

Agente Virtual Sinalizador de LIBRAS

Charles Marcel de Barros, José Mario De Martino (Orientador)

Departamento de Engenharia de Computação e Automação Industrial (DCA)
Faculdade de Engenharia Elétrica e de Computação (FEEC)
Universidade Estadual de Campinas (Unicamp)
Caixa Postal 6101, CEP 13083-970 – Campinas, SP, Brasil

{c023381, martino}@dca.fee.unicamp.br

Abstract – This project aspires to build a LIBRAS signaling virtual agent using the API/Library OpenGL and Cal3D. In the future, it can be an alternative to closed caption of digital television broadcasts, a webpages translator, or maybe a translator of spoken language to sign language for mobile devices such as laptops and PDAs.

Keywords – LIBRAS, Computer Graphics, Animation, 3D, Virtual Agent, OpenGL, CAL3D

1. Introdução

A Língua de Sinais Brasileira (LIBRAS) é uma língua de modalidade gestual-visual que utiliza gestos e expressões faciais para a comunicação, sendo reconhecida como a linguagem natural da comunidade brasileira de surdos. A sua disseminação exige o ensino e treinamento, assim como, o estímulo para a produção de material expresso nesta língua. Vídeo tem sido a mídia preferencial para a produção e a apresentação de material em LIBRAS, entretanto, como o avanço das técnicas de animação tridimensional por computador, a sinalização tornou-se viável também para personagens virtuais. Personagens virtuais são especialmente atrativos em contexto interativos ou caracterizados pela rápida atualização como, por exemplo, na televisão, Internet ou mesmo em terminais de atendimento automático.

Neste âmbito, este projeto ambiciona construir um agente virtual sinalizador LIBRAS que poderá colaborar no desenvolvimento de uma alternativa ao *closed caption* em transmissões da televisão digital ou talvez um tradutor de páginas web [1], uma vez que a leitura de textos extensos pode ser complicada ou demasiadamente desgastante para os surdos.

Pessoas que foram alfabetizadas antes de adquirir a surdez conseguem ler com fluência, porém o mesmo não costuma ocorrer quando são alfabetizadas após terem adquirido a surdez. Neste último caso, estudos mostraram que a performance da leitura é bastante defasada quando comparado às pessoas de audição normal [1]. Desta maneira, um agente virtual sinalizador LIBRAS poderá ajudar a comunidade surda não alfabetizada ou semi-alfabetizada na língua portuguesa a ter acesso a conteúdo até então não disponível em LIBRAS.

2. Proposta

Um agente virtual pode ser entendido como uma estrutura virtual tridimensional articulada que representa uma figura humana. Tipicamente a sua superfície é aproximada por uma malha poligonal tridimensional, a qual é deformada durante a animação através do deslocamento apropriado dos vértices dos polígonos que a compõem. Para o controle da animação, esta superfície deformável é associada a um esqueleto de controle, composto de um conjunto de segmentos (ossos) e juntas (articulações).

O princípio de *keyframing* (poses-chave) é comumente adotado como estratégia de animação em agentes virtuais, consistindo basicamente na especificação de um conjunto de poses principais, responsáveis por definir o movimento e a interpolação entre poses extremas para o estabelecimento de poses intermediárias. As poses intermediárias descrevem as transições entre as poses-chave, definindo a movimentação durante a animação. Em contraposição a sua simplicidade de implementação, esta estratégia apresenta fortes limitações. O espaço de poses realizáveis é definido pelo número e disparidade das poses-chave utilizadas. Uma pose fora deste espaço não é realizável.

2.1 Animação

Através do princípio tradicional de *keyframing*, criar, manter e gerenciar um conjunto significativo de poses-chave que cubra todo leque de poses é inviável, em vista a grande quantidade de sinais existentes na LIBRAS. Por este motivo este trabalho adotou como solução a especificação das poses-chave por composição em canais independentes [2], responsáveis por

compor as posturas a partir da seleção de primitivas, em contraposição ao processo tradicional de manipulação direta de articulações.

Esta abordagem de decomposição das posturas em primitivas possui a potencial vantagem de facilitar a criação das poses-chave que formam o sinal.

2.1.1 Representação dos sinais

A representação dos sinais é subdividida em três canais que especificam as configurações de dedos utilizadas na LIBRAS, o posicionamento/orientação dos pontos de articulação do braço e a trajetória executada nas poses intermediárias.

O canal de configurações de dedos é responsável por compor todas as possíveis posturas que as mãos podem adotar durante a sinalização LIBRAS, fazendo uso de abstrações como flexão, extensão, abdução, adução de dedos [3]. O canal de posição e orientação como o próprio nome indica, realiza o posicionamento do cotovelo, ombro e pulso no espaço de sinalização, juntamente com a orientação da mão (vertical com palma para cima, horizontal com palma para direita etc).

O posicionamento é realizado com ajuda de métodos de cinemática inversa e marcadores que cobrem todos os pontos relevantes do espaço de sinalização, desta maneira, basta selecionar um marcador para cada ponto de articulação e informar a orientação da mão. Estas informações quando combinadas com a configuração de dedos, permitem o estabelecimento de uma quantidade significativa de possíveis poses-chave e assim abranger uma ampla gama de sinais LIBRAS.

As poses intermediárias são obtidas através da interpolação de dois ou mais poses-chave devidamente seqüenciadas e temporizadas. Em muitos sinais, entretanto, observa-se trajetórias complexas entre determinadas poses-chave que seriam difíceis de reproduzir somente com os dois canais mencionados acima. Desta forma se faz necessária a utilização de um canal para modelar trajetórias, como por exemplo, transladar à mão de um ponto de sinalização para outro realizando movimento espiral.

2.1.2 Composição de Frases

Uma vez criado um conjunto de sinais é possível evocá-los seqüencialmente e assim compor frases. Este processo pode ser manual ou automático, sendo que neste último é necessário

que a seqüência de sinais e palavras que compõem a frase, seja fornecida por algum mecanismo como, por exemplo, sistemas de reconhecimento de voz, transmissão de legendas via sinal de televisão (*closed caption*) etc. Vale ressaltar que, assim como é utilizado um mecanismo interpolação na transição intra-sinal, o mesmo ocorre nas transições entre sinais. Observa-se, entretanto, que o principal objetivo da interpolação entre sinais durante a composição de uma frase é proporcionar uma transição suave entre os sinais, evitando assim um comportamento robótico não desejado.

3. Conclusões

Trabalhos de computação gráfica direcionados a LIBRAS são bastante escassos, ainda que algumas tentativas venham sendo feitas na área de tradução, nenhuma demonstrou ter uma representação visual suficientemente adequada. Grande parte das soluções apresentadas se mostraram inviáveis, devido ao dispendioso processo de gerar um vasto banco de dados de sinais, animando-os diretamente um a um com o auxílio de um programa de modelagem 3D convencional, ou ainda, insatisfatórias em termos de inteligibilidade, por se tratar de pictogramas 2D.

O trabalho aqui apresentado visa obter um diferencial na qualidade de apresentação dos sinais através da utilização de avatares 3d animados, como também viabilizar o dispendioso processo de criação dos sinais LIBRAS utilizando a estratégia de animação por composição em canais.

Referências

- [1] J.R Kennaway, J.R.W. Glauert, I. Zwitterlood. Providing Signed Content in the Internet by Synthesized Animation. *ACM Trans. Comput.-Hum. Interact.*, 14(3):15, September 2007.
- [2] C. Vogles, D. Metaxas. Handshapes and Movements: Multiple-Channel ASL Recognition. Em: *Lecture Notes in Computer Science 2915*: 247-258, , Springer, February 2004.
- [3] B. Yi, C. Harris Jr, L. Wang, Y. Yan. Real-Time Natural Hand Gestures. *Computing in Science & Engineering*, 7(3):92-96, IEEE Computer Society, May/June 2005..