

Sistema Gerador de Apoio a um Dicionário Temático Visual-Gestual Baseado em Realidade Virtual

José Remo Ferreira Brega
Departamento de Computação
Unesp - Bauru
remo@fc.unesp.br

Anthony Ferreira Lamarca
Departamento de Computação
UFMT - Barra do Garça
anthonyferreiralamarca@gmail.com

Mário Popolin Neto
Departamento de Computação
Unesp - Bauru
mariopopolin@gmail.com

Diego Roberto Colombo Dias
Departamento de Computação
UFSCar - São Carlos
diegocolombo.dias@dc.ufscar.br

Luis Carlos Trevelin
Departamento de Computação
UFSCar - São Carlos
trevelin@dc.ufscar.br

Resumo—A comunicação é uma necessidade humana, sendo a linguagem oral e escrita as maneiras mais comuns de comunicação. Os deficientes auditivos, no Brasil, utilizam a LIBRAS para se comunicar, contudo, a LIBRAS é uma língua pouco conhecida pela população em geral, o que dificulta a comunicação entre deficientes auditivos e não deficientes. Este artigo apresenta o uso de Humanos Virtuais em um sistema de multiprojeção baseado em aglomerados gráficos, o qual proporciona imersão e interação com um dicionário temático visual-gestual de LIBRAS, permitindo a interpretação e edição de movimentos por meio de dispositivos não convencionais de interação.

I. INTRODUÇÃO

A comunicação é uma necessidade humana, sendo a linguagem oral e a escrita as maneiras mais comuns de comunicação. A língua utilizada pelo indivíduo depende do grupo em que ele está inserido, por exemplo, sendo ouvinte, a comunicação se estabelece em termos oral-auditivos [1]. No entanto, para deficientes auditivos a comunicação se estabelece em termos gestual-visuais, no qual gestual significa o conjunto de elementos linguísticos manuais, corporais e faciais necessários para a articulação de um conjunto de movimentos.

A Língua de Sinais não é uma língua universal e, da mesma forma que a língua oral, é diferente em vários países devido à influência cultural, podendo até mesmo apresentar sinais que variam entre regiões e entre comunidades de deficientes auditivos. Assim como cada país tem sua língua oficial, cada país também possui sua Língua de Sinais padrão [2].

Quase que exclusivamente, o uso da Língua de Sinais tem sido feito por deficientes auditivos, de modo que pessoas sem tais necessidades tendem a não conhecer a língua. Este cenário complica a comunicação entre deficientes auditivos e não deficientes, afetando diretamente a inclusão social da comunidade surda. O desenvolvimento de ferramentas que auxiliem o ensino de Língua de Sinais para pessoas não deficientes possui grande apelo social, pois tende a auxiliar o processo de aprendizagem, permitindo que deficientes auditivos possam se comunicar de maneira efetiva com não deficientes.

Com a crescente evolução da Realidade Virtual (RV) surgem novas ferramentas que possibilitam a implementação e

visualização de ambientes tridimensionais, tornando possível o desenvolvimento de aplicações de baixo custo, visando alcançar maior interoperabilidade, que sejam imersivos e interativos. Assim, a RV possibilita o uso de Língua de Sinais em Sistemas de Realidade Virtual voltados a propósitos educacionais, os quais viabilizam alternativas computacionais para o aprendizado Língua de Sinais. Neste estudo é usado como estudo de caso a LIBRAS.

No aprendizado de uma nova língua é comum o uso de dicionários bilíngues da língua materna para a língua de interesse. Em sua pesquisa, [3] desenvolve dicionários temáticos visuais afim de, por meio de temas e ilustrações, facilitar o aprendizado da nova língua. Este trabalho apresenta um sistema de realidade virtual que, por meio de um aglomerado gráfico, proporciona interação e imersão em um dicionário visual-gestual de apoio ao ensino da língua brasileira de sinais, a LIBRAS.

Este artigo é organizado em 6 seções. Esta seção abordou a introdução; a seção 2 apresenta a LIBRAS e as características que a definem, assim como a integração entre LIBRAS e RV; a seção 3 apresenta a descrição de Humanos Virtuais e suas aplicações; a seção 4 apresenta a definição de dicionário temáticos e suas aplicações; a seção 5 apresenta o sistema desenvolvido e a metodologia empregada na pesquisa e desenvolvimento; e por último, a seção 6, a qual apresenta as conclusões.

II. LIBRAS

No Brasil, a Língua de Sinais LIBRAS foi reconhecida em 24 de abril de 2002 como meio legal de comunicação e expressão das comunidades surdas. A LIBRAS é dotada de toda a complexidade e utilidade encontradas nas línguas orais e, assim como elas, possuem gramáticas próprias, com regras específicas em seus níveis linguísticos, fonológico (configurações e movimentos dos elementos que estão envolvidos nos sinais), morfológico (formação de sinais), sintático (sequência dos sinais) e semântico (significado da sequência de sinais).

A. Fonologia

A fonologia analisa os fonemas de uma língua. Entende-se por fonemas unidades mínimas que estabelecem distinção entre

os vocábulos de uma língua. Na Língua de Sinais são utilizadas as partes superiores do quadril para a geração dos sinais, estes sendo produzidos por meio de expressões manuais (movimento das mãos, braços e antebraços) e não-manuais (movimento da face, dos olhos, da cabeça e do tronco). Por exemplo, os sinais TRABALHAR, BRINCAR e CONSERTAR são feitos no espaço neutro e os sinais ESQUECER, APRENDER e PENSAR são feitos na testa. Como pode ser ilustrado na Figura 1, existem palavras que se diferem entre si apenas pela diferença do ponto de articulação, como LARANJA (a) e APRENDER (b).

Quanto a representação computacional dos sinais referentes à constituição de LIBRAS, tem-se três parâmetros primários que são utilizados de forma combinada: a configuração das mãos, o movimento e o ponto de articulação [2].

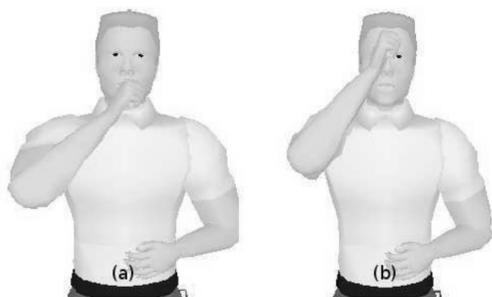


Figura 1. Configuração final do sinal LARANJA (a) Sinal APRENDER (b) [4]

B. Morfologia

A morfologia é o estudo da estrutura dos sinais (léxico) e a forma como os eles são formados. A formação nominal é realizada por meio da alteração do movimento do sinal do verbo.

A representação ilustrada na Figura 2, referente ao sinal TELEFONE (a) é igual à representação do sinal TELEFONAR (b), entretanto no substantivo tem-se o movimento representado por flechas, indicando "ir e voltar" do cotovelo e no verbo tem-se o movimento de uma curva direcional, indicando que o movimento será curvilíneo e unidirecional.

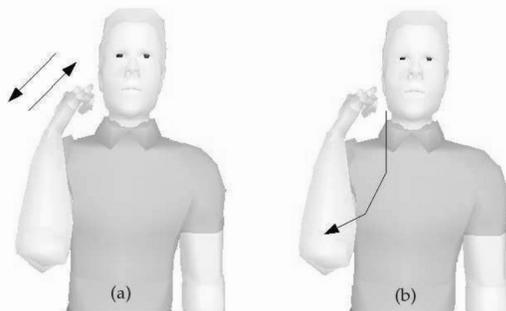


Figura 2. Formação do sinal do substantivo TELEFONE (a) e do verbo TELEFONAR (b) [4]

C. Realidade Virtual no Ensino de LIBRAS

Nos últimos anos vários aplicativos tem sido desenvolvidos com o intuito de auxiliar o processo de ensino-aprendizagem

de Línguas de Sinais. Esta subseção apresenta alguns estudos que combinam o uso de Língua de Sinais e RV.

Secco e Silva [5], apresentam um ambiente interativo para aprendizagem de LIBRAS. O ambiente possibilita a aprendizagem de LIBRAS gestual e escrita para surdos e ouvintes de diferentes faixas etárias. Para tanto, os autores fizeram uso de ferramentas comumente encontradas nos Sistemas Tutores Inteligentes (STI), utilizando também a Aprendizagem Baseada em Problemas (ABL) como estratégia pedagógica para a aprendizagem.

Allbeck e outros [6], desenvolveram um *software* que realiza a tradução da comunicação oral para a língua de sinais utilizando um interpretador virtual. O *software* funciona apenas para italiano.

Ligang e outros [7], propuseram um conjunto de palavras interpretadas em língua de sinais, as quais são representadas por um ambiente VRML. O intuito dos autores foi melhorar a coerência dos gestos por meio de um algoritmo de interpolação.

Coradine e outros [8], desenvolveram o FALIBRAS. O sistema apresenta animações (*Flash*) interpretadas a partir de comandos de voz.

Silva [9] apresenta um sistema *web* para representar gestos em LIBRAS, no entanto, provendo descrições relacionadas a forma como o movimento deve ser realizado. O ambiente valoriza duas estruturas gramaticais importantes para a LIBRAS, configuração de mão e a expressão facial.

III. HUMANOS VIRTUAIS

Humanos Virtuais (HV) são modelos usados como substitutos de humanos em testes ergonômicos baseados em computadores, utilizados em projetos de veículos, trabalho de campo, ferramentas, linhas de produção, para simulação prévia antes da construção, ensino de tutoriais e para representação humana em Ambientes Virtuais (AV) de tempo real [10]. Se necessário, HVs podem apresentar uma aparência bastante realista, baseada em princípios fisiológicos e biomecânicos. A Figura 3 apresenta um modelo de HV.

Na utilização de HVs deve-se levar em consideração os diversos domínios de utilização e os relacionamentos deles com os demais seres humanos pertencentes ao mundo real. Dentre os domínios estão: análise de fatores humanos (altura e comportamento), diferenças entre humanos (características próprias), geração e compreensão de instruções (comunicação de sensações e experiências transportada aos HVs), simulação bio-médica (estrutura físicas e funcionais) e análise de forma e movimento (entendimento sobre as formas geométricas) [11].

Em relação à modelagem, o projeto de HVs deve levar em consideração alguns critérios importantes, tais como: aparência (caricatura até modelo físico), funcionalidade (ações), tempo, autonomia (ações inteligentes), individualidade (personalidade), atributos físicos e biomecânicos (limitações funcionais), movimentação (captura de movimentos) e inteligência (tomada de decisão) [10].

Os HVs podem ser classificados em duas categorias: Agente Virtual e Avatar. Agente virtual é um HV desenvolvido e controlado por software, enquanto um avatar é um HV

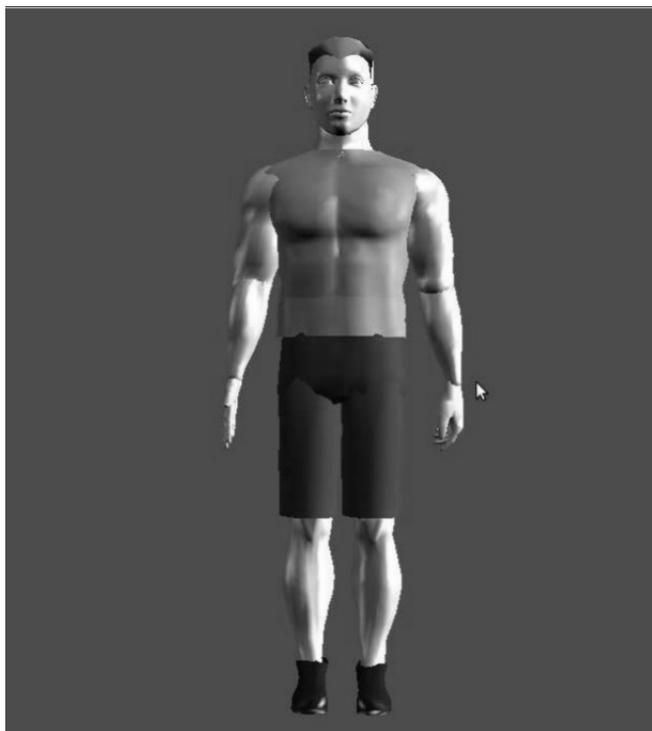


Figura 3. Humano Virtual

controlado por um usuário do mundo real [10]. Segundo [12], avatar é uma representação gráfica de um indivíduo real no AV, combinada com o seu comportamento.

O movimento é um aspecto importante para a representação de HVs, visto que este deve ser tão real e coerente quanto possível [12]. Para isso, na representação dos movimentos, existe a necessidade de especificar um grande número de juntas e detalhamento na modelagem, de modo que fique de acordo com as exigências da aplicação.

A interatividade é utilizada para manipular o HV, sendo ele avatar ou agente virtual. A interação deve especificar movimentos de maneira simples, por meio de rastreamento de sensores, rastreamento de resultado final (cinemática inversa), invocação externa (menus), sintetizadores de voz ou botões de ação (sintetização/interpretação de dados pré-armazenados) [13]. Sendo admissível também a realização do movimento por meio de dispositivos de entrada de dados, como mouses, teclados, luvas, entre outros.

IV. DICIONÁRIO TEMÁTICO

Um dicionário temático é classificado como um modo de organizar vocabulários específicos de determinada ciência, arte ou atividade técnica [14]. O dicionário também é caracterizado por prover informações léxicas sobre um conjunto de palavras. A par da importância capital do léxico de uma língua, encontram-se inúmeras dificuldades em se adquirir, memorizar e utilizar de modo adequado o conjunto de palavras do qual se pode dispor, conforme as necessidades de expressão e/ou comunicação [3].

A aprendizagem do léxico se divide em duas partes: a apresentação e a semantização. Na primeira, ocorre quando

o aluno vai da forma ao sentido ou do sentido à forma. A segunda parte pode acontecer de três maneiras: utilizando-se a Língua Materna; utilizando-se a Língua Estrangeira; utilizando-se meios não-verbais [15].

O ensino do vocabulário é um processo gradativo e que deve ser associado aos contextos das áreas de especialidade. Os alunos devem ser estimulados a aplicar as definições e designações adquiridas em determinadas situações a outros contextos, redirecionando, adequando, deslocando, articulando a linguagem em diferentes ocasiões [16].

A finalidade da pesquisa de [3] foi viabilizar a elaboração de uma série de dicionários temáticos ilustrados bilíngües, cada uma das direções "português-língua estrangeira" visou três fases distintas da aprendizagem: a inicial, a intermediária e a avançada, com o propósito de apresentar uma nova proposta de dicionários ilustrados que estimulem a codificação da comunicação na língua estrangeira e não apenas sua decodificação.

Confirma-se ser vantajosa e positiva a utilização de materiais visuais, as palavras apresentadas com ilustrações eram assimiladas ao final de um mês de pesquisa, enquanto que as sem ilustrações eram assimiladas somente depois de três meses [17]. É notória a melhoria no processo de cognição relacionado ao conteúdo apresentado fazendo uso de representações gráficas, visto a capacidade natural do ser humano em fazer associações.

V. DICIONÁRIO TEMÁTICO VISUAL-GESTUAL

O sistema desenvolvido proporciona ao usuário imersão e interação com um dicionário temático virtual de LIBRAS, por meio de multiprojeções em um aglomerado gráfico. O tema escolhido é representado por modelos 3D, permitindo ao usuário explorá-lo de acordo com a navegação realizada no ambiente tridimensional. Os sinais LIBRAS são interpretados por um avatar, possibilitando ao usuário a interpretação e a edição de gestos por meio de dispositivos não convencionais. A escolha das palavras interpretadas pode ser feita por meio de uma interface de controle ou pela seleção do objeto virtual no modelo 3D que representa o corrente tema. As principais tarefas do sistema são: a navegação no tema escolhido; a interpretação, edição e gravação de movimentos em LIBRAS.

No processo de navegação e interpretação é utilizada uma estrutura de multiprojeção baseada em aglomerados gráficos, garantindo a estereoscopia dos AV e o processamento dos dados distribuídos. O sistema é subdividido em três módulos, cada um executado em um nó dedicado de um aglomerado de computadores. Os módulos foram subdivididos da seguinte maneira:

- Dicionário: módulo responsável pela apresentação da interface gráfica, onde o usuário pode selecionar temas e palavras, modificar configurações do sistema, selecionar dispositivos de interação do sistema e visualizar a descrição da palavra em três idiomas (Português Brasil, Inglês e Espanhol);
- Avatar: módulo responsável pela representação dos gestos de LIBRAS referente a palavra/objeto selecionado pela interface ou pelo cenário virtual da representação do tema; e

- Representação Virtual do Tema ou Palavra: módulo responsável por renderizar o modelo 3D do tema selecionando ou da palavra representada em LIBRAS. O usuário pode selecionar diferentes partes do AV, as quais são representadas por seus gestos de LIBRAS respectivos (tela Avatar).

Os 3 módulos são integrados. A distribuição de dados e sincronização entre eles foi desenvolvida utilizando a biblioteca Glass [18]. Os AV são implementados por meio da API Java3D, o qual é capaz de gerar um ambiente sintético tridimensional por meio da renderização de um grafo de cena. A Figura 4 ilustra a disposição de apresentação dos módulos. A interação com o ambiente é realizada com um WiiRemote, o que permite ao usuário interagir com as três telas utilizando o dispositivo, visto que ele está integrado com os três módulos do sistema.

Na edição de movimentos são utilizados dispositivos não convencionais para o rastreamento corporal. Os controladores Wiimotes e a luva P5-Glove foram utilizados. Após a edição, é possível realizar a gravação dos movimentos gerados em um banco de dados, para posteriormente serem interpretados em LIBRAS no sistema de multiprojeção. A Figura 5 apresenta a modularização do sistema e como eles interagem entre si.



Figura 5. Modularização do Sistema

Os argumentos da Modularização são:

- Argumento 1:** métodos para inicializar o interpretador local com a palavra para a representação em LIBRAS e receber o retorno visual;
- Argumento 2:** comandos remotos para inicializar a multiprojeção e encerrá-la;
- Argumento 3:** instruções de movimentos;
- Argumento 4:** comandos SQL para retornar o caminho do modelo virtual e carregá-lo; e
- Argumento 5:** métodos para inicializar os dispositivos.

Para realizar o rastreamento necessário à criação e edição de gestos, os controladores devem ser anexados aos braços e antebraços do usuário, com o intuito de rastrear os movimentos por meio do acelerômetro (rastreamento mecânico) e infravermelho (rastreamento ótico). A luva é vestida na mão direita,

enviando os movimentos dos seus referidos dedos por meio de sensores (rastreamento mecânico) e de LEDs (rastreamento ótico).

Com a utilização de protocolos de comunicação (SSH e SCP) e uma biblioteca de sincronização e distribuição de dados (Glass), foi possível a unificação das funcionalidades do sistema em um único ambiente, fazendo com que o usuário tenha a percepção de um sistema único.

Na criação/edição dos sinais são utilizados dois dispositivos não convencionais, o controlador Wiimote e a Luva P5-Glove. O controlador é responsável por captar os movimentos dos braços e antebraços esquerdo e direito, já a luva capta os movimentos da mão direita e de seus dedos. As tarefas de edição, realizadas através destes dispositivos, podem aumentar o nível de imersão e interação do usuário com o sistema, sendo estes responsáveis por simular os movimentos da Língua de Sinais, ficando apenas restrita a algumas limitações dos dispositivos. A Figura 6 apresenta a interface do sistema de rastreamento.



Figura 6. Interface do Sistema de Rastreamento

VI. CONCLUSÕES E TRABALHOS FUTUROS

O dicionário temático visual-gestual desenvolvido possibilita a interação com o tema estudado em um sistema de multiprojeção, para a cognição, interpretação e edificação dos sinais LIBRAS. A navegação e seleção no modelo 3D que representa o tema selecionado, pode fazer com que o usuário associe melhor o gesto LIBRAS interpretado pelo avatar. O uso de um ambiente de multiprojeção proporciona um alto grau de imersão, no entanto, o sistema pode ser utilizado em sistemas desktop convencionais.

A edificação de movimentos relacionados à LIBRAS são gerados por meio de dispositivos não convencionais. Os dispositivos permitem que o usuário interaja com o sistema utilizando o próprio corpo, diferentemente de outros sistemas correlatos, os quais permitem apenas o uso de teclado e mouse.

O uso de dicionário temático para ensino-aprendizagem de LIBRAS apresentado neste artigo é uma forma diferente do apoio ao ensino, focando diretamente no tema e nas palavras que o usuário deseja aprender em um dado momento, sendo

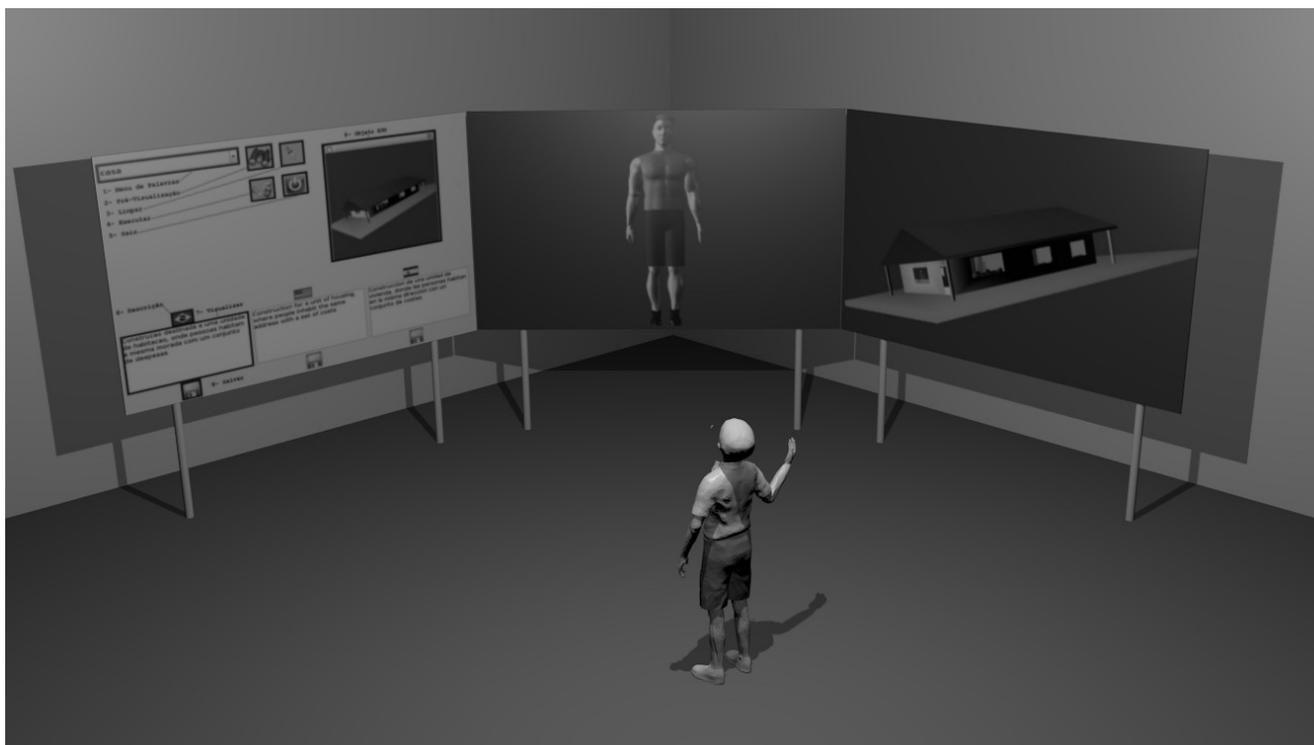


Figura 4. Imersão e Interação com o Dicionário Temático Visual-Gestual de LIBRAS

livre de uma sequência de atividades fixa. Pelas características distintas que um dicionário possui, o sistema pode ser utilizado tanto por pessoas iniciantes, quanto por pessoas experientes em LIBRAS que queiram aprender uma nova palavra.

Futuramente, serão integrados novos dispositivos de rastreamento no sistema, tais como o Kinect e a Data Glove. Estes dispositivos possibilitam rastreamento com um grande nível de precisão. Análises com usuários estão sendo realizadas.

AGRADECIMENTOS

Os autores agradecem ao apoio financeiro provido pela CAPES e XPTA.LAB

REFERÊNCIAS

- [1] M. C. R. Góes, *Linguagem, Surdez e Educação*. Editora Autores Associados, 1996.
- [2] S. A. Marcato, H. V. d. ROCHA, and M. C. M. P. LIMA, “Um ambiente para a aprendizagem da língua de sinais,” *SBC*, 2000.
- [3] C. Xatara. Dicionários temáticos visuais bilíngües – pressupostos teóricos.
- [4] A. R. d. A. Schneider, “Animação de humanos virtuais aplicada para língua brasileira de sinais,” 2008.
- [5] R. L. Secco and M. H. L. F. Silva, “Proposta de um ambiente interativo para aprendizagem em libras gestual e escrita,” in *Anais do Simpósio Brasileiro de Informática na Educação*, vol. 1, no. 1, 2009.
- [6] V. Lombardo, F. Nunnari, and R. Damiano, “A virtual interpreter for the italian sign language,” in *Intelligent Virtual Agents*, ser. Lecture Notes in Computer Science, J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, and A. Safonova, Eds. Springer Berlin Heidelberg, 2010, vol. 6356, pp. 201–207. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-15892-6_22
- [7] L. Tian, L. Dang, and F. Wang, “Study of sign language synthesis technology based on vrm1,” ser. Lecture Notes in Information Technology, 2012, pp. 24–27. [Online]. Available: <http://www.ier-institute.org/2070-1918/Init16/v16/024.pdf>
- [8] L. C. Coradine, F. C. ALBUQUERQUE, P. H. d. S. BRITO, R. L. SILVA, and T. F. L. SILVA, “Sistema falibras: Interpretação animada,” in *LIBRAS, de Palavras e Expressões em Português*. In: *III Congresso Ibero-Americano de Informática na Educação Especial (CIEEE 2002)-Demonstração*. Fortaleza, CE, de, vol. 20, 2002.
- [9] R. L. SILVA, “Ambiente aprenda libras web,” *Trabalho de conclusão de curso apresentado no curso de ciência da computação na Universidade Federal de Alagoas-UFAL*, 2002.
- [10] N. I. Badler, “Real-time virtual humans,” in *Computer Graphics and Applications, 1997. Proceedings., The Fifth Pacific Conference on*. IEEE, 1997, pp. 4–13.
- [11] J. Bates, A. B. Loyall, and W. S. Reilly, “Integrating reactivity, goals, and emotion in a broad agent,” in *In Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*. Citeseer, 1992.
- [12] T. K. Çapın, I. S. Pandzic, N. Magnenat-Thalmann, and D. Thalmann, “Avatar in networked virtual environments,” 1999.
- [13] F. Azuola, N. I. Badler, P.-H. Ho, I. Kakadiaris, D. Metaxas, and B.-J. Ting, “Building anthropometry-based virtual human models,” *Center for Human Modeling and Simulation*, p. 54, 1994.
- [14] S. de Carvalho Pereira Gonçalves, “Dicionário temático infantil de língua portuguesa: Uma diferente proposta de organização microestrutural,” *Revista de Estudos Linguísticos e Literários*, 2009.
- [15] P. Bogaards, *Le vocabulaire dans l'apprentissage des langues étrangères*. Paris: Hatier, Didier, 1994.
- [16] A. Laface, “O dicionário eo contexto escolar,” *Revista Brasileira de Linguística*, vol. 9, pp. 165–169, 1982.
- [17] R. Champagnol, “Conditions d'apprentissage du vocabulaire,” *Les langues modernes*, pp. 66–413, 1972.
- [18] B. B. Gnecco and D. R. C. Dias. (2013, Aug.) Glass library. <http://sourceforge.net/projects/libglass/?source=navbar>.