

Mãos Colaborativas em Ambientes de Realidade Misturada

Claudio Kirner

Universidade Metodista de Piracicaba –UNIMEP
Faculdade de Ciências Matemáticas da Natureza e Tecnologia da Informação
Programa de Pós Graduação em Ciência da Computação

ckirner@unimep.br

Abstract

This paper discusses mixed reality, augmented reality and augmented virtuality and describes the project of a collaborative mixed reality system, which uses the hand as interaction device to manipulate virtual objects. Problems like tracking, calibration and collaboration are discussed and solutions are presented.

1. Introdução

O real e o imaginário sempre fizeram parte da vida das pessoas. Até alguns anos atrás, a única maneira de retratar o imaginário era descrevê-lo ou no máximo desenhá-lo ou representá-lo de maneira restrita.

Com o advento da realidade virtual e o avanço dos equipamentos computacionais, a reprodução do real e a representação do imaginário tornaram-se mais fáceis de serem obtidas, uma vez que foram disponibilizadas interfaces mais intuitivas e rompidos os limites normalmente existente. As pessoas agora podem ativar aplicações computacionais, atuando diretamente sobre elementos tridimensionais conhecidos como: abrir porta, acionar alavanca, puxar gaveta etc.

Além disso, os sentidos e as capacidades puderam ser ampliados como ver/ouvir/sentir/viajar muito longe, a nível de galáxias, ou muito perto, a nível de estruturas atômicas. Nesse contexto, surge a Realidade Aumentada que enriquece os ambientes reais com objetos virtuais, através de algum dispositivo tecnológico. Uma das maneiras mais simples de fazer isto, é usar uma webcam para capturar uma cena, na qual são adicionados objetos virtuais, mostrando-a no monitor.

Assim, o que antes só podia ser descrito, depois pode ser modelado e observado com realidade virtual e agora poder ser incorporado ao nosso ambiente com realidade aumentada. Pode-se agora colocar um vaso imaginário sobre uma mesa real ou ponte projetada no papel sobre um rio verdadeiro. Por outro lado, os ambientes virtuais também podem ser enriquecidos com a presença de elementos reais como pessoas ou objetos, configurando o que se chama de virtualidade aumentada.

Tanto a realidade aumentada quanto a virtualidade aumentada fazem parte de uma categoria mais geral conhecida por realidade misturada.

Essa tecnologia deverá ter grande impacto no relacionamento das pessoas, uma vez que facilita a formalização das idéias, através de novas maneiras de visualização, comunicação e interação com pessoas e informação.

Apesar de todas as áreas do conhecimento deverem usufruir dos benefícios da Realidade Aumentada, ensino/aprendizagem e treinamento deverão particularmente passar por uma grande evolução com novas formas de relacionamento do estudante com professor, colegas e informação, propiciados pela mistura do real com o virtual. Nesse sentido, está em andamento, no Programa de Pós-Graduação em Ciência da Computação da UNIMEP, o projeto "Mãos Colaborativas com Realidade Misturada", que visa permitir que as mãos de várias pessoas interajam diretamente em ambientes virtuais, construindo-os e alterando-os, de forma que os participantes aprendam fazendo, ao mesmo tempo em que geram resultados para serem usados por outras pessoas que visitarem os ambientes,

depois de terminados.

Nos itens seguintes, serão mostrados: uma discussão sobre realidade misturada; os aspectos de colaboração em ambiente de realidade aumentada, a descrição do projeto e suas aplicações.

2. Realidade Misturada

A realidade misturada, abrangendo tanto a realidade aumentada quanto a virtualidade aumentada, pode ser classificada de acordo com suas diversas formas de visualização [5]: a) Realidade aumentada com monitor (não imersiva) que sobrepõe objetos virtuais no mundo real; b) Realidade aumentada com capacete (HMD) com visão ótica direta (see-through); c) Realidade aumentada com capacete (HMD) com visão de câmera de vídeo montada no capacete; d) Virtualidade aumentada com monitor, sobrepondo objetos reais obtidos por vídeo ou textura no mundo virtual; e) Virtualidade aumentada imersiva ou parcialmente imersiva, baseada em capacete (HMD) ou telas grandes, sobrepondo objetos reais obtidos por vídeo ou textura no mundo virtual; d) Virtualidade aumentada parcialmente imersiva com interação de objetos reais, como a mão, no mundo virtual.

Além disso, uma definição mais precisa de realidade misturada envolve: a combinação do real com o virtual; a interação em tempo real; o posicionamento tridimensional do real e virtual (registration) [2].

Para isso funcionar, há alguns problemas tecnológicos que precisam ser bem resolvidos, entre eles: rastreamento de objetos reais; posicionamento e colibração das sobreposições no ambiente tridimensional misturado e interação.

Em geral, as soluções são específicas, exigindo uma delimitação bem definida para as aplicações, o que constitui-se em obstáculos para aplicações de uso geral.

3. Realidade Misturada Colaborativa

A realidade misturada colaborativa baseia-se na existência de espaços reais e virtuais compartilhados acessados por várias pessoas localizadas no mesmo local ou remotas entre si. [3, 4]. No mesmo local, as pessoas visualizam e interagem com os elementos reais e virtuais, cada uma do seu ponto de vista, usando capacete com câmera e rastreadores, por exemplo. Como exemplo tem-se os projetos Studierstube e Magic Book, citados em [1].

No caso de participantes remotos, uma solução é usar ambientes virtuais colaborativos, como espaço compartilhado, inserindo ali objetos virtuais reais como mãos ou pessoas interagindo. O projeto 3D Live [6] apresenta um processo de captura de conteúdo 3D para uso em sistemas de realidade aumentada.

A realidade misturada colaborativa tem elevado potencial para ensino/aprendizagem, devido ao fato de poder trazer elementos virtuais interativos para o ambiente do usuário.

4. Descrição do Projeto

O projeto “Mãos Colaborativas com Realidade Misturada” consiste na elaboração de um espaço remoto compartilhado, onde usuários podem realizar tarefas colaborativas com suas mãos, interagindo com objetos virtuais.

O sistema é composto por um ambiente colaborativo de realidade virtual, representado por um cubo virtual, contendo objetos virtuais que podem ser visualizados e alterados, através da rede. Cada usuário tem ainda um ambiente físico equivalente, composto por um cubo com as mesmas dimensões, onde o usuário pode movimentar sua mão. O interior do cubo é iluminado e rastreado por uma webcam que, além disso, ainda captura a imagem da mão em posições diversas. O mapeamento do cubo físico no cubo virtual faz com que a mão apareça interagindo com os objetos virtuais. Cada usuário pode escolher a posição do seu cubo, de forma que sua mão aparecerá em pontos diferentes no mundo virtual, visualizando no monitor. A figura 1 mostra um diagrama do sistema local de cada usuário.



Figura 1 - Sistema Local do Projeto Mãos Colaborativas com Realidade Misturada.

Versões mais novas do projeto já evoluíram para um cilindro em pé no lugar do cubo, conforme a figura 2, de forma a permitir a participação de um número maior de pessoas, que, podendo escolher o ângulo de entrada, minimizam as possibilidades de colisão das mãos por ocasião da entrada.

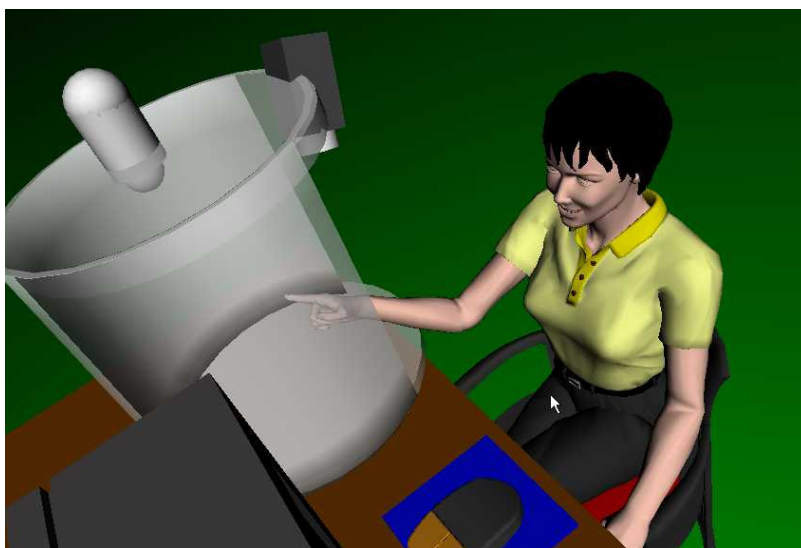


Figura 2 - Versão do projeto com o uso de cilindro.

Para efeito de interação, o usuário acopla uma pequena esfera na ponta do dedo e o rastreamento óptico captura sua posição, através da medida da densidade de pixels e do deslocamento dentro do cubo. A determinação de posição da ponta do dedo permite que o sistema detecte colisão com os objetos virtuais, viabilizando a interação. O uso de uma ferramenta de áudio conferência viabilizará a comunicação verbal entre os usuários.

Como há predominância do mundo virtual, complementado com a participação das mãos reais, o sistema é caracterizado como virtualidade aumentada baseada em monitor com interação.

Com essa infra-estrutura, o projeto está tendo outros desdobramentos como o rastreamento da trajetórias dos movimentos das mãos e armazenamento em banco de dados para análise. Com esses dados, é possível induzir o movimento, mostrando-se trajetórias pré-gravadas, através da inserção de pequenas esferas colocadas nos vários pontos. Outro projeto decorrente é o sistema de autoria colaborativa de mundo virtuais que, a partir de uma biblioteca pré-definida de objetos virtuais no interior do cubo, pode-se com as mãos movimentá-los e alterá-los, gerando cenários virtuais para uso de terceiros.

5. Conclusão

O projeto “ Mãos Colaborativas com Realidade Misturada” encontra-se em fase de desenvolvimento e visa obter uma infra-estrutura de baixo custo e de fácil utilização para o desenvolvimento de aplicações específicas de realidade misturada. Em função de suas características, o sistema está voltado mais para aplicações de virtualidade aumentada uma vez que prioriza o uso das mãos dentro de mundos virtuais. Com isso, espera-se facilitar o trabalho de manipulação de objetos virtuais, usando procedimentos intuitivos, de forma a permitir que não especialistas possam produzir mundos virtuais a partir de bibliotecas de objetos.

Agradecimentos

O autor agradece ao FAP/UNIMEP pelo apoio ao projeto “Construção de um Sistema de Realidade Aumentada para Desenvolvimento Cognitivo da Criança Surda”.

Referências

- [1] AZUMA, R. - A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments. V.6, N.4, Aug. 1997, pp. 355-385.
- [2] AZUMA, R. et. al. - Recent Advances in Augmented Reality. IEEE Computer Graphics and Applications, v.21, Nov/Dec. 2001, pp. 34-37.
- [3] BENFORD, S. et. al. - Understanding and Constructing Shared Spaces with Mixed Reality Boundaries. ACM ToCHI, v.5, N.3, 1998, pp. 185-223.
- [4] BILLINGHURST, M. & Kato, H. - Collaborative Mixed reality. Proc. of the International Symposium on Mixed Reality, ISMR'99, Springer -Verlag, NJ, 1999, pp. 261-284.
- [5] MILGRAM, P. et. al. - Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum. Telemanipulator and Telepresence Technologies, SPIE, V.2351, 1994, pp. 282-292.
- [6] PRINCE, S. et. al. - 3D Live: Real Time Captured Content for Mixed Reality. Proc. of the IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR'02, IEEE/ACM, 2002, pp. 7-13.

Desenvolvimento de Aplicações de Realidade Aumentada com ARToolKit

Daniela Akagui¹, Claudio Kirner²

¹Universidade Metodista de Piracicaba - UNIMEP
Faculdade de Ciências Matemáticas e da Natureza e Tecnologia da Informação

² Universidade Metodista de Piracicaba-UNIMEP
Faculdade de Ciências Matemáticas e da Natureza e Tecnologia da Informação
Programa de Pós Graduação em Ciência da Computação

{dakagui ,ckirner}@unimep.br

Abstract

Augmented reality interface is a growing area, which has great potential to be in educational applications. This kind of interface allows virtual imagery to be superimposed over live video of real world and stimulates the learning based on three-dimensional objects over a scenery. This paper exemplify the use of ARToolKit and new interactive techniques to implement several educational applications.

1. Introdução

Realidade Aumentada [2] é um tipo não convencional de interface, que permite misturar imagens de um ambiente real, obtidas por câmera de vídeo ou por outro processo, com objetos tridimensionais virtuais, enriquecendo a visão do usuário. Nesse ambiente, o usuário tem a sensação de que os objetos reais e virtuais coexistem no mesmo espaço (Figura 1), na medida em que os objetos virtuais são passíveis de visualização e de interação como se existissem no mundo real.



Figura 1- Objetos Virtuais no ambiente real.

Para construir um ambiente de Realidade Aumentada, é necessário a utilização de dispositivos sensoriais que permitam ao usuário integrar, em tempo real, objetos virtuais na cena do ambiente real. O grande avanço tecnológico de periféricos para Realidade Virtual contribuiu muito para o desenvolvimento da Realidade Aumentada, que se utiliza basicamente dos mesmos dispositivos, como:

- Dispositivos visuais: capacete ou *Head Mounted Display* ótico ou de vídeo, óculos estereoscópico ou *stereo glasses* e monitores.
- Dispositivos auditivos: fones de ouvido e microfones externos. Os fones permitem a adição de sons sintetizados por computador e os microfones captam o som do ambiente real.

- Dispositivos táteis e de força: luvas com dispositivos que possibilitam a interação com ações de tato e de força.

Além disso, para aumentar a interação da máquina com o usuário, são utilizados sistemas de rastreamento e posicionamento que acompanham a movimentação, por exemplo, da cabeça do usuário que estiver utilizando um HMD, possibilitando o cálculo do posicionamento dos objetos reais na cena para a renderização do objeto virtual em tempo real no local determinado pelo algoritmo utilizado. Para tal, são utilizados: sistema de posicionamento global (GPS), marcas fiduciais, RFID (*Radio frequency Identification*), sistemas de colisão e outros. Tudo isso visa permitir ao usuário mobilidade e realização de tarefas cotidianas com mais eficiência.

1.1. Áreas de aplicações da Realidade Aumentada: Educação

A Realidade Aumentada vem contribuindo na realização de tarefas de performance humana em diversas áreas como: medicina; montagem, manutenção e reparo de maquinarias complexas; enriquecimento de informações com anotações relativas às cenas do mundo real (*Annotation*); suporte computacional ao trabalho colaborativo (*CSCW*); arquitetura; planejamento de trajetória robótica; entretenimento e educação.

A Realidade Aumentada proporciona uma maneira diferente de aprendizado com o apoio do computador [5, 6]. Ela permite a fácil visualização e manipulação do objeto de estudos, reproduzindo os dados complexos sob a forma de objetos e textos tridimensionais, aumentando a capacidade de percepção do estudante, que passa a ser estimulado pela possibilidade de interação com a interface. Como exemplo: o livro "Sólidos Geométricos com Realidade Aumentada e ARToolKit" torna possível visualizar tridimensionalmente os sólidos nele descritos com ajuda de um dispositivo de vídeo (webcam), conforme a Figura 2.

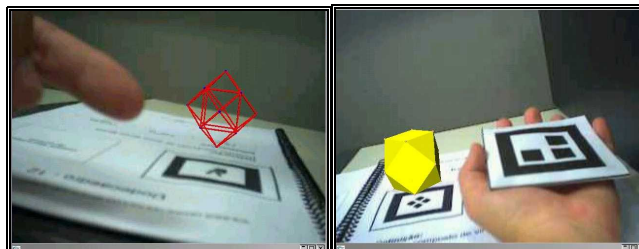


Figura 2- Livro: Sólidos Geométricos com Realidade Aumentada (aplicação dinâmica) por Daniela Akagui e Rafael Santin

Segundo Mark Billinghurst [4], pesquisador de interfaces que misturam o real e o virtual para auxiliar a colaboração, os educadores deveriam trabalhar junto a pesquisadores do campo da Realidade Aumentada para compreender melhor a utilização das aplicações computacionais no ambiente de aprendizagem.

1.2. A Ferramenta ARToolKit: Visão do Usuário

ARToolKit [1, 3] é um software, a partir do qual é possível construir ambientes de Realidade Aumentada. Nesse tipo de ambiente, para sobrepor a realidade com objetos virtuais ou textos 3D, é necessário obter a orientação e a posição dos objetos da cena real para que se possa manipular os elementos tridimensionais. Para obter as cenas do ambiente real, é utilizada uma *webcam* e como método de posicionamento da imagem são usados marcadores ou placas com marcas fiduciais. Os

objetos virtuais são associados às placas e já devem estar no banco de dados do ARToolKit. Quando a aplicação for iniciada, aparecerá na tela a imagem capturada pela *webcam*. Então, a placa colocada no campo da imagem da *webcam* será rastreada e o objeto associado aparecerá sobreposto a ela. O usuário, ao manipular a placa moverá também o objeto virtual.

2. Desenvolvimento de Aplicações Educacionais com Realidade Aumentada baseadas em ARToolKit

O projeto tem como base o estudo e desenvolvimento de ambientes educacionais com a tecnologia de interface baseada em Realidade Aumentada.

Para construir as aplicações com ambientes de Realidade Aumentada, foi utilizado o ARToolKit, que é um software livre. Por isso, esse software está em constante atualização, encontrando-se disponível para download na Internet em vários sites com destaque para o HILT [6].

O uso do software permitiu o desenvolvimento de diversas aplicações simples ou mais elaboradas como: o Livro "Sólidos Geométricos com RA e ARToolKit"; quebra-cabeça; jogo de placas com marcação numérica e outras, que serão descritas em seguida.

2.1. Aplicações simples

Nesse caso, foi construído um marcador e cadastrado. A placa foi associada a um objeto virtual, permitindo o teste do aplicativo, através de visualização e movimentação do objeto virtual dentro do mundo real.

2.2. Aplicações com múltiplos marcadores

Após o cadastramento de diversas placas com símbolos diferenciados, criados previamente, foram feitos testes da capacidade de captura do ARToolKit, ou seja, em quantos marcadores ele conseguiria gerar o objeto virtual associado? A conclusão foi que o aplicativo mostra os objetos das placas visíveis, independente da quantidade. Isto permite o desenvolvimento de aplicações de associações de posição ou de agrupamento.

2.3. Quebra cabeça

Nesta aplicação, um quebra-cabeça comum foi colado num papel, que continha o marcador impresso do lado oposto.

O papel foi cortado de acordo com o formato das peças do quebra cabeça e foi associado um objeto virtual relacionado. Ao montar-se o quebra-cabeça e virá-lo, o usuário expõe o marcador e visualiza o objeto virtual correspondente. Assim, ao terminar o desafio, o usuário vê a figura do quebra-cabeça e, do lado oposto, o objeto ou cena virtual tridimensional correspondente.

2.4. Livro de Sólidos Geométricos com Realidade Aumentada e ARToolKit

Baseado nos estudos de Sólidos Geométricos, algumas figuras pertencentes a esse grupo foram construídas com VRML e associadas a marcadores impressos no livro. Além disso, foi criada uma placa de controle. A aplicação funciona da seguinte maneira: ao colocar o livro em frente a câmera e posicioná-lo de maneira que o marcador e o quadrado do lado que contém as inscrições: "Coloque aqui a placa de controle", fiquem visíveis no campo de captura da Webcam, aparecerá o objeto virtual associado ao marcador. Ao colocar a placa de controle no local especificado, imediatamente o objeto virtual do marcador mudará. O marcador pode disparar três vezes a mudança do objeto: o primeiro objeto é aramado, o segundo é sólido e o terceiro é sólido e tem movimentos de rotação. Todos os

objetos que aparecem devem ser cadastrados e associados a uma placa. Para construir essa aplicação, uma parte do código fonte do ARToolKit foi modificado, permitindo o controle dinâmico dos objetos.

3. Conclusão

A Realidade Aumentada é um campo crescente que pode ser explorada para diversas aplicações adaptadas para cada tipo de área de conhecimento. Ela tem um grande potencial na área de Educação e Entretenimento com o desenvolvimento de jogos educativos. Esta área está sendo beneficiada com o avanço da ergonomia, em função do desenvolvimento tecnológico.

A ferramenta ARToolKit facilita a compreensão das Interfaces com Realidade Aumentada, mas necessita de ajustes, dependendo do tipo de aplicação. Algumas dificuldades de instalação e desenvolvimento de objetos virtuais animados foram superadas e o trabalho continua com o desenvolvimento de mais aplicações para uso efetivo no ensino, buscando aumentar a motivação das pessoas envolvidas.

Agradecimentos

Os autores agradecem ao PIBIC-CNPQ/UNIMEP pela bolsa de Iniciação Científica de Daniela Akagui e ao FAP/UNIMEP pelo apoio ao projeto “Construção de um Sistema de Realidade Aumentada para Desenvolvimento Cognitivo da Criança Surda”.

Referências

- [1] ARToolKit - Download. Disponível em:
<http://www.hitl.washington.edu/artoolkit/download.htm>. Acesso em: 06 mar.2004.
- [2] AZUMA, R.T. - *A Survey of Augmented Reality*. *Presence: Teleoperators and Virtual Environments* 6, 4 (August 1997), 355 - 385.
- [3] BERRE, A.; SYVERINS, T.; SALOMONSEN, M. - **ARToolKit for Dummies**. Disponível em:
<<http://rasmus.uib.no/~st02204/HCI/ARTUserDoc.doc>>. Acesso em: 06 fev.2004.
- [4] BILLINGHURST, M. - **Augmented Reality in Education**. Disponível em:
<http://www.newhorizons.org/strategies/technology/billinghurst.htm>. Acesso em: 06 mar.2004.
- [5] BILLINGHURST, M.; KATO, H., POUPYREV, I. - *The MagicBook: A Transitional AR Interface*. **Computers and Graphics**, November 2001, pp. 745-753.
- [6] WOODS, E.; MASON, P.; BILLINGHURST, M. - MagicMouse: an Inexpensive 6-Degree-of-Freedom Mouse. **Proceedings of Graphite 2003**, Feb, 2003, Melbourne, Austrália, pp. 285-286.

Rastreamento de Trajetórias para Treinamento com Realidade Aumentada

Fabiano Utiyama¹, Cláudio Kirner¹.

¹ Universidade Metodista de Piracicaba – UNIMEP
Faculdade de Ciências Matemáticas, da Natureza e Tecnologia da Informação
Programa de Pós-Graduação em Ciência da Computação

{futiuyama, ckirner}@unimep.br

Abstract

This paper presents the design of a system based on augmented reality to be used as a training tool. People use hands as a manipulation device, so that tracking the hand allows the system to acquire, store in a data bank and recover the path of the hand movements. This path can be analyzed, compared to others, showed as thin tree-dimensional object (cylinder or something like that) and advice the user to do training operations.

1. Introdução

A Realidade Aumentada (RA) consiste numa técnica avançada de interface computacional, que permite a sobreposição de objetos virtuais no mundo real. A RA enfatiza a visualização em conjunto com a interação, pois, com o uso de óculos ou capacete de Realidade Virtual, pode-se visualizar objetos virtuais junto ao mundo real, de maneira altamente realista, incrementando a percepção do usuário no uso de uma interface computacional [1].

Quando a interface de RA utiliza as mãos como elemento de interação, além de permitir a sobreposição de objetos virtuais no mundo real, ela possibilita a manipulação desses objetos com as próprias mãos, tornando viável o desenvolvimento de inúmeras aplicações que beneficiam o treinamento humano, motivando mais o usuário para as tarefas a serem cumpridas.

Este trabalho apresenta uma ferramenta para minimizar as dificuldades das pessoas que necessitam realizar um treinamento. Para isso, estão sendo utilizadas várias tecnologias e linguagens, como banco de dados, VRML e Java. A pesquisa foi gerada a partir de protótipos integrados simplificados, identificados pelos capítulos seguintes deste artigo.

2. Treinamento com Realidade Aumentada

A Realidade Virtual vem ganhando espaço na área de manufatura e treinamento, em função de sua versatilidade na produção e visualização de modelos realistas de produtos e da possibilidade de exercitar a montagem das partes.

No entanto, um problema, que vem dificultando sua larga utilização, é a necessidade de equipamentos especiais para a manipulação de objetos virtuais, como luvas, capacetes, rastreadores, e projetores especiais, dentre outros.

No sentido de minimizar esse problema, as áreas de manufatura e treinamento vêm recentemente usando Realidade Aumentada, que permite o posicionamento e manipulação de objetos virtuais no ambiente real. Ainda aqui, o problema dos equipamentos persiste, mas o uso de rastreamento óptico, visão computacional e a manipulação direta dos objetos virtuais com as mãos, ou dispositivos mais simples como uma vareta com bolinhas, vêm facilitando o trabalho e tornando-o mais acessível e de fácil utilização.

A indústria aeronáutica e automobilística já vem usando Realidade Aumentada para resolver problemas de montagem e treinamento há alguns anos e muitos outros segmentos estão descobrindo

essa aplicação. A área educacional é uma das que tem grande potencial de utilização da Realidade Aumentada, em função do elevado grau de informação visual animada que é possível de ser obtida.

O projeto ART - Augmented Reality for Training [4] tem por objetivo ampliar a capacidade de treinamento e usa visão computacional, animação, monitoração e avaliação das etapas do processo. O sistema usa visualização no monitor, vídeo e som, além da sobreposição de modelos gráficos 3D gerados com OpenGL.

O trabalho "Montagem Virtual usando Técnicas de Realidade Virtual" [3] constitui-se num esforço de pesquisa para criar um ambiente de projeto com montagem virtual. Ele analisa os problemas de montagem virtual para projeto e manufatura e detalha as questões envolvidas com montagem virtual, incluindo o problema da trajetória e do armazenamento das informações em banco de dados.

Outra ferramenta para treinamento com realidade aumentada foi desenvolvida para montagem de mobiliário [7]. O sistema usa marcadores reais para o posicionamento dos objetos virtuais, tem demonstrações animadas e permite o uso de dispositivo manual de ajuste de posição dos objetos virtuais. Além disso, o sistema permite o desenvolvimento de passos de treinamento, através de uma interface de autoria com elementos próprios para a montagem de mobiliário.

Outro sistema interessante, voltado para a área educacional é "Virtual Reality for Lego Mindstorms" [2] que procura simular um kit de montagem real, permitir a montagem de brinquedos mecânicos virtuais com o comportamento dos reais e sua utilização no mundo virtual. Com isto, espera-se que os estudantes aprendam os conceitos envolvidos com o processo de montagem e funcionamento dos brinquedos.

Percebe-se, portanto com essa amostra de trabalhos, que a realidade aumentada pode ajudar em muito as pessoas na tarefa de treinamento e ensino/aprendizagem.

3. Ferramenta para Treinamento com Realidade Aumentada

A seguir, são apresentadas as soluções adotadas para os principais elementos da ferramenta para treinamento.

3.1. Exibindo uma trajetória padrão

Para exibir uma trajetória armazenada no banco de dados, foi utilizada a linguagem VRML, juntamente com Java, que possibilita a realização da troca de dados entre o banco de dados e o mundo virtual. [5]

A aplicação Java se divide em dois arquivos: Servlet e Applet. Servlet é responsável por buscar e armazenar as informações no banco de dados. Applet se comunica remotamente com o arquivo Servlet, enviando as informações deste para o mundo virtual e vice versa.

As trajetórias estão sendo exibidas através de objetos tridimensionais, carregados a partir do banco de dados, como mostram as figuras 1 e 2.

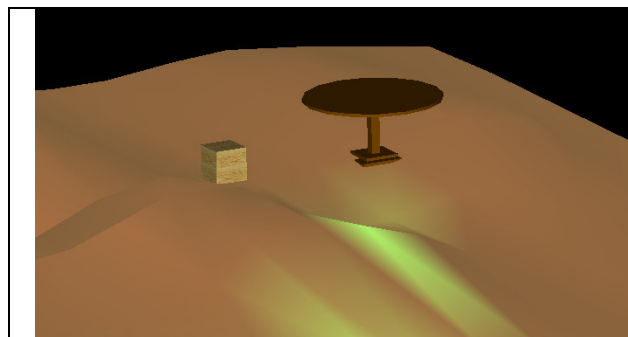


Figura 1: Cubo ao lado de uma mesa

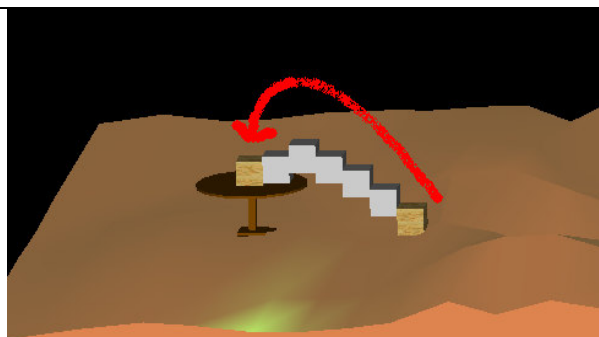


Figura 2: Trajetória do cubo

3.2. Capturando uma trajetória no espaço

Para a captura de uma trajetória no espaço, está sendo utilizado um outro sistema de desenvolvimento de RA, denominado ARToolkit, disponível gratuitamente no site do laboratório HITL da Universidade de Washington. [1]

Este sistema utiliza métodos de visão computacional, que ficam à procura de padrões na imagem do ambiente, capturada em tempo real por uma webcam, obtendo dados de posição e alinhamento.

3.3. Armazenando trajetórias no banco de dados

Para armazenar uma trajetória no banco de dados, está sendo utilizada uma técnica denominada Dead Reckoning. Esta técnica consiste em armazenar somente os pontos necessários de uma trajetória, evitando sobrecarregar o banco de dados [6]. Estes pontos são ligados por segmentos de retas, que irão compor o trajeto.

Na figura 3, observa-se a seleção do ponto da trajetória, no exemplo um cubo, a ser armazenado no banco de dados. Os pontos 1, 2 e 3 foram armazenados no banco de dados, já que suas posições alteram duas coordenadas, X e Y. Já o ponto 4, indicado pela seta azul, se mantém no mesmo segmento de reta do ponto 3. Sendo assim, o banco de dados ignora este ponto, passando a analisar o próximo. Quando aparecer um ponto, onde sua posição se altera em ao menos duas coordenadas, ele é armazenado. Está sendo usada uma margem de erro na avaliação das coordenadas para que pontos próximos de uma reta sejam considerados na reta, minimizando o armazenamento em excesso de pontos. A figura 4 mostra os pontos do trajeto armazenados no banco de dados, identificados pela textura nos objetos.

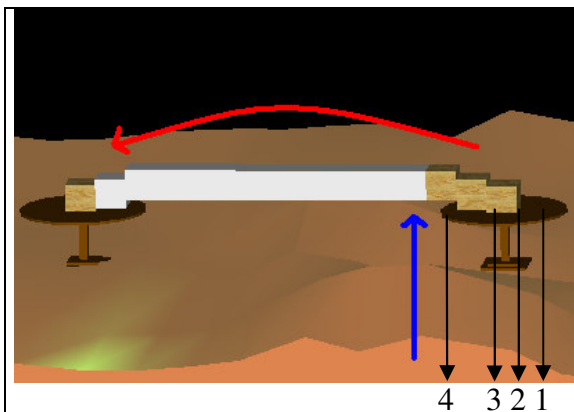


Figura 3: Selecionando o ponto da trajetória de um cubo a ser armazenado no banco de dados

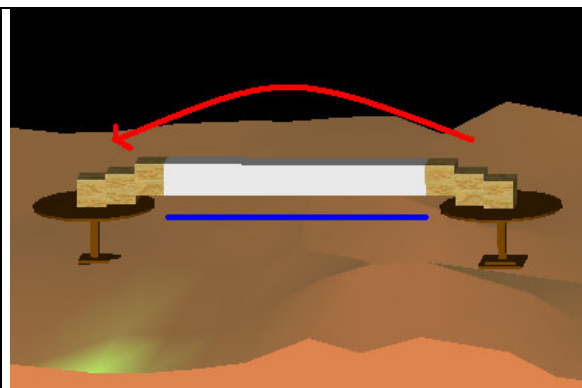


Figura 4: Pontos da trajetória de um cubo que foram armazenados no banco de dados

3.4. Junção dos protótipos simplificados e avaliação

Por fim, foram integrados todos os protótipos anteriores para se gerar uma avaliação do usuário. A trajetória seguida pelo usuário é comparada ao trajeto armazenado no banco de dados, obtendo-se assim uma avaliação quanto ao treinamento. Esta avaliação pode ser realizada, comparando ambos os trajetos ou verificando se o usuário alcançou a rota final proposta para o objeto.

4. Conclusão

A Realidade Aumentada tem contribuído significativamente para o processo de construção de conhecimento de muitos usuários. A possibilidade de visualizar e manipular objetos virtuais, a partir do

seu próprio mundo real, tem levado o homem a se aproximar cada vez mais da máquina. Este artigo abordou o desenvolvimento de uma ferramenta para treinamento, utilizando-se um ambiente aumentado, no qual a própria mão do usuário manipula e interage com o ambiente tridimensional, seguindo trajetórias pré-determinadas.

O que se pretende agora é melhorar e aprofundar a pesquisa. Por exemplo, programar um módulo que faça a captura de uma trajetória, substituindo assim o uso do ARToolkit.

Referências

- [1] AZUMA, R. et al. Recent Advances in Augmented Reality. **IEEE Computer Graphics and Applications**, v. 21, n. 6, p. 34-47, 2001.
- [2] BROWER, T. et al. - **Virtual Reality for Lego Mindstorms**.
< www.princeton.edu/~tbrower/Indie/VirtualRealityforLegoMindstorms.pdf > acesso em 02/04/2004.
- [3] CONNACHER, H. et al. - **Virtual Assembly using Virtual Reality Techniques**.
<www.nist.gov/msidlibrary/doc/cadtech95.pdf>. Acesso em 02/04/2004.
- [4] KUSTABORDER, J.; SHARMA, R. - Experimental Evaluation of Augmented Reality for Assembly Training. **Proc. of the Second IEEE and ACM International Workshop on Augmented Reality, IWAR'99**, Oct. 1999.
- [5] LEA, R.; MATSUDA, K.; MIYASHITA, K.- **Java for 3D and VRML Worlds**.
- [6] MORCRETTE, C. - **VRML Generation Tools for Visualization of Database Content in three dimensions**. Massachusetts Institute of Technology: 1999.
- [7] ZAUNER, J. et al. - Authoring of a Mixed Reality Instructor for Hierarchical Structures. **Proc. of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality - ISMAR'03**, 2003. 10 pp.

Desenvolvimento de Técnicas de Interação para Aplicações de Realidade Aumentada com o ARToolKit

Rafael Santin¹, Claudio Kirner².

¹Universidade Metodista de Piracicaba – UNIMEP
Faculdade de Ciências Matemáticas da Natureza e Tecnologia da Informação

²Universidade Metodista de Piracicaba – UNIMEP
Faculdade de Ciências Matemáticas da Natureza e Tecnologia da Informação
Programa de Pós Graduação em Ciência da Computação.

{rsantin,ckirner}@unimep.br

Abstract

This paper presents new interactive operations developed for augmented reality applications based on ARToolKit. These operations allow the development of innovative interfaces for human-computer interaction. The integrated use of multimedia and virtual reality technologies makes possible the intuitive manipulation of virtual objects, so that the user can superimpose virtual objects over real world and manipulate them using his hands.

1. Introdução

A Realidade Aumentada (RA) consiste numa técnica avançada de interface computacional, a qual permite a sobreposição de objetos virtuais no mundo real. Considerada uma variante da Realidade Virtual, a RA enfatiza a visualização em conjunto com a interação, pois, com o uso de óculos ou capacete de Realidade Virtual, pode-se visualizar objetos virtuais junto ao mundo real, de maneira altamente realista, incrementando a percepção do usuário no uso de uma interface de computador.[5]

Como a interface de RA, além de permitir a sobreposição de objetos virtuais no mundo real, possibilita a manipulação desses objetos com as próprias mãos, torna-se viável o desenvolvimento de inúmeras aplicações que beneficiam o trabalho humano, motivando mais as pessoas, inclusive aqueles com deficiências sensoriais, como os deficientes auditivos.

Em busca do desenvolvimento de aplicações que minimizem as dificuldades encontradas por estes deficientes, criou-se um projeto denominado Sistema de Realidade Aumentada para o Desenvolvimento Cognitivo da Criança Surda [3], dentro do qual insere-se este trabalho. Apesar do projeto ser mais extenso, este artigo restringe-se ao desenvolvimento de técnicas de interação. Para isso, está sendo utilizado um sistema de desenvolvimento de aplicações de RA, denominado ARToolKit, disponível gratuitamente no site do laboratório HITL da Universidade de Washington. [1, 2].

Esse sistema utiliza métodos de visão computacional, que ficam a procura de padrões na imagem do ambiente, capturada em tempo real por uma web-cam. Ao encontrar um padrão, são ativados módulos do aplicativo, responsáveis pela detecção da posição e do alinhamento do objeto virtual a ele associado. Na aplicação ARToolKit, esses padrões são placas com marcas fiduciais (Figura 1), as quais contém símbolos para diferenciar uma das outras, tornando-as singulares. Assim, ao enquadrar uma placa fiducial, já cadastrada no sistema, no campo de captura de vídeo, serão executados módulos que permitirão detectar a posição da placa, conforme o seu movimento e, conseqüentemente, “desenhar” o objeto virtual (estático ou dinâmico) alinhado a mesma, permitindo sua visualização sobreposta com a placa real nos dispositivos utilizados, seja o monitor ou o capacete

de Realidade Virtual (Figura 2). Assim, ao manipular a placa, o objeto virtual realizará os mesmos movimentos dessa, como se a ela estivesse preso [4].

Com isso, é possível desenvolver aplicações altamente motivadoras para que pessoas deficientes possam ser estimuladas a realizar determinadas tarefas por completo e com satisfação.

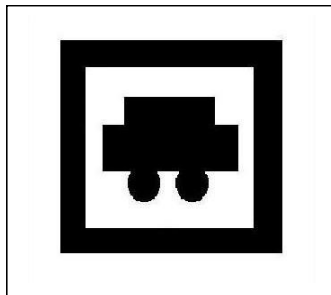


Figura 1: Placa com marcas fiduciais

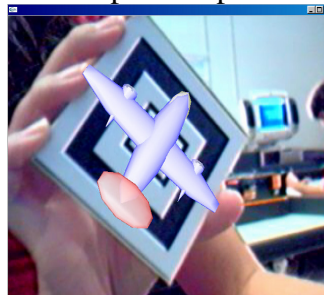


Figura 2: Objeto virtual sobreposto à placa

2 - Modificações do ARToolKit

Como o ARToolKit é uma ferramenta de código aberto, foi possível desenvolver modificações no código referente à interação da aplicação, como a alteração do objeto virtual, a partir da introdução de uma placa adicional de controle. Apesar do projeto estar ligado com aplicações para deficientes auditivos, percebeu-se que ele possui aplicações importantes e interessantes para pessoas normais, motivo pelo qual, decidiu-se implementar também uma modificação para incorporação de som.

2.1. Inserção da Placa de Controle

A alteração dinâmica do objeto virtual, a partir da introdução de uma placa adicional de controle, consiste na utilização de uma placa que possa interferir no objeto virtual de outra placa, de modo a trocá-lo ou alterá-lo, o que viabiliza a associação de mais objetos virtuais a uma só placa, de acordo com a necessidade da aplicação.

Para isso, foi necessário analisar o código do ARToolKit para descobrir como são realizadas as associações das placas com os objetos virtuais, a fim de desenvolver as alterações desejadas.

As associações de placas com objetos virtuais são, sucintamente, realizadas através de um identificador denominado `vrml_id` do tipo inteiro. Assim, foi criada uma variável inteira denominada `contr`, a qual é modificada com a expressão `contr=contr+1`, toda vez que a placa de identificador 0 (placa de controle) aparecer. Como o identificador está sendo somado com o `contr`, segundo a função `vrml_id+contr`, é possível visualizar vários objetos em uma placa, à medida em que a placa de controle for inserida. Como no exemplo da Figura 3, uma placa dessa aplicação poderá ter 5 objetos virtuais diferentes, pois a variável `contr` receberá os valores de 0 a 4, o quais serão adicionados ao índice (`vrml_id`). A troca de objetos pode implicar no aparecimento de objetos diferentes ou na alteração do objeto inicial com uso de escala, translações, rotações, animações, etc.

```
if(object->vrml_id==0) {  
    if(contr==4)  
        contr=0;  
    else  
        contr=contr+1;  
}  
if( object->vrml_id > 0 )  
    arVrml97Draw( object->vrml_id+contr);  
else  
    printf("unknown object type!!\n");
```

Figura 3 : Trecho do código responsável pela alteração dinâmica do objeto virtual

2.2 -Inserção de som

A inserção do som consiste na modificação do ARToolKit para que seja executado o arquivo de áudio, conforme a Figura 4, concomitante com o surgimento do objeto virtual no momento da introdução da placa em frente à câmera.

```
void som(void ) {  
  
    PlaySound("som.wav", NULL, SND_SYNC);  
    ret=0;  
    pthread_cancel (hThread);  
    pthread_exit(NULL);  
}
```

Figura 4: Módulo do som

Porém, ao inserir o módulo de execução de áudio, notou-se que a aplicação ARToolKit não funcionava corretamente, devido à lentidão na atualização da imagem capturada pela câmera. A solução encontrada para esse problema foi a criação de uma thread, como na Figura 5, o que permitiu a execução do módulo de som em “paralelo” de forma concomitante com os outros módulos gráficos. Assim, o problema de lentidão foi resolvido.

```
if (ret==0){  
    pthread_create(&hThread, NULL, (void (*)(void *))som, NULL);// Criação da Thread  
    ret=1;  
}
```

Figura 5: Criação da thread

3- Conclusão

Novas formas de interações em sistemas computacionais estão minimizando o abismo existente entre o homem e a máquina, de forma a facilitar a sua usabilidade. Consoante a esse contexto está a Realidade Aumentada, que possibilita a visualização e a manipulação de objetos modelados por computador no mundo real. Com isso, é possível desenvolver ferramentas que auxiliem o processo de construção do conhecimento do usuário, dada a possibilidade de explorar seus canais sensoriais com o uso de equipamentos adequados, como o capacete de Realidade Virtual, rastreadores e outros dispositivos [3].

Como o ARToolKit é uma ferramenta que auxilia o desenvolvimento de aplicações de Realidade Aumentada e por ser fornecido gratuitamente, ele foi escolhido para a produção de novas técnicas de interação de Realidade Aumentada, como as inserções de áudio e da placa de controle . Com isso, foi possível extrair aplicações mais dinâmicas, através de simples modificações na estrutura dessa ferramenta.

Outras modificações no software estão sendo estudadas, como a possibilidade do uso de um marcador que possa transferir um objeto virtual de uma placa a outra, funcionando como uma pá extraindo o objeto do seu lugar original, movimentando e despejando-o no lugar de destino desejado. Estão sendo desenvolvidas, também, outras aplicações que se beneficiam dessas interações entre o usuário e o sistema.

Agradecimentos

Os autores agradecem ao PIBIC-CNPQ/UNIMEP pela bolsa de Iniciação Científica de Rafael Santin e ao FAP/UNIMEP pelo apoio ao projeto “Construção de um Sistema de Realidade Aumentada para Desenvolvimento Cognitivo da Criança Surda”.

Referências

- [1] ARToolKit - Disponível em <http://www.hitl.washington.edu/research/shared_space>
- [2] AZUMA,R.et al. Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications*, v .21, n.6, p. 34-47, 2001.
- [3] DAINESE,C.A.; GARBIN,T.R.; KIRNER,C. **Sistema de Realidade Aumentada para Desenvolvimento Cognitivo da Criança Surda** In: Symposium on Virtual Reality, 6.,2003, Ribeirão Preto. Anais Proceedings 6 SVR Symposium on Virtual Reality: Ribeirão Preto: SBC, 2003. p.273-282.
- [4] KATO, H.; BILLINGHURST M.; POUPYREV, I. - *ARToolKit version 2.33 Manual*, Nov., 2003.
- [5] PINHO, M.S.; KIRNER,C.- **Uma introdução à Realidade Virtual**. Disponível em <<http://grv.inf.pucrs.br/Pagina/TutRV/tutrv.htm>.> Acesso em 4 Jan. 2004.

Autoria Colaborativa de Mundos Virtuais Educacionais com Realidade Misturada

Silvia C. Galana^{1,2}, Réryka Rubia P.C.L. Silva¹, Claudio Kirner¹

¹ Universidade Metodista de Piracicaba-UNIMEP
Faculdade de Ciências Matemáticas, da Natureza e Tecnologia da Informação
Programa de Pós Graduação em Ciência da Computação

² Faculdades Integradas Rio Branco e Centro Universitário UniSantana - São Paulo -SP
{scgalana, ckirner}@unimep.br , rerykarubia@yahoo.com.br

Abstract

Educational applications based on virtual reality is growing and becoming important to teach and learn. To make those applications real, it is necessary to develop suitable tools, that have characteristics like low cost and easy to be used. This paper presents the development of a collaborative virtual reality authoring tool based on mixed reality, so that remote users working together can manipulate virtual objects with his hands and generate scenes for educational purposes.

1. Introdução

O termo Realidade Virtual (RV) é um fascínio de igual proporção tanto para crianças como para os adultos. Estes últimos, principalmente, lotam qualquer sala de apresentação de um evento onde haja o tema citado. Isto prova que as pessoas, de forma geral, preferem o belo, o inusitado, o diferente, o desafiador, que as motivem e provoquem todos os seus sentidos.

A RV propõe, através de técnicas computacionais, uma “viagem” por um mundo virtual que imita o real ou não. Aliás, a proposta é ser realmente diferente, ousar, e até mesmo “brincar” com o real.

A curiosidade em torno do tema e a forma de utilizá-lo, ao mesmo tempo atrai e repele pretensos usuários. A proposta deste trabalho é justamente atender aos usuários leigos que pretendem criar e manipular ambientes virtuais para fins educacionais, mas não possuem conhecimentos técnicos específicos. Para isto, está sendo desenvolvida uma ferramenta de autoria que utiliza as mãos do usuário como mecanismo de interação e navegação. Os movimentos das mãos serão capturados e rastreados por uma WebCam, dentro de um ambiente de Realidade Misturada de baixo custo.

Este trabalho tem por objetivo inovar a interface utilizada para manipulação de objetos em um mundo virtual, usando os conceitos de realidade virtual e realidade misturada.

2. Realidade Virtual e Realidade Misturada

Em função do uso de novas tecnologias, a Realidade Virtual fornece um novo paradigma para o projeto de interface, no qual o usuário não manipula simplesmente botões ou imagens 2D, mas sim, sente-se imerso em um mundo repleto de objetos tridimensionais [4, 6].

Esta visão de interface permite desenvolver uma nova forma de atuação, ampliando os sentidos e dando um grau de realismo ao virtual até então não conseguido.

Segundo KIRNER [2], pode-se definir Realidade Virtual como uma técnica avançada de interface, onde o usuário pode realizar imersão, navegação e interação em um ambiente tridimensional gerado por computador, utilizando canais multi-sensoriais em tempo real.

O mundo virtual, também conhecido por cenário virtual, é composto por objetos modelados idênticos aos reais ou não. Aliás, a principal característica da Realidade Virtual é justamente não ter comprometimento com o real - pode-se criar à vontade, e proporcionar “ilusão” ao usuário.

A imersão é um sentimento único de sentir-se dentro do ambiente. Para isso, usam-se alguns artifícios como: capacete de visualização (hmd); salas de projeções nas paredes, teto e piso (caverna); som estéreo; captura da posição da pessoa e dos movimentos da cabeça; controles reativos; aromas, etc. A visualização no monitor é considerada não imersiva, pois permite o desvio do foco da atenção.

A interação está relacionada com a capacidade do computador em detectar as entradas do usuário e modificar instantaneamente o mundo virtual e as ações sobre ele. Este processo de interação é altamente estimulante para o usuário, que sente um certo fascínio ao poder manipular o mundo virtual. Certamente, com todas essas características, o grau de envolvimento e a motivação para usar RV são grandes, justificando a curiosidade das pessoas com ou sem conhecimento técnico.

Ao navegar por um mundo virtual, o usuário pode ter sua interação restrita à livre escolha de trajetória, ou plena, tocando em objetos e disparando ações.

A mistura do mundo real com o virtual, usando imagens reais capturadas com uma câmera ou outros dispositivos de visão, é denominada Realidade Misturada. Ela permite a inserção de objetos virtuais no mundo real (Realidade Aumentada) ou a captura de objetos reais e seu transporte para o mundo virtual (Virtualidade Aumentada). Para os dois casos, são utilizadas técnicas de rastreamento, interação, visão computacional, processamento de imagens e realidade virtual.

3. Sistemas de Autoria de Ambientes de Realidade Virtual

Existem vários trabalhos que visam dar ao usuário maior facilidade de manuseio na tarefa de construção de mundos virtuais. Como exemplo, tem-se o CVE-VM [2] que trata do desenvolvimento de um ambiente virtual multiusuário para aplicações colaborativas educacionais para uso pela Internet. O projeto possui uma infraestrutura para o desenvolvimento de ambientes virtuais temáticos, usando o construtivismo, na medida em que facilita a construção de mundos virtuais em grupo. Uma aplicação do sistema foi a criação de museus virtuais. O projeto foi implementado com o uso de JAVA, VRML e interface EAI, além de recursos adicionais para comunicação como Chat de texto e voz (figura 1).



Figura 1 - Interface do Projeto CVE-VM

Analisando-se as aplicações de Realidade Virtual na Educação, percebe-se um grande potencial do computador como ferramenta de ensino-aprendizagem baseada em modelos construtivistas. A relação da Realidade Virtual com o construtivismo reside na noção de imersão, experiências realizadas

em 1ª pessoa. Estas experiências constroem o conhecimento e, neste aspecto, a Realidade Virtual é singular.

As vantagens dessa tecnologia para a educação são significativas e a utilização desse ambiente à distância permite desenvolver o conhecimento para felicidade dos educadores.

Além disso, há também experiências de autoria de mundos virtuais com o uso de realidade misturada, como é o caso do Magic Paddle [3], em que uma espátula real é usada para transferir objetos virtuais e construir cenários tridimensionais. A figura 2 mostra a atuação da espátula no mundo virtual.



Figura 2 - Mobiliário sendo transferido de um catálogo virtual para uma sala virtual, usando uma espátula real (figuras capturadas do vídeo de demonstração Magic Paddle [3])

4 . Proposta da Ferramenta de Autoria de Ambientes Virtuais, usando Realidade Aumentada

O objetivo deste projeto é criar uma ferramenta de autoria de ambientes virtuais com RA, baseada no uso das mãos para a manipulação dos objetos virtuais. O rastreamento da mão, necessário para a interação com os objetos virtuais, é baseado no uso de uma WebCam e de uma pequena esfera colocada na ponta do dedo indicador do usuário. A densidade de pixels da esfera e seu deslocamento no plano permitem a obtenção do posicionamento espacial da ponta do dedo.

O uso desse ponto e da WebCam posicionada na altura de sua cabeça permitirá a interação de sua mão real com os objetos virtuais, dando-lhe a sensação de que sua mão adentrou no cenário virtual.

Os objetos pré-modelados estarão dispostos em prateleiras subdivididas por categorias e sofrerão alterações de brilho ou tonalidade para indicar que foram selecionados através de algoritmos de colisão.

O usuário do sistema poderá criar o seu próprio mundo, a partir destes objetos pré-modelados, usando como interface de comunicação e navegação este ponto capturado e rastreado pela WEBCAM.

Os objetos poderão sofrer alterações previstas em programação como rotação, translação e escala. O acionamento destas funções utiliza teclas de função do teclado convencional. No futuro, este teclado poderá ser também virtual ou com reconhecimento de voz.

Os comandos são: **F1 – pegar; F2 – mover; F3 – soltar; F4 – aumentar; F5 – diminuir; F6 – rotacionar; F7 – parar; F8 – salvar; F9 - animar.**

Ao compor o cenário com os objetos a seu dispor, o usuário poderá manipulá-los de acordo com as funções citadas acima, disparando inclusive uma animação previamente programada para cada objeto.

Ao terminar, ele pode salvar este cenário para que possa continuar a atividade em outro momento.

Trabalhos correlatos que usaram RV foram fundamentais para a compreensão do mecanismo de comunicação entre os objetos modelados graficamente e o usuário.

Ao pesquisar os recursos técnicos disponíveis para a construção do ambiente virtual com RA, foi possível detectar que a maioria dos trabalhos correlatos com RV utilizam a linguagem VRML com JAVA, através da comunicação com EAI.

No entanto, está sendo avaliado o uso de JAVA3D que deve resolver em princípio este problema de comunicação, interação e imersão com uma única ferramenta. Espera-se que essa linguagem traga algumas vantagens integradas como: programação orientada a objetos; multiplataforma; árvores de cenas; bibliotecas de comportamento dos objetos; vários ponto de vista da cena; reconhecimento de arquivos escritos em VRML; não precisa de plug-in específico; threads; renderização da cena que está sendo visualizada somente; som 3D; visão estéreo e suporte para dispositivos não convencionais de entrada como luvas e câmera [3].

5. Conclusão

A principal motivação deste projeto é proporcionar, aos professores leigos, uma ferramenta de trabalho de fácil utilização e de baixo custo, que os ajude a desenvolver suas atividades pedagógicas. Outros membros da equipe estão trabalhando para que este projeto tenha uma dimensão muito maior. A elaboração de um banco de dados, que armazene os movimentos efetuados pelo usuário, através do seu rastreamento, permitirá a análise das ações dos usuários no processo de aprendizagem.

A utilização da linguagem JAVA possibilitará a utilização do sistema em rede pela Internet, facilitando a implementação das características de colaboração.

O reconhecimento de voz permitirá o comando do sistema sem a utilização do teclado.

Referências

- [1] AZUMA, R.T. - A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* 6, 4, August 1997, 355 - 385.
- [2] KAWASHIMA, T. et al. - Magic Paddle: A Tangible Augmented Reality Interface for Object Manipulation. **Proc. of ISMR2001**, 2001, pp. 194-195.
- [3] KIRNER, C.; PINHO, M.S. - **Introdução à Realidade Virtual**. Apostila do Minicurso. XV JAI, XVI Congresso da SBC, Recife, PE, agosto de 1996, 60 pp.
- [4] KIRNER, T.G. et al. - Development of a Collaborative Environment for Educational Applications. **Proc. of Web3D 2001 Symposium**. Paderborn, Germany, Feb., 2001, pp. 61-68.
- [5] PERUZZA, A.P. et al. - **Java 3D: Aplicações Em Realidade Virtual**. In: CARDOSO, A. TEIXEIRA, C.; LAMOUNIER, E. - Ambientes Virtuais – Projeto e Implement. SBC, 2003. 49-78.
- [6] PINHO, M.S. - **Realidade Virtual como Ferramenta de Informática na Educação**. <http://grv.inf.pucrs.br/Pagina/Educa/educa.htm> (acessada em 02/04/2004).
- [7] TRINDADE, J. - *Realidade Virtual – Conceitos e Aplicações*. Escola Superior de Tecnologia e Gestão – Instituto Politécnico da Guarda. Guarda – Portugal. **II WORKSHOP Brasileiro de Realidade Virtual**. Marília – SP. Nov., 1999.

Interface de Comando de Voz Para Sistemas de Realidade Aumentada

Marcio Merino Fernandes
Universidade Metodista de Piracicaba
FCMNTI-Ciência da Computação, Piracicaba-SP
Programa de Pós Graduação em Ciência da Computação
mmfernan@unimep.br

Ednaldo Brigante Pizzolato
Universidade Federal de São Carlos
Departamento de Computação – São Carlos-SP
ednaldo@dc.ufscar.br

Abstract

It can be argued that virtual reality systems constitute the most sophisticated form of man-machine communication. Data manipulation and interactions are often carried out by using multimodal interfaces, such as mouse, special gloves, voice command, and hand gestures. In this short paper we present a research project to develop a framework to integrate multimodal interfaces into existing virtual reality systems, in particular the so called augmented reality systems. The on-going work is concerned with the integration of a voice recognition input interface into augmented reality systems for education.

1. Introdução

Sistemas de Realidade Virtual (RV) podem ser considerados como a forma de comunicação homem-máquina mais sofisticada atualmente disponível, possuindo um alto grau de imersão e interação com o usuário. Uma especialização de sistemas de RV é a chamada *realidade aumentada*, uma tecnologia que faz com que a visão que o usuário tem do mundo real seja *aumentada* com a inclusão de elementos adicionais criados pelo computador. Aplicações para realidade aumentada são encontradas em diversos domínios, como por exemplo medicina (cirurgias), industrial (manutenção de máquinas), arquitetura, decoração, etc. A manipulação de dados e informações multimodais é um fator chave para a coordenação adequada dos diversos elementos compondo um sistema de realidade aumentada, concentrando boa parte dos aspectos técnicos envolvidos. Uma tarefa fundamental no projeto de interfaces para realidade aumentada é o suporte para orientação, navegação e manipulação de objetos em um espaço tridimensional (3D). Interface multimodal é aquela capaz de suportar dois ou mais modos de interação com o usuário, como, por exemplo, mouse e comando de voz, estando normalmente integrada a um sistema multimídia [2]. O objetivo desta nova classe de interfaces é reconhecer de forma mais natural a maneira pela qual os seres humanos se expressam. Exemplos de modalidades incluídas nesse tipo de interface incluem desde dispositivos tradicionais como teclado e mouse, até outros mais sofisticados como reconhecimento de voz e gestos manuais.

Este artigo apresenta um resumo de um projeto em andamento, cujo objetivo é o desenvolvimento de interfaces multimodais que possam ser utilizadas em uma variedade de sistemas de realidade virtual e realidade aumentada, incluindo sistemas para ensino e treinamento, comércio eletrônico, etc. Na primeira fase deste projeto serão desenvolvidas interfaces baseadas no reconhecimento de voz, as quais deverão ser estendidas mais tarde para suportar o reconhecimento de linguagem natural. As interfaces desenvolvidas deverão integrar-se às interfaces gráficas

atualmente empregadas em sistemas baseados na linguagem X3D, ou em sua versão anterior, denominada VRML [6]. Numa fase seguinte outras formas de interface serão consideradas, como por exemplo, reconhecimento de gestos.

2. Arquitetura do Sistema

A suposição de que os ambientes virtuais considerados serão implementados usando X3D induziu a definição de que a interface de navegação deste sistema será baseada na funcionalidade oferecida por *browsers* X3D. Esta decisão visa facilitar a integração das interfaces multimodais desenvolvidas em sistemas existentes. Conforme mostrado na Figura 1, a integração de módulos de interface adicionais é feita através do uso de APIs para conectar a cena X3D a programas na linguagem Java. Isso é feito através da SAI (*Scene Authoring Interface*), a qual permite que aplicações implementadas em Java interajam dinamicamente com a cena 3D. A implementação pode ser na forma de nós do tipo *script*, ou ainda *applets* Java. Dessa forma, comandos externos ao *browser* podem ser enviados via SAI, e também informações de estado sobre a cena podem ser enviadas do *browser* à aplicação externa.

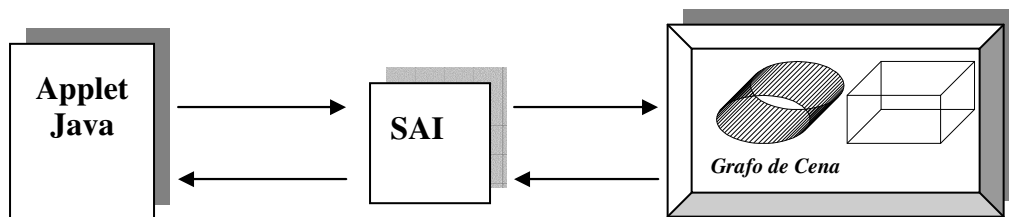


Figura 1. Interação de programa Java com grafo de cena X3D, via SAI.

Conforme mencionado, na primeira fase deste projeto pretende-se integrar um módulo de reconhecimento de fala ao sistema de realidade virtual. Isso pode ser particularmente útil em aplicações do tipo “mãos e olhos ocupados”, possibilitando ao usuário uma forma alternativa de enviar comandos ao sistema. Um exemplo desse tipo de interação ocorre em aplicações nas quais o usuário segura um determinado objeto (usando uma luva especial ou mouse) e solicita informações sobre esse objeto via comandos de voz. Usando essa estratégia o usuário poderia obter as informações desejadas geradas por um sintetizador de voz, sem contudo desviar a atenção da sua tarefa principal.

A integração do reconhecedor de voz a outros dispositivos de interface deverá ser feita conforme o diagrama da Figura 2. Essa integração ocorre externamente ao *browser* X3D, uma vez que em geral este não oferece suporte adequado para essa tarefa. A integração consiste basicamente na seguinte sequência de eventos: a) reconhecimento de comando de voz válido, através de um reconhecedor autônomo; b) transformação do comando no seu equivalente, através de um applet Java; c) envio do comando ao grafo de cena, via SAI; d) modificação correspondente do grafo de cena.

Deve-se notar que a possibilidade de uso de vários dispositivos de interface simultaneamente exige alguma forma de tratamento para a resolução de conflitos, determinação de prioridades, e outras situações resultantes da emissão de comandos concorrentes. Em nosso sistema o tratamento de conflitos será efetuado por um módulo *integrador*, encarregado de ler os diversos comandos de entrada e produzir um comando válido e não ambíguo. Este comando é então enviado ao grafo de cena X3D, via SAI. O integrador de comandos deverá ainda manter informações sobre o estado do mundo virtual, o que é necessário para verificar a validade dos

novos comandos enviados pelo usuário. Um exemplo desse tipo de arquitetura encontra-se descrito em [1].

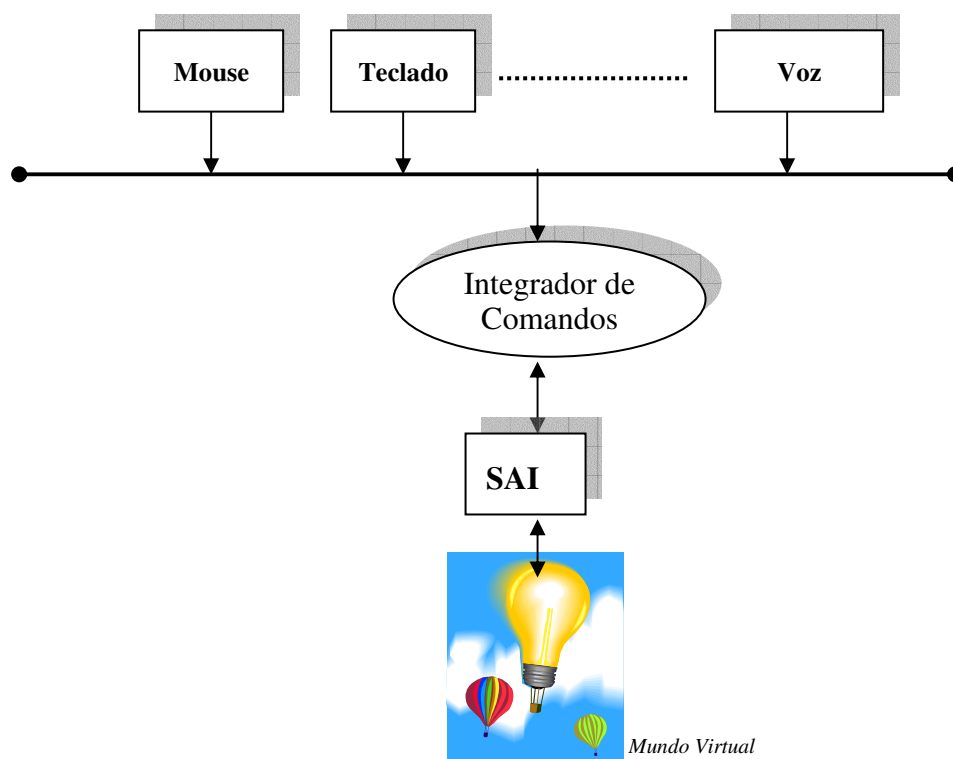


Figura 2. Integração de diversos dispositivos de entrada ao mundo virtual.

3. Interface de Comando de Voz

O reconhecimento de comandos de voz será feito através de um sistema específico para essa função. Reconhecedores de voz automáticos tem sido desenvolvidos há pelo menos 30 anos, porém apenas a partir da metade dos anos 90 alcançaram maturidade suficiente para uso em aplicações reais [3]. Os maiores desafios para a construção de tais sistemas são as estratégias de reconhecimento a serem empregadas, e a construção de uma base de dados contendo um vocabulário grande o suficiente para atender um determinado contexto. Além disso, as esperadas variações de pronúncia devem ser levadas em conta, caso contrário a eficiência do sistema pode ficar comprometida. Tudo isso faz com que a maioria das aplicações utilizando recursos de reconhecimento de fala utilizem pacotes desenvolvidos por terceiros, uma estratégia também adotada neste projeto. Inicialmente, dois sistemas existentes foram avaliados para uso neste projeto:

- IBM Via Voice [5]: Reconhecedor cuja precisão depende do usuário, ou seja, necessita de treinamento para adaptar-se à sua voz e sotaque. Possui um grande vocabulário embutido, e interface para *applets* Java.
- Nuance [4]: Reconhecedor que permite ao usuário modelar a tarefa de reconhecimento em estados, cada um deles sendo responsável por um segmento de conversação. O reconhecimento é feito dentro de um determinado contexto, porém não há a necessidade de treinamento com o usuário final. Também possui interface para *applets* Java.

As características do reconhecedor da Nuance são mais apropriadas para o tipo de interação e uso que ocorre em sistemas de realidade virtual, principalmente no que diz respeito à modelagem de diálogos dependentes do contexto. Por esse motivo, esse sistema foi adotado para o desenvolvimento do nosso primeiro protótipo da interface de voz para realidade virtual.

4. Conclusão

Este artigo descreveu de forma resumida um projeto em andamento para incorporação de interfaces multimodais em sistemas de realidade virtual implementados em X3D. Na sua primeira fase será incorporada uma interface de reconhecimento de voz, capaz de reconhecer e interpretar comandos dentro de um determinado contexto. Pretende-se também utilizar recursos de sintetização de voz, útil como forma adicional de comunicação do mundo virtual com o usuário. Em uma segunda etapa essa interface será aperfeiçoada para interpretar comandos em linguagem natural. Também estamos considerando a implementação de um outro tipo de interface, baseada no reconhecimento de gestos. Nesse caso, dispositivos especiais deverão ser desenvolvidos, possivelmente utilizando-se hardware reprogramável na forma de FPGAs.

Referências

- [1] F. Althof., H. Stocker, G. McGlaun, e M. Lang. A Generic Approach for Interfacing VRML Browsers to Various Input Devices and Creating Customizable 3D Applications. *Web3D-2002 - 7th International Conference on 3D Web Technology*, Tempe, USA, 2002.
- [2] S. Oviatt. Advance in Robust Multimodal Interface Design. *IEEE Computer Graphics and Applications*, USA, Setembro/Outubro 2003.
- [3] E. Pizzolato, M. Rezende. Issues to Consider When Adopting Commercial Speech Interfaces in Virtual Worlds. *SVR 2003 - 6th Symposium on Virtual Reality*, Ribeirão Preto, Brasil, 2003.
- [4] Speech Recognition. Nuance, <http://www.nuance.com/prodserv/prodspeechrec.html>, Abril, 2004.
- [5] Via Voice. IBM, <http://www-306.ibm.com/software/voice/viavoice>, Abril, 2004.
- [6] Web3D Consortium. X3D Documentation, <http://www.web3d.org/x3d>, Abril 2004.

Desenvolvimento de um Sistema de Aquisição de Informações Volumétricas usando Método de Triangulação Laser e Campo de Visão Variável

João Guilherme Darezzo Martins de França¹, Mário Alexandre Gazziro² e José Hiroki Saito³
Universidade Federal de São Carlos – UFSCar
Departamento de Computação
Programa de Pós-Graduação em Ciência da Computação
São Carlos – SP {jgfranca, mariog, saito}@dc.ufscar.br

¹São Carlos – SP, 13575-010, Rua Bispo César Dacorso Filho, 1085 - (16) 271-9873

²São Carlos – SP, 13566-550, Alameda dos Crisântemos, 60 Apto. 11C – (16) 261-6062

³São Carlos – SP, 13571-605, Rua Neide Sampaio, 182 – (16) 3368-4695

Abstract

The most efficient 3D acquisition system uses the principle of laser triangulation or the principle of time of flight – TOF. In the systems based on triangulation, the reach and the depth variation are very limited, but it has a great precision. On the other hand the opposite occurs to the TOF systems where it has a low precision, and a great reach and depth variation [1]. This work describes a project and implementation of a 3D acquisition system, with greater precision and reach than traditional laser triangulation 3D scanners, and it's as versatile as TOF systems. An initial prototype of this system had been developed showing the resolution of 0.19mm, 0.25mm and 0.06mm for x, y and z respectively at 20cm of distance between the sensor and the object. This prototype produces satisfactory results at 20cm to 2m of depth.

1. Introdução

A computação gráfica nos permite criar imagens realísticas do mundo 3D. Mas antes que essas imagens possam ser geradas, modelos gráficos do mundo devem estar disponíveis. Tradicionalmente esses modelos são obtidos por modelagem 3D, um processo muito demorado e limitado em realismo de detalhes. Esse trabalho descreve o projeto e construção de um sistema automático de aquisição de informações 3D de objetos reais usando sensores de distância. O sistema de aquisição de objetos 3D em questão é de alta resolução, e visa aplicações nos mais diversos segmentos civis e governamentais. Os benefícios deste projeto podem ser definidos pelo conjunto de possíveis aplicações e serviços: mapeamento de cavernas; inspeção de falhas estruturais em represas, pontes, torres e prédios; prototipagem rápida e reengenharia; mapeamento topológico [2]; construção de museus virtuais; estudo de como as estátuas foram criadas [3]; cálculo volumétrico, área de superfície e estimação de massa; imageamento da estação espacial internacional [4]; cinema; e animação 3D.

3. Descrição do Projeto e Construção do Sistema de Aquisição de Informações 3D

O sistema é constituído de vários módulos de hardware sob gerenciamento do software de controle que comanda o sensoriamento e o sistema de varredura. Quando termina a operação, o software de controle passa a função para o estimador de profundidade que transforma os dados brutos em uma nuvem de pontos que passa por um processo de remontagem.

3.1 Hardware

O hardware do sistema é composto de uma câmera de vídeo, uma placa de captura de vídeo, um projetor de laser, controlador de movimento, motor com *encoder*, redutor mecânico e um gerador de linha. A placa de captura é controlada por uma API de captura de vídeo e o laser é acionado pelo

controlador de movimento do sistema de varredura. Todos esses módulos são por sua vez acoplados a um microcomputador.

3.2 Varredura

O sistema de varredura deve cobrir toda a superfície da cena desejada. Para tanto, deve existir um controlador de movimento e um motor com *encoder* para cada um dos três eixos de varredura descritos a seguir. A câmera e o laser possuem controles independentes de rotação no eixo-Y (vertical) e rotacionam em conjunto no eixo-X. O controlador de movimento comanda o posicionamento do motor, a velocidade e a aceleração, pela interface RS-232 do microcomputador. O sistema também controla o campo de visão da câmera (FOV - *field of view*).

Inicialmente é feita uma aquisição prévia, com baixa resolução e FOV no máximo com a câmera apontando para o centro da cena para determinar a distância máxima D (entre a câmera e o objeto 3D). Para a determinação dessa distância máxima a varredura é feita apenas pelo laser, mantendo a câmera parada. Em seguida o FOV é ajustado de acordo com a definição desejada pelo usuário (quanto menor for o FOV maior a precisão e menor a variação de profundidade) e então é feita a varredura no eixo-Y (vertical) rotacionando a câmera e o laser em conjunto de modo que o ângulo do laser θ esteja relacionado ao ângulo da câmera α , como mostrado a equação 1:

$$\begin{aligned} \theta &= \tan^{-1}\left(\frac{D}{D/(\tan(\alpha - (\lambda/2)) + b)}\right) & \text{para } \frac{D}{D/(\tan(\alpha - (\lambda/2)) + b)} \geq 0 \\ \theta &= \tan^{-1}\left(\frac{D}{D/(\tan(\alpha - (\lambda/2)) + b)}\right) + 180 & \text{para } \frac{D}{D/(\tan(\alpha - (\lambda/2)) + b)} < 0 \end{aligned} \quad (\text{equação 1})$$

onde b é a linha de base, distância entre a câmera e o laser, e λ é o FOV. A movimentação é feita dessa forma para que o laser seja projetado nos *pixels* mais a direita da imagem (em um sistema com o laser do lado esquerdo e a câmera do lado direito). Como a distância máxima D é conhecida, o sistema nunca vai precisar fazer a aquisição em lugares mais distantes.

Após terminar a varredura de uma linha o sistema rotaciona o conjunto laser-câmera no eixo-X e continua fazendo a varredura no eixo-Y até cobrir toda a área da superfície. Em seguida o sistema determina as regiões que excedem a variação de profundidade d e atualiza a distância máxima D para $D-d$ e ajusta o FOV λ para que o campo de profundidade continue sendo igual ao d anterior. Se esse ajuste não fosse feito a variação de profundidade seria cada vez menor.

O sistema refaz a aquisição nas áreas mais próximas que $D-d$. Esse processo se repete até que não haja mais regiões fora do campo de profundidade. A aquisição é feita por camadas, aproveitando melhor a matriz de CCD (*Charged Couple Device*), obtendo-se maior precisão nas leituras de distância, maior alcance e maior variação de profundidade, além de serem necessários menos aquisições que os sistemas de triangulação laser convencionais, possibilitando maior versatilidade.

3.3 Software de determinação da informação 3D

O software de determinação da informação 3D é uma rotina que tem como entrada a sequência de vídeo, que inicialmente passa por um processo de correção de distorção [5] causado pelos parâmetros intrínsecos da câmera, depois passa por um processamento de imagens, onde a reflexão do

laser é segmentada da imagem, aplicando-se filtragem mediana para eliminação de ruído, detecção de bordas utilizando filtro de *Sobel* [6] ou *Canny* [7] (resultados melhores, porém a um alto custo computacional) e limiarização para tornar a imagem binária. A informação de profundidade é estimada, analisando o deslocamento horizontal de cada *pixel*, criando uma nuvem de pontos.

A câmera que tem comprimento focal f , deve ser localizado na origem a um ângulo α em relação a linha de base e o laser deve ser configurado de forma que emita o padrão de luz em um ângulo θ em relação à linha base (a distância entre o projetor e a câmera é a linha de base b). Um ponto (X,Y,Z) no espaço real é projetado na câmera no *pixel* (u,v) , de forma que as coordenadas reais dos pontos são dados pelas equações 2, 3 e 4, para as coordenadas X, Y e Z, respectivamente:

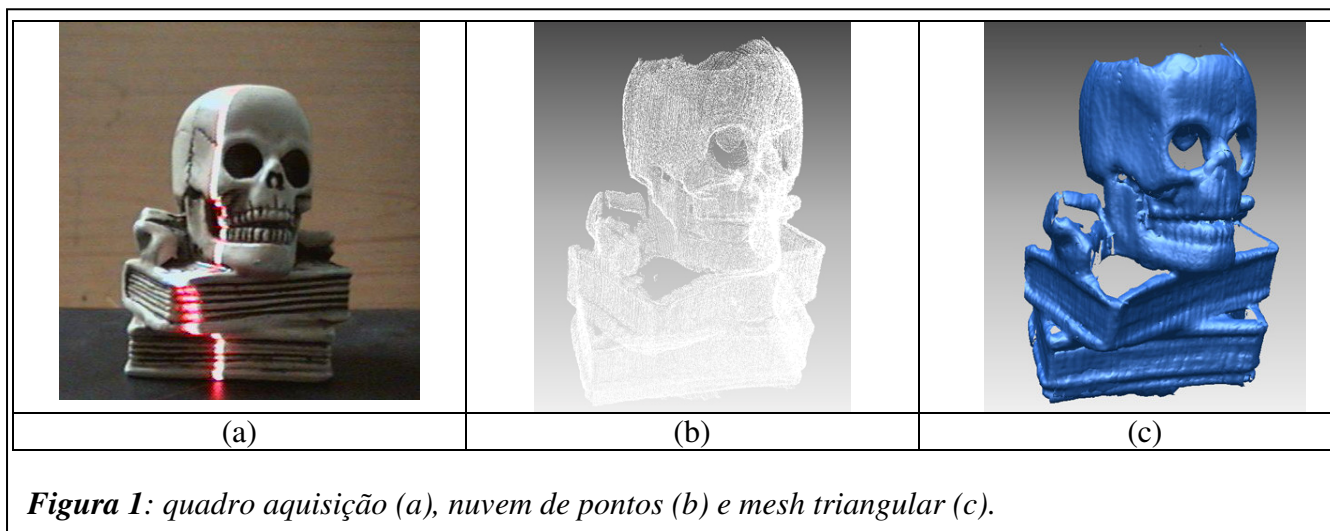
$X = \frac{u \left(\frac{b \cdot \tan \theta \cdot \tan(90 - \alpha)}{\tan(90 - \alpha) + \tan \theta} \right) / \sin(90 - \alpha)}{f \cot(\theta + 90 - \alpha) - u}$	(Equação 2)
$Y = \frac{v \left(\frac{b \cdot \tan \theta \cdot \tan(90 - \alpha)}{\tan(90 - \alpha) + \tan \theta} \right) / \sin(90 - \alpha)}{f \cot(\theta + 90 - \alpha) - u}$	(equação 3)
$Z = \frac{f \left(\frac{b \cdot \tan \theta \cdot \tan(90 - \alpha)}{\tan(90 - \alpha) + \tan \theta} \right) / \sin(90 - \alpha)}{f \cot(\theta + 90 - \alpha) - u}$	(equação 4)

Os parâmetros f , b , θ são conhecidos pela configuração do equipamento e a localização dos *pixels* u,v são encontradas por uma simples detecção de pico percorrendo a linha de varredura da imagem. As equações 2, 3 e 4, foram modificadas em relação às originais de triangulação[8] para poderem admitir rotação da câmera.

4. Resultados e Conclusões

Esse projeto foi desenvolvido em duas fases sendo que, na primeira, a varredura é feita apenas pelo laser e com campo de visão fixo. Esse protótipo já foi construído e testado. Atualmente ele utiliza um motor com *encoder* de 512 pontos e dois canais (que proporcionam 2048 posições) ligados a uma redução mecânica de 3,9:1 obtendo-se resolução de 0.04525° e laser de 650nm 5mw ligado a um gerador de linha de 60°. A resolução do sistema de aquisição é de 640x480 *pixels* a 30 quadros/segundo. A figura 1, mostra um quadro de uma sequência de aquisição de vídeo (a), que é utilizada para a criação de uma nuvem de pontos (b) e posteriormente passa por um processo de poligonização gerando uma *mesh* triangular (c). Esse protótipo produz resultados satisfatórios de 20cm a 2m e estará sofrendo modificações no motor que terá uma redução integrada de 879:1 e câmera com resolução ajustável entre 1280x1024, a 24 quadros/segundo, e 640x480, a 90 quadros/segundo. Como o laser e a câmera não são coaxiais, algumas áreas do padrão de projeção podem ser escondidas por outras superfícies que fazem parte do próprio objeto. Uma melhor definição pode ser feita pelo sensor aumentando-se a linha de base b . Entretanto, o problema de oclusão aumenta. Uma possível solução para esse problema é uma configuração com duas câmeras ou duas fontes de luz e integrar as informações. A segunda fase do projeto consiste do sistema com varredura completa e controle do

campo de visão da câmera. Essa fase está em estado de teste em ambiente virtual para determinação de segmento de aplicações e teste de algoritmos.



Referências

- [1] Rastogi, Pramod K. *“Optical Measurement Techniques and Applications”*, Artech House, Inc. 1997.
- [2] O. Amidi, T. Kanade and R. Miller, *“Vision-based Autonomous Helicopter Research at Carnegie Mellon Robotics Institute.”* *Proceedings of Heli Japan ‘98*. Gifu, Japan, April 1998.
- [3] Marc Levoy, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, Sean Anderson, James Davis, Jeremy Ginsberg, Jonathan Shade, Duane Fulk. *“The Digital Michelangelo Project: 3D Scanning of Large Statues”*, Siggraph 2000, Computer Graphics Proceedings.
- [4] Samson, C., English, C., Deslauriers, A., Chistie, I., and Blais, F. *“Imaging and tracking elements of the International Space Station using a 3d auto-synchronized scanner.”* National Research Council of Canada, April 2002.
- [5] Bouguet, Jean-Yves, *“Camera Calibration Toolbox for Matlab”*, Computational Vision Group - California Institute of Technology Publications, January 2004. http://www.vision.caltech.edu/bouguetj/calib_doc/
Acesso em 04/05/04
- [6] Gonzalez Rafael C., Woods, Richard E. *“Digital Image Processing”*, Addison-Wesley Publishing Company, 1992.
- [7] Jain, Ramesh. Kasturi, Rangachar. Schunck, Brian G. *“Machine Vision”*, McGraw-Hill, Inc. 1995.
- [8] Young, Tzay Y. *“Handbook of Pattern Recognition and Image Processing: Computer Vision Vol.2”*, Academic Press Inc, San Diego, CA, 1994.

VOID: An Endoscopy into a Black-Box

Sandro Canavezzi de Abreu¹, Otto Rössler², Mario Alexandre Gazziro³,
João Guilherme Darezzo Martins de França³, José Hiroki Saito³

¹ Rua Marechal Deodoro, 295 – CEP: 13320-140 – Salto – SP – (11) 4029-2759 – sandroid@mac.com

² Professor at the University of Tübingen. Fellow of the International Institute for Advanced Studies in Systems Research and Cybernetics.

³ UFSCar - Rod. Washington Luís, Km 235 - CEP: 13565-905 – São Carlos – SP – (16) 260-8576 – {mariog, jgfranca, saito}@dc.ufscar.br

Abstract

The “VOID” project (founded by the Langlois Foundation, Canada; and supported by the V2_Lab, at the V2_Org in the Netherlands) is a work in progress that intends to create an interaction that introduces the user to an observational situation that is continuously eroding the position of the user as a super-observer (an observer that doesn’t influence the object observed, as seen in the Classical Objectivity). The user, while indirectly investigating and observing the properties of a supposed internal reality accessed through an interface, is induced to realize that there’s nothing to see there despite the interface itself. The interface is its own observed volatile object, which is continuously disturbed by the presence of the user. In this event without horizon, its borders are the borders of the interface (which could make “palpable” the walls of the Cartesian prison?). Like this, it intends to reverse the situation of an object-centred stage to a contextual and observer-oriented one; objectivity to observerrelativity. From autonomy to covariance, from dictatorship of subjectivity to the world of the machine.

1. The VOID interface

The user will experience an augmented reality interaction with a hermetic object that is configured by 5 main elements:

- An acrylic sphere (whose internal surface is a mirror), 700 mm diameter;
- 8 sound speakers: they are attached outside the sphere, using the sphere as a resonant material.
- An endoscopic prosthesis, for visiting the internal space of the sphere; this prosthesis consists in a stereoscopic camera and light source (to provide images from inside), high sensitive microphone (to receive the sound from inside. The sound will be used as data for the system). This group of elements (the head of the prosthesis) is connected to the outside through an arm. This arm can be pulled/pushed toward inside/outside, can rotate and spin with full degrees of freedom.
- Second camera: static point of view from the interior part of the sphere.
- Projections: The visualization of the internal images will be provided by 2 projections (one is stereoscopic and the other is a simple “mono” projection. They are isolated from each other), showing 2 different aspects from the events inside the sphere.

Technically and briefly described, the dynamics of the installation “VOID” would work like described below.

Introduced inside the sphere there is a tube on whose extremity there is a stereoscopic camera. This camera is filming the internal surface of the sphere, showing its reflexes to the user. The images captured inside are synchronized and calibrated by digital means and then filtered through real-time video effects software. The result is shown on a big stereoscopic projection in front of the user. Attached to the camera there is a light source and a microphone. This microphone is receiving the sound output produced by the 8 speakers symmetrically distributed outside the surface of the sphere. These special speakers use the sphere as a resonance box. Its sound is generated by the presence of the

microphone. By moving the camera/mic/light, the user makes the mic and the speaker react to each other, producing a feedback effect. This feedback is used as material for triggering different mechanisms in the installation. For example: The feedback signal is read by the real-time video software that uses the sound variations (pitches and amplitude) to change the intensity of its effects.

In another room, a second image is projected. It comes from the second static camera based in the bottom of the sphere. This image has no digital distortion. Through internet, users can access a map with information about the activities inside the sphere.

In the process of showing itself to the user, the interface seduces him/her to follow certain time/spatial paradoxes [3]. These paradoxes are created by the intrinsic optical qualities of the materiality of the interface (reflexes), and artificially amplified by the digital processing of some traces from the user presence. In order to do that, the interface relies in a system that processes sound inputs, triggered by the user, that are converted into parameters for a visual output (stereoscopic video projection).

2.1 - Stereovision Technology

The technology for “VOID” is a experiment under development based on an extensive research through documentations that deals with stereo video made in post production (where recorded footage from two parallel cameras receives digitally filters and multiplexing by non-linear edition).

The lack of information concerning **real-time** stereo video images points to the absence of experiments with such technology. I believe that the reason is the difficulty to process a huge amount of data by digital means. Because of that, I decided to produce a hybrid system (analogical and digital) where the part of the process that is constant (RGB filtering and multiplexing) will be done analogically (this part in the systems that I have researched normally is done digitally). The part of the system that needs a dynamic processing, due to the continuous variation of parameters, will be done digitally. With such strategy, I can ensure a real-time interaction between the user and the system.

The anaglyph method for stereo vision consists on addressing separately each image from the left and right cameras to the left and right eyes, respectively [2]. In order to do that, the image that comes from the left camera will be filtered, by throwing away the red signal from the video signal, keeping the other 2 ones (green and blue = cyan). The image from the right camera will be filtered as well, but throwing away the blue and green signals and keeping the red one. Both images are then showed almost simultaneously to the user (the cyan and red images will alternate 60 times per second). By using red/cyan glasses, the eye with the cyan lens will access the red images. The eye with the red lens will access the cyan images. This combination will recover the stereo points of view from the cameras.

The challenge was to develop a analog system that should filter the signals from the cams and put then back together in only one signal, like showed at figure 1. This signal would then be sent to the computer to receive real-time digital effects. By centralizing the digital part of the video processing in only one computer I could avoid losing the sync between the cams. By keeping the filtering/adding process done by analogical electronic circuits external to the computer, I could save processing power from the computer. Like that, the computer only has to deal with the real-time effects (the output from the computer is not an interlaced image, which makes impossible the use of shutter glasses).

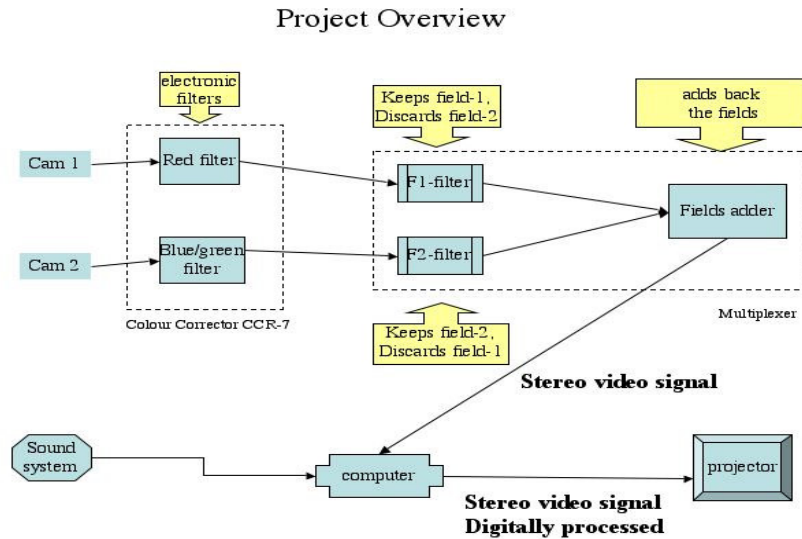


Figure 1 – Project Overview of the electronic components of VOID

2.2 - Spatial Paradoxes

The effect described above is one possibility for expanding the already embedded optical qualities of the concave geometry of the sphere. “Coincidentally”, inside the sphere some very similar spatial and cognitive paradoxes are also observed. The recursive reflexes inside the sphere produce a myriad of strange and counter-sense phenomena: when observed with stereoscopic vision the prosthesis seems to clone itself in several imperfect tridimensional “copies” (distorted, inversed and reversed by the spherical surface). Those copies really seem to be tridimensional ones. It is very often that an observer even can’t say which one is the “original” (PS: there were occasions that I looked directly into the sphere’s surface and it was quite hard to distinguish one image from another). These reflexes have a liquid behaviour: they look like floating in an ethereal fluid, shaping and dissolving themselves continuously. When using the stereo camera, the user could go through these swimming/melding entities that are “reacting” (or actually created) to (by) his movements.

When investigating such environment, I had a very high degree of difficulty to find a spatial orientation. The reflexes inversions (left/right, up/down) produce the following sensation: depending on the position of the camera, I could see an element in the left side of the image that interests me. I want to reach it, so I move the camera to the left, toward that element.

Something very strange happens: everything in the image moves to the right. It means that more left I wanted to go, more toward the right I moved myself. But there is even a better example listed below.

There is a special situation when you move the camera toward the reflexes and the reflexes simply moves back, getting away from you. You want to go closer and move the camera forwards but what actually happens is that you move backwards. This phenomenon happens if you move the camera from the surface (at you’re your back) toward the center of the sphere.

2.3 - Reference Swifting

Those cognitive paradoxes (that can’t be observed in plain or convex mirrors) inspired me to the

following supposition: If the memory would be independent of a spatial localization of the neuronal cells and its particles, like an outsider super-observer [1], when we hypothetically move everything (particles, sub-particles, sub-sub-particles, etc) that is in the left side toward the right side (including the brain's cells) and vice-versa, we could experience the phenomenon I described above: we would move to the right while "moving to the left" (the memory would recall the former state and positions of the particles and would put that in confrontation to what is really happening in that "now"). In "Quantum" terms [5], this reference's swift could also be understood as jumps into parallel possibilities of one's reality. In the case of this interface, it means that the user is not only seeing reflexes or copies of an initial original object but also moving his/her point of view into those copies perspectives. Like that, the images inversions and reversion of the space/object represent the virtual, the potential, the possibility, what haven't come to be realized yet, and maybe won't be [4]. They are not even repetitions provided by the reflexes. They are entities with their own "now", created by the user's presence. Therefore, when you look inside, what you see is not a mirror. There is no mirror. What you see is a rubber-like topologic hypersurface formed by possibilities, potential observers, even yourself as a potential observer, being distorted by the interference of your own observation [6].

This irresponsible inference lead me to the following decision for the interface I will work with: I could digitally intensify even more those analogical reflexes distortion to produce a more radical space/time malleability, making the idea of the "jumps" more tangible for the user. With this I don't want to be explicit, but only reinforce the physical qualities of this internal space, by subtly provoking a cognitive irritation in the user, during his/her attempt to investigate it.

3. Referências

- [1] Adriaansens, A. "The Art of The Accident", NAI Publishers/V2_Organisatie, Roterdã, 1998.
- [2] Ferwerda, J.G. "The World of 3-D, A Practical Guide to Stereo Photography", 1982.
- [3] Gombrich, E.H. "Arte e Ilusão", Ed. Martins Fontes, SP, 1995.
- [4] Mermin, D. "Quantum Mysteries for Everyone", in Science 217435, 1982.
- [5] Penrose, R. "A Mente Nova do Rei", Ed. Campus, 1991.
- [6] Rucker, R. "Infinity and The Mind", Ed. Birkhäuser, Boston, 1982.

Avaliação de Interface de um Sistema de Realidade Aumentada para a Criança Surda com base no Construcionismo

Tânia Rossi Garbin, Carlos Alberto Dainese, Cláudio Kirner, Anais Michele dos Santos,

Márcia Aparecida de Jesus

Universidade Metodista de Piracicaba - UNIMEP

Programa de Pós-Graduação Ciência da Computação

Curso de Psicologia

Curso de Ciência da Computação

{trgarbin, cdainese, ckirner, amsantosa, majesus} @unimep.br

Abstract

The main of this study is to evaluate Augmented Reality interfaces and their contribution to the development of hearing impaired children's cognitive. This work is based on theory of constructionism which primary idea is that the knowledge is not transmitting, but each person should to construct the knowledge through individual and collaborative experiences. It was utilized the Heuristic analysis to evaluate Augmented Reality interfaces.

1. Introdução

A possibilidade de interação entre imagens reais e virtuais que ocorre através da **Realidade Aumentada (RA)** pode oferecer ao usuário maiores informações sensitivas, facilitando a associação e a reflexão sobre a situação. Os sistemas de Realidade Aumentada permitem que o usuário decida sobre os ambientes, compondo cenas de imagens tridimensionais geradas por computador misturadas com imagens reais, aumentando as informações do cenário e oferecendo condições para a imersão no ambiente criado. “A principal característica destes ambientes é que as informações do mundo real são utilizadas para criar um cenário adicionado com elementos sintéticos gerados por computador” [3].

Segundo Valente [6], o computador é uma ferramenta que deve ser explorado para oferecer condições ao aluno desenvolver-se em um ambiente que favoreça a construção do conhecimento voltado para o “*aprender para a vida*” (p.39). Para Massetto [4], os ambientes educativos devem oferecer condições que possibilitem a criação, devem ser espaços agradáveis, a aula deve permitir aplicações práticas e a relação do conhecimento com experiências apoiadas e realidade do aluno.

Papert [5] não questiona o valor da instrução, mas enfatiza que o importante é que as crianças descubram por si novos conhecimentos e, para ocorrerem essas descobertas, é importantes a relação do novo com a experiência anterior. No construcionismo, a aprendizagem ocorre quando a criança está engajada e utiliza, de forma consciente, estratégias de resolução de problemas para a construção significativa.

Tendo por objetivo avaliar interface de ambiente de realidade aumentada para facilitar o desenvolvimento cognitivo da criança surda, esta proposta foi construída e estruturada a partir do referencial teórico do Construcionismo. A idéia central é que o conhecimento não é transmitido, cada indivíduo tem que reconstruir o conhecimento através de experiências individuais e coletivas. Para análise da Interface foi utilizado procedimento de análise Heurística.

1.1 Realidade Aumentada e a Intervenção com Crianças Surda

É muito freqüente a criança surda não ter comando sobre a linguagem oral. Assim, a formulação de perguntas verbais, por parte de profissionais, não é adequada, pois freqüentemente as crianças não entendem. Portanto, não podem demonstrar conteúdos aprendidos ou elaborações associativas e reflexivas. Deve-se, então, proceder observando os comportamentos, identificando como realizam operações complexas que necessitam de conteúdos abstratos, e verificar como ocorrem os processos cognitivos sem a representação da linguagem oral.

Três características podem ser consideradas relevantes em ambientes de Realidade aumentada:

1) é um ambiente virtual gerado sinteticamente por computador, que exige alto grau de realismo; 2) é interativo e 3) o usuário é imersivo no ambiente. A diferença básica entre sistemas de Realidade Virtual e Aumentada está no grau de imersividade. Em sistemas de realidade virtual o senso de imersão está sobre o controle do sistema, em contraste ao de Realidade Aumentada, o mundo real é aumentado e o usuário passa ser um elemento participativo, em que imagens virtuais são misturadas com reais para criar “senso de visão” aumentada.

A interface deve ser adequada considerando as necessidades do usuário. Desta forma, para a criança surda é necessário oferecer ambientes com recursos visuais que garantam e superem as necessidades auditivas. Todas as informações devem ser planejadas para que a criança consiga executar as atividades de forma motivadora. Os reforçadores utilizados devem garantir que a criança tenha interesse e, principalmente, retenha as informações e realize associação e reflexão sobre as atividades.

1.2 Breves Considerações sobre o Construcionismo

Construcionismo é um termo utilizado por Papert [5] para explicar como um processo educativo pode ocorrer de forma a garantir ao aluno a construção de um novo conhecimento. Papert destaca a importância das relações sociais e do engajamento da criança de forma a garantir o envolvimento nas situações novas considerando as vividas anteriormente. Valente [6] considera que no Construcionismo a criança constrói alguma coisa através do fazer, motivado pelo envolvimento afetivo.

Para Schulünzen [8] o ambiente Construcionista, deve ser favorável ao interesse da criança, deve ser um ambiente contextualizado e significativo. O problema deve emergir da necessidade dos usuários que devem decidir resolvê-lo com o auxílio do computador. Nesta abordagem, o profissional deve ter preparo para utilizar a tecnologia e aproveitar os recursos que as ferramentas podem oferecer de forma a garantir flexibilidade intelectual, capacidade de criar, inovar e, principalmente, enfrentar o desconhecido para promover o desenvolvimento.

2. O estudo sobre a interface do ambiente

Os métodos de avaliação de interface podem, ou não, envolver usuários. A Avaliação Heurística permite uma avaliação global da interface, sendo eficaz na detecção de erros e facilitando sua melhoria. Neste modelo destacam-se algumas regras: o sistema deve usar a linguagem do usuário e não termos técnicos; deve-se seguir um padrão no uso de palavras, ações ou situações; o design deve prevenir os erros; as informações contidas no sistema devem ter relevância para evitar competição de informações e conseqüente ambigüidade; ajudar os usuários a reconhecer, diagnosticar e corrigir erros, etc.

A opção por este método de avaliação está relacionada à investigação dos Sistemas de Realidade Aumentada, pois é uma área nova e que necessita uma ampla investigação **em todas as etapas do desenvolvimento**. Assim, avaliamos que seria mais adequado **não envolver o usuário criança surda**.

O ambiente de Realidade Aumentada foi desenvolvido utilizando o software Artoolkit, que é um aplicativo disponível na Internet. Para sua utilização é necessário cadastrar, através de uma câmera e vídeo (Webcam) algumas imagens de marcadores e associar, a cada um, um objeto virtual.

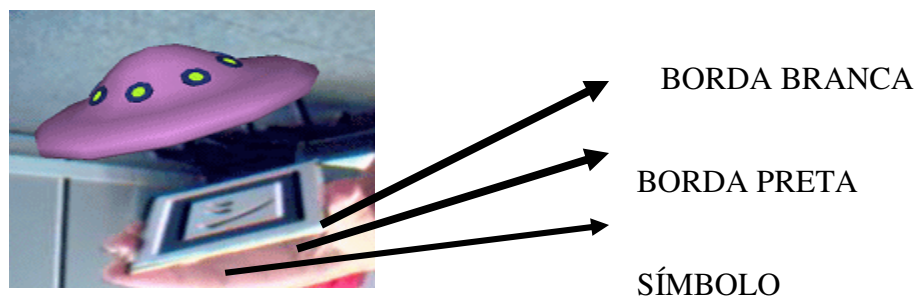


Figura 1- Visualização de objeto virtual e da placa no monitor.

Assim, quando se coloca a imagem de um marcador na frente da Webcam, o aplicativo faz o rastreamento (“leitura”) desta imagem e sobre esta se sobrepõe o objeto virtual associado a ela anteriormente. Para esta associação é necessário que o aplicativo capte a imagem por completo.

Objetivando estudar e avaliar a interface foram realizados dois experimentos. Inicialmente, foram utilizados os objetos virtuais do próprio aplicativo. As figuras virtuais utilizadas foram: a figura estática representada por um boneco de neve (Fig. 2), disco voador animado (3D), que fica suspensa e gira em círculo e um E.T. (3D) que balança a cabeça e mexe os braços. Após essa etapa inicial de reconhecimento do aplicativo, foram realizados testes com objetos (figuras) virtuais retirados da Internet. Estes objetos foram associados a imagens que não apresentavam conexão alguma entre eles. Os objetos utilizados foram: Homem Aranha, figura plana, tipo foto e estática; Dobby (elfo doméstico do filme Harry Potter), figura plana, estática e desfocada; Yugioh (personagem de desenho infantil), figura plana, estática, tipo recortada.

Posteriormente, as placas (imagens) continham determinado número, de 1 até 5. Estes números foram associados a figuras virtuais de personagens de desenhos infantis, retirados da Internet, de forma que o número relacionava-se a quantidade de personagens virtuais.

Nesta etapa foram utilizados placas de madeira nas mesmas dimensões do quadrado do Bloco Lógico (brinquedo infantil) para facilitar a sensação tátil da criança, e os símbolos (numerais) foram fixados aos blocos. (Fig. 3 e 4).

3. Discussão e trabalhos futuros

Constatamos que os fatores movimento e tipo de imagem podem tornar o ambiente mais agradável ao usuário. As figuras de personagens de filmes infantis, por apresentarem em suas características imagens estáticas e em duas dimensões, determinaram dificuldades na atenção e menor interesse para estimulação. Verificamos que as figuras em três dimensões animadas determinaram maior facilidade para ocorrer o processo de atenção e maior motivação quanto à interação.

Concluimos que o estudo dos fenômenos da percepção humana é importante no design de interfaces, isso porque o “usuário deve ‘perceber’ as informações apresentadas na interface através dos sinais que a constitui”[1].



Figura 2- Boneco de Neve

Figura 3- Homem Aranha

Figura 4. objetos relacionados ao numeral.

Verificamos que nos Sistemas de Realidade Aumentada é possível explorar a percepção visual e tátil através de experiências com equipamentos que garantam a imersão, mesmo sem a utilização de periféricos como óculos, capacete e luva utilizados em experimentos de realidade virtual, e podem garantir que o processo de abstração ocorra facilitando as fases do processo cognitivo. Outra vantagem do Sistema de Realidade Aumentada está relacionada ao custo, pois é possível o desenvolvimento de ambientes relevantes e interessantes para a criança utilizando materiais acessíveis e disponíveis nos laboratórios de informática das escolas.

A partir destes experimentos, deverão ser construídos ambientes interativos que garantam a abstração em atividades reflexivas para resolução de problemas complexos de ciências.

Referências

- [1] BARANAUSKAS, M. C. C.; ROCHA, H. V. *Design e Avaliação de Interfaces Humano-Computador*. Campinas – SP: NIED/UNICAMP, 2003
- [2] BATTAIOLA, A.L., ELIAS, N.C., DOMINGUES, R.G., ASSAF, R. e RAMALHO, G.L. Desenvolvimento de Interface de um Software Educacional com Base em Interface de Jogos. In: V Symposium on Human Factors in Computer Systems. Fortaleza – CE: SBC. *IHC 2002 V Symposium on Human Factors in Computer Systems*. Fortaleza – CE: SBC. 2002. p.214-225.
- [3] DAINESE, C.A., GARBIN, T.R. e KIRNER, C. Sistema de Realidade aumentada para o Desenvolvimento da Criança Surda. In: VI SVR – Symposium on Virtual Reality:, 2003. Ribeirão Preto - SP. *SVR 2003 VI SYMPOSIUM ON VIRTUAL REALITY*. Ribeirão Preto: SBC, 2003. p.273-281
- [4] MASETTO, M.T. Aula na Universidade. In: Ivani Fazenda (Org) *didática e Interdisciplinariedade*. Campinas: Papirus, 1998. p. 179-192.
- [5] PAPERT, S. *Amáquina das Crianças: Repensando a Escola na Era da Informática*. Trad. Sandra Costa. Sandra Costa. Porto Alegre: Artes Médicas, 1994
- [6] VALENTE, J.A. Aprendendo para a vida: o uso da informática na educação especial. In Valente, J.A. *Aprendendo para a vida: os computadores na sala de aula*. São Paulo. Cortez Editora, p.29-42, 2001
- [7] VALENTE, J.A. Informática na Educação Especial. In: Valente, J.A. *Liberando a mente: computadores na educação especial*. Campinas, S.P. Graf. Central da UNICAMP, p.62-79, 1991
- [8] SCHLÜNZEN, E. T. M. Mudanças nas Práticas Pedagógicas do Professor: Criando um Ambiente Construcionista, Contextualizado e Significativo para Crianças com Necessidades Especiais Físicas. São Paulo, 2000. Tese (Doutorado em Educação: Currículo) – PUC-SP.