo alcance.

Do ponto de vista prático, a visualização deve estar diretamente relacionada à aplicação, pois nas aplicações de chat que pretendem reproduzir comunidades e cidades reais, a visão assíncrona é muito mais natural para o usuário, pois este quer estar livre para escolher sua direção, independentemente da visão de outro usuário.

As aplicações com câmeras web sempre disponibilizam uma *selfview*, que é uma pequena janela que mostra a visão da própria câmera para que a pessoa

sempre veja como a sua imagem está sendo enviada para a pessoa remota. No caso dos videochats, a visualização é síncrona no sentido de que a pessoa remota deverá visualizar a mesma imagem que a pessoa local vê na sua própria webcam, para que possam ter uma interação em tempo real. A aplicação também precisa ter um controle do envio da imagem local e recebimento da imagem remota para não congestionar a aplicação do usuário, e este aspecto pode ser visto como assíncrono, pois o envio deve ocorrer após o recebimento da imagem, ou vice-versa. É claro que dependendo do hardware, da rede usada, haverá um pequeno retardo na transmissão, e também porque as aplicações com webcams trabalham com uma baixa taxa de transmissão de quadros (fps).

2.1.4

Linguagens para desenvolvimento de ambientes virtuais

OpenGL (29) é amplamente usado como API (*Application Programming Interface*) para gráficos 3D. É independente da plataforma e também do sistema operacional, sendo portanto usado em PCs, estações de trabalho e supercomputadores. Permite fazer o *rendering*, executar funções gráficas 2D e 3D, transformações geométricas, processamentos de cor, iluminação, sombreamento, mapeamento de textura, fog e efeito de *motion blur*.

Java3D (31) é uma API baseada em descrição de cena, para desenvolvimento de *applets* e aplicações 3D em redes, usada em algumas aplicações disponíveis como citado na seção 2.2.

Virtual Reality Modeling Language (VRML) (1) é um padrão Internacional para multimídia e mundos virtuais na Internet. VRML acrescenta interação, gráficos estruturados, e dimensões extras (z e tempo) às experiências online na internet. Aplicações variam desde simples efeitos 3D em páginas Web a aplicações científicas, de entretenimento e educação. É um padrão independente da máquina.

Funciona como um codificador de informações 3D para um arquivo a ser baixado na internet. Os arquivos VRML são acessados via Web através do Netscape Navigator ou Internet Explorer, mas precisam de plug-ins para serem visualizados. Um arquivo VRML contém uma cena gráfica com a descrição hierárquica da geometria 3D, textura e iluminação e ao ser executado, renderiza a cena gráfica e permite a navegação do usuário no ambiente virtual.

Mundos baseados em VRML sempre foram olhados como pobres na comunidade de ambientes virtuais pois o usuário interage apenas com objetos do mundo virtual de forma limitada e não interage com outros usuários. Para se obter maiores desempenhos e habilidades nestes mundos é necessário que eles sejam integrados a partes desenvolvidas em outras linguagens o que torna

muito mais difícil o seu desenvolvimento.

Um *script* de interface permite a comunicação com um programa de um ambiente externo e que este manipule uma cena gráfica VRML. Desta forma, códigos em Java e JavaScript podem excluir e adicionar nós a uma cena gráfica, gerar eventos, e receber notificações de eventos.

C e *C++* são linguagens largamente utilizadas no desenvolvimento de mundos virtuais. Mundos desenvolvidos em *C*, por exemplo, se auto suportam e podem gerar seus próprios browsers e quando executados em máquinas de alto desempenho sempre geraram ambientes mais convincentes e com melhor qualidade de rendering que os escritos em VRML. Por essa razão tem sido difícil justificar o desenvolvimento de grandes plataformas em VRML (54).

2.2

Plataformas para ambientes virtuais

Primeiramente, serão apresentados alguns softwares de uso popular e depois alguns projetos de pesquisa para o desenvolvimento de tecnologias mais sofisticadas.

2.2.1

Aplicações gráficas disponíveis na Internet

As plataformas apresentadas a seguir encontram-se disponíveis na Internet e podem ser usadas a partir da instalação de um browser disponibilizado no site da aplicação. Nem todas são de uso totalmente gratuito.

Estas plataformas gráficas, *3D* ou *2D* não possuem as convencionais salas de chats, elas possuem vários mundos como se fossem cidades dentro de um grande mundo virtual. Como usuário, é possível desempenhar diferentes atividades, como em uma cidade normal. Algumas plataformas disponibilizam sua tecnologia permitindo o uso de seu servidor para o desenvolvimento de mundos que farão parte da plataforma ou também para atividades acadêmicas.

Worlds Chat Space Station (64)

Este é um dos primeiros mundos virtuais lançados na internet, em 1995. Em Worlds Chat Space Station é possível explorar um mundo virtual *3D*, interagir com muitas pessoas representadas por avatares, Figura 2.1.

Existe uma galeria de mais de 150 avatares onde o usuário pode escolher o seu visual, ao entrar no programa. A interface mostra um mapa do ambiente *3D*, no qual o usuário pode clicar em um ponto e se teletransportar. Os avatares podem passar pelas portas, mas não pelas paredes, apesar de poderem

passar através de outros avatares. Isso pode causar certos problemas sociais na interação.

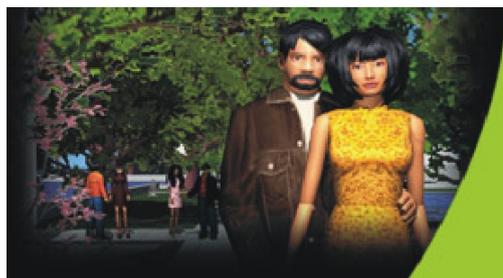


Figura 2.1: Worlds Chat Space

Worlds chat oferece as seguintes habilidades:

- Comunicação voz a voz
- Transmissão de áudio e vídeo
- Chat via texto
- Jogos com multi-jogadores
- Comércio eletrônico 3D
- Avatares adaptáveis
- Conjunto de animações para os avatares, com expressão de gestos

Esta plataforma foi desenvolvida em Java e C++, é um conjunto de tecnologias integradas em uma arquitetura cliente-servidor que oferece aos desenvolvedores de mundos virtuais os componentes necessários (*browser*, ferramentas e servidor) para criar, distribuir, e manter ambientes virtuais multi-usuários na Internet ou em redes TCP/IP privadas.

WorldsShaper é a ferramenta de autoria da *Worlds.com* escrita em Java, para permitir a construção de recursos avançados tais como estados compartilhados, portais e voz. O *WorldsServer* oferece uma arquitetura robusta, flexível com capacidade escalável que oferece a vantagem de desenvolvimento de pequenos protótipos de mundos e depois aumentar sua capacidade à medida que a popularidade deste mundo for crescendo.

ActiveWorlds (28)

ActiveWorlds (AW) é uma plataforma gráfica de chat 3D, onde as pessoas são representadas por avatares que podem percorrer os diversos ambientes virtuais disponíveis e interagir com outros avatares, com bots e com objetos virtuais. Esta plataforma promove uma experiência de realidade virtual não imersiva através da web, permitindo que o usuário visite e converse em ambientes gráficos 3D que podem também ser construídos pelos próprios usuários. O Universo do Active Worlds possui uma característica de comunidade online, com milhares de usuários distribuídos por milhões de quilômetros quadrados de um território virtual.

Além de chat como função principal, o AW oferece as seguintes possibilidades:

- Explorar mais de 1000 mundos virtuais específicos
- Construir sua própria casa 3D na Internet
- Fazer compras online no shopping 3D virtual e conversar com os vendedores
- Fazer amigos e conversar com pessoas, através de seus avatares, localizadas em todo o mundo
- Brincar com jogos interativos 2D e 3D, assim como participar de jogos coletivos como bingo, por exemplo
- Escolher entre um vasto conjunto de identidades de avatares, todos com animações pré-gravadas.

A figura 2.2 mostra alguns avatares interagindo em um ambiente virtual do AW.



Figura 2.2: ActiveWorlds 3D Chat - Avatares (28)

Esta plataforma também inclui programas educacionais, disponibilizando seu servidor para ambientes de educação à distância. O ambiente é totalmente gerado por computação gráfica. As imagens da Figura 2.3 abaixo mostram o crescimento da área do AlphaWorld, que corresponde a uma matriz de coordenadas norte, sul, leste, oeste, com um nível zero no centro, por onde as pessoas entram no AW e podem teletransportar-se ao fornecer as coordenadas do ponto.

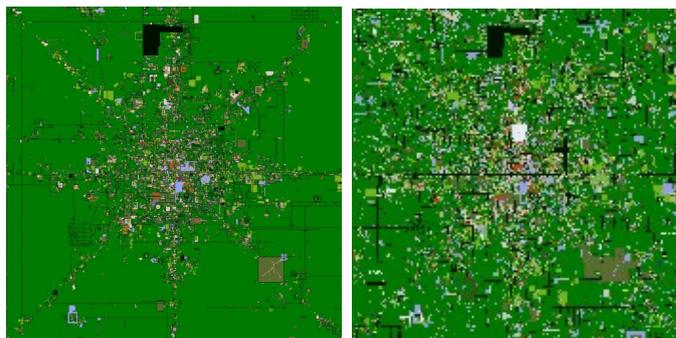


Figura 2.3: Construções ao longo dos eixos Norte-Sul e o equador do AlphaWorld (28)

Para gerar estas imagens do AlphaWorld, Activeworlds escreveu um programa para rastrear a base de dados e examinar cada “célula” do AlphaWorld e associá-la a um pixel na imagem - dando a esta célula a cor predominante nesta área.

Existem duas categorias de usuários, os turistas com direitos limitados e os cidadãos que são usuários registrados e com amplos direitos. É dado o direito de construção de casas dentro da área do AW, em determinadas regiões, a partir de um banco de dados de objetos; as animações dos avatares, pré-gravadas, tentam expressar emoções, mas muito pouco realistas.

Moove Roomancer (27)

Moove Roomancer é um browser de chat 3D para encontrar e visitar amigos online. O avatar tem um aspecto humano realístico. Ao se conectar, o usuário recebe uma casa e dispõe de diferentes tipos de mobília para escolher. O avatar que é chamado aqui de ator é capaz de mover e agir de uma maneira mais próxima da humana. Existem vários avatares diferentes e cada um possui um tipo de “personalidade” e movimentos próprios. O avatar é capaz de usar os móveis, sentar em cadeiras ou sofás, deitar, dançar, cumprimentar com as mãos, encolher de tamanho ou crescer, em tempo real.



Figura 2.4: Moove Roomancer

A Figura 2.4 mostra um ambiente do Roomancer e alguns avatares. O ponto forte dessa plataforma é a simulação eficiente de movimentos dos avatares nesses espaços virtuais, baseada na linguagem de programação DOORS (*Dynamic Object Oriented Runtime System*), desenvolvida para essa aplicação.

A plataforma permite a comunicação com HTML; em qualquer momento é possível alternar do mundo virtual 3D para uma página de internet. Um clique sobre um objeto dará informações sobre o mesmo, assim como um clique sobre uma folha de papel em uma mesa do ambiente 3D permite que a pessoa escreva um e-mail. A plataforma já está preparada para usar som 3D.

O sistema usa tecnologia *Peer-to-Peer*, portanto, não tem um servidor controlando as comunicações entre os usuários.

No Moove Roomancer o usuário clica no ponto destino para mover-se até ele, e apenas as coordenadas do destino são transferidas. A interpolação do movimento, no entanto, é dependente do desempenho do hardware local, conseguindo diminuir a carga de transferência de dados quando comparado aos sistemas *driving through* que usam *joystick* ou teclado para locomoção, comum nos sistemas em VRML ou jogos.

O programa possui alguns bots. Edgar é um bot com visual humano, dá boas vindas aos novos usuários, tem uma alta capacidade de aprendizagem, ótima memória e é versado em lingüística. Edgar fala inglês e alemão fluentemente e está aprendendo outras línguas. Outro bot é o Boomer, um basset. Boomer é um agente inteligente que vai buscar coisas interessantes para os usuários, na Web: fotos ou atores prediletos, arquivos de música, etc e exibe-os imediatamente.

Blaxxun (55)

Blaxxun é uma plataforma para várias aplicações diferentes, que pode ser acessada via HTML, Java ou com um plug-in.

Mundos virtuais em VRML não possuem características multi-usuários (no futuro, é provável que isso seja incorporado ao VRML). A Blaxxun combina mundos VRML com sua tecnologia e gera vínculos de ligação para ter multi-usuários em mundos VRML, o que é uma forma de driblar essa limitação.

O grupo da Blaxxun desenvolve vários projetos de pesquisa para expandir sua tecnologia. Segue abaixo uma descrição sumária de dois deles.

- Padronização MPEG-4, aplicações Cliente e Servidor MPEG-4 (56).

O projeto SoNG, Figura 2.5 pesquisa e desenvolve novas formas para o enriquecimento de mídia de comunicação e sua transmissão na internet. O protótipo final vai mostrar uma aplicação 2D e 3D em um estúdio para transmitir conferências ao vivo interativamente para um grande número de participantes.



Figura 2.5: TV Interativa - protótipo

Consultores online vão poder atender a clientes diretamente pela internet e a maior ênfase do projeto está na influência de padrões, especialmente para o padrão ISO MPEG-4. O projeto lançou a iniciativa de uma plataforma aberta para produzir uma completa cadeia ponto a ponto de produtos relacionados ao MPEG-4. Esta visão amplia o potencial para aplicações como TV interativa e computadores móveis e também pode mudar a maneira de interagir com esses sistemas.

SoNG Blaxxun está desenvolvendo um dos primeiros tocadores de MPEG-4 2D e 3D integrados capaz de renderizar, ao vivo, transmissões de vídeo e áudio assim como ambientes gráficos 2D e 3D.

- INVITE-EU: Ambiente de treinamento virtual distribuído inteligente (30).

INVITE é um projeto que começou em fevereiro de 2000 e foi estimado para 3 anos de duração. Desenvolve um ambiente colaborativo de aprendizado para ensino à distância através de ambientes virtuais distribuídos e compartilhados.



Figura 2.6: Treinamento Virtual

INVITE será uma plataforma para tele-aprendizado síncrono que tem como objetivo propiciar processos sociais de aprendizado através dos ambientes virtuais. Utiliza a imagem de câmera dos participantes integrada ao mesmo ambiente virtual, como mostra a Figura 2.6. Os resultados pretendem ser relevantes nos fatores sociais e cognitivos para situações de aprendizado colaborativo.

2.2.2

Aplicações de vídeo disponíveis na Internet

Esta seção ilustra algumas aplicações que usam a imagem da câmera como forma de visualização das pessoas. Os programas apresentados abaixo estão todos disponíveis através da internet e têm um número muito representativo de usuários.

iSpQ Videochat (44)

Esta plataforma oferece vídeo ao vivo e interações de voz para até 5 pessoas ao mesmo tempo em vídeoconferência. Usuários podem também enviar mensagem com uma foto instantânea da webcam e incluir mensagem de voz opcionalmente. A figura 2.7 mostra a tela para interação de um grupo.

Outra característica interessante é o V-Mail (Visual Mail): a pessoa tira uma foto instantânea, escreve uma mensagem e o iSpQ envia para o endereço de e-mail especificado. Qualquer recipiente de e-mail verá a mensagem com a



Figura 2.7: iSpQ - Quatro pessoas em vídeochat (44)

foto da pessoa. Esta plataforma tem em torno de 2.000.000 de usuários, que é um número muito significativo.

A taxa de frames por segundo (fps) obtida em uma interação de videochat varia muito em função da capacidade do computador cliente e do tipo de conexão que está sendo usada, mas é possível utilizar a plataforma a partir de uma conexão de modem e de um PC pentium 2 ou equivalente. iSpQ é projetado para atingir 15 fps, a uma resolução de 320x240 pixels, quando em videoconferência; no entanto, cada usuário obtém uma taxa diferente dependendo da capacidade de seu computador, do tipo de câmera, do tamanho de imagem selecionado, e da conexão de Internet que estiver sendo usada. Para obter uma taxa maior de transmissão (fps) ou transmissão de mensagens mais rápida, recomenda-se diminuir a resolução da imagem para 160x120 pixels.

MSN Messenger (11)

É um programa de mensagem instantânea bastante popular, que permite também a videoconferência entre participantes; o MSN vem instalado com o sistema Windows da Microsoft.

Durante a videoconferência, ou videochat, é possível mudar o tamanho da tela que exibe a imagem de câmera em tamanhos pequeno, médio ou grande. Além de voz e imagem ao vivo, os usuários podem adicionar *emoticons* às suas conversações. O programa permite enviar arquivos, enviar mensagens para celulares, fazer ligações telefônicas e jogar, Figura 2.8.

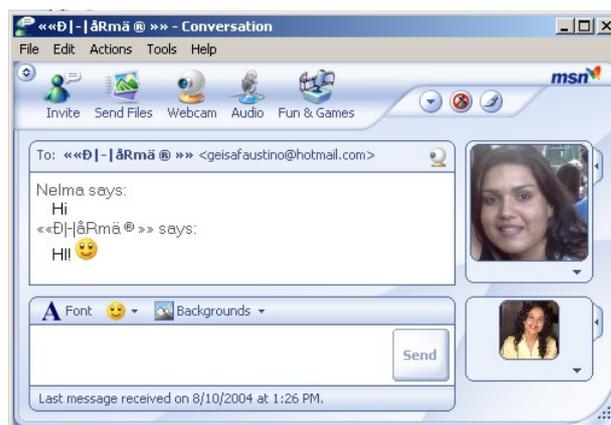


Figura 2.8: Tela do MSN para conversação (11)

Netmeeting

É uma aplicação para videoconferência que faz parte do sistema Windows. Funciona através de conexão do tipo *peer-to-peer*, permitindo o uso de vídeo e voz, e também oferece um quadro branco comum aos participantes para que os mesmos possam escrever e desenhar pontos comum da discussão, no caso de uma reunião. É também possível fazer transferência de arquivos e compartilhar programas através do Netmeeting.

2.2.3

Projetos de Pesquisa para videoconferência

Os projetos apresentados aqui são desenvolvidos em ambientes acadêmicos visando atingir uma maior interação das pessoas através de ambientes virtuais e com uso de imagem de câmeras, integrando diferentes técnicas de computação visual.

Immersive Meeting Point (im.point) (51)

O *Heinrich-Hertz-Institute*, na Alemanha, desenvolveu o protótipo de um sistema de videoconferência 3D imersiva. Segundo eles, Im.point representa um conceito avançado de uma nova geração de sistemas de videoconferência. A mesa de conferência local é integrada a uma cena de conferência virtual incluindo outros participantes remotos representados em uma grande tela de exibição e com uma perspectiva de visualização adequada. A representação 3D da cena virtual e dos participantes remotos, em tamanho natural, mediam a impressão de uma reunião face-a-face.

O protótipo chamado *Immersive Meeting Point* (im.point) usa o conceito de realidade mista para reproduzir este encontro virtual. O sistema integra

objetos 3D realistas dos conferencistas em um ambiente virtual compartilhado.

Usando o padrão MPEG-4, a nova tecnologia de codificação de vídeo é aplicada e demonstrada em tempo-real.



Figura 2.9: A mesa de conferência

Eles constroem uma imagem 3D dos conferencistas para colocá-los em um ambiente virtual e renderizar a perspectiva desejada desta representação para visualização. Esta representação 3D é baseada em mapas de disparidade que precisam ser estimados para cada entrada de vídeo, quadro a quadro. Então, usando as técnicas de rendering com base em imagem (image-based rendering), eles geram imagens realistas da cena, adaptadas ao ângulo de visão dos conferencistas locais. Os conferencistas são exibidos através desta “imagem real”, que é a imagem capturada de câmeras e reprocessada, e não por meio de avatares (Figura 2.9).

Características do sistema:

- abordagem de realidade mista, consistente transição entre mundo real e virtual
- ambiente virtual compartilhado permitindo características de telecolaboração
- transmissão de vídeo e áudio 3D para a representação natural dos conferencistas
- contato olho-no-olho realístico pela análise 3D e rendering dos ângulos visuais virtuais
- parte superior do corpo em tamanho real representando gestos e expressões corporais
- arquitetura de software flexível e modular para futuras extensões

The office of the future and Group Teleimmersion (52)

“Este trabalho é baseado em uma aplicação unificada de visão computacional e computação gráfica em um sistema que combina e constrói, com base na exibição de imagens panorâmicas, sistemas de exibição de mosaicos, modelagem baseada em imagens (image-based modeling), e ambientes imersivos”.

“A idéia básica é usar técnicas de visão computacional, em tempo real, para extrair dinamicamente a profundidade por-pixel e informações de refletância para superfícies visíveis no escritório, tais como paredes, móveis, objetos e pessoas. Usando esta informação de profundidade é possível então exibir (projetar) imagens nas superfícies, renderizar as imagens dos modelos capturados das superfícies, ou interpretar mudanças nas superfícies. A primeira parte do projeto se concentra nos aspectos de exibição e captura. A abordagem é capturar, reconstruir, e então transmitir modelos dinâmicos com base em imagens, através da rede, para exibição em uma localidade remota”.

O projeto do grupo *tele-immersion* irá transmitir dados de vídeo para serem reconstruídos como uma única imagem contínua, para ser vista em tela de projeção em tamanho natural de uma pessoa para múltiplos participantes entre duas localidades remotas.

A implementação inicial do sistema incluiu 3 telas de projeção vizinhas e 8 câmeras de captura montadas acima das telas de exibição, além de computadores, servidor de câmera, software de rendering, ligação de interface de câmeras para vídeo digital (Figura 2.10).

Para aquisição e implementação do rendering, as câmeras são conectadas em pares aos PCs, que estão interconectados por Ethernet de 100Mbps. As câmeras são sincronizadas por um cabo controlado de um PC e totalmente calibradas. Os PCs atuam como servidores de vídeo e codificam os dados do vídeo para JPEG em resolução completa VGA, ou uma menor região de interesse especificada pelo algoritmo. Os dados do vídeo são enviados para rendering em um PC que comanda duas telas de exibição usando uma placa de vídeo NVIDIA GeForce4. O algoritmo de rendering mistura estas imagens de câmera usando técnica baseada em campo de iluminação para criar uma imagem contínua, de alta resolução, com largo campo de visão de uma perspectiva distante da tela.

2.3

Discussão

Para propor uma melhoria tecnológica para estas aplicações, é necessário considerar também popularização x sofisticação.



Figura 2.10: Telas de exibição e um conjunto horizontal de câmeras

Para proporcionar interações entre pessoas localizadas remotamente, duas questões abrangentes devem ser consideradas: *qual é a melhor tecnologia para isto?* e *qual a melhor forma de visualizar as pessoas?* A segunda pergunta é extremamente importante para atingir o objetivo da aplicação e obter satisfação dos usuários; e mais importante: a forma de representação escolhida para a pessoa determina diretamente a tecnologia que será desenvolvida para a aplicação, e quais serão as características do sistema.

A visualização constitui, sem dúvida, um dos principais aspectos das interações não mediadas (naturais), portanto merece especial atenção nas interações virtuais. Atualmente, sem considerar as aplicações para Telepresença, pelo apresentado neste capítulo, pode-se concluir que existem duas formas gerais de representar as pessoas: a imagem de câmera de vídeo e sua representação gráfica chamada *avatar*.

Muitas pessoas sonham com os hologramas. Quão bom seria poder ver a outra pessoa em tamanho natural e com a aparência real! Mas isto está um pouco distante tecnologicamente, e ainda muito caro para se tornar popular, apesar de muitos pesquisadores estarem investindo nesta área. Para se chegar ao ideal, muitos desenvolvimentos tecnológicos precisam ser feitos.

É importante observar que a tecnologia de desenvolvimento de aplicações com avatares gráficos é mais tradicional, e por isso tem uma maior variedade

na internet; até porque a tecnologia de webcam é mais recente e as primeiras webcams tinham imagens bem precárias e exigem, ainda hoje, mais recursos para transmissão das imagens, tanto do computador do usuário quanto da tecnologia de rede. Mas nos últimos anos, a webcam vem se popularizando e com isso, as aplicações gráficas tendem a incorporar essa mídia nos seus ambientes, assim como as aplicações típicas de voz passaram a inserir a visualização através de webcam, a exemplo do Skype (36).

Seria possível enriquecer essas interações entre as pessoas melhorando/variando algumas características da forma de visualização e sem um hardware de alto custo?

Responder a esta questão pode levar a uma solução intermediária entre o ideal e as atuais opções de sistemas com imagem de vídeo ao vivo ou avatares. Este questionamento levanta a necessidade de avaliar quais fatores são importantes para que a tecnologia corresponda às expectativas.

A exemplo dos projetos complexos e sofisticados, fica clara a necessidade de investigar *Presença x Tecnologia* para investir na melhoria de parâmetros que sejam relevantes e possam efetivamente surtir resultados positivos para o uso da aplicação. Isso vai ao encontro dos interesses da área de pesquisa em *Presença*, cujos fundamentos serão apresentados no capítulo 3, assim como as motivações obtidas das informações provenientes desta área de pesquisa.