

# UMA FERRAMENTA BASEADA EM REALIDADE AUMENTADA PARA VISUALIZAÇÃO DE PRÓTESES EM PACIENTES DE CIRURGIA

Isabela Santos Gonçalves<sup>1</sup>; Claudio Eduardo Góes<sup>2</sup>

<sup>1</sup>Bolsista PROBIC, Graduando em Engenharia da Computação, Universidade Estadual de Feira de Santana, e-mail: [belecomp@gmail.com](mailto:belecomp@gmail.com)

<sup>2</sup>Orientador, Departamento de Ciências Exatas, Universidade Estadual de Feira de Santana, e-mail: [cegoes@gmail.com](mailto:cegoes@gmail.com)

**PALAVRAS-CHAVE:** realidade aumentada, visão estéreo, malha poligonal.

## INTRODUÇÃO

Realidade Aumentada (RA) é definida usualmente como a composição de objetos virtuais tridimensionais (3D), gerados por computador, com um ambiente real, em tempo real, transmitindo ao usuário a idéia de que objetos virtuais coexistem com objetos reais no mesmo espaço [BIMBER & RASKAR (2005)]. O principal desafio é, a partir da imagem da cena real saber onde e como inserir o objeto virtual, respeitando posição, oclusão, iluminação entre outras restrições.

Esse tipo de abordagem já é utilizada em diversas áreas, das quais a medicina merece maior destaque, em função do uso intensivo de imagens para fins de exame, planejamento cirúrgico e pesquisa. Nesse contexto, a RA pode melhorar a noção espacial do médico, auxiliar no diagnóstico e ainda no planejamento e execução de cirurgias guiadas por imagem [PEREIRA (2003)]. Este projeto está focado na simulação de cirurgias craniofaciais, onde a RA pode ser aplicada na composição de imagens do paciente, capturadas por câmeras de vídeo, em tempo real, com imagens médicas (TC e MRI) e modelos 3D (extraídos dos exames ou de próteses).

Na primeira etapa deste trabalho, foi desenvolvido um protótipo capaz de identificar a face do paciente em um vídeo. Esse protótipo utiliza as bibliotecas ARToolKit [ArToolKit (2010)] e OpenCV [BRADSKI & KAEHLER (2004)], para suporte das técnicas de RA e Visão Computacional, respectivamente, utilizadas. O método de Viola-Jones [JONES & VIOLA (2003)], implementado na biblioteca OpenCV, foi usado para a identificação de faces (frontais) e seus principais elementos (olhos e nariz). A partir desses elementos, iniciou-se o processo de composição, utilizando a biblioteca ARToolKit.

Nos testes realizados para a ferramenta desenvolvida usando o método de Viola-Jones, percebeu-se que esse tipo de método é bastante sensível não só as características da base de dados de treinamento (tamanho e iluminação do ambiente, por exemplo), mas também aos parâmetros intrínsecos do próprio classificador. No caso da aplicação alvo, seria preciso ainda definir identificadores capazes de lidar com variações da orientação da face, fora da posição frontal.

A partir da experiência adquirida na primeira etapa do projeto, o objetivo dessa nova etapa é investigar o uso de uma abordagem geométrica para o problema de identificação e alinhamento de características faciais. A ideia básica é ajustar um modelo “neutro” de face 3D a imagem do paciente (inspirado em uma das etapas do trabalho de [BLANZ & VETTER (1999)]). Para auxiliar o processo de ajuste, será investigado o uso de técnicas de visão estéreo [TRUCCO & VERRI (1998)], como ferramenta auxiliar a obtenção de informações tridimensionais da face do paciente. A partir do mapa de profundidade gerado, o alinhamento poderia ser feito utilizando características geométricas dos modelos, ambos em 3D.

## MATERIAIS E MÉTODOS

Foi dada continuidade aperfeiçoando a aplicação desenvolvida no projeto anterior. Naquele projeto, já era possível o alinhamento e o acompanhamento de um modelo de máscara “neutro” sobre o rosto de uma pessoa, porém ainda eram necessários alguns ajustes, possibilitando a máscara acompanhar as mudanças de escala, rotação e translação que podem ocorrer no rosto da pessoa -durante o vídeo.

Após isso, foi realizado um estudo sobre as técnicas e algoritmos de visão computacional, mais precisamente a Visão Estéreo. Tal estudo permitiu a compreensão das principais características desta técnica, bem como suas dificuldades.

Na sequência, através de estudos mais aprofundados, constatou-se que a biblioteca OpenCV, a qual já estava sendo utilizada no projeto, possui funções que auxiliam no desenvolvimento de aplicações com visão estereo, mais especificamente na geração de mapas de profundidade.

Após esses estudos, juntamente com outros participantes do grupo de pesquisa a qual este projeto está vinculado, desenvolveu-se um módulo que a partir da imagem capturada por duas câmeras paralelas com deslocamento, gerando um mapa de profundidade, onde é possível diferenciar o rosto da pessoa do restante da imagem.

Com os resultados obtidos dos mapas de disparidade, iniciou-se a implementação de um módulo para gerar uma malha poligonal a partir do mapa de profundidade. Com isso é possível construir um modelo 3D para a face e assim ter todas as informações tridimensionais da face do paciente, eliminando a utilização dos classificadores desenvolvidos anteriormente para a detecção dos pontos da face.

É importante ressaltar que as implementações foram feitas, utilizando a linguagem C/C++ e as bibliotecas OpenGL (suporte gráfico 3D) e OpenCV (suporte a manipulação de imagens).

## RESULTADOS

Primeiramente, para melhorar o alinhamento e o acompanhamento da máscara, as coordenadas iniciais da face, dos olhos e do nariz são armazenadas, ou seja, as coordenadas de quando a pessoa está totalmente de frente e com o rosto alinhado na vertical. Então, a cada quadro é feito um cálculo para verificar as mudanças que ocorreram em relação às posições iniciais, mudanças tais como:

1. Rotação: houve a inclinação do rosto;
2. Translação: deslocamento lateral ou vertical do rosto;
3. Escala: a câmera está mais próxima ou mais afastada do rosto;

Essa verificação é válida para quando o rosto continua na posição frontal. Esse cálculo é feito com base nas novas posições dos olhos e do nariz. Com isso, aplica-se a coordenadas iniciais da face, a transformação correspondente a mudança de orientação, deslocamento e/ou escala. Dessa forma, têm-se o exemplo da escala da máscara desenhada acompanhando a escala do rosto, como pode ser visto na figura 1.

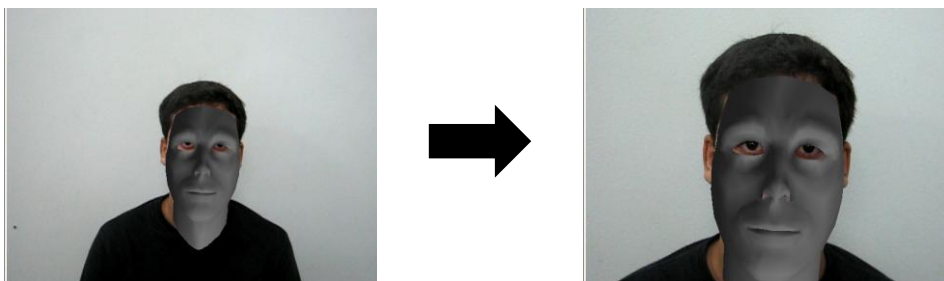


Figura 1 – Aplicação na máscara da escala correspondente à medida que a câmera se aproxima ou se afasta do rosto

Na etapa seguinte, foi desenvolvido um módulo de visão estéreo. Este módulo funciona a partir da imagem capturada por duas câmeras paralelas com deslocamento, afim de mimetizar a visão humana, gerando um mapa de profundidade. Os primeiros resultados obtidos, não foram satisfatórios, pois, não era possível distinguir claramente o fundo do objeto de interesse (neste caso a face de uma pessoa).

No entanto, depois de melhores ajustes na calibração da câmera e nos parâmetros intrínsecos e extrínsecos, os resultados melhoram. A figura 2 apresenta um exemplo de como os resultados melhoram, onde é possível distinguir melhor o que é o rosto da pessoa.

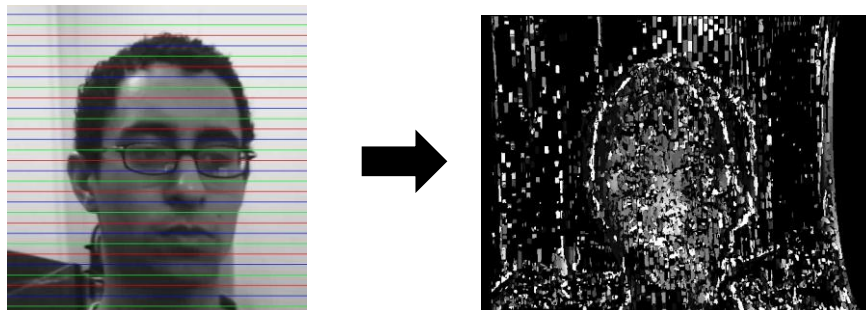


Figura 2 – Aprimoramento dos resultados da malha de disparidade obtida para o rosto de uma pessoa

Posteriormente, foi criado outro módulo para gerar uma malha poligonal a partir das informações do mapa de profundidade. Com isso é possível construir um modelo 3D para o rosto e assim ter todas as informações tridimensionais da face do paciente. Eliminando a utilização dos classificadores desenvolvidos anteriormente para a detecção dos pontos da face.

Esse módulo recebe uma imagem em escala de cinza ou então a matriz do mapa de profundidade gerado pela visão estéreo. Para ambos os casos aplica-se um algoritmo desenvolvido para a criação de triângulos a partir das informações da entrada, gerando assim a malha poligonal.

Para o desenho da malha, utiliza-se a biblioteca gráfica OpenGL, que utiliza um modelo de iluminação baseado em Phong. Neste modelo a iluminação dos objetos dependem da direção de incidência dos raios luminosos [SHREINER et al (2005)]. Essa direção é feita com base na normal das faces que compõe o objeto. Portanto, o passo seguinte ao algoritmo de triangulação, foi realizar o cálculo da normal para cada triângulo gerado.

Na figura 3 é mostrado o exemplo da malha poligonal para uma imagem de um mapa de disparidade obtido a partir do módulo de visão estéreo, com isso pode-se verificar a aplicação do módulo da geração da malha com o outro módulo desenvolvido. É importante ressaltar que para melhorar a visualização, coloriu-se a malha, onde as regiões de menor altura são representadas pela cor verde e as de maior altura pela cor vermelha.



Figura 3 – Exemplo da geração da malha poligonal para a imagem do mapa de disparidade obtida do módulo de visão estéreo

Além disso, no módulo de geração de malha foram desenvolvidas outras funcionalidades para melhor visualização da malha gerada. É possível girar a malha em todos os eixos (x, y e z), assim como aumentar ou diminuir o zoom.

## CONCLUSÕES

Em relação as melhorias propostas para o projeto anterior o objetivo foi alcançado. A utilização das coordenadas da face, olhos e nariz quando a pessoa está na posição frontal como referência para a composição do objeto sintético permitiu um melhor ajuste da máscara ao rosto da pessoa. Isso pode ser percebido principalmente quando ocorrem no vídeo aspectos como rotação, translação e escala em relação a tal rosto.

No que diz respeito ao que foi desenvolvido na segunda etapa do projeto, como apresentado nos resultados, foram desenvolvidos dois módulos separados, o de visão estéreo e outro para geração de malhas poligonal, porém eles ainda não foram integrados para atingir o objetivo principal do projeto, a composição de objetos sintéticos. No entanto, no decorrer do projeto como os estudos em relação à visão estéreo, surgiu a idéia de usar o *Kinect* [QUEIROZ (2010)] para a geração do mapa de disparidade ao invés das duas câmeras e da biblioteca OpenCV. Com isso espera-se obter um mapa de disparidade superior aos já obtidos, com menos ruídos e mais robusto. Com essa nova idéia e com a integração dos dois módulos, acredita-se obter uma ferramenta robusta, onde a composição do objeto sintético seja bastante precisa e que não haja limitações em relação à orientação da face do paciente, suprimindo, portanto as deficiências do protótipo desenvolvido anteriormente.

## REFERÊNCIAS

- BIMBER, O., RASKAR, R. 2005. Spatial Augmented Reality Merging Real and Virtual Worlds. A K Peters LTD, 3th edition.
- PEREIRA J.M.L.B. 2003. A realidade aumentada na engenharia biomédica: estado da arte. Escola superior de tecnologia.
- LAMB, P. 2010 [online]. Artoolkit Home Page. Homepage: <http://www.hitl.washington.edu/artoolkit/>.
- BRADSKI, G. R., KAEHLER, A. 2004. Learning OpenCV Computer Vision with the OpenCV Library. O' Reilly, 2th edition.
- JONES, M., VIOLA, P. 2003. Fast multi-view face detection. In MRL Technical Report TR2003-96, Cambridge.
- BLANZ, V., VETTER, T. 1999. A Morphable Model for the Synthesis of 3D Faces. SIGGRAPH, p:187-194.
- TRUCCO E., VERRI A. 1998. Introductory Techniques for 3d Computer Vision. Prentice-Hall.
- SHREINER, D. et al. 2005. OpenGL(R) Programming Guide: The Official Guide to Learning OpenGL(R), Version 2. 5th. ed. Reading, MA: Addison-Wesley Professional.
- QUEIROZ M. 2010. Um cientista explica o Microsoft Kinect. Homepage: <http://webholic.com.br/2010/11/09/um-cientista-explica-o-microsoft-kinect/>