

EXPERIMENTO DE FUSÃO DE DADOS EM REALIDADE AUMENTADA

Prof. Dr. Jorge Centeno¹
Prof. Msc. Glauber Acunha Gonçalves²
Eng. Agrim. Fernando César Silveira Abreu³
Eng. Cart. Patrícia de Castro Pedro⁴

¹ UFPR - Departamento de Geomática, centeno@ ufpr.br

² UFPR - Departamento de Geomática/ UFRG glauberac@ ufpr.br

³ UFPR - Curso de Pós-Graduação em Ciências Geodésicas, {abreu} { present_cp}@ ufpr.br

RESUMO

A Realidade Aumentada é uma área de pesquisa, dentro das técnicas da realidade virtual, que vem despertando crescente interesse devido a seu grande potencial de aplicação na solução de problemas que demandam a visualização de informações espaciais. Seu potencial como ferramenta para a visualização de informação cartográfica deve ser avaliado, sendo que neste particular encontra-se muito pouco trabalho. Neste artigo são apresentados os primeiros resultados de um projeto orientado à aplicação de realidade aumentada para a visualização de dados em cartografia. O projeto tem como objetivo o desenvolvimento de um protótipo de Realidade Aumentada para fins de visualização de informações cartográficas que sirva de apoio para trabalhos de campo. Neste particular, o artigo aborda a questão da fusão de dados artificiais e uma imagem natural.

Palavras chaves: realidade aumentada, processamento digital de imagens, fotogrametria digital

DATA FUSION IN AUGMENTE REALITY

ABSTRACT

Augmented Reality is topic, within the techniques of virtual reality, with increasing potential of application in the solution of problems that demand the visualization of spatial information. Its potential as tool for the visualization of cartographic information must be therefore evaluated. In the present paper we present the first results of a project aimed to the application of augmented reality for the visualization of spatial data in cartography. The project aims at the development of an prototype of Augmented Reality for visualization of cartographic information, as a tool that can be used in field works. The paper deals with a special topic within the problem, which is the the fusion of artificial data and a natural image.

1 INTRODUÇÃO

Trabalhos de campo, visitas a locais abertos, demandam a consulta de mapas atualizados. A identificação de locais e objetos existentes é realizada mediante a comparação da realidade visível com a informação disponível em um mapa ou banco de dados. Isto é normalmente feito usando mapas impressos ou imagens projetadas em uma tela de computador. Cabe ao usuário efetuar a correspondência entre as duas representações, estimando, da melhor maneira possível, e dentro da limitação dos recursos disponíveis, sua posição e orientação, para compatibilizar o mapa e a realidade. Esta tarefa é ainda mais complicada quando o usuário visita um local e conta com um mapa de objetos ou atributos dos mesmos, não visíveis, como por exemplo um mapa contendo futuras construções ou construções destruídas. Nesta situação, o usuário encontra dificuldade em compatibilizar sua visão com o mapa, o que dificulta a leitura do documento cartográfico e a tomada de decisão.

Uma maneira de contornar este problema é através do emprego de técnicas de Realidade Aumentada (Augmented Reality – AR). A Realidade Aumentada (AR) é uma atividade de pesquisa nova que explora várias aproximações para incrementar a visão do mundo real mediante a visualização de informação adicional junto com a cena natural. As técnicas de Realidade Aumentada permitem aos usuários examinar a realidade, objetos tridimensionais, junto com informação adicional previamente selecionada, relevante para a

solução de um determinado problema. Originalmente, a Realidade Aumentada encontrou campo de aplicação no desenvolvimento de jogos de computador de simulação. Nesta aplicação, não havia grandes exigências quanto à precisão do sistema, nem um grande compromisso com a realidade. O desenvolvimento de novos sensores de movimento, computadores com maiores recursos e melhores dispositivos de visualização abriu a possibilidade de estender a aplicação de AR para a solução de problemas reais. Hoje, as pesquisas apontam para a aplicação dos métodos de AR em muitos setores de indústria, por exemplo, para consertar ou produzir aparelhos sofisticados, para a inspeção de locais perigosos ou após um desastre, bem como para visualizar os eventuais resultados de simulações ou projetos ainda inexistentes.

No presente artigo são abordadas técnicas de fusão de dados para a visualização de imagens „híbridas“, superpondo dados vetoriais a uma imagem real, captada por uma câmara.

2 REALIDADE VIRTUAL E REALIDADE AUMENTADA

2.1 REALIDADE VIRTUAL

Realidade virtual é a aplicação de um conjunto de tecnologias digitais utilizadas com a finalidade de gerar um ambiente artificial onde o usuário possa ser “submergido” através de técnicas de visualização ou outros estímulos e com alto grau de interatividade. Este ambiente é composto por elementos, geralmente gráficos, gerados em computadores que são muito próximos da aparência dos objetos na cena real. Assim, uma cena artificial é gerada e projetada e o usuário, ao apresentá-la fica com a impressão de estar “dentro” de um ambiente tridimensional, com características muito próximas da realidade que ele conhece.

A origem do termo é atribuída a Jaron Lanier (Aukstakalnis e Blatner 1992) quem a conceituou como “um ambiente tridimensional interativo, gerado em computador, no qual uma pessoa é submersa”. Há três pontos de chave nesta definição:

- Trata-se de um ambiente artificial (uma cena tridimensional) gerado por computadores, o que demanda equipamentos de alto desempenho gráfico para prover o nível adequado de realismo.
- O sistema deve permitir a interatividade em tempo real.
- O sistema deve possibilitar a “submersão” do usuário, ou seja, a sensação de estar “dentro” dele e poder se deslocar dentro do ambiente e mudar o ponto de vista como ocorre numa cena real.

Exemplos de sistemas de realidade virtual encontram-se hoje disponíveis ferramentas para o desenvolvimento de projetos de engenharia e arquitetura e de documentação ou reconstrução de sítios históricos. Uma crítica à realidade virtual é o fato de que mesmo os modelos mais sofisticados não conseguem reproduzir perfeitamente a realidade em qualquer paisagem e que, na análise de uma cena natural pelo olho humano, o ambiente provê uma riqueza de informação que é difícil de modelar usando computadores. A cena apresentada na realidade virtual é geralmente uma simplificação da realidade.

2.2 REALIDADE AUMENTADA

A realidade aumentada encontra-se no meio, entre o “mundo real” e a “realidade virtual”. A realidade aumentada resulta da combinação da cena real, vista pelo usuário, com dados superpostos, que não são evidentes na análise visual comum. A natureza da informação varia em função da aplicação, mas sua principal característica é o fato de ser uma informação que o ser humano não percebe na cena e que, através da superposição dos dados se torna visível e ajuda na compreensão da cena. A situação ideal seria aquela na qual o usuário não conseguisse discriminar entre a cena real e a informação superposta, de maneira que para ele os elementos artificiais pareçam naturais. No entanto, tal grau de perfeição nem sempre é desejado nas aplicações de AR. Existem aplicações onde é desejado que a informação superposta seja evidenciada, destacada, de maneira a facilitar o desempenho do usuário dentro da cena natural. Isto pode ser feito, por exemplo, superpondo o nome das ruas durante a vista a um local.

Enquanto a realidade virtual se esforça em atingir a total submersão do usuário na cena artificial, a intenção da Realidade Aumentada é que o usuário não perca o vínculo com a realidade e que a perceba normalmente, apenas acrescida de novas informações. Como as imagens de AR resultam da fusão da realidade com dados artificialmente superpostos, é necessário um mecanismo que seja capaz de produzir esta fusão de maneira eficiente e em tempo real.

Milgram (Milgram e Kishino 1994; Milgram, Takemura et al. 1994) propõe a classificação dos sistemas de realidade aumentada usando três características: “Fidelidade de reprodução”, “Intensidade da sensação de presença” e “Grau de correspondência espacial com a realidade”.

- A *fidelidade de reprodução* se refere à qualidade da cena artificial ou a representação dos objetos gerada pelo computador. Pode variar de simples textos superpostos na imagem, passando por linhas ou os cantos dos objetos, até elementos tridimensionais com texturas realistas. Um dos principais condicionadores desta fidelidade é a capacidade do sistema de visualização atualizar a cena em tempo real.
- A *intensidade da sensação de presença* mede o nível de imersão do usuário dentro da cena exibida.

Esta propriedade é relacionada com a tecnologia de visualização. Há vários tipos de sistemas de visualização, com diferentes graus de submersão. Em um sistema de realidade aumentada, este conceito é ambíguo, pois não se pretende isolar completamente o usuário da realidade, mas sim deixá-lo perceber a realidade junto com um ambiente gerado artificialmente. Duas alternativas podem ser utilizadas: na primeira, o usuário percebe a cena com seus próprios olhos e na sua frente são projetadas as imagens artificiais superpostas. A alternativa seria projetar capturar a imagem com ajuda de uma câmara e projetar o resultado da fusão na tela de um computador.

- O grau de correspondência espacial com a realidade diz respeito à compatibilidade entre o sistema tridimensional da realidade (cena real) e o sistema tridimensional dos dados virtuais (imagem virtual). Do ponto de vista da realização, superpor uma imagem artificial a uma imagem real é uma tarefa relativamente simples. O problema principal consiste em manter esta correspondência na medida em que a interação do usuário é efetivada. A cada movimento do usuário, os dois sistemas devem ser deslocados paralelamente. Ou seja, mudanças de escala ou ângulo de visão efetuadas no espaço do usuário devem ser simuladas no espaço virtual, de maneira de que os objetos em ambos sistemas casem perfeitamente. Isto requer conhecimento detalhado da relação entre os sistemas de referência, a posição do usuário e sua orientação. Em situações onde a realidade é representada como uma imagem num monitor, este casamento é mais simples, pois é possível derivar informação da imagem, como cantos ou perfis (Leebmann, 2001). A situação é menos favorável quando os objetos artificiais são projetados na frente do usuário e a fusão se processa no cérebro. Neste caso, para obter um grau realista de submersão, o sistema deve ser informado a respeito do movimento do usuário e determinar o efeito deste movimento na cena que exibida.

O sistema proposto consta de um dispositivo para captura de imagens, que pode ser o olho do usuário ou uma câmara. Para fins de facilitar o entendimento será considerado o caso de uma câmara de vídeo ligada diretamente ao computador. Este dispositivo é encarregado de efetuar a projeção da cena tridimensional num espaço bidimensional no plano da imagem. Além dos parâmetros da orientação interior, é necessário conhecer também a orientação exterior do dispositivo, ou seja, sua posição no espaço objeto tridimensional e as eventuais rotações em torno dos três eixos. A orientação exterior da câmara pode ser determinada com auxílio de dispositivos auxiliares, como sistemas de posicionamento global (GPS) e rastreadores inerciais (trackers) sensíveis a pequenas variações angulares.

Dependendo do tipo de aplicação, o usuário precisa de acesso a vários tipos de dados, por exemplo: a aparência final de um edifício, quando a finalidade é construção e os usuários são arquitetos e engenheiros, ou a posição de determinados objetos, como hidrantes, quando a aplicação está associada a atividades de resgate em catástrofes. O sistema de AR precisa ter acesso à informação pertinente e apresentá-la adequadamente, por exemplo como um conjunto de linhas (wireframe), como objetos realisticamente representados, ou como textos flutuantes.

A integração realista de objetos virtuais em uma real cena requer que os objetos virtuais se comportem em modos fisicamente plausíveis: eles podem ocultar outros objetos ou serem escondidos por eles. Eventualmente, podem projetar sombra e não devem deslocar-se de maneira diferente em relação aos objetos reais. Os sistemas de AR precisam de descrições muito detalhadas do real ambiente para simular tais restrições físicas e sua atualização é uma tarefa complexa.

3 METODOLOGIA

Antes de projetar informação sintética sobre uma imagem, deve ser garantida a existência da correspondência tridimensional entre os objetos artificiais e a cena percebida pelo usuário. O registro geométrico de uma imagem virtual e uma cena real é necessário para representar as duas imagens simultaneamente. O mundo artificial pode ser representado usando um sistema euclidiano tridimensional, projetado no display. O ajuste requer o conhecimento da localização do usuário no espaço 3D euclidiano e os parâmetros exatos da calibração da câmara, quando uma câmara é usada (Azuma 1993; Azuma e Bispo 1994; Tuceryan et al, 1995).

Quando uma câmara fotográfica está sendo utilizada, as imagens geradas por este equipamento podem ser ajustadas com ajuda de pontos observados na cena. Este é um problema conhecido como "matching" e é resolvido procurando na cena feições, linhas ou pontos, que existam no modelo virtual. Através das coordenadas destes pontos em ambos sistemas, o ajustamento e as transformações necessárias podem ser estimados. O ajuste geométrico acurado requer que feições marcantes sejam identificadas na imagem e que as mesmas sejam convertidas ao espaço 3D da cena. Para isto, no entanto, torna-se necessário conhecer a posição destas feições com precisão. Neste experimento, a orientação exterior da câmara foi calculada através de pontos de controle levantados na área de estudo.

Determinar a posição com relativa exatidão é uma necessidade para garantir o eficiente funcionamento do sistema AR; pois ele deve calcular a posição relativa dos objetos e apresentá-los ao usuário. Os erros tolerados variam em função do equipamento utilizado. Erros maiores são tolerados em dispositivos de visualização transparentes, pois o olho é encarregado de compensar estes erros. Já em dispositivos onde a

fusão se realiza na tela da unidade de visualização, os erros tornam-se mais evidentes e a exatidão dos dados geométricos deve ser maior.

O mecanismo físico de geração de uma imagem fotográfica pode ser descrito em função dos feixes de luz que incidem no plano da imagem após atravessar o sistema de lentes. Considerando que tais feixes descrevem uma trajetória retilínea, chega-se a condição de colinearidade (Moffitt e Mikhail, 1980), conforme o conjunto de equações 1. Este equacionamento matemático é afetado, na prática, por outros fatores, como a qualidade geométrica das lentes. Por este motivo, o equipamento deve ser calibrado.

$$\begin{aligned} x_c &= f \cdot \frac{m_{11}(X - X_c) + m_{12}(Y - Y_c) + m_{13}(Z - Z_c)}{m_{31}(X - X_c) + m_{32}(Y - Y_c) + m_{33}(Z - Z_c)} \\ y_c &= f \cdot \frac{m_{21}(X - X_c) + m_{22}(Y - Y_c) + m_{23}(Z - Z_c)}{m_{31}(X - X_c) + m_{32}(Y - Y_c) + m_{33}(Z - Z_c)} \end{aligned} \quad (1)$$

A estimativa dos parâmetros para cada fotografia é obtida através de pontos de controle, com coordenadas fotogramétricas e coordenadas tridimensionais de campo conhecidas. No experimento aqui descrito, uma vez tomadas as fotos, utilizou-se o modelo fotogramétrico na calibração da câmara, correção dos erros sistemáticos da imagem e, finalmente, na determinação, através de uma recessão espacial, dos parâmetros de orientação exterior. Uma vez determinados, tais parâmetros são inseridos no modelo de coordenadas homogêneas para obter a projeção de qualquer ponto na cena.

Com isto, pontos, superfícies e objetos tridimensionais podem ser exibidos sobre a imagem, real, passando a ser visualizadas em conjunto com ela. Assim, gera-se uma imagem híbrida, onde parte da imagem é derivada da imagem digital real e outra parte corresponde a elementos artificiais superpostos. Algumas técnicas de processamento digital de imagens, como a fusão de imagens por componentes de cores IHS podem ser necessárias nesse processo final para dar realismo à cena.

4 EXPERIMENTO

O experimento de fusão de dados artificiais e imagens reais foi executado nas quadras de futebol do Centro Politécnico Jardim das Américas, da UFPR, em Curitiba. Uma câmara digital Sony DSC P50 , de 2.2 Mpixels foi utilizada para captar 5 imagens, com as quais se procedeu a calcular a recessão espacial. Para tanto, um levantamento topográfico com o uso de uma estação total, por irradiação, foi planejando e executado. Esses pontos compuseram a estrutura de dados das coordenadas planialtimétricas dos pontos de controle em um referencial geodésico local.

A estimativa dos parâmetros de orientação permite o posicionamento dos objetos virtuais dentro da cena e sua projeção no plano de imagem. Quatro tipos de objetos virtuais foram introduzidos. O primeiro é um cartaz de propaganda a ser colocado na margem do campo de futebol. O segundo elemento artificial é a demarcação das linhas do campo, pouco evidentes na imagem original. O terceiro elemento é uma pintura na parede de um edifício que aparece no fundo da imagem. Finalmente, o último elemento é uma pintura no teto do outro prédio. Estes elementos são visíveis na figura 1.

Cada um destes elementos foi então modelado de acordo à geometria derivada das fotografias e eles foram superpostos à imagem. Para diminuir o efeito de artificialidade dos objetos virtuais na cena, uma estratégia de fusão foi adotada, ponderando a intensidade da elemento virtual segundo as seguintes regras:

- Valor 0 – preserva o valor original na posição projetada.
- Valor 1 – substitui o valor original pelo valor da imagem virtual na posição projetada.
- Valores entre 0 e 1 – efetua a fusão dos valores originais e da imagem virtual.

O resultado dessa projeção pode ser visualizado nas figuras 1 e 2. Um detalhe da superfície da placa inserida na imagem pode ser visto na figura 3.



FIGURA 1 – Vista da fusão de elementos virtuais e reais



FIGURA 2 – Vista da fusão de elementos virtuais e reais



Figura 3 – Modelo criado para a superfície da placa publicitária

Os elementos presentes nas figuras anteriores foram visualizados de maneira a harmoniza-los com a cena real. Adotando o valor 1 para o fator de ponderação das imagens, é possível evidenciar um determinado elemento. Na figura 4, a demarcação da pequena área foi evidenciada, escolhendo uma cor saturada e adotando o fator de ponderação igual a 1.



FIGURA 4 – Vista da fusão de elementos virtuais com fator de ponderação igual a 1

5 CONSIDERAÇÕES FINAIS

A integração realista de objetos virtuais em uma real cena requer que os objetos virtuais se comportam em modos fisicamente plausíveis: eles podem ocultar outros objetos ou serem escondidos por eles. Eventualmente, podem projetar sombra e não devem deslocar-se de maneira diferente em relação aos objetos reais. Os sistemas de AR precisam de descrições muito detalhadas do ambiente para simular tais restrições físicas e sua atualização é uma tarefa complexa. No caso mais geral, esta tarefa é difícil de ser atingida, pois em aplicações práticas, o conhecimento total da cena é impossível. No entanto, em ambientes controlados, esta opção é viável, como foi mostrado no experimento. O grau de realismo que é dado aos elementos artificiais pode ser utilizado como elemento de comunicação. Por um lado, os esforços orientados à integração do elemento na cena, de maneira a misturá-lo completamente com os outros objetos de uma forma realista, preservando as cores e efeitos de iluminação naturais, são interessantes para dar um aspecto real à cena projetada no display. Por outro lado, estes esforços mascaram a presença destes elementos, o que pode não ser desejado. Em aplicações práticas, pode haver interesse em enfatizar a localização dos elementos virtuais. Por exemplo, na inspeção de redes de abastecimento subterrâneas. Neste caso, os elementos artificiais deverão ser representados de maneira a garantir alto contraste com a cena natural.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- AUKSTAKALNIS, S. e BLATNER, D. (1992) "Silicon Mirage - The Art and Science of Virtual Reality", Peachpit Press Inc., Berkeley, USA
- AZUMA, R. (1993) Tracking Requirements for Augmented Reality. *Communications of the ACM*, 36, 7. 50-51.
- AZUMA, R. e BISPO, G. (1994) Improving Static and Dynamic Registration in an Optical See-Through HMD. *Proceedings of SIGGRAPH '94* (Orlando, FL, 24-29 July 1994), Computer Graphics, Annual Conference Series, 1994, 197-204 + CD-ROM.
- LEEBMANN, J. Application of an augmented reality system for disaster relief. *International archives of photogrammetry, remote sensing and spatial information sciences*. Vol. XXXIV-5/W10.
- MILGRAM, P. e KISHINO, F. (1994) *A Taxonomy of Mixed Reality Visual Displays*, IEICE Transactions on Information Systems, vol. E77-D, no. 12.
- MILGRAM P, TAKEMURA H, UTSUMI A e KISHINO F. *Augmented Reality: A Class of Displays on the Reality Virtuality Continuum*. Telemanipulator and Telepresence Technologies. SPIE, 1994.
- MOFFITT, F.H e MIKHAIL, E.M. (1980). *Photogrammetry*. Harper&Row Publishers, Inc., New York, NY.
- TUCERYAN, M., D. S. G. et al. (1995). "Calibration Requirements and Procedures for a Monitor-Based Augmented Reality System." *IEEE Transactions on Visualization and Computer Graphics* 1 (3): 255-273.