

UNIVERSIDADE DO MINHO
DEPARTAMENTO DE INFORMÁTICA

Ensino da Língua Gestual

Assistido por Personagens 3D Virtuais

Leonel Domingues Deusdado

Dissertação submetida à Universidade do Minho para obtenção do grau de
Mestre em Informática.

Toda a actividade científica sob orientação do
Prof. António José Borba Ramires Fernandes

Abril de 2002

Abstract

The Portuguese Sign Language (LGP) faces a low diffusion level, among the deaf as well as the hearing communities. The communication of a deaf person, to be effective, requires the knowledge of LGP, not only on the speaker's side, but also from the listener's.

It is also essential that both speak the *same* sign language. Due to the shortage of information, or lack of availability of the language, it is common to see people of distinct regions, using different gestures for the same words.

It is in this babylon that, in linguistic terms, the deaf people live nowadays. As communication is an essential factor to learning, it is fundamental to explore all the possible ways of taking the LGP to a large number of people.

The presented work, focuses on the use of the new technologies in the diffusion and teaching of the sign language. Even though there is an agreement on the need of human intervention in the teaching of a sign language, it is also true that there is no dispute in relation to the added value that the new technologies represent as support material in the teaching.

In this dissertation, it is given special emphasis on the technologies that use images and/or models based in three-dimensional graphics to create virtual signers.

Resumo

A Língua Gestual Portuguesa (LGP) debate-se com um nível de divulgação demasiado baixo, tanto entre a comunidade surda como na comunidade ouvinte. A comunicação por parte de um surdo, para ser eficaz, envolve o conhecimento da LGP, não só da parte do emissor como também do receptor.

É também fundamental que ambos falem a *mesma* língua gestual. Devido à falta de informação, ou falta de disponibilização da mesma, é frequente verificar-se que, na prática, pessoas de regiões distintas, utilizam gestos diferentes para as mesmas palavras, sem que cada um entenda a variação do seu vizinho.

É nesta babelónia, em termos linguísticos, que os surdos actualmente vivem. Sendo a comunicação um factor crucial para a aprendizagem, torna-se fundamental explorar todas as formas possíveis para fazer chegar a LGP ao maior número de pessoas.

O trabalho apresentado, debruça-se sobre a utilização das novas tecnologias para divulgação e ensino da língua gestual. Embora exista um consenso sobre a necessidade de intervenção humana no ensino de uma língua gestual, também é verdade que não há contestação em relação ao valor acrescentado que as novas tecnologias representam como material complementar desse ensino.

Nesta tese é dado um especial ênfase às tecnologias que utilizam imagens sintetizadas e/ou modelos baseados em gráficos tridimensionais para produzir gestualistas virtuais.

Agradecimentos

Ao meu orientador, Professor António José Borba Ramires Fernandes, pelo tema de investigação e pela oportunidade que me facultou em trabalhar numa área que tanto me motivou.

Ao meu pai, à minha mãe e aos meus irmãos que permitiram que pudesse chegar até aqui, e que tanto se esforçaram e me motivaram para olhar sempre em frente.

A todos os meus amigos e namorada que me souberam compreender e me suportaram durante os tempos mais difíceis.

À Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Bragança, pelas facilidades concedidas durante o período de elaboração desta tese.

Índice

Introdução	1
1.1 A Língua Gestual.....	1
1.1.1 A Língua Gestual na Europa	2
1.1.2 Questões mais Frequentes a Propósito da Língua Gestual	2
1.2 A Comunidade Surda.....	3
1.3 A Revolução Silenciosa.....	4
1.4 O Futuro “Ideal”, as Possibilidades Tecnológicas Perante as Pessoas Surdas	6
1.5 Objectivos da Tese	8
1.6 Resumo dos Capítulos Seguintes	10
Estado da Arte	11
2.1 Integração e Trabalho de Apoio	11
2.1.1 Formação de Pessoas Surdas: Análise da Situação	12
Dificuldade no Processo Formativo	12
Necessidades Específicas.....	12
2.2 Testes de Suportes Multimédia Para a Educação e Supressão de Barreiras de Comunicação Para Pessoas Surdas	13
2.2.1 Symbol	13
2.2.2 CD_Rom Interactivo de Visita ao Palácio Real de Madrid	14
2.2.3 Critérios e Adaptações Especificas Para Pessoas com Necessidades Especiais	14
2.3 Representação de ASL na World-Wide Web com Soletração Através de Dedos	15
2.3.1 O Modelo da Mão.....	16
2.3.2. Painel de Controlo	17
2.3.3 Animação	17
2.3.4 Soletrar Através dos Dedos	18
2.4 Modelos de Markov Para o Ensino Interactivo de Gestos.....	19
2.4.1 Objectivos do Projecto.....	19
2.4.2 Construção do Sistema.....	19
2.5 Projecto Mudra - Multimédia e Animação 3D Para Pessoas Surdas.....	21
2.5.1 O Projecto.....	21
2.5.2 As Vantagens.....	22
2.5.3 Descrição Técnica.....	23
2.5.4 A Mostra de Material 3D	24
2.6 “Avatar“ Gestualista na Web	25
2.6.1 Gravação e Visualização dos Dados a Partir dos Movimentos dos Gestos	26
2.6.2 Dos Dados Extraídos dos Movimentos Para a Página Web.....	27
2.7 O Projecto “Simon”	28
2.7.1 Futuros Desenvolvimentos.....	30
2.8 Modelação e Animação de Língua Gestual Representada por meio de Linhas.....	30
2.8.1 Modelação de Língua Gestual	31
2.8.2 Expressões Faciais	32
2.8.3 Definição da Postura	33
2.8.4 Sintetização de Gestos	33
2.8.5 Animação de Gestos	34

2.8.6 Visualização	34
2.8.7 Exemplos de Expressões Faciais	35
2.8.8 Exemplos de Gestos	35
2.8.9 Discussão e Conclusões Retiradas	36
Tecnologias Gráficas de Apoio.....	37
3.1 Computação Gráfica	37
3.1.1 Origens da Computação Gráfica.....	37
3.1.2 O Crescimento da Computação Gráfica.....	38
3.1.3 O que é a Computação Gráfica.....	38
3.2 Interação em Ambientes Tridimensionais.....	40
3.2.1 Parâmetros do Processo Interactivo	41
Resolução da Interação	41
Graus de Liberdade.....	41
3.3 Interação em Ambientes Virtuais	42
3.3.1 Interação Indirecta sobre Objectos.....	43
Manipulação com Controlos Comuns	43
Manipulação com Controlos Específicos Para Manipulação 3D	45
3.3.2 Interação Directa Sobre Objectos	46
Interação com Manipuladores Tridimensionais	48
3.3.3 Cinemática e Cinemática Inversa	49
3.3.4 Navegação	52
Classificação das Metáforas de Navegação	53
Características Típicas do Movimento de uma Câmara Virtual	53
A Aplicação	55
4.1 Construtor de Movimentos para Língua Gestual	55
4.1.1 Os Modelos.....	55
Poser	56
Construção dos Modelos.....	57
Simplificação dos Modelos (mãos)	60
4.1.2 Construtor de Movimentos.....	62
Pulse3D	62
Descrição do Ambiente de Trabalho	63
A Biomecânica da Mão.....	64
Definição de <i>Joints</i> e <i>Bones</i> na Articulações dos Modelos	65
Especificação da Interface	68
Implementação e especificação da Interface (manipulador 3D + rotação + translação) para o Construtor3D	69
4.2 Abecedário 3D.....	70
4.2.1 O Modelo.....	71
4.2.2 Exportação.....	71
4.2.3 “Morpher Modifier” no 3DStudioMax.	72
4.2.4 Display no Pulse 3D – Página Web	74
4.3 Animações Vídeo	75
4.4 Gestuário 2D	76
Testes de Usabilidade.....	77
5.1 Perguntas e Respostas.....	77
5.2 Resumo Geral.....	87

5.2.1 A Aprendizagem de Língua Gestual Recorrendo à Informática	87
5.2.2 O Abecedário3D	87
5.2.3 O Construtor 3D	88
5.2.4 O Gestuário2D.....	88
5.2.5 As Animações Vídeio Propostas	88
5.2.6 A Aplicação no Geral.....	88
Conclusões e Trabalho Futuro	89
6.1 Conclusões	89
6.2 Trabalho Futuro.....	91
Bibliografia	93
Anexos	C
Respostas Individuais aos Inquéritos	C
Conteúdo e Instalação do CDROM.....	B

Lista de Figuras

Figura 1 – Painel de controlo (mão direita).....	17
Figura 2 – Tradução ASL de cadeia de letras e números (mão direita).....	18
Figura 3 – Posição final da mão para o gesto “C”	19
Figura 4 – Interface do sistema de demonstração para reconhecimento de gestos.....	20
Figura 5 – Humanóide 3D	22
Figura 6 – Características para a interação 3D	24
Figura 7 – Conversões entre o Servidor e Cliente	25
Figura 8 – Captura de movimentos com ajuda de um intérprete de língua gestual	26
Figura 9 – Reprodutor de língua gestual	28
Figura 10 – Gestualista virtual Simon.....	28
Figura 11 – Captura de movimentos a partir de um intérprete em língua gestual.....	29
Figura 12 – Editor de linhas com várias possibilidades para conseguir uma expressão facial específica.	32
Figura 13 – Modelo do esqueleto da parte superior do corpo com as suas articulações específicas.	33
Figura 14 – Janela de diálogo de composição de gestos e biblioteca de armazenamento de gestos.....	33
Figura 15 – Dois tipos de linhas são válidos nas expressões faciais, neste caso gerando um ar de surpresa.....	35
Figura 16 – Expressão facial sorridente através de gráficos de linhas.....	35
Figura 17 – Dois tipos de representação dos gestos com ênfases em diferentes partes do corpo para captar a atenção do espectador.....	36
Figura 18 – Relacionamento da computação gráfica com outras áreas	39
Figura 19 – Controlo dos Graus de Liberdade.....	42
Figura 20 – Écran com excesso de controlos.....	44
Figura 21 – Redução do número de controlos	44
Figura 22 – Interface com ícones de transformação	44
Figura 23 – <i>Widget</i> tipo Relógio.....	45
Figura 24 – <i>Widget</i> 3D a Escala.....	46
Figura 25 – Disposição de vistas ortográficas	46
Figura 26 – Disposição incorrecta de eixos coordenados em vistas ortográficas.....	47
Figura 27 – Controlo de Movimentação de objectos do GLEEM	48
Figura 28 – Movimento em diferentes sistemas de referência	48
Figura 29 – Eixos do sistema de referência do objecto no Cinema 4D	49
Figura 30 – Figura articulada por <i>Joints</i> e <i>Bones</i>	49
Figura 31 – Cinemática do modelo de um braço humano.....	51
Figura 32 – <i>Rendering</i> do modelo do tronco superior humano	51
Figura 33 – Modelo movimentando a mão para junto do objectivo, utilizando a cinemática inversa, com imposição dos estrangimentos dos ângulos de junção do ombro e cotovelo.....	52
Figura 34 – Componentes de um “observador virtual”	52
Figura 35 – Alternativas para movimento de câmara	53
Figura 36 – Lista de opções Poser	56
Figura 37 – Vista Geral do ambiente de trabalho – Interface – do Poser	57
Figura 38 – Modelo da mão utilizada em experiências por Ouhaddi e Horain	58

Figura 39 – Comparação de parâmetros entre a fotografia e o modelo	58
Figura 40 – Modelo completo de onde foram retiradas as duas mãos	59
Figura 41 – Modelos das mãos resultante	59
Figura 42 – O modelo 3D de humanóide completo	60
Figura 43 – Modelos com redução do número de polígonos	60
Figura 44 – Modelo de mão inicial -1449 polígonos	61
Figura 45 – Modelo de mão simplificado - 197 polígonos	61
Figura 46 – Modelo de mão final otimizado no Pulse3D.....	62
Figura 47 – Vista geral do Interface Pulse3D.....	64
Figura 48 – Esqueleto e articulações da mão.....	65
Figura 49 – Movimentos ao nível das articulações Metacarpo-Falanges	65
Figura 50 – <i>Bones</i> e <i>Joints</i> do polegar da mão direita	66
Figura 51 – Pontos de possível articulação do modelo representativo da mão direita	66
Figura 52 – Estrutura hierárquica de todos os componentes e funções do modelo mão direita	67
Figura 53 – Pontos de possível articulação do modelo de corpo inteiro.....	67
Figura 54 – Restrições de movimentos nos pontos articulados.....	68
Figura 55 – Modelos finais de interação para a mão com os respectivos controladores	70
Figura 56 – Interface do Construtor 3D	70
Figura 57 – Corte do modelo necessário à animação.....	71
Figura 58 – Interface da aplicação Maximum Pose.....	72
Figura 59 – Animação do tipo <i>Morphing</i>	73
Figura 60 – Momento de redefinição dos estados de Morpher no 3DStudioMax.....	73
Figura 61 – Estrutura hierárquica do Abecedário3D	74
Figura 62 – Representação da letra F	74
Figura 63 – Animação gestual do verbo amar	75
Figura 64 – Imagem 2D do gesto representando o número zero.....	76

Capítulo 1

Introdução

1.1 A Língua Gestual

A língua gestual é uma língua de movimento e de espaço, das mãos e dos olhos, da comunicação abstracta assim como do contar de uma história icónica, mas acima de tudo, é a língua da comunidade surda. Não se trata de uma nova língua nem de um sistema recente de comunicação desenvolvido por pessoas ouvintes, pelo contrário, trata-se de uma forma de comunicação que ocorre entre pessoas que não ouvem. É uma língua que até há bem pouco tempo tinha sido algo ignorada e por esse motivo parecia subvalorizada no seu potencial. É diferente, muitas vezes impressionantemente diferente das línguas faladas, mas partilha características e processos gramaticais com muitas outras línguas faladas. A verdade é que o mundo da comunicação em língua gestual não é um mundo triste, cinzento, um mundo de silêncio. É um mundo alegre, colorido e às vezes mesmo muito barulhento. É a primeira língua para os surdos, aquela que em primeiro lugar e naturalmente utilizam para comunicar.

A língua gestual existe em Portugal, há incontáveis gerações. Mas por ser falada por uma minoria, ela, pura e simplesmente, aos olhos do cidadão comum, não existia. Mas isso não obsta a que, no nosso país, haja efectivamente duas línguas autóctones: a Portuguesa e a Gestual Portuguesa. Isto para não falar do Mirandês e dialectos afins.

A necessidade de remover as barreiras que impedem o reconhecimento das línguas gestuais tem estado presente em diversos documentos de cariz internacional. Assim, vários congressos de surdos a nível internacional recomendaram quer o reconhecimento das línguas gestuais, quer da profissão de intérprete. Também as Nações Unidas se pronunciaram sobre o assunto, integrando esta problemática na resolução nº 48/96 que definiu as "normas sobre igualdade de oportunidades para as pessoas com deficiência". Finalmente, o Parlamento Europeu aprovou uma resolução sobre línguas gestuais que prevê a abolição pelos estados-membros dos obstáculos à utilização das línguas gestuais e sublinha a importância do intérprete de língua gestual. Também a Constituição

Portuguesa passou a ter, após a última revisão constitucional referência à língua gestual Portuguesa. Trata-se de uma língua própria, diferente da língua Portuguesa e das línguas gestuais dos outros países, isto é, das outras comunidades de surdos.

A aprendizagem da língua gestual Portuguesa é por isso um instrumento indispensável para a verdadeira integração dos surdos portugueses e para que usufruam da sua plena cidadania. As crianças e os jovens surdos são talvez os principais afectados por não estar ainda garantido o ensino e o acesso generalizado da língua gestual Portuguesa e a sua utilização nas mais diversas situações na sociedade.

1.1.1 A Língua Gestual na Europa

Contrariamente à crença popular, a língua gestual não é internacional. Cada país tem a sua própria língua gestual. Nos países onde existem duas ou mais línguas faladas, por exemplo, a Bélgica e Espanha, existem o mesmo número de línguas gestuais. As pessoas surdas tal como as pessoas ouvintes, são educadas na língua da sua região e raramente o serão numa outra língua do país. Por exemplo, na Bélgica, as pessoas surdas flamengas são educadas em escolas de surdos flamengas e não nas escolas de surdos de expressão francófona, e vice - versa. Em Portugal isto também tende a acontecer, daí ser fundamental também um unificador da língua gestual Portuguesa para que os surdos não se sintam “fora” cá “dentro”. As línguas gestuais não dependem das línguas dos seus respectivos países. Isto pode ser comprovado pelo facto de a Grã-Bretanha, Irlanda e os Estados Unidos da América serem todos países falantes de Inglês, mas terem línguas gestuais completamente diferentes. De facto, a língua gestual Americana (ASL¹) está mais próxima da língua gestual Francesa (LSF²) do que da língua gestual Britânica (BSL³). A explicação é simples: as raízes da ASL encontram-se em França, e não na Inglaterra.

É quase impossível determinar o número de utilizadores de língua gestual na União Europeia, pois não foram recolhidos dados nesse sentido, e existem muito poucas pesquisas quantitativas para nos reportamos. O número de gestualistas surdos deverá ser mais ou menos o mesmo da incidência de pessoas surdas pré-linguais, 1 por 1000/1500, apesar de nem todos os surdos pré-linguais gestualizarem e, pelo contrário, alguns surdos pós-linguais são gestualistas.

1.1.2 Questões mais Frequentes a Propósito da Língua Gestual

1. A língua gestual é "universal"?
NÃO. Cada país, por vezes cada região, tem uma língua gestual própria, testemunha de uma cultura e dos modos de vida que são exclusivos da comunidade surda desse país ou região.
2. É possível exprimir um pensamento totalmente abstracto na língua gestual?

¹ American Sign Language

² Langue des Signes Française

³ British Sign Language

SIM. Do mesmo modo que nas línguas orais.

3. É difícil aprender a língua gestual?

SIM... e... NÃO. As pessoas ouvintes não estão habituadas a exprimirem-se numa base totalmente visual, nem de utilizarem as mãos e o corpo para tal. No entanto o nível de dificuldade depende igualmente das capacidades de cada pessoa.

4. Os surdos "soletram" todas as palavras para comunicar?

NÃO todas. Os surdos utilizam os gestos que são as imagens simbólicas dos conceitos exprimidos. É o conjunto dos parâmetros do gesto assim como a expressão da face e o movimento do corpo que criam uma imagem precisa da ideia exprimida. No entanto há palavras que por serem pouco usuais ou muito específicas necessitam de ser soletradas

5. Como se chama a língua gestual em Portugal?

Chama-se língua gestual Portuguesa, abreviadamente designada por LGP. Erroneamente ainda há muitas pessoas que continuam a designá-la só por linguagem gestual.

6. Que diferença existe entre a língua gestual Portuguesa e o Português gestual?

A língua gestual Portuguesa é uma verdadeira língua com um léxico e uma gramática próprios. É a comunicação das pessoas surdas, cada língua gestual é portadora da cultura das pessoas que a utilizam. O Português gestual é um método deturpado de comunicação que utiliza o vocabulário da língua gestual Portuguesa numa estrutura gramatical e numa ordem pertencente à língua oral Portuguesa. Serve unicamente para mostrar a ordem dos gestos numa frase em Português mas não transmite de maneira válida o pensamento de alguém, visto não respeitar as regras gramaticais da língua gestual Portuguesa e a necessidade duma lógica visual.

7. Existe um gesto para cada palavra em Português?

Como nas línguas orais, a língua gestual tem por vezes várias traduções diferentes (segundo os sentidos) para uma palavra de Português ou, ao inverso, um só gesto permite traduzir várias palavras de Português.

1.2 A Comunidade Surda

Para dissipar quaisquer dúvidas, convém esclarecer, desde já, que nem todos os surdos pertencem à comunidade surda, e que esta inclui também pessoas ouvintes. A comunidade surda é um conceito antropológico e linguístico: são os falantes da língua gestual portuguesa, é uma minoria linguística do nosso país. Os filhos ouvintes de pais surdos também fazem parte dela, porque aprenderam a língua gestual, desde miúdos, para comunicar com os pais [1].

Para classificarmos um indivíduo como “surdo”, devemos levar em consideração, aspectos anatómicos-fisiológicos e também significados qualitativos, pois a surdez possui graus muito diferentes, nuances extremamente subtis. Dentre as muitas e

diferentes formas de classificação de surdez, usaremos a que Oliver Sacks (1986) utilizou no seu livro “Vendo Vozes”:

1. “Audição deficiente, pessoas que só conseguem ouvir alguma coisa com o uso de acessório auditivo e bastante paciência por parte daqueles que lhes falam. Muitos de nós terão pais ou avôs nessa categoria.
2. Extremamente surdos, muitos em decorrência de uma doença ou lesão no ouvido nos primeiros anos de vida, mas no caso deles, como também acontece com as pessoas de audição deficiente, ainda é possível ouvir e falar, especialmente com os novos aparelhos auditivos.
3. Profundamente surdos, que não tem qualquer esperança de ouvir alguma fala. As pessoas profundamente surdas não podem conversar de maneira usual, podem interpretar a leitura labial, usar a língua gestual ou fazer as duas coisas.”

Calcula-se que existam cerca de 100 a 150 mil pessoas que não ouvem bem em Portugal. Há as pessoas que nasceram surdas, outras perderam o ouvido, por vários motivos, por acidente, por trabalhar expostos ao ruído intenso durante vários anos, etc. Mas alguns deles, por variadíssimas razões (incompreensão, falta de informação ou preconceito) nunca chegaram a entrar em contacto com a LGP. Portanto, não fazem parte da comunidade surda.

Há cerca de 15 mil pessoas, em Portugal, que têm como língua primeira a língua gestual Portuguesa, o seu modo de expressão, o seu principal traço de identidade pessoal e de grupo. A língua gestual tem para eles exactamente o mesmo papel que tem para nós o Português. Daí que, naturalmente, prefiram 'juntar-se aos seus'. Não se trata de um *apartheid*, mas de uma comunidade diferente, que fala a mesma língua. E tal como os surdos podem sentir-se 'noutro país' quando comunicam com um ouvinte, também um ouvinte se sentirá 'estrangeiro' se entrar, por exemplo, no bar da Associação Portuguesa de Surdos, onde a língua gestual é falada por todos. A leitura labial também é importantíssima, é talvez a primeira coisa que em qualquer escola de surdos se ensina a uma criança surda, é fundamental para que qualquer pessoa surda possa comunicar com pessoas ouvintes, ou pelo menos conseguir entendê-los. Em Portugal o ensino da leitura labial não se generaliza muito a surdos que ensurdecem na idade adulta. Como facilmente se entende a língua gestual é mais rica, mais esclarecedora, mais eficaz, mas o desejável é o complemento destes dois tipos de comunicação.

1.3 A Revolução Silenciosa

Portugal e a Finlândia são os únicos países europeus que reconhecem os surdos como minoria linguística na Constituição. A Lei Fundamental Portuguesa consagra, no art. 74, ponto 2, alínea h, a necessidade de o Estado 'proteger e valorizar a língua gestual portuguesa, enquanto expressão cultural e instrumento de acesso à educação e de igualdade de oportunidades'.

Pela primeira vez, é reconhecido o direito de uma criança surda frequentar a escola com possibilidade de ter o ensino em língua gestual Portuguesa. E o salto qualitativo é por demais evidente, visto que, de acordo com os estudos existentes, só assim ela poderá, mais facilmente, aprender o Português. Em Copenhaga, numa escola onde se concentram parte dos alunos surdos da Dinamarca, e cujo corpo docente é constituído por professores surdos e professores ouvintes, os miúdos aprendem como segunda língua, e sem dificuldade nenhuma o Dinamarquês, sendo por consequência bilingues. Fazem todo o seu percurso escolar juntos, comunicam em língua gestual e tiram apontamentos em Dinamarquês, com toda a naturalidade. Em Portugal, esta é ainda uma realidade nova, mas os estudos feitos nos países escandinavos concluem que as crianças surdas que têm a língua gestual do seu país como língua primeira tem resultados semelhantes às crianças ouvintes.

Em Portugal, os professores surdos escasseiam. Ora, sem um professor que fale em LGP ou um intérprete, nenhum aluno surdo consegue ultrapassar as dificuldades inerentes às matérias leccionadas no ensino secundário. Para já nem falar no ensino superior. Faltam, portanto, os intérpretes. Existem alguns intérpretes que não escolheram, por assim dizer, a aprendizagem, porque são filhos de pais surdos.

Barbara Schmidbauer (PSE/D), em nome da Comissão do Emprego e dos Assuntos Sociais, questionou a Comissão Europeia com a pergunta oral B4-489/98 nos seguintes termos:

- O projecto Linguas Gestuais demonstrou que o principal problema com que os surdos se vêem confrontados consiste na ausência de reconhecimento oficial das linguagens gestuais por parte dos Estados-Membros da UE que lhes permita afirmarem os seus direitos enquanto cidadãos da UE de pleno direito. Estará a Comissão disposta a publicar uma Comunicação ou uma Directiva relativa ao reconhecimento oficial das linguagens gestuais utilizadas pelos surdos em cada um dos Estados-Membros?
- Dada a ausência de intérpretes qualificados em línguas gestuais em muitos países da UE, que medidas adoptou a Comissão Europeia visando garantir o financiamento da UE a programas no domínio da educação e da formação profissional destinados à formação de formadores e de intérpretes de línguas gestuais?
- Atendendo à importância de que para os surdos se reveste a informação visual, tenciona a Comissão Europeia introduzir legislação-quadro relativa à transmissão, em línguas gestuais, de um número mínimo de programas televisivos?
- Existem nos Estados-Membros da UE sete sistemas diferentes de "telefone e teletexto", os quais não são compatíveis entre si, o que gera inúmeros problemas aos utilizadores surdos que telefonam a amigos surdos em outros Estados-Membros da UE. Tenciona a Comissão introduzir legislação-quadro que garanta a compatibilidade dos equipamentos de telecomunicação por texto e de videofone destinados aos surdos em toda a Europa?

- Tenciona a Comissão Europeia adoptar medidas que garantam um "design" universal das aplicações multimédia, por forma a que os surdos não sejam excluídos das novas aplicações?
- Que medidas tenciona a Comissão adoptar, a fim de garantir que os funcionários actantes nas instituições da UE sejam formados nas questões relativas à surdez e que as reuniões públicas realizadas pelas instituições da UE sejam acessíveis aos surdos?

O Comissário Franz Fischler, durante a sua intervenção, confirmou o facto de a Comissão estar consciente da importância da língua gestual, tendo afirmado ainda que a Comissão tudo fará para sensibilizar os cidadãos para se prestar um maior apoio às pessoas que fazem uso da língua gestual como meio de comunicação.

1.4 O Futuro “Ideal”, as Possibilidades Tecnológicas Perante as Pessoas Surdas

O que é que aconteceu realmente ao longo destes últimos anos? Os meios audiovisuais já estão a ser efectivamente usados na educação e formação das pessoas surdas? Já existem conteúdos pedagógicos em CD-ROM's, *sites* WWW, Cassetes Vídeo, ou noutro suporte audiovisual destinados ao ensino da língua gestual Portuguesa de modo a alargar as possibilidades de comunicação, aprendizagem, integração e socialização de todas as pessoas surdas? É um dos objectivos desta tese, procurar caminhos novos para o futuro, na aprendizagem de LGP, para essas novas gerações de cidadãos que só tem a comunicação visual como suporte para acederem à informação e ao conhecimento: as pessoas surdas.

Actualmente, uma nova forma de uso do computador no processo de ensino-aprendizagem, tem sido a educação à distância. Esta modalidade de ensino (dito virtual) remonta a sua origem ao final do século XIX, quando era feita através de cartas. No século XX esta forma de ensino foi gradualmente incrementada, incorporando os avanços tecnológicos, começando com a rádio, passando à televisão, e chegando ao uso do computador e da Internet. É exactamente este último avanço que tem sido o foco de diversos estudos e pesquisas nestes últimos anos. As suas características acarretam certamente vantagens no apoio ao ensino de pessoas com deficiências auditivas.

Está em desenvolvimento desde 1996, num consórcio formado sobretudo por universidades, uma rede ultra rápida de comunicação, chamada de Internet2. Esta rede seria bem parecida com a Internet que se tem hoje, com a diferença de ser bem mais veloz. Isto permitiria o uso de vídeo-conferência em tempo real, sem aqueles problemas (muito desconfortáveis) que muitas vezes acontecem hoje em dia devido à falta de largura de banda. Esta rede já está em funcionamento em algumas universidades dos EUA como a Universidade de Gallaudet e a Universidade de Georgetown, que têm aproveitado para ensinar língua gestual através da Internet. E isto é só o começo, pois as potencialidades desta nova rede são realmente muito grandes.

No entanto, algo que está actualmente na vanguarda da tecnologia é a realidade virtual, com imagens em três dimensões (3D), que se aproximam bastante da nossa realidade concreta. A realidade virtual apresenta um grande potencial de interacção entre as pessoas, e pode fazer uso, além dos recursos áudio-visuais a que estamos acostumados, de outros sentidos para uma interacção mais humana e realista. O que se percebe é o grande potencial pedagógico e integrador desta nova realidade. Várias pesquisas estão em curso, e muitas outras ainda são necessárias para, de facto, poder fazer-se uso das potencialidades que esta nova tecnologia parece oferecer.

O telefone de voz tornou-se num importante meio de comunicação, aproximando as pessoas, mas esta tecnologia não pode ser usada directamente por todos. Pessoas com deficiências auditivas precisam de serviços de telecomunicações alternativos, os quais têm que ser disponibilizados em iguais termos. A conversação pessoal via telecomunicações pode agora dar um grande passo em direcção à acessibilidade para todos. Os futuros serviços de telecomunicações podem oferecer aos utilizadores a possibilidade de comunicar simultaneamente através de vídeo, texto e voz. Um dos principais serviços que uma grande parte dos países europeus já possui é o serviço de intermediação telefónica e conversores de texto-voz e voz-texto para cidadãos com deficiência de audição e fala. Baseado num serviço de intermediação entre o emissor e o receptor, constituído por operadores que transcrevem de voz para texto ou de texto para voz consoante a necessidade dos interlocutores da comunicação, dão assim a possibilidade de se estabelecer uma comunicação entre uma pessoa surda e um ouvinte. A 3ª geração de telemóveis (UMTS) vem aí e promete abrir uma nova era de comunicações a surdos e a pessoas com dificuldades na fala, já que se poderá alargar as potencialidades da Internet e da multimédia a um dispositivo móvel de reduzidas dimensões.

Nos últimos quinze anos a comunidade surda assistiu com expectativa a esta saga prodigiosa da evolução das tecnologias de informação. Para eles aquilo que encontravam de mais fascinante nas tecnologias de informação era a comunicação com o computador se poder processar no modo ideal para uma pessoa surda: através da comunicação visual.

Também é com as pessoas ouvintes que devemos contar, pessoas que vivem fascinadas na aprendizagem desta língua ou aquelas que por necessidade têm de aprender a comunicar em língua gestual. São pessoas que adoram ou necessitam praticar essa forma de comunicação mas que se vêem por vezes frustradas na concretização desse objectivo.

Uma das coisas que mais tem impressionado é a atitude da juventude actual. São muitos os jovens que querem aprender a comunicação em língua gestual. São jovens do chamado ensino integrado que querem aprender a comunicar com os seus colegas surdos. E, para o comprovar, apresenta-se um "estudo de caso" feito junto das escolas de instrução da Guarda Nacional Republicana. Há quase mil agentes actuais e futuros da Guarda Nacional Republicana muito interessados em conhecer um conjunto de expressões em língua gestual que os habilitassem a comunicar com uma pessoa surda em momentos específicos que acontecem na comunicação com os agentes das forças de segurança: um acidente, um crime, uma transgressão às regras de trânsito ou um simples pedido informação.

A verdade é que tudo o que possa contribuir para dar resposta as necessidade das pessoas surdas é sempre bem vinda. A resposta poderá encontrar-se naquilo que nós designamos como os meios de comunicação interactivos dos quais "et pour cause" a Internet surge como meio privilegiado. A Internet está na moda. Para a comunidade surda sempre houve um relacionamento difícil com as modas do momento, mas a Internet é apenas uma faceta surgida a "posteriori" daquilo que já pensavam há 25 anos : potenciar para a efectiva integração social das pessoas surdas os meios audiovisuais, na sua componente "visual". Portanto, não é uma moda mas apenas uma "peça de roupa" que já imaginavam há um quarto de século em meio "virtual".

Aprender uma língua gestual é um processo difícil, trabalhoso e exigente em termos de vontade e de perseverança. Contudo, e tal como acontece com as línguas faladas, esse processo exige uma prática constante num meio que mesmo não sendo real, pode ter feição virtual através animações virtuais WWW, CD-ROM's ou cassetes. E porque o tempo corre, e é preciso desenvolver uma estratégia que possa responder aos anseios de muitas pessoas no tempo mínimo útil, que se concebe aquilo a que designamos como a produção de suportes digitais interactivos em 3D de comunicação/formação em língua gestual.

É esta a realidade que se impõe com toda a evidência. Ou seja, com a estratégia que se tem seguido até hoje nunca se conseguirá uma eficiente formação na comunicação em língua gestual [2]. E, por tabela, a efectiva integração das pessoas surdas na comunidade sempre continuará como uma miragem algo distante e inacessível. E, pelo contrário, uma estratégia que tire partido das tecnologias multimédia e da Internet já abre perspectivas mais vastas para essa integração. É uma estratégia que não só envolverá as pessoas surdas, mas também as suas organizações. Ou seja, a estratégia conducente à produção de um Gestuário digital interactivo em 3D acessível através da Internet. Isto é um objectivo perfeitamente ao alcance do estado actual da arte nas tecnologias da informação.

Mas, as possibilidades tecnológicas e da Internet não se ficam por aqui. Actualmente, perante os recursos mais rápidos na computação gráfica e processamento de imagem, já nos encontramos num estádio em que a comunicação visual se poderá realizar de um modo efectivo e acessível. As modernas animações tridimensionais e a conversão da voz em texto abrem às pessoas surdas possibilidades que ainda há bem poucos anos só eram imagináveis no mundo da ficção científica. Aquela façanha do computador HAL, de ser capaz de interpretar pela leitura labial, que Stanley Kubrick imaginou no seu filme "2001 Odisseia no Espaço" já é hoje possível.

1.5 Objectivos da Tese

A possibilidade do uso das novas tecnologias, principalmente da Internet, dos CD-ROM, e todos os serviços associados a estas, para a supressão ou redução de barreiras linguísticas e sociais, que ainda vigoram entre uma sociedade dita "normal" e uma sociedade com dificuldades auditivas com uma língua própria, permanece ainda nos nossos dias como um claro objectivo.

Os novos desenvolvimentos em áreas como a computação gráfica, bases de dados distribuídas, redes de computadores, paralelamente ao desenvolvimento galopante do hardware, renovam as esperanças de obter um sistema global de aplicações integradas, capazes de se adaptar e preencher carências de utilizadores com necessidades especiais sob a forma de ajuda linguística ou de comunicação num mundo dominado pelas maiorias.

Pretende-se com esta dissertação implementar um sistema de apoio à língua gestual Portuguesa, virado sobretudo para o ensino de pessoas com deficiências auditivas que se iniciam nesta língua. Este tipo de aplicação não se poderá resumir à comunidade surda, o objectivo é também dar a conhecer esta língua aos ouvintes, pois para uma comunicação entre surdos e não surdos é necessário um esforço também dos segundos. O sistema contará com as potencialidades multimédia da Internet, com a apresentação de ambientes gráficos tridimensionais com personagens virtuais. A aplicação será proposta através de quatro directrizes principais:

- 1- Construtor 3D – Será a componente mais ambiciosa de toda a aplicação. Em geral permitirá a construção de gestos completos de LGP. Uma página Web interactiva será proposta com um modelo completo 3D de um humanóide, e mãos adicionais (direita e esquerda). Assim sendo, o utilizador comporá os gestos com a ajuda de manipuladores tridimensionais, rotações e translações dos modelos, que após a sua composição serão associados a palavras, letras ou números em língua gestual, para numa fase futura, serem armazenados numa base de dados. Esta componente permitirá assim memorizar a construção dos gestos pormenorizadamente, bem como organizar e armazenar uma série de construções de gestos próprios a utilizar em várias traduções aquando da aprendizagem da LGP.
- 2- Abecedário 3D – Serão construídas animações 3D realistas para todas as letras do abecedário LGP, com a possibilidade de visualizar a animação de vários pontos de vista. Assim sendo, e a partir do modelo do braço e mão de um humanóide 3D, serão armazenadas as animações de todas as letras, e apresentadas numa página Web, permitindo assim a aprendizagem ou aperfeiçoamento de gestos da língua gestual. Pretende ser uma animação bastante perfeita com modelos e cinemática proporcionada por aplicações externas.
- 3- Animações Vídeo – À imagem do que acontece com tantos sites, animações vídeo serão adicionados para descrever gestos em LGP. Mas as animações serão construídas a partir de modelos 3D animados e não a partir de imagens reais. Serão numa primeira fase animados somente os verbos.

- 4- Gestuário 2D – Imagens 2D serão apresentadas a partir do Gestuário [3] e associadas a um texto explicativo sobre a posição dos dedos e mãos. Será também associada uma chamada de atenção para possíveis confusões ou erros com gestos semelhantes.

1.6 Resumo dos Capítulos Seguintes

O texto que se segue toma uma sequência lógica desde a observação de estudos e projectos para o apoio de ensino de língua gestual até à apresentação da aplicação desenvolvida com os respectivos resultados, testes de usabilidade e conclusões.

Assim, no capítulo 2 é apresentada uma vista geral do estado da arte no que respeita à integração e trabalho de apoio na formação de pessoas surdas, respectivamente no desenvolvimento de aplicações específicas direccionadas para pessoas com dificuldades auditivas no campo das interfaces, ensino e aperfeiçoamento de língua gestual.

No capítulo 3 é descrito em linhas gerais a interface do ambiente gráfico onde se incluem aplicações virtuais multimédia e o modo de interacção destas com utilizadores, com especial atenção à manipulação de objectos tridimensionais, cinemática e cinemática inversa como recurso à construção de modelos articulados, navegação em mundos virtuais e diferentes interfaces para uma possível e correcta utilização das tecnologias gráficas 3D.

No capítulo 4 é apresentada a descrição do trabalho desenvolvido no CD-ROM que acompanha esta dissertação, começando com a realização de um construtor de movimentos para língua gestual. Desde a construção dos modelos até à aplicação propriamente dita passando pela biomecânica e articulação dos modelos, especificação da interface e manipuladores bem como do uso de cinemática inversa para controlar de forma natural e humana os modelos apresentados. É também dedicado à descrição do restante trabalho prático para complementar esta tese que resulta de uma aplicação global, interactiva, 3D, distribuída na rede (www) ou CDROM, para ajudar o ensino e aperfeiçoamento de LGP a pessoas surdas e ouvintes que a queiram aprender para poder comunicar-se com pessoas com dificuldades auditivas que comuniquem neste tipo de língua. Através de um numerário, abecedário, vídeos e restantes palavras em língua gestual num gestuário 2D, apresentar-se-à a construção da respectiva aplicação. Será também uma oportunidade de uniformizar a língua gestual que dentro de um mesmo país poderá conter certas diferenças que representam os hábitos de uma sociedade numa dada região.

O capítulo seguinte é reservado a testes de usabilidade com base nos resultados obtidos a partir de inquéritos sobre a integração do trabalho prático numa instituição de ensino de LGP.

Por último são retiradas as conclusões e apresentadas sugestões para um possível trabalho futuro.

Capítulo 2

Estado da Arte

2.1 Integração e Trabalho de Apoio

As tecnologias de comunicação e informação desenvolvem-se a uma velocidade difícil de seguir, inclusive por parte dos profissionais de informática. Isto significa, entre outras coisas, que se não é fácil fazer previsões no se que refere à incidência sobre a sociedade em geral, menos ainda no relativo às necessidades de pessoas com deficiências. O futuro está cheio de possibilidades e infelizmente também de problemas.

Em determinadas sociedades de alto nível tecnológico, por exemplo os Estados Unidos da América, o reconhecimento da necessidade de acesso das pessoas com ausência de algumas capacidades a um emprego remunerado, deu lugar a leis com as quais se pretende impedir a discriminação, nos casos em que já existem soluções técnicas.

O que se tem hoje, em termos tecnológicos, e ao nível da Internet, é algo que possibilita ao surdo uma independência quase total em relação ao ouvinte. Esta é uma característica extremamente positiva para o surdo, uma vez que não precisa mais contar com a boa vontade de um ouvinte. O surdo, através da Internet, começa a ter a possibilidade de se desenvolver de forma mais independente e auto-determinada, seguindo os ditames da sua própria consciência. Numa sociedade com uma linguagem eminentemente verbal, parece provável que o grupo de pessoas com deficiências auditivas possam ser os que tenham, talvez, uma maior necessidade de aceder a sistemas que facilitem a aquisição de conhecimentos e a interpretação da realidade.

Já é uma realidade, desde a década de 60, o uso do computador no suporte ao processo ensino-aprendizagem. Neste contexto, há aqueles que acreditam na substituição do professor pelo computador, sem perdas para o processo de aprendizagem, como também há os que não vêem ganho algum com o uso do computador na "escola". O que se tem percebido desta (longa) discussão, e dos resultados que já foram obtidos, é que o computador, usado como actividade complementar, pode trazer grandes ganhos a este processo.

2.1.1 Formação de Pessoas Surdas: Análise da Situação

Dificuldade no Processo Formativo

A pessoa surda enfrenta no seu processo formativo uma série de dificuldades centradas em três eixos principais:

1. A sua própria surdez que o situa em inferioridade de condições quando o principal meio de comunicação é a fala.
2. O professor ou qualquer outro agente mediador, por falta muitas vezes de um sistema de comunicação partilhado.
3. Os materiais, principalmente textos escritos, supõem um bom domínio do idioma correspondente para extrair informação dos mesmos.

Necessidades Específicas

A pessoa surda necessita para a sua formação de um meio de comunicação eficaz. Para isso, podem contribuir as aplicações informáticas específicas quando as aplicações convencionais, com ou sem adaptações, resultem insuficientes. O outro objectivo prioritário de qualquer programa informático educativo, pensado para pessoas com surdez, deveria ser o permitir o acesso à informação em condições adaptadas à sua deficiência sensorial e às peculiaridades da sua comunicação. Este objectivo não se pode limitar a adaptações curriculares, mas também, à adaptação do contexto educativo.

Dentro do campo das tecnologias de informação, poder-se-ia pensar numa série de aplicações informáticas dirigidas ao melhor aproveitamento das possibilidades educativas que, em princípio, ninguém nega a pessoas com surdez. No entanto, para que estas aplicações sejam efectivas deverão ter em linha de conta alguns princípios elementares. De outro ponto de vista, as tecnologias de informação oferecem possibilidades de desenvolver certas aptidões que se encontram geralmente minimizadas em alunos com surdez. No entanto, este ponto introduz elementos complexos que se referem concretamente à aquisição e uso do conhecimento. Enquanto que as condições anteriores se referiam a sistemas de adaptação sobre esquemas de ensino já estabelecidos, neste caso tratar-se-ia de programas desenhados para o ensino de determinadas destrezas.

Os tempos recentes têm testemunhado um aumento de interesse no campo da comunicação “homem-máquina” que implica a linguagem falada acompanhada de gestos e expressões faciais. Cassel [4] introduziu um sistema onde agentes virtuais pudessem comunicar automaticamente com outros agentes através de canais sincronizados que representam a fala, gestos e expressões faciais. Os agentes têm o seu próprio “conhecimento do mundo” e actuam de acordo com esse conhecimento. O sistema simula um comportamento real da comunicação humana. A geração de gestos consiste basicamente em símbolos com pouco ou nenhum movimento dos dedos, mas em qualquer dos casos já o bastante para anotar evoluções neste campo.

Thalmann, Noser e Haung [5] desenvolveram seres humanos virtuais autónomos para actuar em filmes, jogos e Televisão Interactiva. Os seres humanos virtuais

trabalham como objectos independentes e podem enfrentar situações complexas. Uma característica especial é o comportamento dos agentes virtuais. Podem representar conclusões sobre o seu conhecimento baseados na percepção do entorno virtual. As mãos, equipadas com sensores, reconhecem tamanhos e objectos, os agentes podem pegar nesses objectos, movê-los e largá-los. Um complexo comportamento com base em movimentos de braços e mãos simula realisticamente o comportamento humano.

Um dos projectos pioneiros no treino e ensino de língua gestual com o apoio de modelos 3D virtuais foi o sistema de diálogo para sintetizar e compreender a língua gestual Japonesa, descrita e desenvolvida por Lu e Igi em 1997 [6]. Este sistema de treino, ajuda a comunicação de pessoas surdas com a comunidade falante. Um conjunto de palavras e frases são predefinidas e capturadas através de gestos para posterior utilização. A desvantagem deste sistema é a representação pouco detalhada dos modelos, principalmente da mão e da cara. Após este projecto, muitos outros se seguiram, alguns dos mais importantes que abarcaram muito de perto a génese desta dissertação serão enunciados a seguir.

Não se pretende assim, um método de visualização que compita com pessoas reais na produção de língua gestual, já que uma qualidade tal só muito dificilmente poderá ser alcançada.

Em continuação, serão descritos vários projectos envolvidos na investigação e testes de suportes computacionais para o ensino de língua gestual.

2.2 Testes de Suportes Multimédia Para a Educação e Supressão de Barreiras de Comunicação Para Pessoas Surdas

Foram analisados suportes multimédia por um grupo de profissionais em centros de integração onde estudam alunos com deficiências auditivas [7]. Os programas testados foram sobretudo aplicações adaptadas à aprendizagem para pessoas surdas. Foram eles:

- Symbol – Base de dados, para a aprendizagem de vocabulário, destinado a pessoas com transtornos para se comunicar oralmente [7].
- CD_Rom interactivo de visita ao Palácio Real de Madrid [7].

Teve-se em linha de conta uma dupla vertente, a técnica e a adaptabilidade dos utilizadores, com o fim de analisar a possível distância entre ambos, e acima de tudo para analisar o que se consideram elementos básicos, como em qualquer teste.

2.2.1 Symbol

Em geral, tira pouco partido da tecnologia multimédia que utiliza, há poucas imagens em movimento, narração e sons, ajustando-se basicamente às oportunidades que oferece um computador. Neste sentido poder-se-iam introduzir muitas melhoras no que respeita à interacção, movimento, etc. A ideia de uma base de dados de imagens com

possibilidade de associar a cada imagem o seu significado em língua gestual, tal como a aplicação faz, pode contribuir a melhorar os requisitos básicos do desenvolvimento verbal na comunidade surda.

Como ideia e formato, para trabalhar na aprendizagem com pessoas de deficiência auditiva, é um bom produto, já que permite a observação dos modelos operativos do aluno e têm uma forma flexível de apresentação da informação. O tipo de cenário utilizado, uma casa, faz que talvez não seja idóneo para jovens em formação, pelo que trabalhando noutro entorno, poderia dar um melhor resultado.

A utilização de imagens reais desperta muito interesse entre os utilizadores. O modelo da aplicação pode constituir um entorno adequado para a formação de conceitos, entretenimento lógico, metalinguístico, memorização do léxico, etc, no entanto poder-se-iam sugerir as seguintes melhorias:

- Seria conveniente pensar na aplicação deste programa a campos semânticos específicos.
- A navegação pelo sistema (selecção do caminho a seguir em cada momento) não é de todo intuitiva e é relativamente fácil esquecer alguns passos. Sente-se falta de um sistema de ajuda que permita escolher a opção desejada em cada momento.

2.2.2 CD_Rom Interactivo de Visita ao Palácio Real de Madrid

Em geral, como ferramenta de apoio à docência, é um elemento motivador e que ajuda a que os alunos participem na busca de informação, o que no caso das pessoas com deficiência auditiva, é muito importante, já que supõem um entretenimento cognitivo.

A sua eficácia para pessoas com deficiência auditiva seria muito maior se pudesse dispor de resumos em linguagem clara e fácil dos textos contidos no programa.

É um programa que resulta atractivo pela motivação e aquisição de novas formas de conhecimento, mas para se adequar às necessidades curriculares das pessoas com deficiência auditiva, teria que sofrer modificações sobretudo no que respeita à língua.

O seu aproveitamento para pessoas com deficiência auditiva passaria, então, por acrescentar um glossário de fácil acesso dentro da estrutura do programa para as palavras complexas mais frequentemente utilizadas nos diferentes textos, esse glossário deveria usar uma linguagem o mais simples possível e incluir a sua tradução em língua gestual.

2.2.3 Critérios e Adaptações Específicas Para Pessoas com Necessidades Especiais

Segundo Vanderhiden [8], devemos partir da ideia de que não se devem gerar sistemas especiais fechados para pessoas com ausências de algumas capacidades, salvo quando não se possa resolver de outra forma. Para este, existem uma série de considerações básicas a ter em linha de conta:

- Utilizar uma perspectiva de sistema aberto.
- Cooperação com as utilidades de acessibilidade do sistema operativo e programas específicos.
- Minimizar as habilidades necessárias para o manuseamento do programa.
- Facilitar manuais e formação acessível.

Certamente, que se os requisitos atrás mencionados fossem levados à prática, a adaptação destas ferramentas às pessoas surdas seria muito menos complexa e nalguns dos casos desnecessária. Não obstante, convêm fazer uma série de recomendações de forma a que estes sistemas se possam utilizar como se fossem “feitos à medida”.

Em geral e no caso de pessoas com deficiências auditivas:

- É necessário um formato visual para todos aqueles avisos/sinais sonoros do programa.
- Eleição de uma linguagem adaptada ao nível compreensivo do utilizador evitando o calão informático.
- Maior frequência de imagens.
- Possibilidade de complementar algumas explicações mediante a língua gestual.
- A necessidade de qualidade de interfaces é ainda maior para pessoas com deficiências auditivas, pela sua incidência no processo de aprendizagem.

2.3 Representação de ASL na World-Wide Web com Soletração Através de Dedos

Este trabalho apresenta técnicas que possibilitam alargar o vocabulário de gestos em ASL (American Sign Language) através de VRML 2.0 com propósitos educacionais.

O VRML (Virtual Reality Modeling Language) [9] é uma linguagem standard na descrição de modelos gráficos tridimensionais na WWW. Dado que os modelos 3D em VRML proporcionam uma documentação visual mais efectiva quando comparada com os tradicionais 2D, através de imagens ou vídeos [10], a construção de gestos em VRML não ajudará unicamente na aprendizagem de ASL aos surdos, como também proporciona um acesso universal na aprendizagem deste tipo de material.

Depois de um primeiro olhar, o conceito de construção de modelos 3D para gestos ASL em VRML é simples e intuitivo. No entanto, descobrir como ampliar um vocabulário estático de desenhos de mãos em VRML, tal com o criado por Geitz, et al [10], num vocabulário mais alargado, já não é trivial. A seguir, são apresentadas técnicas que sistematicamente resolvem este problema.

O trabalho foi realizado assumindo que os gestos ASL podem ser descritos em termos de características linguísticas de gestos com as mão (modelo das mãos, posições, orientações e movimentos). A ideia baseia-se em sistemas de gestos de Stokoe, Casterline e Croneberg [11]. Dado que cada movimento têm um ponto de partida e um ponto final, o problema de descrição de gestos dinâmicos pode reduzir-se ao seleccionar

de vários gestos estáticos como *key-frames*, e depois, construir a animação entre eles para a formação do gesto. A expressão facial e linguagem corporal não foram consideradas neste trabalho.

Neste estudo é descrita a colocação em prática de uma interface acessível via Web para gestos com base em ASL. O trabalho teve início com o modelo tridimensional simples da mão e foram adicionadas especificações para todos os graus de liberdade na mão para possibilitar a construção de gestos estáticos. Na interface, os graus de liberdade de uma mão podem ser alterados pela manipulação de um painel de controlo. Tanto o controlo entre os dedos e pulso, como a localização e orientação do modelo da mão são passíveis de ajustamento individual.

Depois, foi utilizada a função de animação *key-frame* para a composição gestos dinâmicos através da especificação e integração de gestos estáticos pré-seleccionados em *key-frames*.

O Web site proporciona um protótipo de um conceito de professor virtual. Aqui os utilizadores poderão introduzir uma cadeia de caracteres ou números e ver a sequência de animação correspondente em ASL gerada dinamicamente em VRML.

2.3.1 O Modelo da Mão

O primeiro passo para a representação de gestos em VRML foi a construção de um modelo 3D de uma mão. Para guardar o ficheiro com o tamanho de menor dimensão possível, foi decidido preservar os graus de liberdade na mão mas ignorar outros detalhes, tais como a textura.

O modelo da mão consiste nas seguintes entidades: antebraço, palma, e três “ossos” para cada dedo e polegar, estes últimos construídos por um cilindro e duas esferas associadas as extremidades dos cilindros para representação de cada uma das articulações. Os cilindros e as esferas são nodos geométricos em VRML (primitivas). Para simplificar ainda mais a implementação deste modelo, foi ignorada a parte de ligação de “carne” entre o polegar e o dedo indicador, o que torna sem dúvida o modelo um pouco menos real. Dado que o modelo é de facto simples, foi conseguido um tamanho de ficheiro muito reduzido (cerca de 9KB).

Os graus de liberdade principais na mão podem ser caracterizados como o seguinte [12]: cada dedo pode fazer a flexão distal, flexão média, flexão próxima e desvio próximo. O polegar é similar. No total temos 20 graus de liberdade para os dedos e polegar. Além disso, o punho pode flectir em dois sentidos e o antebraço tem 6 graus de liberdade: x,y,z para desvio lateral, desvio frente-trás e rotação. O resultado final são 28 graus de liberdade no modelo completo da mão, o que representa a totalidade dos graus de liberdade existente numa mão humana, e assim sendo, um modelo com capacidades de emular muito razoavelmente os movimentos desta extensão humana.

2.3.2. Painel de Controle

O meio proporcionado para especificações dos graus de liberdade no modelo da mão é através de um painel de controle. A Figura 1 é uma imagem do painel de controle construído em VRML. Cada controle do tipo roda é usado para ajustar um grau de liberdade, cada controle do tipo bola é usado para ajustar dois graus de liberdade ao mesmo tempo. Para uma fácil compreensão, as rodas e as bolas foram dispostas nas posições correspondentes ao ângulo de visão do modelo da mão onde cada grau de liberdade poderá ser ajustado com alguma facilidade.

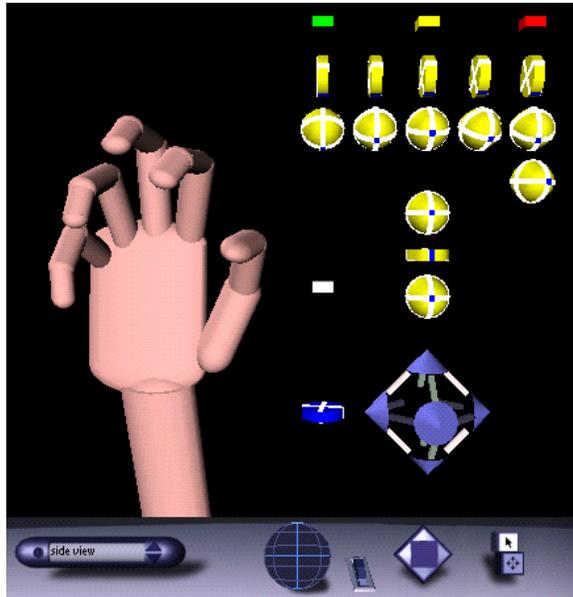


Figura 1 – Painel de controle (mão direita)

Após a definição de um gesto, podemos clicar no botão de controle para o envio dos valores de graus de liberdade remotamente para uma *script* CGI, de nome GestureMaker. O GestureMaker responderá com um novo ficheiro VRML contendo todas as especificações. Este foi o método usado para a produção de todos os gestos estáticos do alfabeto ASL incluindo também números (do 0 ao 9). Nesta especificação, foi escolhida arbitrariamente uma das possíveis variações para cada gesto [13], no entanto seria de todo interessante adicionar as diversas variações.

2.3.3 Animação

Com modelos de mãos estáticos, gestos dinâmicos poderão ser criados através de animação tipo *key-frame*. Foram usados nodos para interpolação linear de valores para campos de translação e rotação. Os valores de graus de liberdade para a sequência de *key-frame* são passados à *script* GestureMaster, o qual retorna um ficheiro VRML para interpolar gestos dinâmicos sobre o modelo.

Por exemplo, a letra “Z” no alfabeto ASL é um gesto realizado com o uso do dedo indicador para desenhar um “Z”. Quatro pontos (instâncias do modelo) são usados como *key-frames* para a criação do gesto dinâmico para o “Z”. A letra “J” usa o dedo midinho para desenhar o “J”, são portanto necessários 3 pontos que depois de interpolados apresentam a sequência e animação pretendida.

2.3.4 Soletrar Através dos Dedos

Após a criação de gestos dinâmicos de letras ou números, foi dado o passo seguinte, a criação da animação entre as várias letras e números. A Figura 2 demonstra através do web site a animação da mão direita traduzindo gestos formados por cadeias de letras e números introduzidos numa caixa de entrada. Os utilizadores podem também visualizar dígito a dígito animações de letras ou números.



Figura 2 – Tradução ASL de cadeia de letras e números (mão direita)

O problema óbvio no sistema é a colisão e penetração de dedos e polegar em várias entradas de textos, como por exemplo a passagem da letra “A” para a letra “B”. Neste caso, foi necessário inserir caminhos de movimentos durante estes casos específicos e especiais. Mas o ideal seria o desenvolvimento de um algoritmo geral para detectar e solucionar este problema.

2.4 Modelos de Markov Para o Ensino Interactivo de Gestos

2.4.1 Objectivos do Projecto

Christopher Lee e Yangsheng Xu [14] desenvolveram em 1996 um sistema de reconhecimento de gestos baseado em modelos ocultos de Markov (HMM⁴), que podem interactivamente reconhecer gestos e a capacidade de aprender e realizar novos gestos *online*. Além disso, é capaz de fazer modificações após reconhecimento interactivo do exemplo de alguns gestos. Este sistema demonstrou reconhecimentos de confiança em 14 gestos diversos depois de somente um ou dois exemplos de cada um deles. A interface do sistema tem ligação com uma *Cyberglove*⁵ para a captura de gestos do alfabeto da língua gestual (Figura 3). O sistema foi implementado como parte de uma interface interactiva para, por exemplo, programação e tele-operações num robô.



Figura 3 – Posição final da mão para o gesto “C”

As motivações por detrás deste tipo de interacções, para além da aplicação à língua gestual, incluem o aumento da efectividade da tele-operação, dispondo as pessoas para interactivamente ensinar novas tarefas a robôs ou refinar as suas habilidades e permitir um controlo do sistema mais eficaz e semi-autónomo como acontece com os sistemas dos aviões e plantas industriais.

2.4.2 Construção do Sistema

Este trabalho teve como objectivo a construção de um sistema que possa não só interagir com o utilizador no reconhecimento de gestos, mas também que possa aprender novos gestos, modifica-los de modo interactivo e *online*. O resultado foi assim a geração

⁴ Hidden Markov Models

⁵ Luva de Dados Virtual do tipo “*Cyberglove*”

automatizada e um treino interactivo de um conjunto de HMM que representam gestos humanos. Usando este caminho como objectivo, foi construído e testado um sistema capaz de reconhecer palavras e letras da língua gestual usando tecnologia virtual baseada numa *Cyberglove* para a entrada de dados (Figura 4).

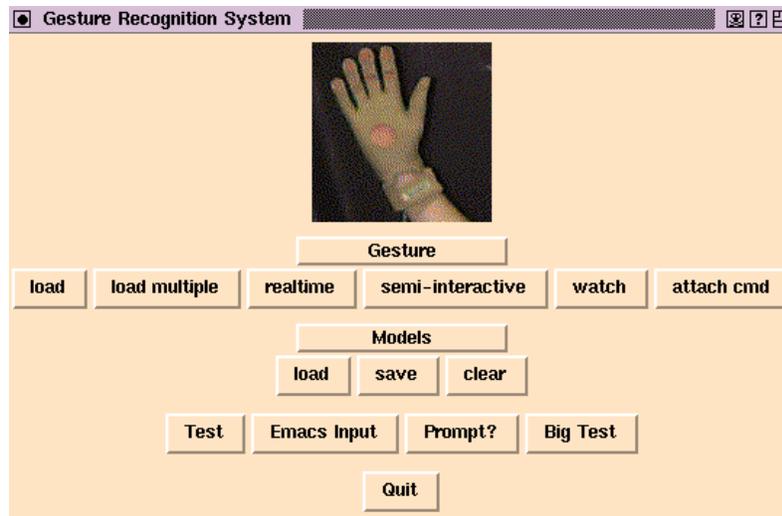


Figura 4 – Interface do sistema de demonstração para reconhecimento de gestos.

Neste sistema, cada tipo de gesto é representado por um modelo de Markov, uma lista de sequências de observações e uma acção opcional a ser realizada sobre o reconhecimento do gesto. O conceito de treino interactivo é baseado geralmente nos seguintes procedimentos:

1. O utilizador faz uma série de gestos.
2. O sistema automaticamente divide o conjunto de dados de entrada em imagens e gestos separados, e tenta imediatamente classificar cada gesto.
 1. Se o sistema está certo da sua classificação, ele realiza a acção associada a esse gesto (se alguma foi especificada). Esta acção poderia passar o resultado da classificação a um nível superior de HMM, ou enviar o comando a um robô.
 2. Se o sistema de alguma forma se mostra inseguro na classificação de um determinado gesto, pergunta ao utilizador para confirmar essa classificação. O utilizador ou:
 - Confirma a classificação do sistema, ou
 - Corrige a classificação, ou
 - Adiciona um novo tipo de gesto à base de dados de gestos do sistema.
3. O sistema adiciona todos os “símbolos” do gesto encontrado à lista de sequência de exemplos desse próprio modelo de gesto, a seguir actualiza os parâmetros da reentrada do modelo de HMM em acumulação com exemplos de sequência já existentes.

Encontra-se nesta implementação um reconhecimento e actualização automática de gestos com base em HMM e algoritmos de Baum-Welch, suficientemente rápidos durante o normal uso do sistema o que representa um sistema com carácter verdadeiramente interactivo.

2.5 Projecto Mudra - Multimédia e Animação 3D Para Pessoas Surdas

2.5.1 O Projecto

Com o intento do projecto intitulado "3D Techniques for Computer-Based Training in Sign Language", um protótipo de hardware/software para gravação, editoração e visualização de gestos foi desenvolvido, com base em modelos em três dimensões e animação por computador [15]. Este projecto de investigação e desenvolvimento começou em Abril de 1998 e durou até Março de 2000 no departamento de Media Technology and Design da Universidade Politécnica de Upper na Áustria. Este projecto envolveu a colaboração de algumas empresas de mídia austríacas e foi suportado pelo Ministério da Ciência daquele país.

A comunicação entre aqueles que podem ouvir e a comunidade surda não poderá ser melhorada sem formação e o conhecimento da língua gestual. O primeiro passo nessa direcção poderá ser criado oferecendo, a quem pode ouvir, a oportunidade para apreender a língua gestual em larga escala. Os gestos podem ser apresentados com verdadeira qualidade através de tecnologias multimédia, produtos de treino deste género estão a ser desenvolvidos um pouco por todo mundo incluindo todos os diferentes tipos de língua gestual.

Até agora, gestos únicos eram apresentados através de sequências de vídeo como programas de treino. O objectivo principal desta ideia/projecto é o desenvolvimento de técnicas para gravação de gestos e a sua visualização através de animações 3D via computador. Técnicas de captura de movimentos foram implementadas através de sensores e de seguida armazenados numa base de dados.

Este projecto permite assim, armazenar gestos usando captura de movimentos com ajuda de hardware, seguindo a posição e movimentos das mãos, braços e do corpo superior, ocorrendo a visualização já durante a gravação dos movimentos. Os dados dos movimentos obtidos são aplicados ao software de animações 3D onde aí poderão também ser editados e corrigidos. Após isso, os dados da animação deverão ser exportados para um formato de um ficheiro específico onde são tratados por um motor 3D na aplicação MUDRA. As características interactivas são proporcionadas ao utilizador do programa através de várias opções de câmaras e ângulos de visualização 3D.

Resumindo, os três aspectos mais importantes que resultam do projecto são:

- Captura de movimentos

- Visualização, edição e correcção de dados
- Integração das animações 3D numa aplicação multimédia

A verdadeira meta deste projecto é posta em prática pela implementação de técnicas mencionadas mais adiante na gravação, visualização e geração de animações gestuais gráficas. A gravação da multiplicidade de gestos que serão necessários ao projecto MUDRA, não foi tarefa deste projecto.

2.5.2 As Vantagens

- Em contraste com os vídeos onde os gestos a estudar são apresentados apenas de um ponto de vista fixo, na visão através de animações interactivas 3D pode escolher-se arbitrariamente o ângulo de visão. Neste sentido um gesto poderá ser mostrado através de um humanóide nos vários ângulos (Figura 5), o que facilita consideravelmente a aprendizagem de todos os gestos.

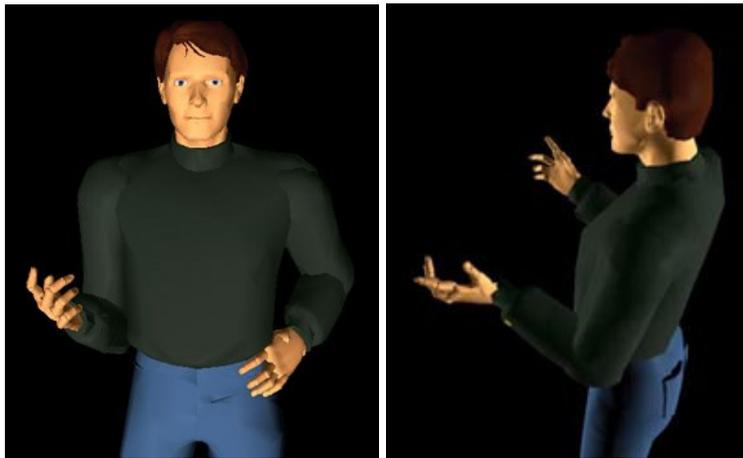


Figura 5 – Humanóide 3D

- Os dados da animação poderão ser interpolados. Assim, a mesma taxa de *frames* poderá ser mantida para qualquer velocidade “*playback*” escolhida.
- Além disso, o utilizador poderá alterar a escala de representação em tempo real da animação para um tamanho arbitrário. Quando o faz, não há decréscimo de qualidade como acontece no caso dos vídeos.
- Os dados que descrevem os polígonos do humanóide 3D serão unicamente guardados uma vez para qualquer animação. As animação 3D consomem menos espaço de memória do que as sequências em vídeo já que para cada *frame* da animação somente os ângulos de junção do corpo humano necessitam de ser conhecidos. As técnicas 3D poderão ser incorporadas em aplicações tendo como base a Internet.

- Outros problemas tais como cenas adequadas de iluminação são muito comuns nas produções em vídeo, mas não com o uso de animações por computador, por exemplo, fontes de luz poderão ser adicionadas ou removidas arbitrariamente pela animação.
- As técnicas 3D aqui utilizadas poderão também á posteriori ser usadas como um intérprete para língua gestual em transmissões televisivas.

2.5.3 Descrição Técnica

As origens das técnicas da captura de movimentos para a gravação digital dos movimentos através de sensores foi encontrada no campo do exército e aviação. Para a animação por computador várias técnicas foram primeiro refinadas.

No armazenamento de dados e conseqüente reprodução das animações gestuais 3D com movimentos tão exactos, são necessários meios que geralmente não são necessários para animações em vídeo ou jogos de computador. Neste projecto, no entanto, uma modelação correcta da cinemática humana e a colocação exacta das partes constituintes do corpo é essencial. Assim como o ajuste do humanóide 3D com as proporções do actor, bem como a “calibragem” dos sensores utilizados.

Para a gravação dos movimento de mãos e dedos são utilizadas duas *Cybergloves*. A versão destas luvas de dados utilizadas neste projecto tem capacidade de medir até 18 ângulos de junção de dedos e pulso através de sensores em cada uma das mãos. Para conseguir a posição e orientação dos braços e a parte superior do corpo são usados três sensores do sistema *Frank Polhemus*. Um quarto sensor é utilizado para medir os movimentos da cabeça. Cada um destes 6 sensores trabalha com 6 graus de liberdade em cada ciclo de medida, determinando os campos gerados pela voltagem electromagnética com a ajuda de um transmissor.

No início de 1999 o *Cyberglove* era o único método comercial (para o propósito deste projecto) suficientemente exacto de luvas de dados. Mas, manteve-se também o contacto com a Universidade Técnica de Berlim, onde uma *Sensorglove* estava em desenvolvimento num projecto cujo objectivo seria o reconhecimento de gestos. Infelizmente esta luva de dados ainda estava numa fase de protótipo (naquela época), sem nenhum software lógico de calibração e visualização disponível. A decisão pela *Frank Polhemus* foi baseada no resultado de vários trabalhos científicos. Outras tecnologias com sensores foram abandonadas com problemas de exactidão e multiplicidade de vistas.

Até há bem pouco tempo, a visualização em tempo real da animação de um humanóide era observada basicamente através da última geração de hardware, especialmente através da captura paralela (luvas virtuais) de movimentos e os dados adicionais resultantes dos sensores que deveriam ser processados posteriormente. Uma animação suficientemente realista da face humana realizada com deformações 3D, ainda não é possível de tratar em tempo real num equipamento comum.

2.5.4 A Mostra de Material 3D

A calibragem das duas *Cybergloves* deu lugar ao uso do software que acompanhava as luvas de dados. É um programa com visualização dos movimentos das mãos esquerda e direita, com um painel de comandos tipo barras de deslocamento, que podem ser usados para a calibragem de cada um dos sensores em particular de uma das luvas. Assim sendo, o valor dado pelo sensor dos *Cyberglove* poderá ser ajustado ao tamanho da mão do actor. Os dados calibrados são gravados num ficheiro para cada luva de dados.

A captura de movimentos foi suportada pelo software de animação 3D “*Softimage3D*” num Pentium III sobre Windows NT. O sistema foi equipado com um acelerador gráfico OpenGL (*FireGL 1*). Os dados recebidos pelos sensores são encaminhados por três portas série e aplicados a um humanóide 3D em tempo real pelo nomeado *Softimage Channel Drivers*. Os *drivers* da *Cyberglove* usam a biblioteca *VirtualHand* fornecida com a *Cyberglove*.

O humanóide 3D para a representação dos gestos – de nome “Daniel”- consiste num conjunto de polígonos com capacidade de rotações sobre eixos de junção das “peças”. Não há deformações nesta malha de polígonos. A posição e orientação da parte inferior dos braços, parte superior do corpo e cabeça são determinados pelos sensores *Polhemus*. A parte alta dos braços é orientada e adequadamente escalada durante a captura de movimentos para a construção de uma correcta conexão entre os ombros e os cotovelos.



Figura 6 – Características para a interacção 3D

Foram feitos muitos avanços para poder ajustar as proporções do actor com os elementos do humanóide 3D, e os sensores *Polhemus* conseguiram fazer uma representação correcta no humanóide 3D. Este avanço e a cinemática da parte superior dos braços realizada no *Softimage* permitiram representar uma multitude de expressões.

Houve como não poderia deixar de ser um cuidado com a exactidão durante a gravação dos movimentos, para que representassem verdadeiros “gestos” com a maior exactidão possível. Um manual de correcção das trajectórias foi obtido, e implementado devido ao elevado número de gestos a serem guardados.

2.6 “Avatar” Gestualista na Web

Este trabalho é parte do projecto ViSiCAST [16], subsidiado pela União Europeia. O projecto tinha o objectivo de anunciar previsões meteorológicas a partir de um humano virtual (Avatar), guardando as traduções semi-automáticamente a partir de texto para língua gestual. Com a ajuda do computador, palavras escritas são analisadas e transpostas numa sequência de gestos. Posteriormente, estes gestos são exibidos, usando animação por computador a partir de um humanóide. As áreas de foco da aplicação seriam a televisão, www/multimédia e as chamadas transacções cara a cara [17,18]. Assim sendo, a aplicação estaria disponível também numa página da world-wide-web. Para isso, foi desenvolvido um *Plug-in* para os *Browsers* e gerado o conteúdo inicial em língua gestual, conteúdo inicial este que focava como foi dito atrás a previsão do estado do tempo.

Tecnicamente há duas formas de animar língua gestual: com base na captura de movimentos e baseada na criação de gestos sintéticos. No projecto ViSiCAST a parte de gestos totalmente sintetizados estava ainda em fase de desenvolvimento [19,20]. Por outro lado, o anúncio da previsão meteorológica é realizado com base na captura de movimentos reais. Para a aplicação web, previsões meteorológicas reais foram usadas como fluxo de *Input*. Antes da previsão de um determinado dia em particular poder ser gesticulada na página web, teriam de se suportar algumas conversões como se mostram na Figura 7.

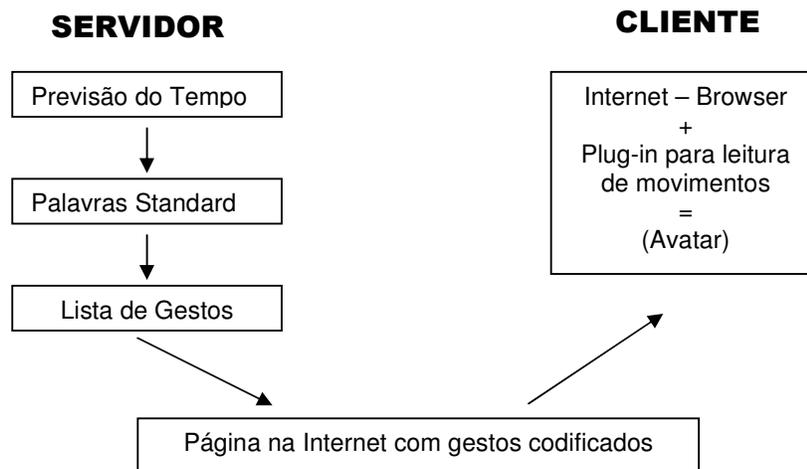


Figura 7 – Conversões entre o Servidor e Cliente

As secções apresentadas descrevem as conversões das previsões via modelos de línguas standard para listas de gestos e as conversões das listas de gestos via página web para a animação no écran do utilizador.

2.6.1 Gravação e Visualização dos Dados a Partir dos Movimentos dos Gestos

Como foi mencionado atrás, a aplicação para anunciar a previsão do estado do tempo na web através de um “Avatar” é baseada na captura de movimentos [21]. Por conseguinte, todos os gestos potencialmente relacionados com uma possível previsão meteorológica foram capturados e gravados. A captura de um gesto significa que os movimentos de um gestualista real são recolhidos através dos seguintes periféricos de Input: um sistema óptico para a recolha de expressões faciais, um fato magnético para a definição das posturas e luvas de dados para o movimento das mãos e dos dedos (Figura 8). Com estes periféricos no corpo, o gestualista realiza os gestos necessários à aplicação, para este caso, gestos como: sol, chuva, zero graus, outros números, partes do dia, etc.



Figura 8 – Captura de movimentos com ajuda de um intérprete de língua gestual

Foram também construídos gestos para três tipos de línguas gestuais: língua gestual Holandesa, Alemã e Britânica, cada uma a partir de um gestualista nativo do país correspondente. O equipamento construído originalmente era próprio unicamente para a captura de movimentos globais. Para os gestos, a qualidade dos requisitos de captura foram muito maiores do que os usados com esta técnica em aplicações prévias. Para a geração de língua gestual animada, fizeram-se testes para a velocidade e direcção dos gestos como base para testar o suporte animado do significado inerente ao gesto. O

equipamento sofreu assim melhoramentos, e na versão final, os movimentos corporais, posturas, movimento das mãos e expressões faciais foram capturadas com uma resolução elevada (mais elevada do que a normalmente usada nos jogos de vídeo). Os gestos resultantes poderiam assim ser facilmente reconhecidos aquando da sua exibição no écran do computador.

O parceiro no projecto responsável pelo equipamento de captura, também desenvolveu um software que tornasse possível o posterior processamento dos gestos. A parte da visualização significava então também, que o programa do humano virtual “Avatar”, estava dotado com dados gravados a partir de três sistemas (captura, sintetização e edição). Esta visualização era assim necessária para um posterior processamento manual. Alguns gestos eram guardados num único ficheiro, o que derivava que alguns dados teriam de ser “cortados” para obtenção de gestos em separado. Além do mais, era necessário rever os dados e editar aqueles que se achasse necessários. Podia por exemplo, acontecer que o gestualista humano, ele próprio, não realizasse o gesto na perfeição ou o sistema das luvas de dados não registasse o declive necessário em certos movimentos de dedos, porque, por exemplo, as luvas poderiam estar um pouco largas em relação às mãos do gestualista.

O Avatar desenvolvido pelo ViSiCAST foi chamado de “Visia”. Os gestos capturados (e possivelmente editados) podem ser assim re-especificados, um a um, separadamente ou numa sequência perfeitamente fluente. O software de visualização em tempo real tinha possibilidade de misturar duas sequências capturadas, o que resultava numa performance gradual dos gestos. Isto não seria possível com a junção de *clips* de vídeo. No sistema de previsão meteorológica ViSiCAST, há também a secção de construção e geração de frases gesticuladas incluindo algumas variáveis na construção de gestos tal como temperaturas, estados do tempo e direcção do vento.

2.6.2 Dos Dados Extraídos dos Movimentos Para a Página Web

Tecnicamente, o vasto conjunto de ficheiros de movimentos e o software descrito eram suficientes para visualizar sequências de gestos do Avatar “Visia”. No entanto, esta tecnologia não era “amigável” para com o utilizador como também não era a forma correcta para a sua integração na Internet. O problema do envio de informação através da Internet é assim resolvido em dois passos. Por um lado, o software Avatar e os ficheiros de movimentos são guardados num PC. Por outro lado, a lista de carregamentos dos movimentos é codificada numa linguagem especial (SIGML – Signing Gesture Markup Language, desenvolvida no ViSiCAST) e incluída numa página HTML normal. De modo a visualizar e controlar o Avatar de uma forma correcta e amigável para o utilizador, foi criada uma interface simples e de uso fácil. O software para a visualização gestual foi criado para correr como um *Active X control*. Isto permitiu a criação de versões compatíveis com as aplicações Microsoft Windows, em particular com o Internet Explorer. Com base em experiências com pessoas surdas, com limitações literárias ou não, foi desenhada uma página web para a previsão do estado do tempo através então de um humanóide que reproduziria língua gestual (Figura 9).



Figura 9 – Reprodutor de língua gestual

Na Figura 9 pode ver-se o avatar “Visia”, encontram-se presentes três botões, similares aos encontrados num reprodutor de vídeo. As funções são: Início dos gestos, possibilidade de fazer pausa num dado ponto da previsão e parar toda a animação. Uma barra de deslocamento poderá ser usada também para saltar para um ponto particular de modo a resumir ou repetir uma parte da previsão.

Concluindo, foi o primeiro protótipo no anúncio da previsão do estado do tempo a partir de uma aplicação do género, resultando da combinação de técnicas de conhecimento de vários campos: animação virtual humana, e captura de movimentos em língua gestual.

2.7 O Projecto “Simon”

No Reino Unido, a ITC (Independent Television Commission) comissionou em 1999 o projecto da Televirtual [22] para o desenvolvimento do Simon, um gestualista virtual (Figura 10) para tradução de textos impressos e capturas televisivas em língua gestual, para resolver alguns problemas de agregação de língua gestual a programas televisivos.



Figura 10 – Gestualista virtual Simon

A apresentação com gestualistas humanos em directo é algo dispendioso para os canais televisivos, e os serviços gestuais requerem ainda facilidades específicas ao nível de estúdios e de pessoal.

No Reino Unido, o reconhecimento crescente das necessidades inerentes à comunidade surda resultaram numa legislação em relação às televisões de forma a passarem a fornecer a tradução em língua gestual, incrementando ao longo do tempo, o número de programas acompanhado por esta tradução. O problema para os canais televisivos poderia ser resolvido então com um novo sistema de geração de língua gestual, capaz de resolver a maioria das edições televisivas. Seria perceptivelmente mais barato de produzir para os canais televisivos, necessitaria significativamente de menor capacidade de transmissão do que um segundo sinal vídeo MPEG-2 de um gestualista humano num quarto do écran, proporcionaria então gesticulação automática a todos os programas legendados ou não, resolvendo assim as necessidades da comunidade surda.

O software consiste em dois módulos básicos: traduções linguísticas a partir de Inglês impresso em língua gestual, e a animação de um humano virtual. O software de animação permite que o Simon possa traduzir em tempo real. Um dicionário de gestos permite ao sistema o acompanhamento do movimento físico, expressão facial e posições do corpo, que são guardadas como dados de movimento capturado (não em imagens ou em vídeo) num disco rígido. O sistema pode chamar esses movimentos físicos (em qualquer ordem), e fazer interpolação entre eles de forma a criar seqüências naturais e graduais dos gestos. Os dados de captura de movimentos – que incluem informação de mãos, corpo e cara – são aplicados a um modelo 3D de um humano virtual com alto grau de detalhe. Este modelo inclui representações muito realistas e exactas das mãos e cara de forma a criar a máxima expressão e realismo na caracterização dos gestos. Outros modelos gráficos foram construídos com personagens muito mais novas e até “desenhos animados” de forma a poderem ser escolhidos pela audiência televisiva consoante a idade.

Os dados dos movimentos foram capturados a partir dos gestos de intérpretes experientes em língua gestual, usando três formas de captura avançada de movimentos (Figura 11):



Figura 11 – Captura de movimentos a partir de um intérprete em língua gestual

1. *Cybergloves VR* como luvas de dados para a captura do movimento das mãos.
2. Um fato especial magnético foi utilizado para adquirir as posições do tronco superior, braços e posições da cabeça.
3. Foi utilizado o dispositivo óptico *Facetrack* na captura de expressões faciais e posições dos lábios.

2.7.1 Futuros Desenvolvimentos

Ainda em 1999 houve um debate sobre a data a partir da qual seria razoável os operadores de televisão começarem com as transmissões de serviços gestuais. Naturalmente, esta data dependeria também da capacidade das *set-top-boxes* a utilizar já terem a tecnologia disponível, e também, a realização de um acordo para a standardização gestual a transmitir. Era também esperado que a comunidade surda preferisse usar a British Sign Language (BSL) em prol da Sign Supported English (SSE). Mas o sistema estava ainda unicamente funcional para SSE, um dos desenvolvimentos futuros seria a sua conversão para BSL.

Além disto, havia três etapas bem definidas para o presente e futuro do projecto:

1. Etapa dos movimentos corporais. Melhorar na captura de movimentos corporais. Um pequena quantidade de dados comprimidos a partir desta captura representariam os movimentos transmitidos no sinal digital de televisão até às *set-top-boxes*, onde o Simon imitaria os movimentos com exactidão.
2. Reconhecer que o sistema de tradução linguístico pode ser empregue num estúdio produzindo automaticamente língua gestual.
3. Etapa para a tradução de língua gestual a partir unicamente de áudio.

2.8 Modelação e Animação de Língua Gestual Representada por meio de Linhas

É agora apresentado um projecto com uma abordagem diferente, mas com os mesmos objectivos, o ensino de língua gestual através de modelos computacionais, que não se fiquem unicamente pelos vídeos de personagens reais ou imagens estáticas.

Godenschweger e Wagener [23] consideram que um dos objectivos desde há muito para ajudar as pessoas surdas é o desenvolvimento de sistemas automáticos para produzir língua gestual. Este grande problema pode decompor-se em pontos mais específicos:

- Obter um modelo de linguagem capaz de ser gestualizado.

- Modelar e descrever os movimentos corporais necessários na composição de gestos.
- Obter gráficos 3D via computador capazes de simular esses movimentos gestuais.

Não houve progressos satisfatórios na solução de nenhum deste três pontos, mas o trabalho está em curso em cada uma destas áreas.

Outro problema que ocorre na animação de seres humanos virtuais está na sincronização dos movimentos e gestos humanos. A dificuldade de apresentar uma postura certa no momento certo, conduziu na utilização de agentes reais para ajudar na captura de movimentos através de dispositivos especiais. Assim se garante uma especificação estética realista dos movimentos do agente virtual. No entanto este método de registo somente pode ser usado para o caminhar e movimentos gestuais. Outras tarefas adicionais para o actor virtual, como por exemplo a expressão facial, requerem novos processos de captura de movimentos mais caros e que consomem mais recursos, por isso, se mantém o grande problema na geração de língua gestual.

Foram propostos algoritmos para a geração de expressões de língua gestual através de gráficos de linhas. É usada a vantagem da natureza abstracta das linhas, onde não é necessária uma apresentação muito detalhada. Nos gráficos de linhas, somente são visualizados os contornos e algumas outras linhas e é possível focar certas áreas usando diferentes estilo de linhas [24]. Neste caso uma apresentação através da traçagem de linhas leva-nos longe na simplicidade, modelos de menor detalhe mas com uma imagem estética fácil de entender pelas pessoas surdas.

O uso de gráficos de linhas para a língua gestual não é um campo de investigação desconhecido. Nos livros de texto, as ilustrações são representadas por norma em gráficos de linhas. A razão, aqui, é que o desenho através de linhas de mãos e dedos, podem ser facilmente reconhecidos e distinguidos.

Os algoritmos aqui descritos foram desenhados para serem usados em numerosas aplicações. Por exemplo, um dicionário pode ser construído e editado para representar gestos. Ainda, um sistema de ensino para tradutores de língua gestual poderia ser desenvolvido. Este tipo de algoritmo é especialmente atractivo para aplicações via Internet devido a velocidade de transmissão.

2.8.1 Modelação de Língua Gestual

Foi desenhado e implementado um sistema no qual a língua gestual é gerada através de simples modelos 3D e apresentada em forma de gráficos de linhas. As animações gestuais são construídas através de um conjunto estático de gestos actuando em diferentes *frames*. A transição entre os gestos estáticos é produzida através de técnicas de interpolação. A especificação de um gesto estático é aplicada ao esqueleto subjacente do modelo 3D. Através do modelo deformável, as linhas específicas principais de certa direcção ou vista são extraídas e apresentadas.

Uma importante característica da língua gestual é a expressão facial. Sem a expressão facial, os gestos poderão ser interpretados de modo erróneo ou então não na

sua totalidade. É descrito então uma nova técnica para modelar expressões faciais que é muito simples, tendo-se assim conseguido apresentar de uma forma bastante rica as linhas para representar expressões faciais aceitáveis.

2.8.2 Expressões Faciais

O método de modelação utiliza um modelo fixo em 3D de uma cabeça sobre a qual são editadas as correspondentes e relevantes expressões faciais da língua gestual, isto é, a boca, olhos, pestanas e as possíveis rugas para acentuar certas posições e movimentos. Não existe limite no número de linhas que poderão ser editadas ou associadas ao modelo da cabeça.

A Figura 12 apresenta a aplicação desta implementação posta em prática, na qual a especificação das linhas faciais está representada. Um utilizador poderá adicionar, eliminar ou mover as linhas associadas assim como pontos de controlo específicos de forma a obter várias aparências e expressões da cara.

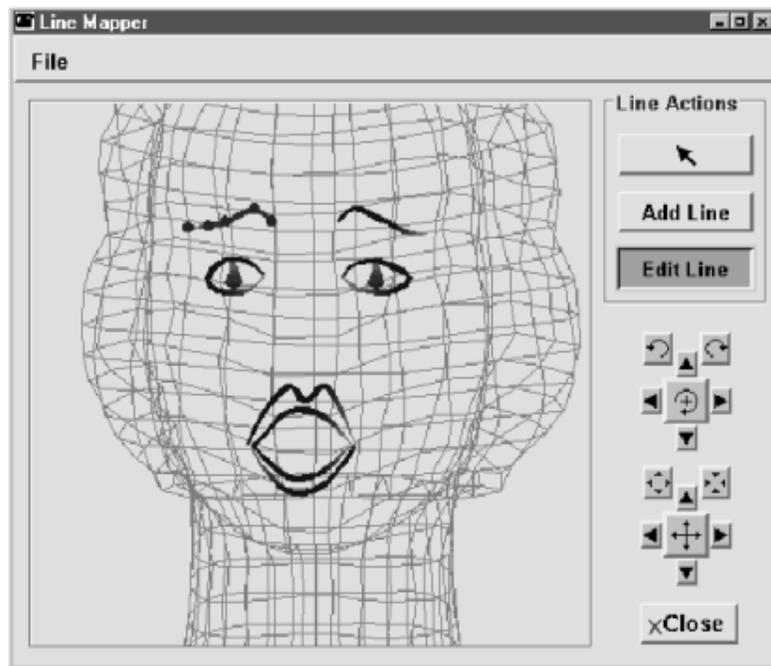


Figura 12 – Editor de linhas com várias possibilidades para conseguir uma expressão facial específica.

Uma definição para a expressão facial é associada a uma descrição verbal e guardada numa biblioteca de expressões. Quando se compõe um gesto, é possível fazer referência a esta expressão facial a qualquer momento. A transição entre as diferentes expressões é interpolada suavemente. Aqui, a especificação Bézier é usada para o esquema de interpolação.

2.8.3 Definição da Postura

Na língua gestual, a postura é por norma construída com certos gestos da mão, dos movimentos dos braços e cabeça. Há imensos gestos para a mão esquerda e direita que podem ser usados numa representação de língua gestual.

A postura nesta aplicação é definida pelos gestos da mão esquerda e direita e pelo movimento superior do corpo. A especificação do movimento das mãos e corpo é-nos dada por ângulos de rotação de um esqueleto que se aproxima ao esqueleto humano (Figura 13). Cada especificação das mão e do corpo é associado a uma descrição verbal e guardada numa biblioteca.

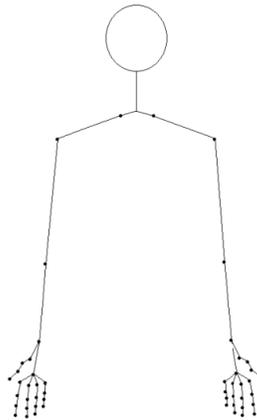


Figura 13 – Modelo do esqueleto da parte superior do corpo com as suas articulações específicas.

2.8.4 Sintetização de Gestos

Um gesto nesta aplicação é composto pelos movimentos das mão (ambas as mãos), o movimento do corpo e pelas expressões faciais. A Figura 14 mostra uma janela com vários gestos armazenados, onde a composição dos gestos se poderá realizar facilmente.

Gesture	right Hand	left Hand	Body	facial E.	pre. time	fol. time
Sign_a	sign_a	neutral	rArmUp	neutral	0.6	0.7
Sign_a	sign_a	neutral	rArmUp	neutral	0.6	0.7
Sign_b	sign_b	neutral	rArmUp	neutral	0.4	0.8
Sign_c	sign_c	neutral	rArmUp	neutral	0.7	0.7
Sign_d	sign_d	neutral	rArmUp	neutral	0.6	0.9
Sign_e	sign_e	neutral	rArmUp	neutral	0.5	0.7
Sign_f	sign_f	neutral	rArmUp	neutral	0.8	0.8

Figura 14 – Janela de diálogo de composição de gestos e biblioteca de armazenamento de gestos

Dois valores de tempo deverão ser adicionados para a definição dos gestos, os períodos antes (*pre. time*) e depois (*fal. time*) influenciam a interpolação durante a transição entre dois gestos. O período precedente define o tempo de espera do início do movimento mas já com a postura inicial, o período “depois” define durante quanto tempo este gesto continuará imóvel após a conclusão do gesto. Na janela de diálogo, cada gesto poderá ser modificado trocando a referência do nome das acções, ou os valores de tempo. Após a modificação dos nomes de referência sobre as acções a janela de *rendering* é imediatamente alterada e reestruturada, e o novo gesto estará apto a ser examinado.

2.8.5 Animação de Gestos

A composição dos gestos é uma combinação gesticular de realização estática e dinâmica acompanhada de expressões faciais. Os movimentos são realizados suavemente com velocidade variável e tempos controlados.

Neste sistema, a geração de gestos animados consiste em definir um conjunto de gestos usados nos diferentes *frames* para a composição da animação. Para a definição dos movimentos entre cada um destes *frames* assim como a transição entre as diversas expressões faciais, foram utilizados esquemas lineares de interpolação. Um utilizador poderá escolher a velocidade de saída e entrada entre cada ponto do gesto definindo a inclinação das tangentes de entrada e saída. Para uma animação de gestos aceitável foram calculados e idealizados 15 *frames* por segundo.

Por vezes nalguns gestos mais simples unicamente são necessários dois *frames* (inicial e final) para a identificação correcta do gesto. No entanto, na definição de gestos mais complicados o utilizador é forçado a incluir vários *frames* de gestos entre os 15 *frames* para conseguir obter uma animação e apresentação conveniente do gesto em questão.

É necessário sublinhar que todos os passos de progressão descritos nesta secção funcionam num modelo analítico ou descritivo de gestos e sinais.

2.8.6 Visualização

São agora descritos os modelos para visualização e animação. Foi utilizada uma ferramenta de *rendering* [24] para produzir o desenho das linhas para a descrição dos modelos. Esta ferramenta permite que o utilizador varie significativamente a qualidade das linhas, em particular no que diz respeito à largura, saturação, esboço, aparência final das linhas, etc. Isto, alternadamente permite ao utilizador acentuar certos aspectos dos gestos, em particular até ao uso excessivo de ênfases para propósitos de ensino.

Nesta secção temos resultados de exemplos do trabalho até aqui apresentado. Na primeira parte são mostradas algumas expressões faciais com várias linhas e estilos de linhas. Na segunda parte, são mostrados alguns gestos que se podem utilizar como *keyframes* para a animação gestual.

2.8.7 Exemplos de Expressões Faciais

O exemplo seguinte mostra-nos dois estilos diferentes de linhas para expressões faciais. O primeiro exemplo reflecte a expressão facial de surpresa que pode ser visto na Figura 15. Para a simulação de rugas à volta da boca aberta, linhas adicionais foram associadas ao modelo principal. Junto com olhos abertos e sobrancelhas levantadas, obtém-se um olhar de assombro.

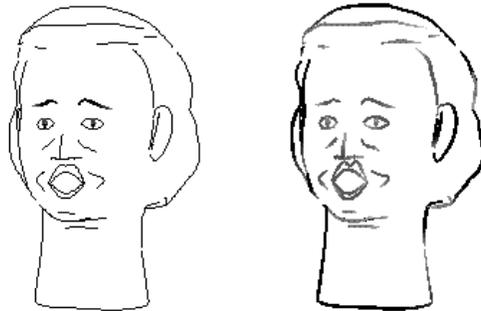


Figura 15 – Dois tipos de linhas são válidos nas expressões faciais, neste caso gerando um ar de surpresa.

Na Figura 16 é apresentada uma face sorridente. Neste exemplo é utilizado o mesmo número de linhas associadas do caso anterior. Uma expressão completamente diferente poderá ser encontrada trocando levemente alguns aspectos das linhas.

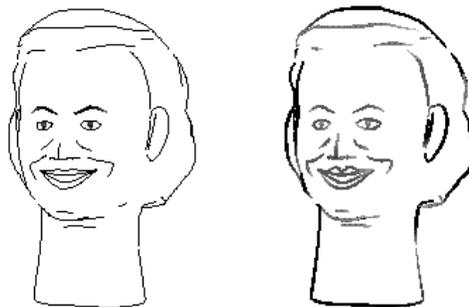


Figura 16 – Expressão facial sorridente através de gráficos de linhas

2.8.8 Exemplos de Gestos

Na Figura 17-a) é apresentada uma representação do gesto Alemão “*Niemals*” (Port. nunca). Neste exemplo os dedos da mão direita são acentuados para conduzir a atenção dos espectadores a uma característica importante do gesto, que neste caso é a posição dos dedos e mão direita.

No exemplo seguinte da Figura 17-b) é apresentado o gesto Alemão "Baum" (Port. árvore). Neste gesto o modelo está representando a copa da árvore e seguidamente o tronco. Além de gesticular, a boca também se mexe para dizer a palavra "Baum". Aqui o realce é fixado na cara traçando aí linhas mais densas que as restantes

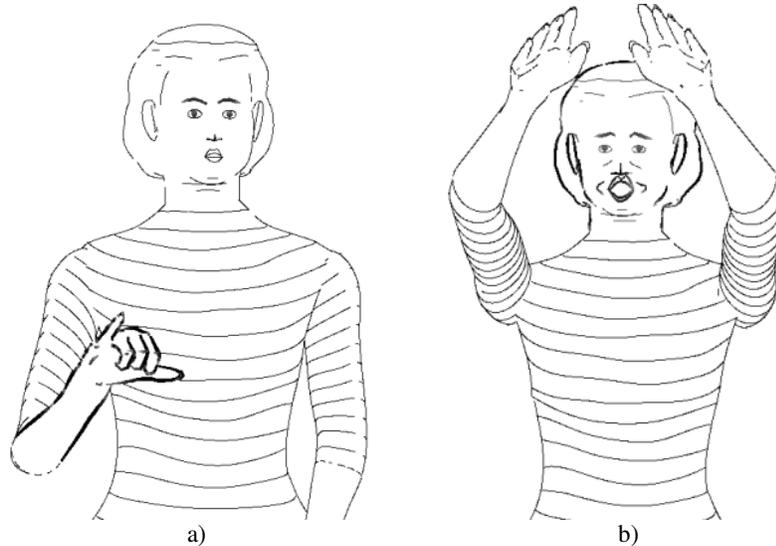


Figura 17 – Dois tipos de representação dos gestos com ênfases em diferentes partes do corpo para captar a atenção do espectador.

2.8.9 Discussão e Conclusões Retiradas

Neste trabalho foi apresentado o desenho e a posta em prática de um sistema de modelação e visualização de linguagem gestual usando gráficos de linhas. Este sistema está preparado para variar a visualização de modo a acentuar diferentes partes de gestos individuais variando o estilo de linhas utilizado.

Foram apresentadas imagens e animações a professores de língua gestual (um deles surdo) e foi recebido um *feed-back* de observações muito encorajador. Em particular a possibilidade para variar a visualização para o acentuar de partes importantes consideradas como elementos chave na aprendizagem e ensino por parte dos professores da língua gestual. O facto de serem utilizados gráficos de linhas foi bem recebido já que este estilo se presta naturalmente para representações de ênfase localizadas e é um tipo de representação facilmente impresso.

O trabalho futuro deverá prosseguir em duas direcções. Primeiro, melhorar o sistema de interacção com o utilizador e o sistema de definição e acentuação de gestos. Segundo, idealizar métodos para integrar este modelo e sistema de visualização num outro ambiente, por exemplo, saber que características de ajuda serão necessárias para o suporte via ensino.

Capítulo 3

Tecnologias Gráficas de Apoio

3.1 Computação Gráfica

3.1.1 Origens da Computação Gráfica

A computação gráfica está presente em todas as áreas, desde os mais inconsequentes jogos electrónicos até ao projecto dos mais modernos equipamentos para viagens espaciais, passando também pela publicidade, com os mais incríveis *spots* publicitários feitos por computador e pela medicina onde a criação de imagens de órgãos internos do corpo humano possibilita o diagnóstico electrónico de males que noutros tempos somente seria possível com intervenções cirúrgicas complicadas e comprometedoras.

Parece existir consenso entre os pesquisadores da história da computação gráfica de que o primeiro computador a possuir recursos gráficos de visualização de dados numéricos foi o "Whirlwind I" (furacão), desenvolvido pelo MIT⁶. Este equipamento foi desenvolvido em 1950, com finalidades académicas e também possivelmente militares pois, logo de seguida, o comando de defesa aérea dos EUA desenvolveu um sistema de monitorização e controlo de voos (SAGE - Semi-Automatic Ground Environment) que convertia as informações capturadas pelo radar em imagem num tubo de raios catódicos (na época uma invenção recente) no qual o utilizador podia apontar com uma caneta óptica.

⁶ Massachusetts Institute of Technology

Acontece que nesta época os computadores eram orientados para fazer cálculos pesados para físicos e projectistas de mísseis não sendo próprios para o desenvolvimento da computação gráfica.

Em 1962, surgiu uma das mais importantes publicações de computação gráfica de todos os tempos, a tese do Dr Ivan Sutherland ("Sketchpad - A Man-Machine Graphical Communication System"), propunha uma forma de interacção muito semelhante ao que hoje chamamos interfaces WIMP – Window-Icon-Menu-Pointer.

Esta publicação chamou a atenção das indústrias automobilísticas e aeroespaciais americanas. Os conceitos de estruturação de dados bem como o núcleo da noção de computação gráfica interactiva levaram a General Motors (GM) a desenvolver o precursor dos primeiros programas de CAD⁷. Logo de seguida diversas outras grandes corporações americanas seguiram este exemplo sendo que no final da década de 60 praticamente toda a indústria automobilística e aeroespacial utilizava software CAD.

3.1.2 O Crescimento da Computação Gráfica

Dois factores, entretanto, foram fundamentais para o desenvolvimento da computação gráfica tal como hoje a conhecemos:

a) O desenvolvimento da tecnologia de circuitos integrados durante a década de 70 que permitiu a redução de preços e a conseqüente popularização das máquinas.

b) O fim da ideia de que os fabricantes de computadores devem fornecer apenas a máquina e o sistema operativo e que os utilizadores devem escrever seus próprios programas de aplicação. A popularização dos programas de aplicação prontos e integrados (editores de texto, editores gráficos, processadores de imagem, bases de dados, etc) permitiram a popularização da computação gráfica na medida em que possibilitaram que o utilizador comum sem conhecimento ou tempo para desenvolver aplicações gráficas (nem sempre tão simples de serem programadas) pudessem utilizar as facilidades das mesmas.

3.1.3 O que é a Computação Gráfica

Segundo a ISO ("International Standards Organization") a computação gráfica pode ser definida como um conjunto de métodos e técnicas utilizados para converter dados num dispositivo gráfico, via computador.

Se tomarmos como base a definição da ISO, duas áreas têm uma estreita relação com a Computação Gráfica; são elas:

a) Processamento de Imagem: envolve técnicas de transformação de imagens. As transformações visam, em geral, melhorar características visuais da imagem como por

⁷ Computer Aided Design

exemplo aumentar o contraste, alterar a focagem ou ainda aumentar/reduzir cores e reduzir eventuais distorções.

b) Reconhecimento de Padrões: também conhecida como análise de imagens, busca isolar e identificar os componentes de uma imagem a partir da sua representação visual.

A Figura 18 ilustra o relacionamento entre a Computação Gráfica, o Processamento de Imagem, o Reconhecimento de Padrões e o Processamento de Dados convencional, segundo a visão da ISO.

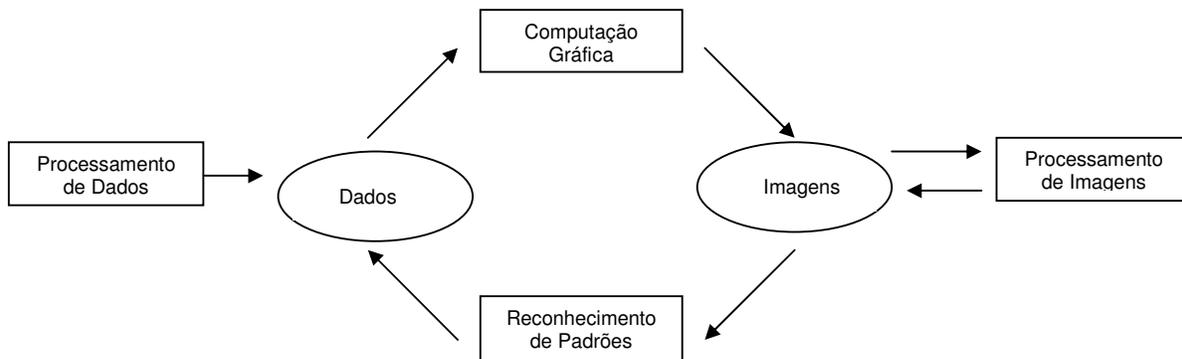


Figura 18 – Relacionamento da computação gráfica com outras áreas

Já Rogers [25] classifica a computação gráfica em passiva e interactiva. Como computação gráfica passiva entende-se o uso do computador para definir, armazenar, manipular e apresentar imagens gráficas. O computador prepara e apresenta dados armazenados sob a forma de figuras e o observador/utilizador não interfere nesse processo. Exemplos deste tipo de actividade podem ser simples como a geração automática de um gráfico de barras a partir de uma tabela, bem como a simulação do movimento de um veículo espacial a partir dos dados recolhidos em campo.

Computação gráfica interactiva também utiliza o computador para preparar e apresentar imagens. Neste caso, entretanto, o observador/utilizador pode interagir em tempo real com a imagem.

A manipulação de imagens em tempo real apresenta como principal problema o número de cálculos envolvidos para se trabalhar com imagens relativamente complexas.

Para contornar este tipo de problema podem ser adoptadas soluções tais como:

- Utilizar máquinas mais rápidas;
- Melhorar os algoritmos tornando-os mais eficientes;
- Construir módulos de "hardware" dedicados a certos tipos de operações;

- Reduzir a complexidade da imagem. Nesse caso corre-se o risco de se produzirem imagens de qualidade insuficiente conforme a aplicação.

3.2 Interacção em Ambientes Tridimensionais

O ambiente tridimensional é um cenário, dinâmico ou não, armazenado no computador e exibido, em tempo real, através de técnicas de computação gráfica. Estes ambientes podem ser divididos em duas categorias básicas: ambientes tridimensionais imersivos e ambientes tridimensionais não-imersivos.

Nos ambientes imersivos as imagens são exibidas ao utilizador de tal forma que o faça acreditar que ele está imerso neste ambiente. Normalmente, esta exibição, a fim de atingir a sensação de imersão, é realizada através de dispositivos especiais.

Nos ambientes não-imersivos, as imagens são exibidas no écran de um computador, projectadas em telas, etc.

Pinho [26] definiu que, na interacção virtual não-imersiva, toma-se o utilizador como alguém que observa e interage com um cenário tridimensional através do écran de um computador. Na interacção virtual imersiva considera-se que o utilizador está a utilizar dispositivos de realidade virtual capazes de impedir que este veja o mundo real que o cerca, passando a visualizar apenas o ambiente gerado pelo sistema gráfico.

O processo de interacção num ambiente virtual, seja ele imersivo ou não, é considerado actualmente como um sistema contínuo onde cada acção do utilizador deve ser respondida com uma acção do controlador do sistema do ambiente virtual [27].

- A ideia de interfaces entre seres humanos e sistemas deste tipo é bastante antiga. Entretanto, algumas premissas básicas aplicam-se totalmente à interacção em ambientes virtuais.

As técnicas de interacção, neste contexto, permitem mapear entradas do utilizador (posições de mãos, braços, cabeça, comandos de voz, etc) capturadas por dispositivos, em acções e comandos nos ambientes virtuais.

Apesar dos sistemas gráficos interactivos serem capazes de exhibir imagens de alta qualidade em tempo real a partir de modelos tridimensionais, ainda persiste o problema de como interagir em ambientes virtuais de uma maneira natural e livre de erros. Por isto, o projecto de uma interface homem-máquina deve possuir não só o conhecimento das capacidades motoras do ser humano, mas também, deve ter conhecimento de como se realizam as tarefas tridimensionais.

A interacção com objectos em ambientes tridimensionais é realizada com objectivo de efectuar operações de selecção, manipulação e navegação no ambiente tridimensional.

Por manipulação entende-se a tarefa de alterar algum parâmetro de um objecto tridimensional que acarrete mudança das suas propriedades, por exemplo: tamanho, posição ou orientação. Por selecção entende-se o processo de definir, de entre os objectos num ambiente tridimensional, sobre qual deles se deseja operar. A navegação, por sua vez, é o processo de mudar a posição do observador no ambiente tridimensional.

Neste capítulo analisar-se-á o processo de interacção em ambientes virtuais.

3.2.1 Parâmetros do Processo Interactivo

As operações básicas do processo interactivo em ambientes tridimensionais possuem dois parâmetros comuns, que devem ser considerados na construção das metáforas de interacção. São eles a resolução e o número de graus de liberdade dos processos interactivos.

Resolução da Interacção

A resolução trata do tamanho ou intervalo obtido entre dois quaisquer valores consecutivos na definição de um parâmetro como rotação ou posição. Numa análise prévia, pode parecer que quanto menor o intervalo, melhor será a qualidade da ferramenta de interacção. Entretanto, este é um aspecto que depende muito da aplicação e dos seus objectivos. O alinhamento de objectos entre si, por exemplo, pode tornar-se bastante difícil se a resolução disponível no processo interactivo for muito grande. Imagine-se dois objectos cujas alturas sejam controladas por dois *sliders*⁸ que o utilizador pode manipular com o rato. Neste caso, a dificuldade surge quando o utilizador tenta colocar os dois objectos numa mesma altura dentro de um cenário. Caso os intervalos entre os dois valores consecutivos sejam muito pequenos, o utilizador terá dificuldade para definir valores iguais (a não ser que utilize uma interface textual).

Outro exemplo é a traçagem de uma linha perfeitamente horizontal. Caso a ferramenta de interacção dê ao utilizador um controlo muito fino do movimento dos extremos da linha, tornar-se-á bastante difícil criar dois pontos com a mesma coordenada Y.

Em casos como estes, frequentemente, as ferramentas de interacção disponibilizam recursos como “grades de realinhamento”. Nestas, ocorre uma espécie de redução forçada da resolução do movimento, tornando a escolha de duas posições alinhadas entre si um processo mais rápido e simplificado.

Graus de Liberdade

O segundo parâmetro, comum a todas as técnicas de interacção com objectos tridimensionais, é o número de graus de liberdade da ferramenta ou da técnica de interacção. O número de graus de liberdade define o número de eixos de coordenadas que podem ser manipulados simultaneamente durante o processo interactivo. Um rato, por exemplo, tem intrinsecamente dois graus de liberdade no que se refere ao seu movimento. Uma metáfora de interacção construída utilizando-se de um rato pode, entretanto, ter apenas um grau de liberdade desde que o movimento num dos eixos seja ignorado.

A necessidade do controlo dos graus de liberdade de um processo interactivo surge, por exemplo, quando se precisa mover um objecto apenas numa direcção, como no caso

⁸ Barras de Deslocamento

de encaixar um rectângulo entre dois outros já existentes. Na Figura 19 pode observar-se um caso em que a diminuição dos graus de liberdade pode auxiliar no deslocamento.

Considere-se que o objectivo é levar com o rato o rectângulo R do ponto A, ao ponto E, sem colidir com as paredes do mapa durante o trajecto. No percurso entre os pontos A e B, o objecto R não precisa mover-se no eixo Y, pois se o fizer, poderá colidir com os blocos já posicionados. No ponto B os graus de liberdade devem ser novamente restaurados. Ao iniciar o movimento em direcção ao ponto C o movimento deve ficar restrito ao eixo Y e assim sucessivamente até que o utilizador consiga posicionar o objecto no local desejado.

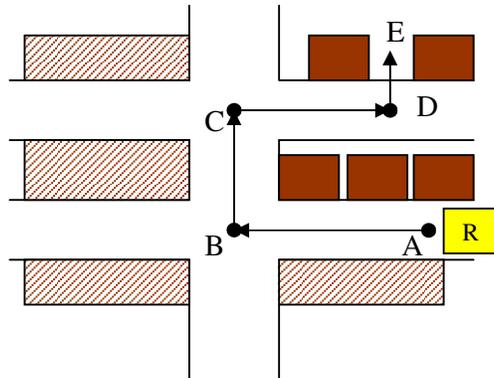


Figura 19 – Controlo dos Graus de Liberdade

3.3 Interação em Ambientes Virtuais

Os principais sistemas a utilizarem a interacção com ambientes tridimensionais não-imersivos são os editores gráficos tridimensionais, seja para modelação de objectos ou para a criação de animações. Nestes sistemas o utilizador geralmente observa os objectos tridimensionais projectados numa janela no écran de um computador.

Por meio de comandos com o rato, teclado ou outro dispositivo, pode seleccionar-se e editar as propriedades destes objectos e também alterar parâmetros como a posição ou o ângulo de visão do observador, configurado-se, nestes casos, um processo de navegação.

Noutras palavras, em ambientes tridimensionais não-imersivos a interacção com um objecto é realizada com objectivo de efectuar operações de selecção, rotação, translação e variação de escala sobre estes objectos ou então para realizar alterações na configuração do observador do ambiente virtual. Em todos estes casos, a imagem está exibida sobre uma écran de um computador, através de uma projecção bidimensional (ortográfica ou perspectiva).

As metáforas de interacção em ambientes não imersivos podem ser classificadas de acordo com a forma como o utilizador altera os valores que definem os parâmetros de rotação, escala, translação e navegação. Esta manipulação pode ser feita directa ou

indirectamente. A manipulação directa é aquela em que o utilizador opera directamente sobre a projecção do objecto no écran. A manipulação indirecta é aquela em que o utilizador opera sobre algum tipo de “entidade intermediária” que permite a geração de novos valores a serem usados na modificação dos parâmetros do objecto.

3.3.1 Interacção Indirecta sobre Objectos

A forma mais comum (e mais fácil de implementar) para a manipulação de objectos tridimensionais em ambientes não-imersivos é o uso de manipulação indirecta.

Nas metáforas desta categoria o utilizador actua sobre objectos de interface bidimensionais, conhecidos como *widgets*, e os valores obtidos destes objectos são usados para a manipulação dos objectos tridimensionais.

Na manipulação indirecta por objectos de interface, o controlo do objecto é feito através da actuação do utilizador sobre botões, *sliders* ou outros elementos, que o utilizador manipula no écran do computador.

A grande vantagem desta abordagem é o facto dos utilizadores estarem bastante habituados com o uso destes controlos bidimensionais, o que facilita em muito o processo de aprendizagem do uso da interface.

Nesta categoria há duas abordagens: a utilização de controlos comuns de interfaces 2D e a criação de controlos específicos para a manipulação tridimensional.

Manipulação com Controlos Comuns

Cabe aqui especificar o que se considera como sendo um “controlo comum”. Um controlo comum de interface é aquele disponível na maioria das interfaces gráficas, como é o caso de *sliders*, botões, menus, etc.

Para o aproveitamento, por exemplo, de um *slider* numa aplicação tridimensional, o projectista do software precisa apenas capturar o número devolvido por este objecto e utilizá-lo para alterar a posição, a rotação ou a escala de um objecto 3D. Neste caso, pode ter-se um *slider* para cada eixo e para cada atributo a ser manipulado. Isto, entretanto, pode causar um certo grau de “poluição visual” na montagem do écran de interacção, pois serão necessários muitos controlos num mesmo écran (Figura 20).

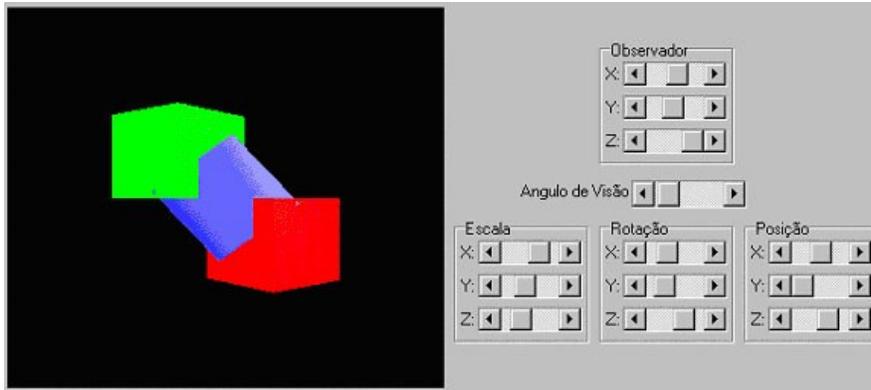


Figura 20 – Écran com excesso de controlos

Para minorar este problema é possível adoptar um esquema de selecção da qual a transformação geométrica (rotação, escala ou translação) está a ser trabalhada, reduzindo desta forma, o número de *sliders*. Esta selecção pode ser feita em menus ou seleccionadores mutuamente exclusivos como *radio-buttons* (Figura 21).



Figura 21 – Redução do número de controlos

A fim de dar um aspecto tridimensional ao processo interactivo e assim facilitar a identificação das funcionalidade de cada controlo, podem ser adoptadas estratégias que, apesar de simples, provocam um resultado bastante interessante do ponto de vista da qualidade da interacção. Uma destas estratégias é o posicionamento dos *sliders* sobre eixos de coordenadas e a representação das transformações por ícones (Figura 22).

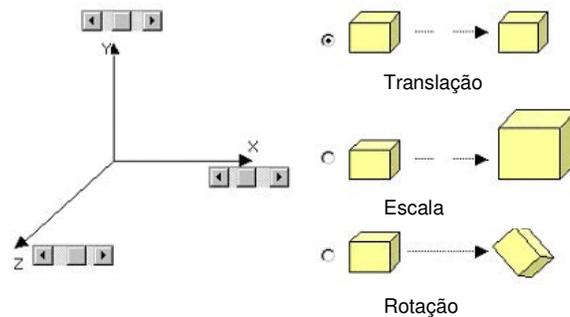


Figura 22 – Interface com ícones de transformação

Além dos *sliders* podem utilizar-se outros tipos de *widgets*, como botões ou simplesmente cursores. No caso de botões, a cada clique, o objecto pode sofrer um certo grau de modificação nos seus parâmetros de escala, rotação ou translação.

Manipulação com Controlos Específicos Para Manipulação 3D

Nesta categoria, encontram-se os *widgets* de interface criados (ou adaptados) para actuar especificamente na interacção com objectos tridimensionais.

Um exemplo interessante é o uso de *widgets* do tipo relógio, que na verdade não foram criados para uso em interfaces de aplicações tridimensionais, mas são bastante úteis para este caso, em especial para a rotação de objectos. Nesta metáfora, imagens de relógios de ponteiros são criadas e dispostas de tal forma que, pelo movimento do rato, pode alterar-se a posição dos ponteiros. A implementação deste método, é feita através da criação de uma linha imaginária entre a posição do rato e o centro do relógio (Figura 23).

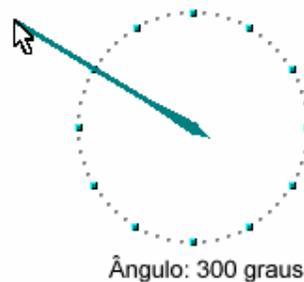


Figura 23 – *Widget* tipo Relógio

Para dar um aspecto visual mais afeito ao processo tridimensional, uma alternativa é colocar as imagens dos relógios sobre as três faces visíveis de um cubo. Com isto, em cada face, o utilizador pode manipular o objecto ao redor de um dos eixos coordenados.

Para efectuar a translação de um objecto pode-se imaginar um conjunto de ícones que se desloquem ao longo de eixos coordenados. Estes ícones, a partir do movimento do rato, deslocam-se presos aos eixos, actuando de forma muito semelhante a um *slider*.

Para a escala, pode utilizar-se a ideia de “aumentar e encolher” um objecto “pegando nele” através de algum manipulador. Uma ideia interessante neste caso é a inclusão de vectores nas faces visíveis de um cubo. Nestes vectores o utilizador pode clicar e arrastar de forma a “encolher ou aumentar” um objecto (Figura 24).

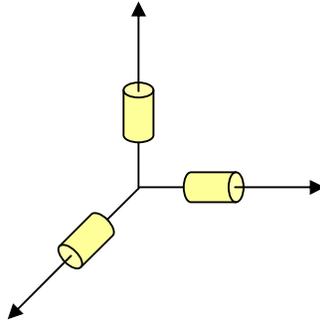


Figura 24 – Widget 3D a Escala

3.3.2 Interação Directa Sobre Objectos

Considera-se manipulação directa como sendo o processo de alteração dos parâmetros de rotação, escala e translação de objectos tridimensionais através do apontamento sobre a imagem da sua projecção. Geralmente as aplicações disponibilizam um conjunto de vistas ortográficas e uma vista em perspectiva.

Os casos de interacção sobre vistas ortográficas são semelhante ao processo de interacção bidimensional, comum a editores de desenho. Nestes, o utilizador clica sobre o objecto desejado e com o movimento do rato pode posicionar, rotacionar ou alterar o tamanho de um objecto.

Na Figura 25 pode observar-se a organização física de três vistas ortográficas sobre as quais um utilizador pode interagir a fim de modificar parâmetros de um objecto.

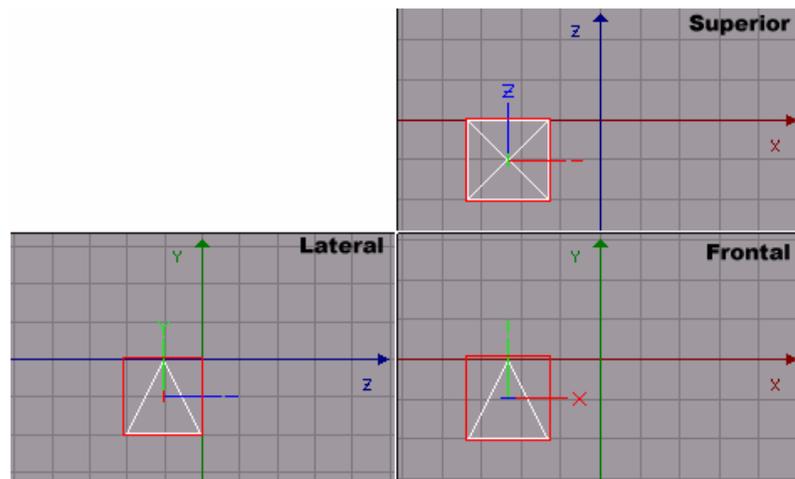


Figura 25 – Disposição de vistas ortográficas

Um aspecto importante na montagem de uma écran com vistas ortográficas é a disposição de cada uma destas vistas sobre o mesmo. É aconselhável que o posicionamento seja feito de forma que os eixos coordenados iguais, exibidos em vistas diferentes, sejam alinhados de tal forma que o utilizador possa facilmente perceber a relação entre eles.

Na Figura 26 pode observar-se um exemplo de disposição que não propicia este relacionamento.

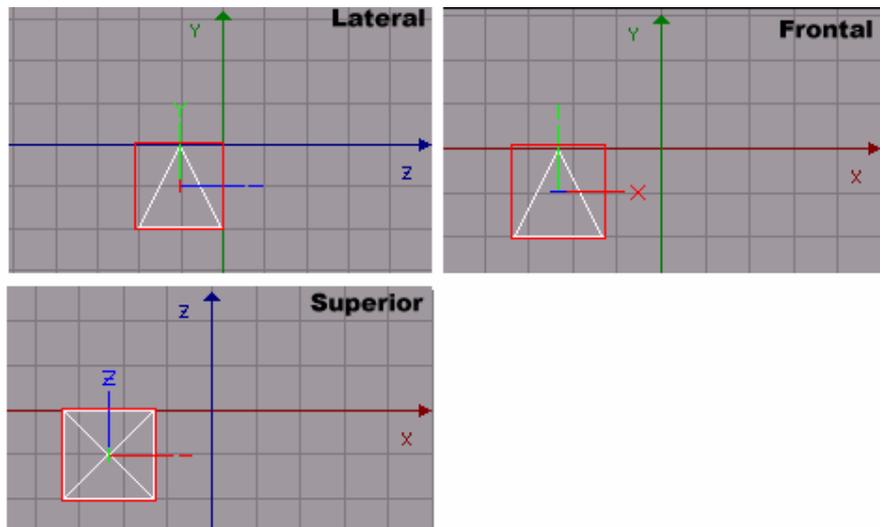


Figura 26 – Disposição incorrecta de eixos coordenados em vistas ortográficas

Para realizar uma translação o utilizador pode clicar sobre um objecto, mover o rato e posicioná-lo onde desejar. Cabe ressaltar que, em cada vista ortográfica, o objecto só pode ser movido em duas direcções, de acordo com a vista escolhida. O diferencial entre a manipulação de vistas e o uso de ambientes gráficos 2D é que, ao efectuar a movimentação ao longo de um dos eixos coordenados, numa das vistas, a outra vista que possui este eixo também deve ser alterada.

Para a realização de uma escala, o utilizador clica sobre um objecto e ao mover o rato são calculadas as distâncias por ele percorridas, ao longo dos eixos coordenados, entre o ponto inicial (antes do início do movimento) e o ponto final. Estas distâncias podem então ser usadas como factores de escala para alterar o objecto.

Na verdade, o que a maioria dos editores faz, é utilizar como factor de escala a razão entre a distância calculada e o tamanho do objecto nesta mesma direcção. Com isto, se o rato se mover ao longo do eixo X em 5 unidades, e o objecto tiver 10 unidades neste mesmo eixo, então o tamanho final do objecto será de 15 unidades, ou seja, será aplicado um factor de escala de mais 50%.

Na rotação, por sua vez, pode tomar-se a posição do rato, no momento da selecção do objecto, e a posição do rato ao longo de seu movimento. A partir disto, pode-se, com estes dois pontos criar um vector e a seguir calcular o ângulo deste vector com o eixo horizontal.

Interacção com Manipuladores Tridimensionais

Estes manipuladores são objectos que, colocados sobre o centro de referência do objecto a ser manipulado, podem permitir um controlo mais fácil, preciso e intuitivo dos parâmetros do objecto em questão.

Disponíveis em bibliotecas como Open Inventor [28] e GLEEM (*OpenGL Extremely Easy-to-use Manipulators*) [29], procuram simular “alças” através das quais o utilizador pode girar, mover ou aumentar um objecto 3D.

Para mover um objecto, o GLEEM por exemplo, possui um controlo, formado por setas tridimensionais, que permite ao utilizador mover o objecto na direcção que desejar, clicando sobre um dos eixos e deslocando o rato na direcção escolhida. Este controlo é apresentado na Figura 27.



Figura 27 – Controlo de Movimentação de objectos do GLEEM

O uso destes manipuladores afecta o objecto usando sempre seu próprio sistema de referência ao invés de usar o sistema de referência do universo. Ou seja, quando um deslocamento em X for efectuado este acarretará um movimento do objecto ao longo de seu eixo X actual e não ao longo do eixo X do universo. Na Figura 28 pode observar-se as diferenças entre deslocamentos nestes dois sistemas de referência.

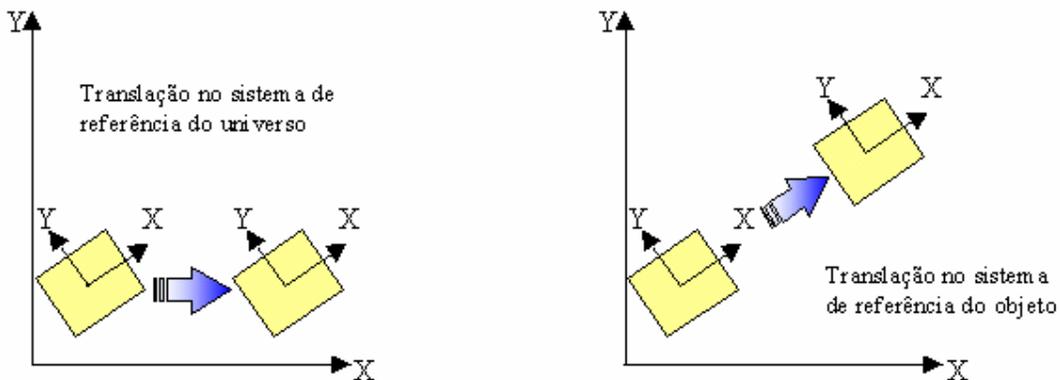


Figura 28 – Movimento em diferentes sistemas de referência

Para que esta operação seja intuitiva é preciso, portanto, que os manipuladores sejam exibidos e alinhados com os eixos do sistema de referência do objecto. Na Figura 29 pode observar-se um exemplo disto no software Cinema4D.

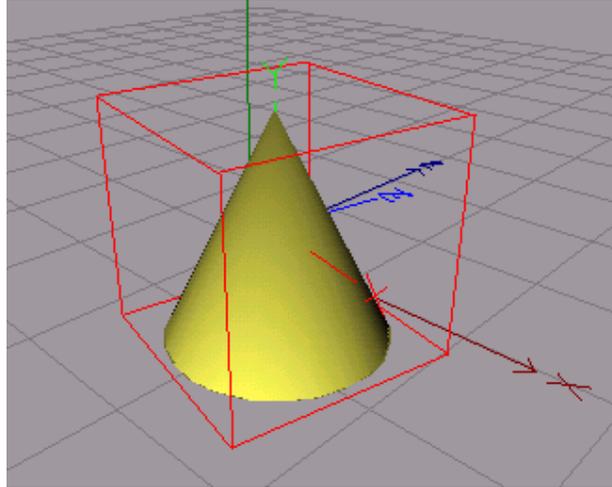


Figura 29 – Eixos do sistema de referência do objecto no Cinema 4D

3.3.3 Cinemática e Cinemática Inversa

Numa animação por computador, uma figura articulada é frequentemente modelada por um conjunto de segmentos rígidos⁹ conectados por junções¹⁰. Abstractamente, uma junção é um ponto de contracção na relação geométrica entre dois segmentos adjacentes (Figura 30).

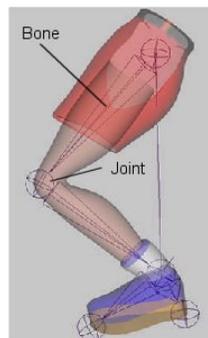


Figura 30 – Figura articulada por *Joints* e *Bones*

⁹ Do inglês *Bones*

¹⁰ Do inglês *Joints*

Esta relação é expressa por um número de parâmetros chamados de ângulos de junção. Assim sendo, os segmentos são conectados para a formação de uma estrutura em árvore, resultando uma colecção de ângulos de junção de todas as junções, correspondendo exactamente à configuração de uma figura. Esta correspondência fornece uma imediata representação pelo computador das configurações da figura articulada dados os conjuntos de ângulos de junção, esta configuração de correspondências é directa e linear para o computador, no entanto, o problema de encontrar o conjunto de ângulos de junção que correspondam a uma dada configuração (Cinemática Inversa¹¹) já não é tão trivial.

O problema que a cinemática inversa tenta solucionar é extremamente importante numa animação por computador porque é frequente numa animação espacial, com ângulos de junção, que o animador procure dar-lhes expressão de uma forma simples e ordenada. Naturalmente, o problema recebe considerável atenção na animação por computador como também noutras áreas tal como a robótica, mas os vários algoritmos existentes, por norma, reflectem aspectos muito particulares do problema e falham no fornecimento de uma solução global, eficiente e robusta para posicionamentos e figuras altamente articuladas num sistema de animação interactiva [30].

As várias aplicações da cinemática inversa para figuras articuladas são as seguintes:

- Manipulação Interactiva
- Controlo da animação
- Acessibilidade no ambiente espacial
- Evitar colisões

Assim, podemos dizer que a cinemática é o processo de cálculo da posição no espaço dos extremos de uma estrutura conectada dados os ângulos de todas as junções. Este processo é relativamente fácil e para ele há uma só solução. A cinemática inversa faz o contrário. Dados os pontos finais de uma estrutura, que ângulos estarão associados a estas junções para um posterior arquivo dos pontos finais. É de cálculo difícil, e há geralmente muitas ou até uma infinidade de soluções. Como já foi referido, este processo é particularmente útil na animação por computador e robótica. Por exemplo, se o software souber onde está um determinado modelo (corpo) em relação ao seu ombro, simplesmente necessita de calcular o ângulo das junções para conseguir o movimento pretendido.

De alguns estudos e projectos de investigação resultaram estruturas biomecânicas do tronco superior simuladas por uma árvore de cinemática [31,32]. Uma árvore apropriada de cinemática consiste numa série de nodos, cada um deles com um elemento superior na hierarquia (pai), e um número arbitrário de filhos. O nodo, representa a junção dos corpos, contém a transformação desde o eixo local do nodo pai até o próprio nodo. A componente de translação mantém-se inalterada na representação de segmentos rígidos do corpo conectados por cada junção. A componente rotação, deverá ser

¹¹ Do inglês *Inverse Kinematics*

modificável para alterar a orientação destes segmentos. O eixo local de cada nodo poderá ser calculado via cinemática tal como se apresenta na Figura 31:

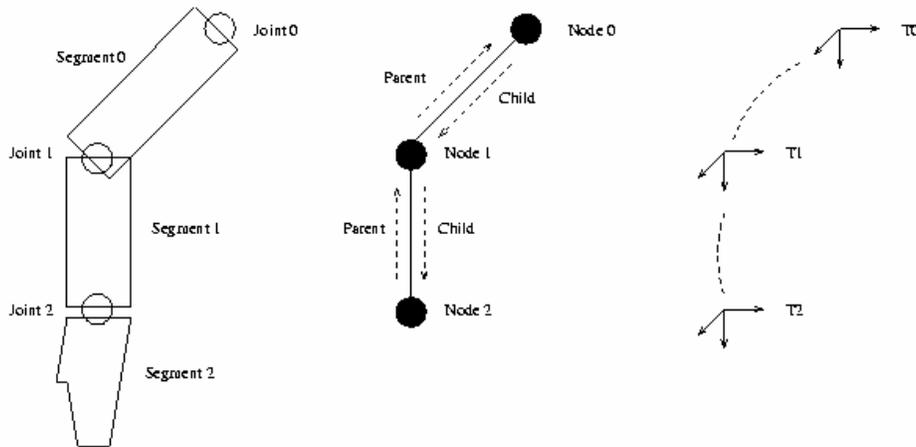


Figura 31 – Cinemática do modelo de um braço humano.

Cada nodo contém o seu quadro de coordenadas em relação ao quadro de coordenadas do seu pai. Consequentemente o eixo local de qualquer nodo poderá ser calculado usando o quadro de coordenadas dos seus antecessores (Neste exemplo, $T_2=T_1.T_0$).

De modo a promover uma aparência realista do tronco superior humano, cada nodo é considerado um modelo poligonal associado representando a superfície dos segmentos do corpo directamente seguido das junções. Assim um nodo, poder ser renderizado¹² transformando o modelo poligonal associado até ao local das coordenadas de um nodo específico.

A árvore inteira poderá ser completada atravessando toda a estrutura a partir do nodo raiz conforme se apresenta na Figura 32.

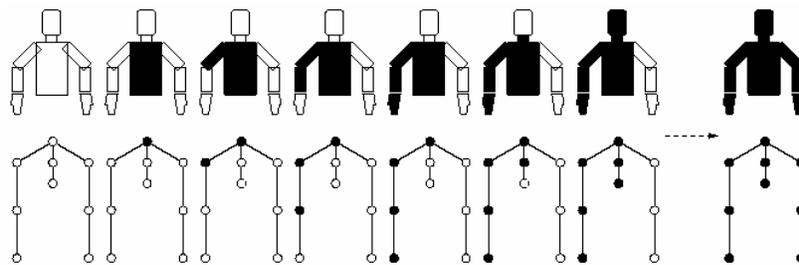


Figura 32 – *Rendering* do modelo do tronco superior humano

¹² Do Inglês *Rendered*

Uma outra tarefa que complementa a cinemática inversa é a imposição de constrangimentos de movimentos em estruturas (esqueletos) que visem representar movimentos semelhantes aos realizados por humanos, principalmente ao nível da rotação nos ângulos de junção das articulações. Deverão assim, ser especificados limites angulares sobre cada um dos eixos de cada ponto de junção, de forma a impedir movimentos que humanamente seriam impossíveis.

Por exemplo, o antebraço, gira aproximadamente de 30 a 180 graus sobre o ponto de junção situado no cotovelo. Este limites angulares são assim especialmente importantes, quando usados em conjunto com a cinemática inversa. Significa assim, que posições não naturais poderão ser facilmente evitadas aquando do reposicionamento dos pontos finais após o calculo dos ângulos associados aos pontos de junção.

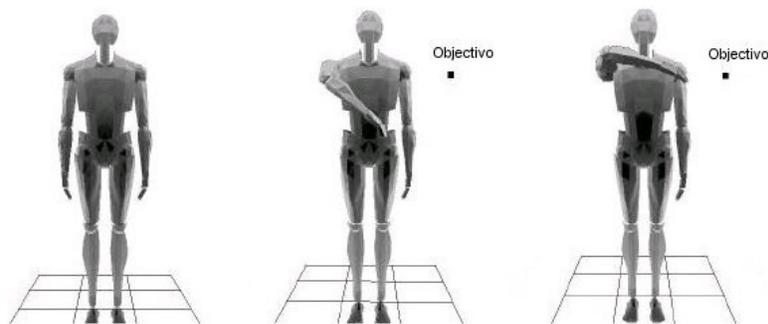


Figura 33 – Modelo movimentando a mão para junto do objectivo, utilizando a cinemática inversa, com imposição dos constrangimentos dos ângulos de junção do ombro e cotovelo.

3.3.4 Navegação

A navegação num ambiente tridimensional, caracteriza-se, em última instância, pelo deslocamento de um observador virtual, ou câmara, dentro do modelo geométrico do cenário que se está a visualizar.

Um observador de um ambiente gráfico também denominado por câmara, é composto de uma posição, de um ponto de interesse e de um ângulo de visão (Figura 34), podendo ainda recorrer a um vector vertical sobre a posição da câmara.

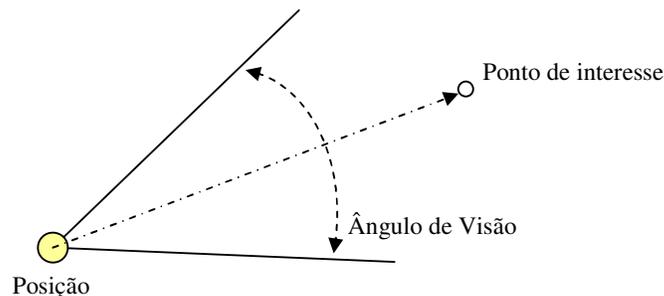


Figura 34 – Componentes de um “observador virtual”

A posição define o local de visão do observador no ambiente, o ponto de interesse define o local para onde o observador está a olhar, e o ângulo de visão a porção do ambiente que pode ser visualizada. As metáforas apresentadas a seguir são, portanto, técnicas para a alteração destes parâmetros.

Classificação das Metáforas de Navegação

As metáforas de navegação em ambientes não imersivos podem ser classificadas segundo dois critérios: semântica da câmara no processo interactivo e forma de controlo da navegação.

No primeiro, a classificação é feita diferenciando-se, ou não, a câmara em relação aos demais objectos do cenário. Nesta classificação a câmara pode ser entendida de duas formas:

- Como apenas mais um objecto e, portanto, usar as mesmas metáforas de interacção já descritas para objectos;
- Ou pode ser considerada como um objecto peculiar, com parâmetros e funções próprias e, portanto, necessitar controlos e metáforas específicos.

O segundo critério de classificação (forma de controlo da navegação), semelhante aquele usado para a manipulação de objectos, divide as metáforas de navegação em metáforas de controlo directo e metáforas de controlo indirecto. Nesta última, usam-se controlos de interfaces para manipular a câmara virtual, na primeira, efectuam-se as alterações directamente sobre a imagem da câmara.

Características Típicas do Movimento de uma Câmara Virtual

Por ser um objecto de características específicas, as operações sobre uma câmara virtual requerem algumas preocupações especiais.

Ao mover a posição do observador virtual, por exemplo, é importante atentar para o que fazer com o ponto de interesse. Uma alternativa é mantê-lo fixo, independente do movimento da posição. Por outro lado, pode-se movê-lo acompanhando os movimentos da posição. Na Figura 35 pode observar-se o efeito destas duas alternativas a partir de uma situação inicial comum. A mesma preocupação deve ser considerada quando houver movimentação do ponto de interesse.

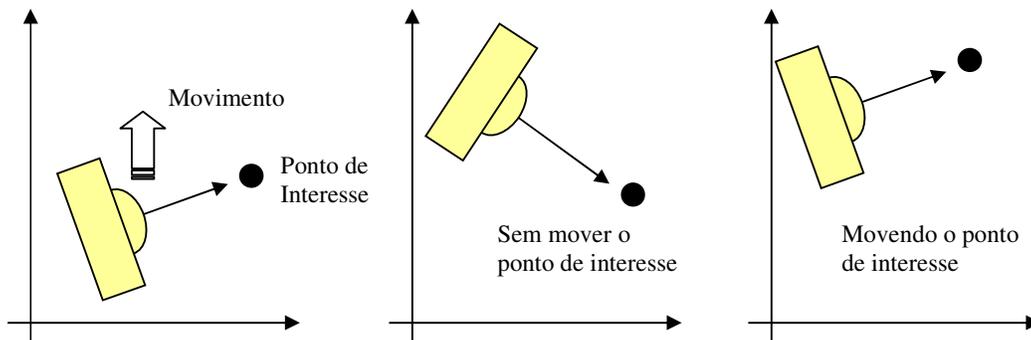


Figura 35 – Alternativas para movimento de câmara

Outra operação típica de uma câmara virtual é o movimento que simula o movimento da cabeça de uma pessoa ao redor do pescoço. Neste caso, a câmara deve ser girada ao redor dos seus três eixos, tomando sempre a configuração actual como referência para novos movimentos e não os eixos coordenados do universo.

Capítulo 4

A Aplicação

4.1 Construtor de Movimentos para Língua Gestual

4.1.1 Os Modelos

Os gestos são uma forma natural de comunicação entre as pessoas. Também podem ser usados em comunicação homem-máquina. A análise gestual através de uma componente artificial representa um tópico de interesse por forma a superar problemas de ensino de língua gestual para pessoas surdas e ouvintes interessados ou necessitados deste tipo de informação. Estes gestos poderão ter como base, imagens, vídeos ou modelos 3D capazes de imitar gestualistas na sua tarefa de ensino ou tradução. Existem assim duas diferentes aproximações para que a língua gestual possa ter “expressão” via computador: a análise do aspecto 2D [33] ou uma análise baseada num modelo 3D [34]. Os métodos 2D baseiam-se num limitado número de vistas, escondendo assim por vezes, detalhes importantes para a definição de movimentos. Por outro lado, uma aproximação 3D permite-nos uma total modelação dos gestos, com possibilidade de adicionar várias vistas e ainda alterar o tamanho do objecto sem perda de qualidade exceptuando quando aplicadas texturas. Assim sendo, temos o caminho aberto à animação por humanóides com uma base gestual de interacção homem-máquina.

A fixação e focagem na parte de interesse do corpo humano poderá resultar num ganho em termos de detalhe e flexibilidade. Experiências com um sistema de “focos de atenção” activa para promover interfaces naturais em ambientes interactivos [35] foram tópicos de interesse para esta tese, já que se assemelharam ao conceito criado para este trabalho prático de separação de modelos de partes de interesse do corpo humano actuando num mundo “ordenado” e interactivo. Assim, os modelos 3D que se acharam ser necessários para a construção de uma aplicação capaz de gerar interactivamente e em tempo real gestos e posições em língua gestual numa página Web seriam três: mão direita, mão esquerda e um humanóide de corpo inteiro.

A interacção passará por estes três modelos, por forma a conseguir todos os movimentos necessários na construção de gestos com um detalhe e uma rapidez

aceitável, dado que na construção de gestos temos vários focos de interesse (dedos, mãos, pulso, braço, cabeça etc), estes focos de interesse terão de estar suportados com alguma qualidade e sobretudo detalhe ou especificação para que possa haver uma compreensão inequívoca de todos os gestos. Detalhe este que seria mais difícil de conseguir, se enveredássemos por um só modelo (capaz de suportar todos estes focos de interesse) já que deveria ter uma imagem de grande plano no écran (zoom), ou das partes de interesse do gesto, de modo a poderem ser manipulados com alguma facilidade e clareza, além disso perderíamos em rapidez devido à conseqüente complexidade inerente a um modelo destas características.

Foi encontrado na aplicação Poser o modelo de onde derivaram os objectos mãos (direita e esquerda) com as características necessárias para o construtor de movimentos para língua gestual. Quanto ao modelo de corpo inteiro, com focos de atenção sobretudo nos braços, ante-braços e cabeça foi encontrado um modelo no site www.3Dcafe.com, leve e com definição bastante elevada nas partes de interesse (tronco superior e cabeça).

Poser

O Poser é uma ferramenta de uso fácil no desenho e animação de figuras. Com o Poser foram também produzidas ilustrações e animações 3D de figuras humanas em acção nas mais variadas situações e posições. Estes modelos podem ser seleccionados com um nível de detalhe variado, cores e texturas. Todos os modelos no Poser são possuidores de cinemática inversa permitindo a sua articulação tal como acontece no mundo real, como por exemplo, mover uma mão e ser seguida pelo respectivo braço. Usando os diversos controlos de animação, poder-se-á refinar e formular a animação de uma figura. O Poser adiciona também o movimento do caminhar humano de uma forma minimamente realista. Com o Poser poderemos fazer o *rendering* de figuras a qualquer resolução e completá-las com uma máscara *anti-aliasing* de forma a criar uma definição bastante razoável. Foi desenhado para ser compatível com outras aplicações gráficas 3D tal como Ray Dream Studio, Painter 3D, Bryce 3D, entre outras, exporta a geometria 3D em numerosos formatos de ficheiros, incluindo OBJ, DXF, 3DS e 3DMF.

Têm como base de trabalho figuras 3D (modelos), humanas ou de animais, de corpo inteiro ou de partes do corpo como mãos, cabeça e cabelo, esqueletos e manequins. Dentro das figuras humanas podemos encontrar homens, mulheres, crianças, nus, com roupa normal ou de trabalho, em baixa ou alta resolução. Como o próprio nome da aplicação indica as posições para futuras animações são essenciais, das mais variadas posições de corpo inteiro, posições faciais para representação das mais diversas expressões da cara, até posições das mãos e dedos para língua gestual Americana (ASL) se encontram nos arquivos de gestos. Na construção de movimentos, tudo o que se têm de fazer é definir uma posição, mover-nos para um ponto temporal diferente, e definir outra posição, a partir daí, o Poser criará o movimento através de interpolação de valores, preenchendo os espaços intermédios entre as diferentes posições de uma forma suave e consistente.



Figura 36 – Lista de opções Poser

Outros objectos 3D para possíveis interacções como escadas, esferas, caixas, cilindros, etc, também fazem parte do conteúdo do Poser. Os diferentes tipos de luzes, câmaras e materiais estão também presentes para a construção de cenas animadas com grau de realismo superior.



Figura 37 – Vista Geral do ambiente de trabalho – Interface – do Poser

Construção dos Modelos

Os modelos de mãos existentes noutras aplicações do mesmo género são do tipo da utilizada por Ouahdi e Horain em 1999 [36]. Como se pode ver na Figura 38, este tipo de objecto (mão 3D) é basicamente construído à base da junção de cilindros com esferas para a construção dos dedos e de caixas associados aos cilindros para idealizar palma e punho.

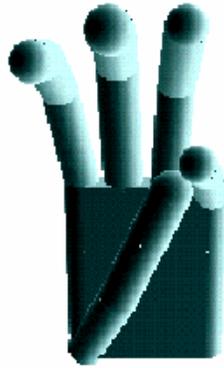


Figura 38 – Modelo da mão utilizada em experiências por Ouhaddi e Horain

A forma de armazenar e especificar os modelos em questão era normalmente projectada sobre duas camadas. A primeira camada descreve o modelo articulado como um gráfico de objectos geométricos com atributos por cores, translações, rotações e grupos dessas primitivas. Este modelo implementa 26 graus de liberdade. Seis são globais (translações e rotações) em torno do pulso. Os 4 graus de liberdade de cada dedo são: abdução/adução e flexão/extensão para a articulação entre o metacarpo e falanges, flexão/extensão para as articulações entre as falanges e falanginhas e falanginhas e falangetas. Os constrangimentos biomecânicos levam normalmente a uma redução de número de graus de liberdade, neste caso para 21.

A segunda camada é a descrição da forma ou figura dos componentes articulados. Os parâmetros do modelo necessitam de ser ajustados como forma de assegurar uma robustez desejável e os caminhos a percorrer pelas articulações. A mão real é então fotografada, extraída a sua silhueta e identificada nas diferentes partes da mão, os parâmetros morfológicos dos dedos e da palma da mão são depois estimados a partir das dimensões das áreas correspondentes.



Figura 39 – Comparação de parâmetros entre a fotografia e o modelo

Para a componente prática desta tese, apresentada no CDROM, e no caso específico do construtor de movimentos de língua gestual, havia intenção de ir um pouco mais além na qualidade dos modelos apresentados. Deveriam parecer-se mais com a realidade, principalmente as mãos teriam de ser parecidas às humanas tanto na forma como nas articulações, sabendo desde o início que condições como o número de polígonos, definição de texturas e cores deveriam ser alvo de atenção visto que os modelos resultantes não deveriam ser pesados para evitar desconforto na altura da sua interacção e consequente manipulação.

Foram assim, retirados os modelos para as mãos (direita e esquerda) a partir da “amputação” das mesmas num modelo do Poser de corpo inteiro (Figura 40) com um

número de polígonos não excessivo, mas com uma qualidade inequívoca e natural parecido às extensões humanas. O processo de “retirar” as mãos ao modelo de corpo inteiro, não foi mais do que ir apagando (*delete*) as partes deste que não teriam interesse, já que era composto de várias “peças”, sendo duas delas as mãos resultantes.

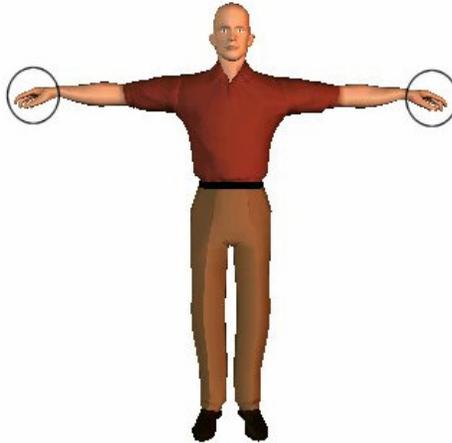


Figura 40 – Modelo completo de onde foram retiradas as duas mãos

Desta separação resultaram dois objectos (Figura 41) que representariam as mãos com muita definição já que contavam com um número de polígonos razoável (1449 polígonos), haveria no entanto que proceder posteriormente à sua simplificação de forma a tornar mais leve a aplicação onde se incluíriam.

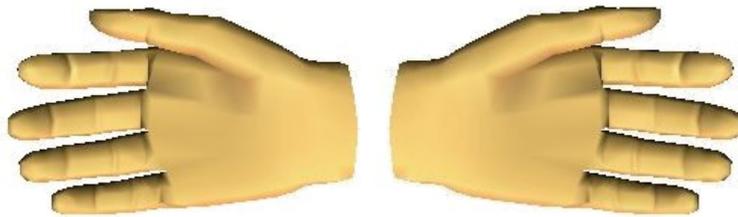


Figura 41 – Modelos das mãos resultante

O modelo de corpo inteiro, como já foi dito atrás, foi encontrado na Web. É um modelo representando um homem, simples, e que ocupa pouco espaço em memória. Para este modelo não seria necessário grande definição, já que unicamente necessitaríamos de mover os braços e cabeça por forma a expressar a posição do tronco superior do humanóide aquando da construção dos gestos. A interacção com as mãos seria realizada nos modelos encontrados acima, que no interface aparecerão aos lados do humanóide, tudo num mesmo écran. O modelo de corpo inteiro já combina uma estrutura de polígonos e um mapa ou imagem da textura aplicada sobre o anterior, de forma a reduzir assim a complexidade geométrica do modelo como se comprova na Figura 42.

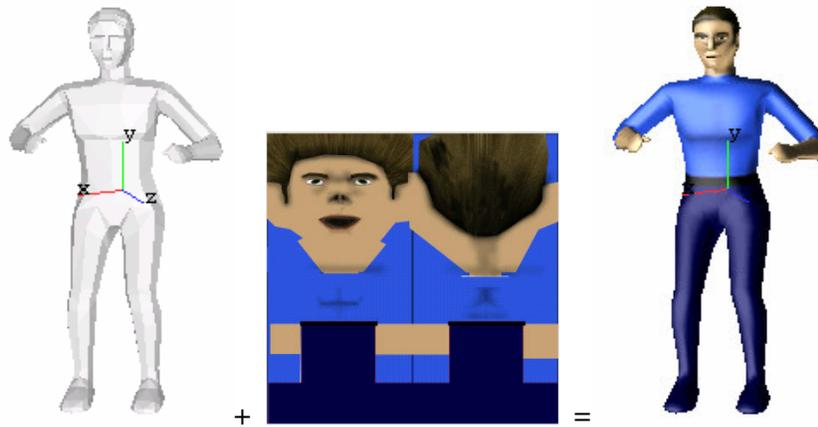


Figura 42 – O modelo 3D de humanóide completo

Simplificação dos Modelos (mãos)

Para a simplificação de modelos 3D existem diversas aplicações que basicamente trabalham com a possibilidade de redução de polígonos (Figura 43) e/ou a conseqüente conservação de texturas através da construção de um mapa (imagem 2D) para aplicar no modelo alterado de forma a manter aparentemente as mesmas características visuais.

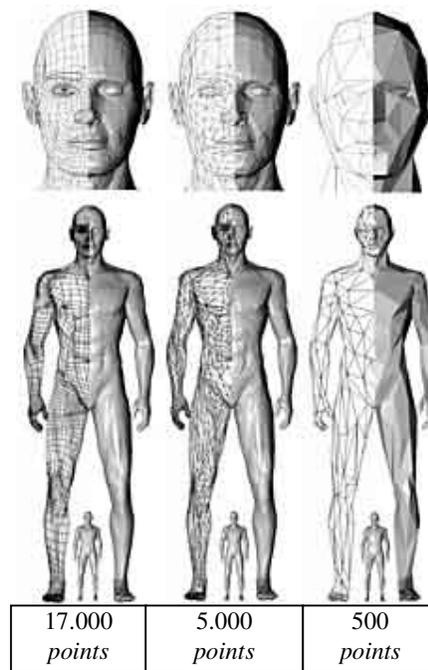


Figura 43 – Modelos com redução do número de polígonos

A futura interação natural pretendida com a antecedente simplificação dos modelos, principalmente das mãos, era tida como fundamental. Seria o objecto com manipulação mais frequente e específica, sendo que não necessitaríamos de tão elevado

número de polígonos como nos mostra a Figura 44 (1449 no modelo retirado do Poser) para animação de gestos com alguma qualidade.

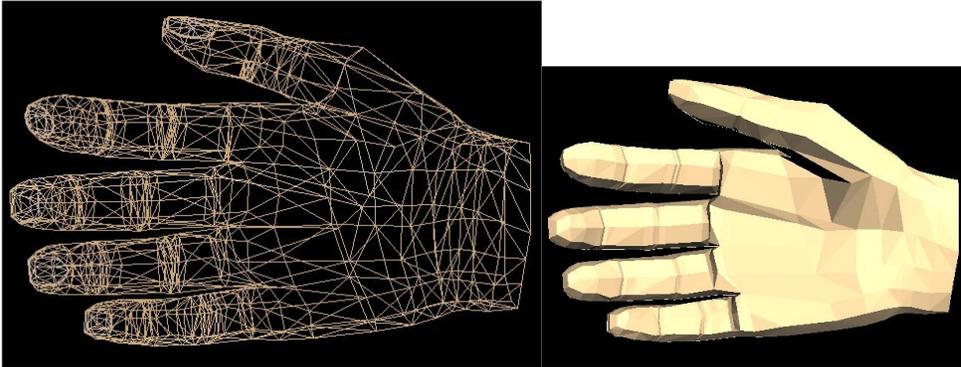


Figura 44 – Modelo de mão inicial -1449 polígonos

É certo, que haveria perda no campo de detalhe e definição de zonas de articulação, perda esta colmatada posteriormente com técnicas de alisamento e suavização de contornos (*smoothing, anti-alising e wing*). No que sem dúvida ganharíamos e muito, era em velocidade e realismo na interacção com o objecto, permitindo movimentos fluidos provocando assim um controlo natural e um consequente agrado na utilização do construtor de movimentos para língua gestual. Como a aplicação tinha o objectivo de trabalhar na Web, os ficheiros que contêm as especificações do modelo deveriam ser também de menor dimensão possível para evitar tempos de espera no seu carregamento.

Assim, foi utilizado um *plug-in* do 3DstudioMax – Polygon Cruncher – para redução de polígonos e vértices nos modelos das mãos, antes primeiro teve de se exportar e importar os ficheiros correspondentes aos modelos a trabalhar, da aplicação Poser para o 3DstudioMax. O modelo resultante desta redução e encontrado como aceitável por mim (Figura 45) era 86% mais simples ao nível de polígonos (redução de 1449 para 197 polígonos) e consequentemente ao nível de armazenamento na forma de ficheiro. Com esta redução, não houve grande alteração na aparência geral da mão, continuava a parecer-se com uma mão humana, e muito importante, ao nível dos dedos, este mantiveram-se separados, com dimensões e morfologia própria o que salvaguardava uma interacção e gestualização ao nível do formato e aparência aceitável.

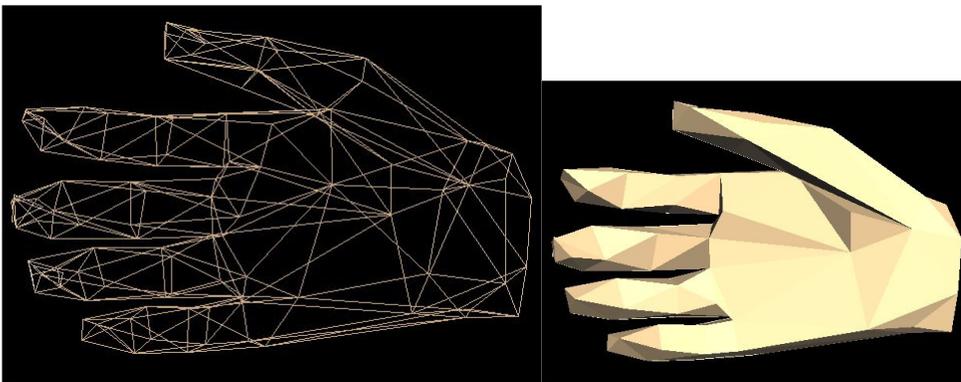


Figura 45 – Modelo de mão simplificado - 197 polígonos

A Figura 46 mostra-nos a versão final dos modelos das mãos (neste caso a direita) utilizando já técnicas de suavização de contornos (*smoothing* e *anti-aliasing*) e arredondamento de segmentos (*wing*). Estas técnicas foram implementadas já no software de animação Pulse3D que se descreve a seguir.

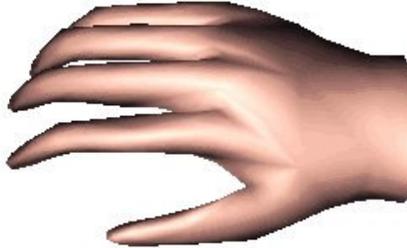


Figura 46 – Modelo de mão final otimizado no Pulse3D

4.1.2 Construtor de Movimentos

O que se pretendia criar, era então, uma aplicação 3D interactiva, leve quanto baste, capaz de ser implementada para a Web e de fácil manipulação capaz de construir gestos correspondentes a letras, números ou palavras em língua gestual. Aplicação esta, que contaria com os três modelos referidos anteriormente, com manipuladores directos sobre estes objectos de forma a criar movimentos que imitem os da vida real. Estes objectos possuiriam também capacidade de rotacionar sobre si, de forma a alterar o ângulo de visão e proporcionar uma manipulação e especificação eficiente.

Havia então que escolher a tecnologia capaz de implementar este tipo de aplicação. Sabendo das mudanças contínuas no mundo dos gráficos 3D, foram assim, feitas pesquisas a aplicações Web3D. Web3D é um termo genérico, que descreve o *display* de uma tecnologia de computação gráfica em *sites* Web. As tecnologias Web3D remontam do já antigo VRML (*Virtual Reality Modeling Language*) até às mais comuns usadas nos dias de hoje para quem pretenda ter Web interactiva 3D com alto nível de realismo e rapidez de interacção. Estamos a falar de tecnologias como Pulse3D, Anfy3D, Shockwave 3D, RealActor entre outras. A escolha recaiu na aplicação Pulse3D, que na altura se mostrou eficiente para o objectivo pretendido.

Pulse3D

A tecnologia Pulse3D combina um conjunto de técnicas como: *authoring*, *playback*, *server-side*, *file format*, e *scripting*, o que torna possível dar qualidade às vistas, conteúdo interactivo, velocidade e aspecto visual em animações, por forma a se posicionar no topo dos formatos 3D para a Web. O Pulse3D surpreende pela facilidade no uso e programação, pela capacidade de organizar um mundo 3D e pelos resultados oferecidos num ambiente Web. O Pulse3D é composto nos seguintes módulos:

- Pulse Creator – Esta é a componente de importação, exportação e manipulação de modelos, sons, texturas e comportamentos. Geração de objectos e figuras com capacidades interactivas e em tempo real de comportamentos, luzes, sons e animações em geral.
- Pulse Player – De fácil e rápido *download* gerando um *plug-in* para o *browser* com instalação própria, constitui a plataforma que permite aos utilizadores a visão e interação dos objectos e das animações Pulse pela rede Internet.
- Server-side, File format, e outras inovações incluindo o uso de :
 - Disponibilidade de streaming
 - Compressão de dados
 - Encriptação
 - Formação de pacotes
 - Gestão de direitos digitais
- Pulse Producer – Consiste num *plug-in* de exportação, que converte cenas de animação criadas do 3Dstudio Max para ficheiros Pulse com formatação própria e otimizados para a Web.
- Pulse Exporter para o Maya – Consiste num *plug-in* de exportação, que converte animações Maya em projectos do tipo Pulse Creator.
- PulseScript – Linguagem de *script* nativa do Pulse, contendo uma vasta biblioteca de *scripts* prontos a usar, estendendo a habilidade de criação de combinações de interactividade mais elaborada.

Descrição do Ambiente de Trabalho

O Pulse Creator permite-nos criar um sem número de animações ou comportamentos¹³ para cada objecto numa determinada cena. Basicamente os passos gerais na construção de uma animação interactiva idealizada no Pulse Creator contém os seguintes itens:

- Preparação e importação dos modelos para o Pulse Creator
- Se necessário, re-orientação da geometria
- Optimização da geometria
- Criação de partes do esqueleto para um determinado modelo
- Configuração e definição da hierarquia dos componentes do modelo
- Modificação da geometria do modelo
- Adição de luzes e texturas
- Criação de estados alterados morfológicamente (*morphers*)
- Adição de interactividade e som
- Colocação das câmaras na cena
- Publicação do conteúdo

¹³ Do Inglês *Behaviors*

Conforme é apresentado na Figura 47, a zona básica de trabalho têm a seguinte aparência:

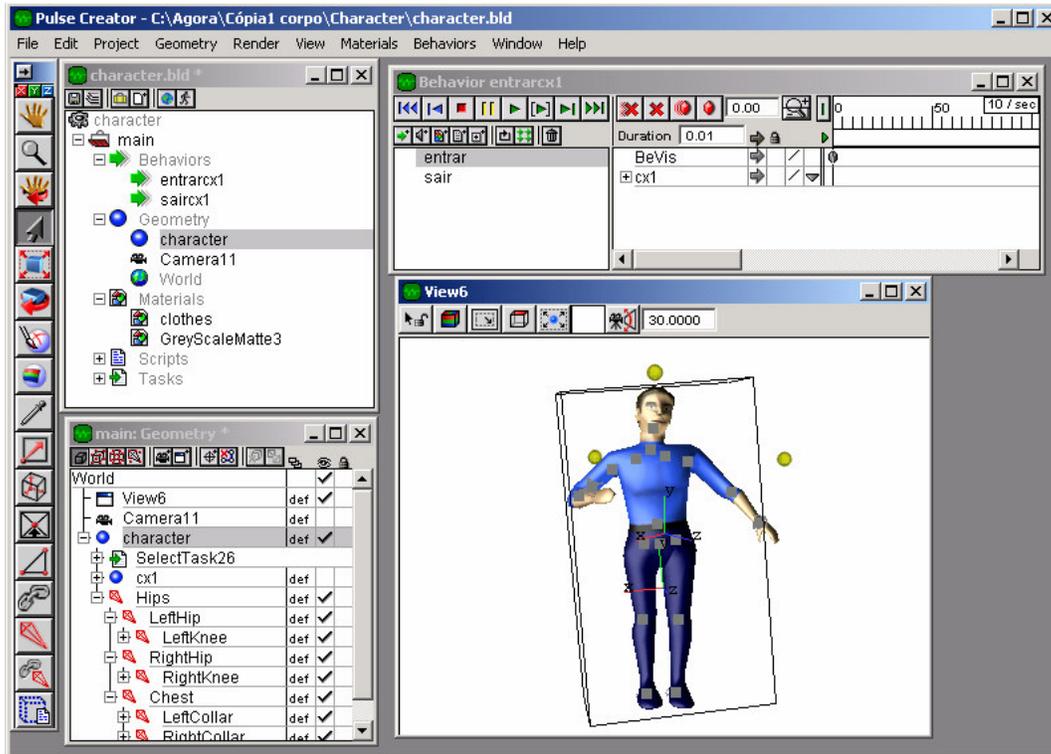


Figura 47 – Vista geral do Interface Pulse3D

Esta é composta principalmente por:

- Build Window – Adição de geometrias, *behaviors*, materiais, audio e *scripts*. Criação de *packages*. Fechar e gravar os projectos.
- Geometry window – Edição e criação da geometria de objectos, *bones*, vistas, câmaras e *morphers*, organizados de forma hierárquica. Atribuição de *scripts* aos objectos.
- View window – Edição e *display* dos modelos.
- Behavior Editor – Animação e adição de interactividade através da construção de *behaviors*.

A Biomecânica da Mão

O esqueleto de uma mão comporta os ossos do punho (carpo), do metacarpo e os ossos dos dedos. As articulações em cada mão poderão ser identificadas como indica a

Figura 48 pelas articulações situadas na base do pulso (6) entre o carpo e metacarpo, as articulações de cada dedo sobre o seu metacarpo (5) e as articulações entre falanges (1,2,3,4).

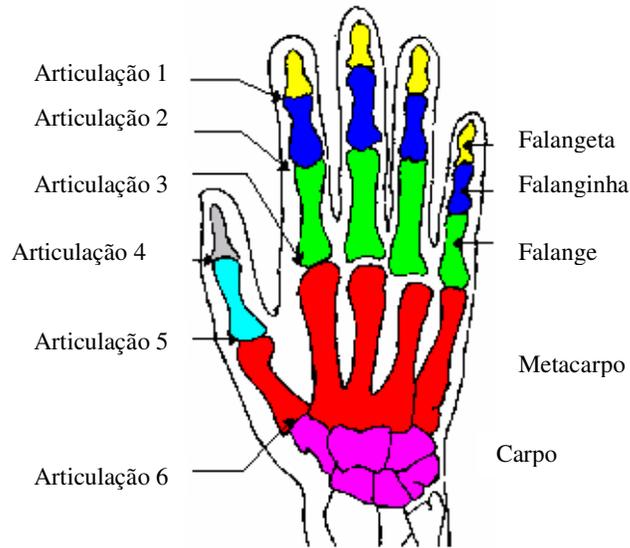


Figura 48 – Esqueleto e articulações da mão

A mão é um objecto articulado com 28 graus de liberdade. Seis deles são definidos ao nível dos punhos: 3 rotações e 3 translações. Vinte e dois são associados às diferentes articulações dos dedos. Dois graus de liberdade (abdução/adução e flexão/extensão) para as articulações metacarpo-falanges (Figura 49), um outro (flexão/extensão) ao nível de cada articulação inter-falange. A articulação carpo-metacarpo no pulso possui 3 graus de liberdade: abdução/adução, flexão/extensão e uma pseudo rotação entre os ossos do Carpo e a base do metacarpo do pulso, devido ao relaxamento ou não dos ligamentos.

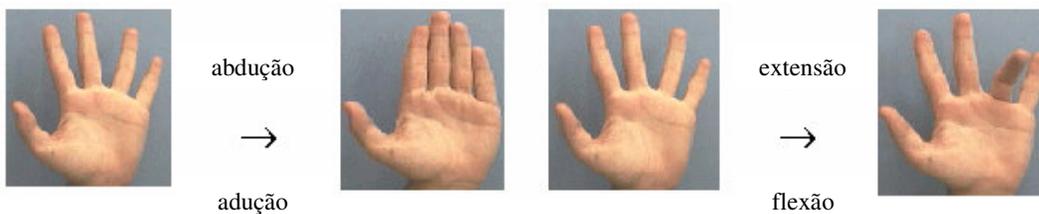


Figura 49 – Movimentos ao nível das articulações Metacarpo-Falanges

Definição de *Joints* e *Bones* na Articulações dos Modelos

A partir do modelo da mão importado do 3DstudioMax, foram criados no Pulse3D, pontos de junção para possíveis articulações (*joints*), pontos estes que têm ligação ao objecto global através de um *link* a vértices constituintes do mesmo objecto. A Figura 50 mostra-nos a divisão do polegar da mão direita em 3 *bones* e 4 pontos de junção (pol 1,

pol 2, pol 3 e pol 4), cada um destes está ligado a número de vértices específico, criando assim uma zona articulada e semi-independente em relação à zona posterior ou anterior. Como se pode ainda observar na Figura 50 o ponto de junção “pol 3” têm associados 5 vértices (vert 29, 30, 31, 35 e 36) do modelo global, conseguindo-se assim segmentos rígidos ou zonas autónomas de articulação nos objectos (*bones*) por forma a uma posterior manipulação das mesmas .

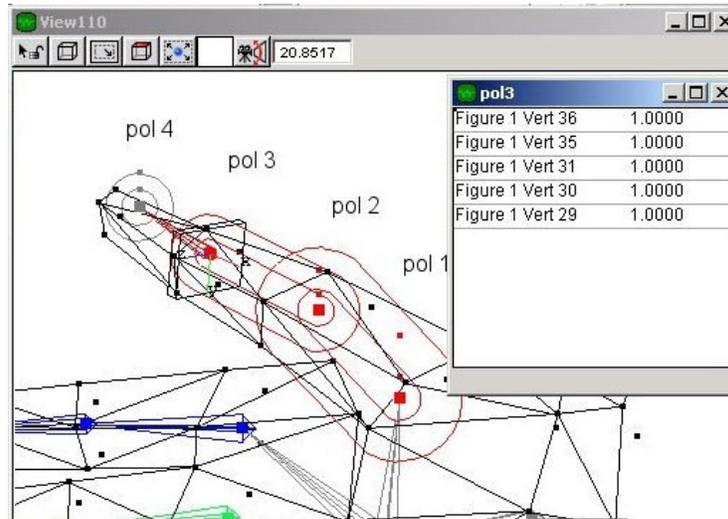


Figura 50 – *Bones e Joints* do polegar da mão direita

No caso das mãos (Figura 51) estas subdivisões deram origem a 19 pontos de transformação mas só 12 deles foram activados para a articulação das mãos, destes 12 pontos articulados 7 deles estão aptos unicamente para extensões e flexões e os 5 restantes acresce também as possibilidades de adução e abdução, o que no total representa 17 graus de liberdade para este modelo. Todos os pontos de articulação activos, são limitados por valores máximos e mínimos de rotação de forma a não gerar movimentos ou deslocações humanamente impossíveis.

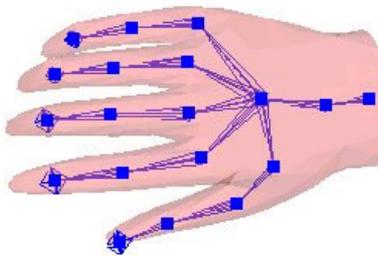


Figura 51 – Pontos de possível articulação do modelo representativo da mão direita

Estas estruturas foram armazenadas em ficheiros Pulse3D de forma hierárquica como se pode ver na Figura 52 neste caso representando a mão direita.

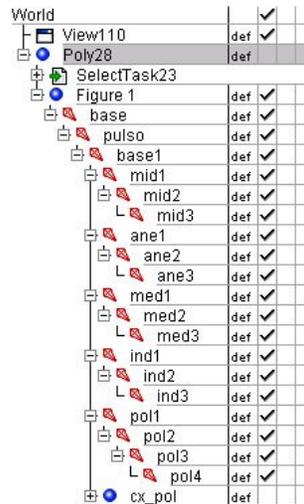


Figura 52 – Estrutura hierárquica de todos os componentes e funções do modelo mão direita

Também no modelo de corpo inteiro foram definidos pontos de articulação criando uma estrutura “óssea” para futuras rotações e translações, foram identificados assim 18 pontos, mas apenas 8 deles (do tronco superior) foram activados, devido ser unicamente necessário mover a parte superior do corpo na construção de gestos, respectivamente mãos, braços, antebraços, ombros, pescoço e cabeça.

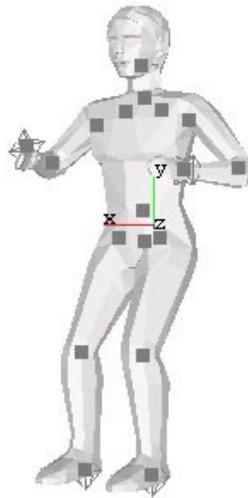


Figura 53 – Pontos de possível articulação do modelo de corpo inteiro

O Pulse3D contém especificidades de cinemática inversa o que proporciona (como já foi referido no Capítulo 3) por exemplo deslocar um dedo e a mão acompanhar este movimento partindo do cálculo dos ângulos de junção dados os pontos finais. No entanto seria ainda necessário anular movimentos humanamente impossíveis. A activação dos pontos de articulação passaria por definir para cada um destes, valores de rotação

máxima e mínima por forma a limitar os movimentos. A Figura 54 mostra-nos a restrição destes valores para dois pontos de junção com diferentes graus de liberdade. Para movimentos no ponto articulado do indicador “ind 1” foram definidos dois graus de liberdade “abdução/adução e extensão/flexão”, com valores de rotação mínimos de -10 e -22 respectivamente até valores máximos de 30 e 50 , isto nos eixos “[0]” que corresponde ao “x” para a abdução/adução, e no eixo “[1]” que corresponde ao “y” para a extensão/flexão. Para o ponto de junção “ind2” foi unicamente possibilitado um grau de liberdade “extensão/flexão”, com valores de rotação entre 0 e 90 , e utilizado unicamente o eixo “y” para futuros movimentos.

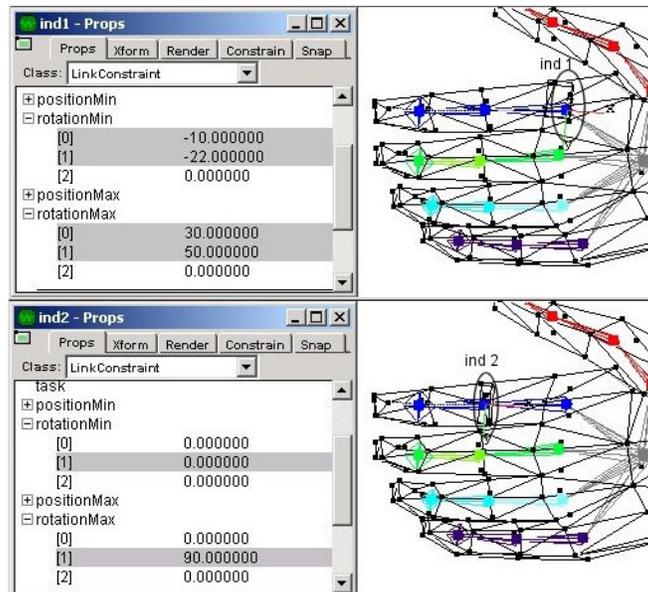


Figura 54 – Restrições de movimentos nos pontos articulados

Especificação da Interface

A animação não é a arte do “desenho do movimento” mas a arte do “movimento que é desenhado”. O que acontece entre cada *frame* é mais importante do que existe em cada *frame*. A animação é “arte gráfica que ocorre no tempo” [37].

As aplicações gráficas dinâmicas têm de ser capazes de responder a eventos assíncronos de *input* independentemente quando este acontecem, assim, os desenhadores terão que desenvolver os seus programas com pequenos conhecimentos sobre quando e em que ordem estes eventos ocorrem. Por conseguinte, a qualquer momento durante a execução de uma aplicação dinâmica o estado geral terá de ser mantido explicitamente numa estrutura de dados global.

Antes de mais, há que expor algumas considerações gerais de como qualquer tipo de interface deste género deverá servir o utilizador [38,39]:

- A construção de interfaces é um processo interactivo, de experimentações e de manipulação de múltiplos parâmetros.
- As construções de interfaces devem aumentar a produtividade, diminuindo o ciclo de prototipagem, tornando, dessa forma, mais fácil a sua manutenção e alteração, aumentando assim a coerência e a normalização dos interfaces.
- O utilizador tem de sentir que está no comando.
- O utilizador é o que tem mais consciência das suas necessidades e o software deve responder a todas essas necessidades sem criar novas dificuldades.
- O programa deve fomentar o pensamento reflexivo, deve aumentar as capacidades do utilizador, obrigando-o a pensar.

Fixando-nos agora em aplicações do tipo interactivo, o objectivo do conjunto de interfaces é associar as típicas capacidades interactivas associadas a modernos e potentes gráficos, nomeadamente um rato, o sistema de janelas, e os vários tipos standard de *widgets* interactivos, como caixas de entradas de texto, barras de deslocamentos e botões.

Outros estudos e investigações, como por exemplo a criação de uma gama de ferramentas para interfaces levada a cabo por Turner et al [40] chamadas de ferramenta de quinta dimensão tinham o objectivo de encapsular num modo orientado aos objectos, alguns dos novos dispositivos de entradas 3D tal como uma *SpaceBall* ou uma *DataGlove*.

Implementação e especificação da Interface (manipulador 3D + rotação + translacção) para o Construtor3D

Após algumas experiências com vários tipos de controlos ou manipuladores (directos e indirectos) para interagir com o utilizador de forma a permitir modelar as diferentes posições e combinações de gestos nos três modelos à disposição, sobressaiu o manipulador tridimensional (do género do GLEEM abordado no Capítulo 3) pela forma directa e natural como interage com o modelo ou articulação associada. Estes manipuladores não são mais do que esferas colocadas junto do conjunto a rotacionar nas várias direcções e associados com o centro de referência da articulação ou conjunto de articulações a serem manipulados, o que permite um controlo fácil, preciso e principalmente intuitivo dos parâmetros do objecto ou conjunto de objectos a serem manipulados. Adicionalmente, e aqui ao contrário do manipulador GLEEM, a mudança de eixos coordenados é feita através de três controladores indirectos ou seja um botão para cada um dos eixos X Y e Z.

Nos modelos das mãos, como se pode ver na Figura 55, temos 5 esferas de diferentes cores, cada uma associada a um dedo que o controla directamente em todos os seus graus de liberdade definidos, existe ainda neste modelo uma sexta esfera que controla as movimentações do pulso. No modelo de corpo inteiro temos 3 esferas associadas aos braços e cabeça de modo a poder construir posições específicas gerais do tronco superior associadas aos respectivos gestos. Cada um dos modelos tem possibilidade também de rotacionar através de interacção directa do rato sobre o modelo,

no caso das mãos e como já foi dito atrás, o tipo de rotação nos diferentes eixos poderá ser alterado através dos botões X, Y e Z. Existem ainda dois botões para cada um dos modelos (ON e OFF) que activam e desactivam o movimento e a visibilidade de todos os controladores directos 3D.

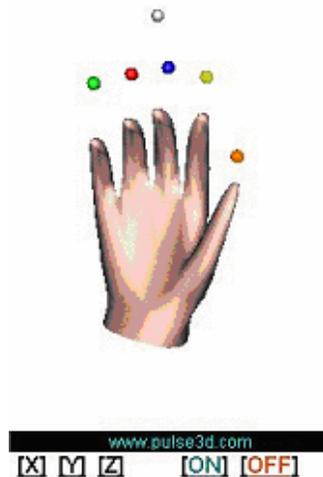


Figura 55 – Modelos finais de interação para a mão com os respectivos controladores

Assim, e resumindo, uma das partes constituintes da aplicação prende-se com um construtor de movimentos para língua gestual. Conforme referido e se comprova pela Figura 56, a interface é apresentada através de 3 modelos manobráveis pela interação do rato em manipuladores directos (esferas), posicionados no extremos da parte do objecto a movimentar.

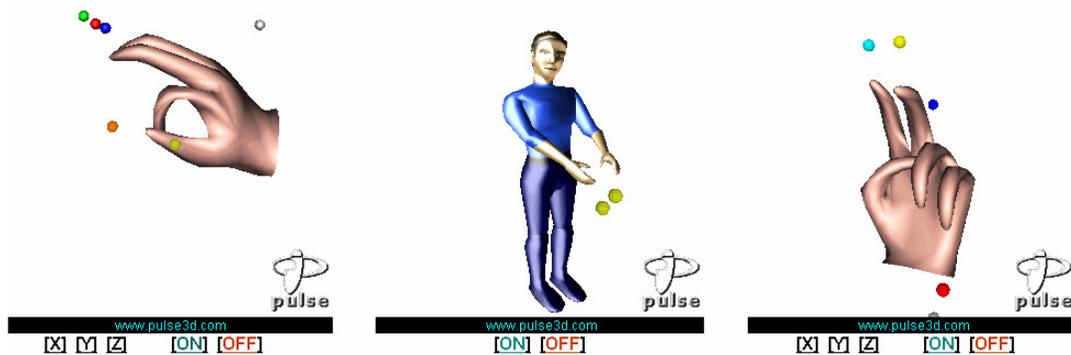


Figura 56 – Interface do Construtor 3D

4.2 Abecedário 3D

Os primeiros gestos e talvez os mais importantes a serem ensinados a quem se pretende iniciar em língua gestual são sem dúvida as letras que compõe o abecedário. Assim sendo, seria também objectivo desta aplicação a feitura e apresentação de uma base de dados dos gestos que representassem todas as letras do abecedário em língua

gestual de uma forma inequívoca e inovadora via computador e Internet. Esta apresentação teria de ir mais além de simples vídeos ou imagens 2D. As animações a três dimensões seriam opção já que permitiria além do elevado realismo da animação uma escolha diversificada de ângulos de visão que se mostraram eficazes no aperfeiçoamento e clarificação dos gestos apresentados.

4.2.1 O Modelo

Como para a definição das letras era desnecessário um modelo de corpo inteiro já que os gestos correspondentes ao abecedário se limitam à mão direita ou esquerda, e pensando logicamente no minimizar do tamanho dos ficheiros resultantes para aumentar a rapidez e naturalidade dos movimentos, decidiu-se definir um modelo do ante-braço, braço e mão para a composição de todos os gestos. O primeiro passo seria encontrar um modelo 3D nestes moldes com elevada resolução, principalmente ao nível dos dedos por razões óbvias. Após uma procura de modelos com estas características revelou-se serem raros e faltosos de qualidade, pelo que se decidiu retirar a parte de interesse a um modelo de corpo inteiro com uma definição muito aceitável e um número de polígonos não demasiadamente elevado (Figura 57), modelo este originário do Poser, com maior definição comparando com o modelo usado no Construtor3D, que adicionalmente continha uma mais valia que era a sua composição em diferentes “peças” de modo a permitir a animação na própria aplicação Poser.

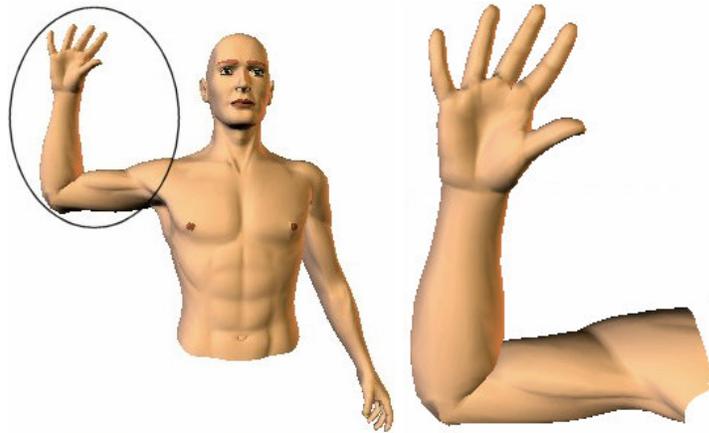


Figura 57 – Corte do modelo necessário à animação

Foram assim definidas as animações de todas as letras no Poser, cada uma com uma média de 30 *frames* resultando 26 ficheiros correspondentes a 26 letras. O passo seguinte seria converter estes ficheiros para o 3DStudioMax de modo a poderem ser exportados para o Pulse3D.

4.2.2 Exportação

Como a aplicação Poser não possuía capacidade de exportar objectos mais animação correspondente para o 3DStudioMax, após umas trocas de ideias no fórum de

software sobre Web3D no site <http://www.renderosity>, descobriu-se uma aplicação shareware de nome “Maximum Pose” capaz de fazer a ponte entre estas duas aplicações (Figura 58).

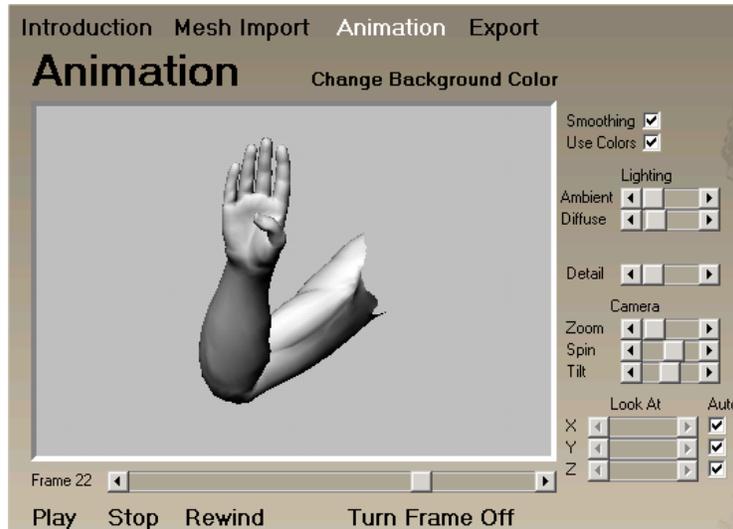


Figura 58 – Interface da aplicação Maximum Pose

Assim sendo, foram animados os modelos no Poser, exportados para o formato Wavefront-obj criando um ficheiro para cada *frame* da animação, compondo em média 30 ficheiros por animação. Estes ficheiros foram importados pela aplicação Maximum Pose através da opção “Mesh Import”. Ainda no Maximum Pose, a animação poderia ser visualizada *frame a frame*, podendo especificar ainda algumas características da animação. Após todas as especificações, foi exportado o modelo e a animação para cada letra do abecedário em dois ficheiros, respectivamente, um do tipo MS-3DStudioScript (contendo o movimento) e outro do tipo 3DS (contendo o modelo), importados posteriormente pelo 3DStudioMax.

4.2.3 “Morpher Modifier” no 3DStudioMax.

Já no 3DStudioMax, e após a importação e conversão dos ficheiros do tipo *script* e 3DS para ficheiros proprietários do 3DstudioMax do tipo “Max”, todas as animações teriam de ser convertidas para o tipo *Morpher* para posteriormente poderem ser convertidas para o Pulse3D. *Morphing* é uma técnica de animação semelhante ao *twening* na animação 2D, ou seja, um objecto *Morph* combina dois ou mais objectos pela interpolação dos vértices do primeiro objecto combinando com este os vértices do objecto seguinte. Quando esta interpolação ocorre durante um determinado espaço de tempo, resulta então uma animação do tipo *Morphing* (Figura 59).



Figura 59 – Animação do tipo *Morphing*

As posições dos vértices dos objectos 1 e 2 são combinados como extremos de uma interpolação no terceiro objecto, gerando assim uma animação do tipo *Morphing*.

Assim e para obter o grau de realismo conseguido durante a animação dos gestos correspondentes a cada letra do Gestuário bem como a transição entre os mesmos, teriam de ser definidas 30 posições (objectos) para cada letra que seriam interpoladas através da técnica de *Morpher Modifier*, já que, como foi dito, seria o único modo de converter estas animações para o Pulse3D.

Dado que o trabalho de definir um a um, 30 estados diferentes para cada uma das 26 letras seria ingrato e moroso, pensou-se em elaborar uma *script* no 3DStudioMax que pudesse ajudar a esse fim. Assim se construíram animações prontas a serem exportadas para o Pulse3D como se pode ver na Figura 60.

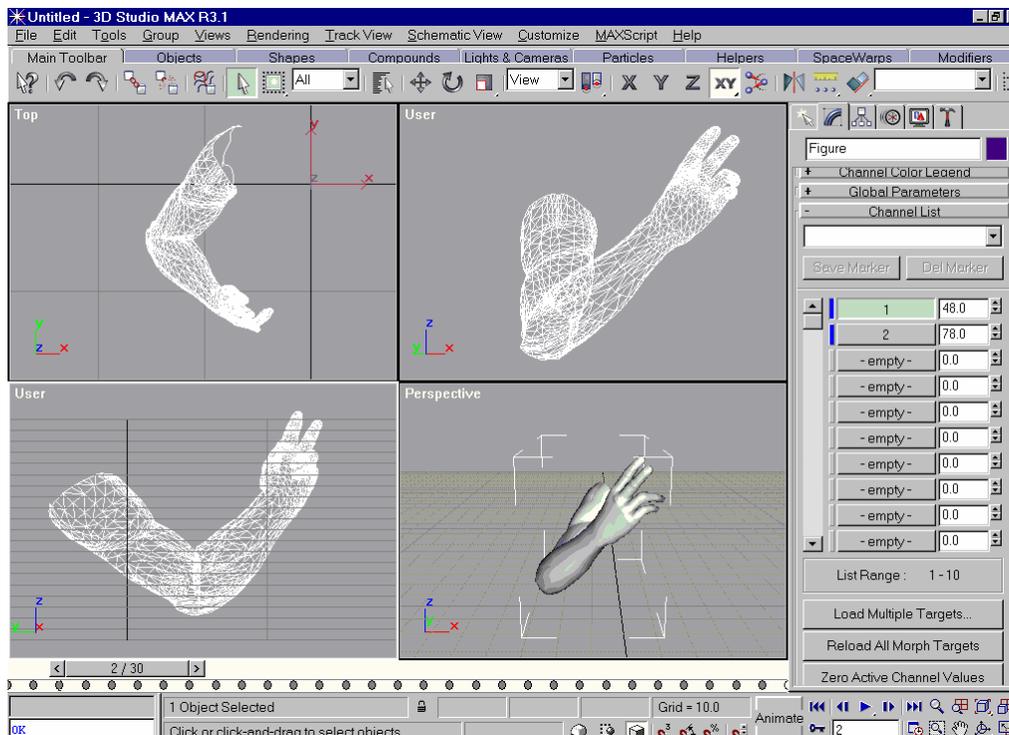


Figura 60 – Momento de redefinição dos estados de Morpher no 3DStudioMax

4.2.4 Display no Pulse 3D – Página Web

Depois de importadas as animações resultantes para o Pulse3D com a ajuda do *Plug-in - Pulse Producer*, foram organizadas e armazenadas num único ficheiro Pulse as 26 animações correspondendo às letras do abecedário (Figura 61), cada letra contendo em média 30 estados do tipo “*morph*” (derivados do 3DStudioMax) que possibilitaram o bom desempenho final da animação.

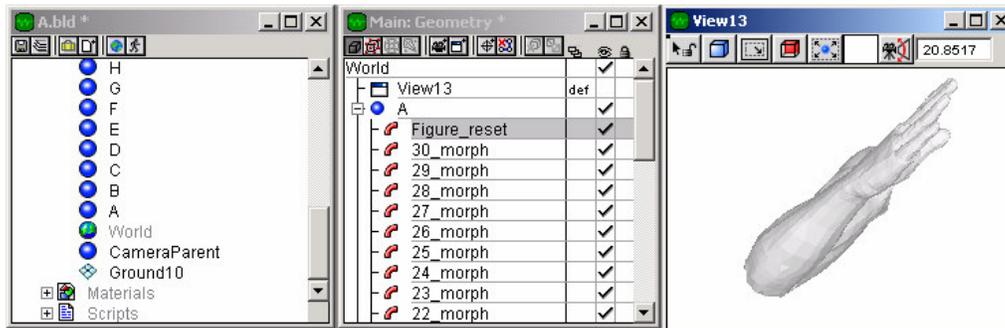


Figura 61 – Estrutura hierárquica do Abecedário3D

Foram ainda definidas no Pulse3D, *behaviors* e *scripts* necessárias para conseguir a visibilidade do objecto e animação da letra escolhida pelo utilizador enquanto que as outras ficariam ocultas, mas em memória de modo a evitar estados de espera, transições lentas e sobreposição das animações.

Após a definição de toda a estrutura hierárquica e conseguidos os ajustes efectivos, foram gravados os ficheiros web que deram origem a aplicação do Abecedário3D resultando na aplicação interactiva Web apresentada na Figura 62. São ainda apresentadas 4 vistas diferentes da animação (de frente, de cima, do ponto de vista do humanóide e uma vista rotativa sobre o braço: Cam1, Cam2, Cam3 e Cam4 respectivamente) de forma a animar o modelo em ângulos distintos para uma perfeita compreensão dos movimentos constituintes de um determinado gesto.

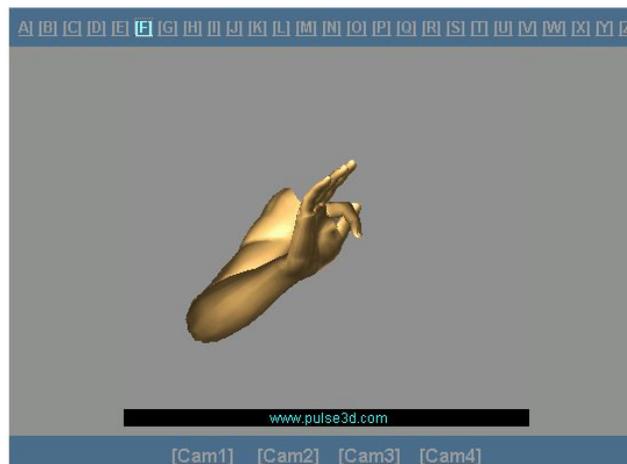


Figura 62 – Representação da letra F

4.3 Animações Víde

As animações vídeo de gestos representando todo o tipo de palavras, letras ou números em língua gestual é outro meio possível para partilhar na rede ou num CDROM os movimentos e expressões de alguns dos gestos mais comuns e de uso mais frequente. Por norma, este tipo de animação é realizado utilizando pessoas reais que são gravadas digitalmente quando implementam o respectivo gesto gerando um ficheiro de vídeo, colocado num destino específico da aplicação Web. A utilização de imagens reais desperta muito interesse entre os utilizadores tal como foi verificado no Capítulo 2 quando se analisou a Base de Dados - Symbol - para a aprendizagem de vocabulário, destinado a pessoas com transtornos para comunicar oralmente.

No trabalho prático implementado para complementar esta tese, decidiu-se também pelo uso de animações vídeo para representação de verbos e algumas palavras de uso mais frequente ordenadas pelas letras de “A” a “Z” conforme se pode verificar na Figura 63, mas estas animações foram geradas a partir de personagens 3D virtuais e não a partir de imagens reais. Este tipo de animação, raramente encontrado em aplicações do género na representação de gestos, é também bem aceite pelos utilizadores com vantagens acrescidas, como a possibilidade e facilidade de implementar um aumento ou redução no tamanho da imagens sem perda de qualidade e principalmente pela possibilidade de gerar gestos de ângulos ou vistas diferentes após a construção do modelo e a animação de um respectivo gesto.



Figura 63 – Animação gestual do verbo amar

Os modelos utilizados foram retirados da aplicação Poser e posteriormente alterados de forma a poderem ser manipulados e animados ainda dentro do Poser, daí resultaram os diferentes ficheiros de vídeo que fazem parte também desta aplicação.

4.4 Gestuário 2D

Formular uma definição rigorosa e completa da linguagem humana não é tarefa fácil, apesar do que poderão pensar os falantes como utilizadores normais desse instrumento. A língua gestual Portuguesa possui um léxico rico, variado e potencialmente infinito, formado por unidades significantes arbitrárias, embora se tenha de reconhecer a importância da analogia icónica como princípio da formação inicial de alguns gestos. Assim sendo, apresenta-se um gestuário2D de alguns dos gestos mais utilizados, neste caso específico a numeração. No caso apresentado, aliado à sequência figurativa representando a estilização dos gestos realizados pelo surdo gestualista, aparece a descrição verbal de cada uma destas sequências tentando ir o mais longe possível na individualização de cada configuração, cada orientação, cada localização, cada movimento e até cada expressão corporal. O surdo gestualista é sempre apresentado em “plano americano”, de modo a que facilite a compreensão da globalidade do gesto, sem perder nada da individualidade de cada um dos parâmetros, bem como da localização dos braços e das mãos relativamente ao corpo e da inevitável expressividade.

É também apresentado uma chamada de atenção do utilizador para um sem número de particularidades dos gestos, de semelhanças, de variações e para tudo quanto se entenda ser útil para uma melhor compreensão e memorização correcta do gesto, bem como para alguns aspectos da gramática da própria língua gestual.

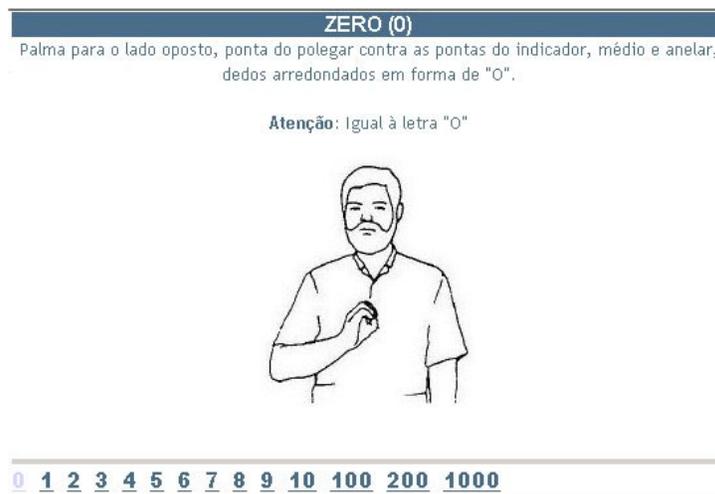


Figura 64 – Imagem 2D do gesto representando o número zero

Os dados para esta componente foram retirados do Gestuário de língua gestual Portuguesa [3].

Capítulo 5

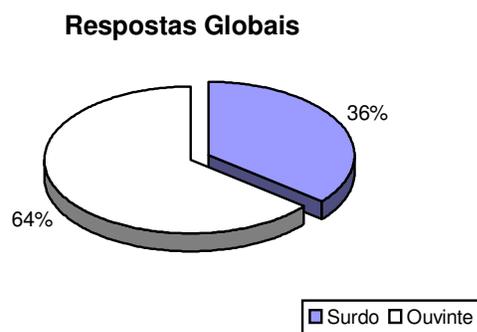
Testes de Usabilidade

5.1 Perguntas e Respostas

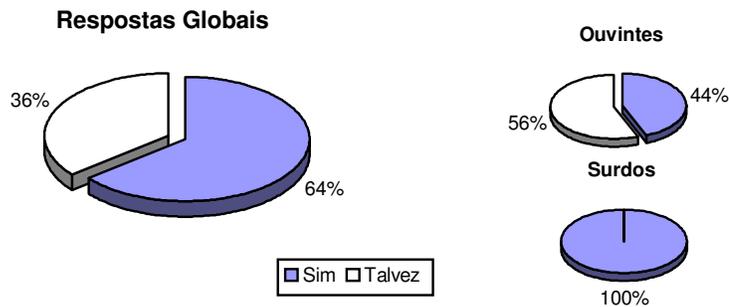
O inquérito que se apresenta em anexos foi proposto à Associação de Pais para a Educação de Crianças Deficientes Auditivas (APECDA) – Mazagão- Aveleda- Braga entre as datas de Novembro de 2001 e Janeiro de 2002, depois de apresentada e testada a aplicação, obtiveram-se respostas num total de 14 inquéritos. Este, é composto por 24 perguntas que são apresentadas a seguir bem como as respostas e um pequeno comentário aos resultados a cada uma delas.

Achou-se por bem, dividir e concluir os resultados a partir de uma análise comparativa entre duas secções: as respostas dos inquiridos ouvintes e as respostas dos inquiridos surdos, que totalizam 9 e 5 respectivamente.

Como se caracteriza?

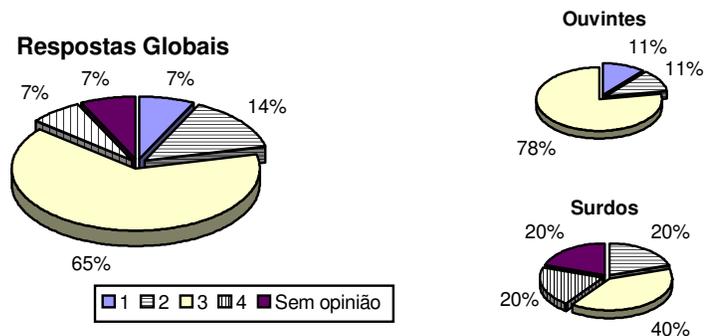


Acha que a Informática/Internet pode ajudar à aprendizagem da Língua Gestual Portuguesa (LGP)?



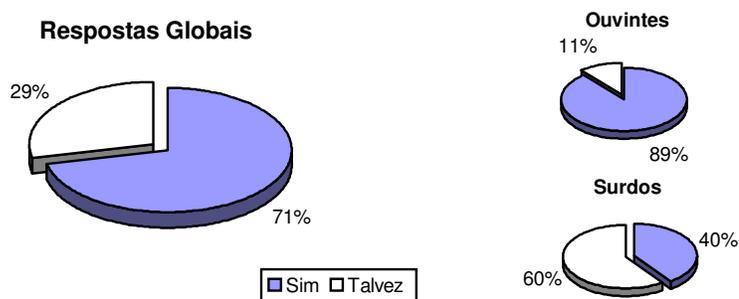
Na globalidade 64% dos inquiridos apontam o “Sim” com resposta a esta questão. Os ouvintes inquiridos dividem-se entre o “Sim” e o “Talvez” enquanto que há uma total concordância da parte dos inquiridos surdos. Poderemos verificar assim que principalmente as pessoas surdas vêm na informática em geral e na Internet em particular uma via muito aceitável para a aprendizagem da LGP.

Exprima de 0 (min) a 4 (max) qual acha ser seu grau de conhecimento de LGP?



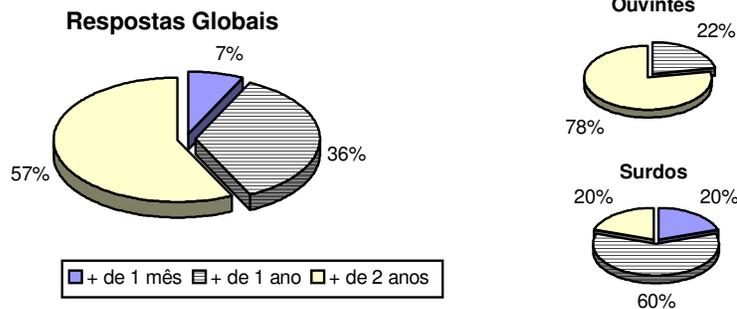
65% dos inquiridos revelou ter um conhecimento de LGP de 3 numa escala de 0 a 4. Os resultados obtidos deixam transparecer no geral que os inquiridos surdos têm um conhecimento mais profundo de LGP em relação aos ouvintes, ouvintes estes, claro está, que se interessam e necessitam de aprender LGP. As respostas são em média de “3” para os surdos e “2,7” para os ouvintes.

Existe falta no mercado de aplicações informáticas que apoiem a LGP?



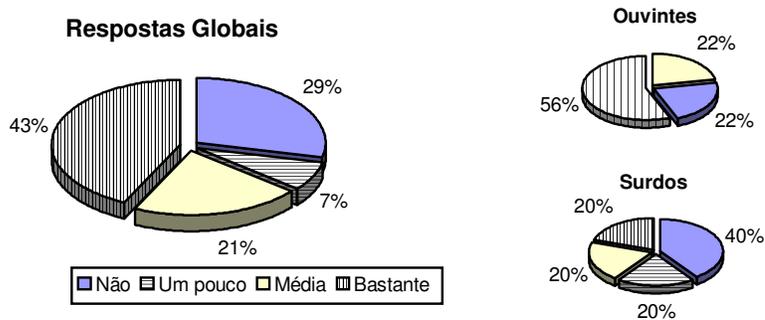
No geral obtiveram-se 71% de respostas “Sim” e 29% de “Talvez”. De notar que a maioria das respostas “Talvez” se fez sentir da parte dos inquiridos surdos.

No seu entender, quanto tempo acha necessário para se conseguir aprender o básico de forma a estabelecer uma conversa normal em LGP?



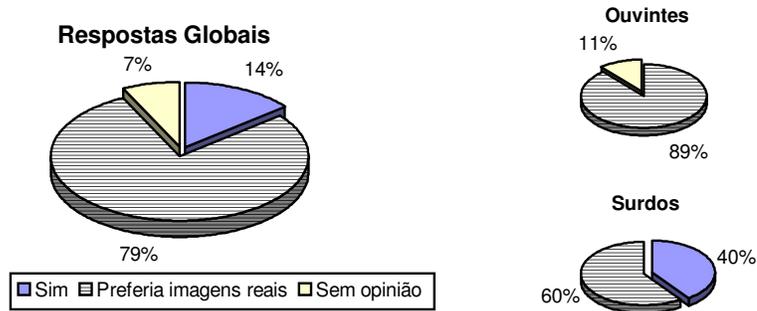
No bolo total, a maioria dos inquiridos (57%) define como mais de 2 anos como o tempo necessário de forma a estabelecer uma conversa normal em LGP. Percebe-se sensivelmente que os inquiridos surdos apontam para um período de aprendizagem bastante inferior comparativamente com as respostas dos ouvintes.

Acha difícil a utilização da aplicação no seu conjunto?



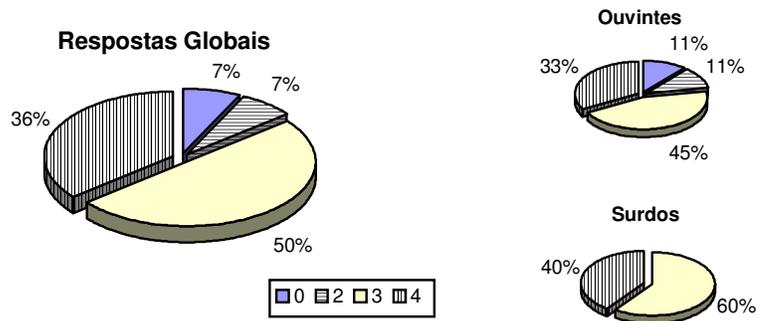
A resposta maioritária com 43%, da totalidade dos inquiridos foi de “Bastante”, de notar que apenas um inquirido surdo subscreveu esta resposta, donde se poderá deduzir que os inquiridos surdos se adaptaram mais rapidamente ao interface facilitando assim a utilização da aplicação. A maioria das respostas das pessoas surdas que testaram e responderam a este inquérito não acharam difícil a utilização desta aplicação no seu conjunto.

Acha que a utilização de personagens 3D tal como se apresentam na aplicação ajuda e motiva a aprendizagem de LGP?



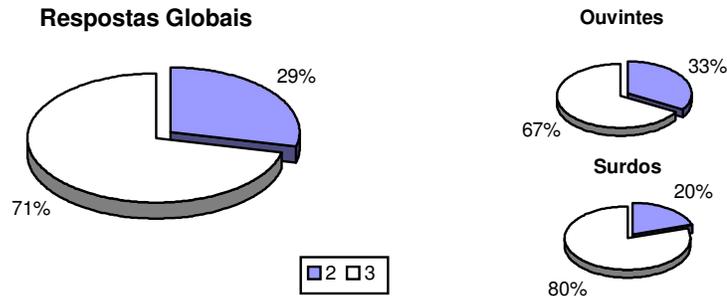
Ambos os lados (Surdos e Ouvintes) acham que a utilização de imagens reais ajudaria mais na motivação da aprendizagem de LGP (79% do total das repostas). Com o realce de serem os inquiridos surdos os únicos a optarem por personagens 3D em 40% das suas repostas.

Exprima de 0 (min) a 4 (max) como considera útil o abecedário 3D da aplicação na aprendizagem e aperfeiçoamento das letras LGP?



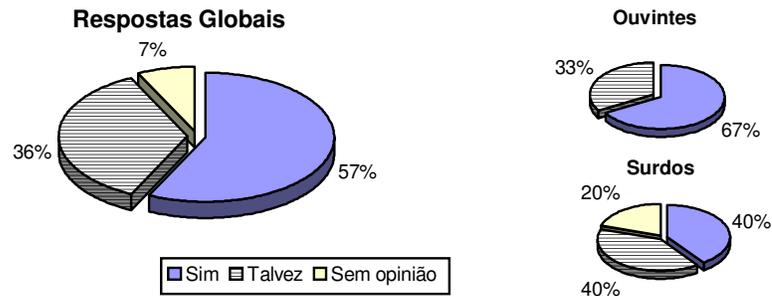
Esta questão obteve como resposta maioritária o “3” com 50% das respostas globais. Entre surdos e ouvintes esta resposta manteve-se como maioria, em segundo lugar aparece a resposta “4” em ambos os casos, as repostas dos inquiridos surdos ficaram-se por aqui enquanto que houve duas respostas de inquiridos ouvintes abaixo do “3”.

Exprima de 0 (min) a 4 (max) a perfeição dos gestos e do modelo apresentado no abecedário 3D.



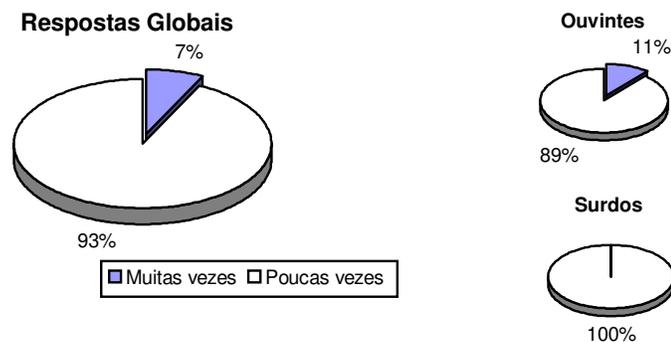
Obtiveram-se duas únicas faixas de respostas: o “2” e o “3”. Sendo maioritária a resposta “3” com 71% dos inquiridos a optarem por esta. Nesta questão não houve desvios de respostas consideráveis entre surdos e ouvintes.

A construção de um determinado gesto tal como acontece no construtor 3D é viável à posterior memorização do mesmo?



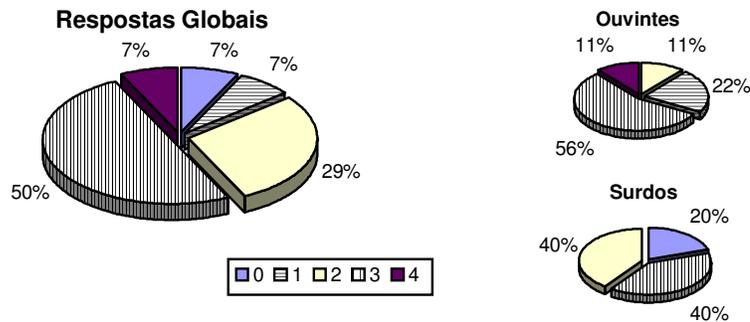
Para esta questão, 57% dos inquiridos responderam que “Sim”, 36% “Talvez” e 7% não têm opinião sobre a resposta. Não se notaram diferenças entre as repostas de inquiridos surdos e ouvintes. Ambos consideram assim, importante a função do Construtor de movimentos para posterior memorização dos gestos.

Consegue construir gestos no construtor 3D capazes de representar letras, números ou palavras de forma inequívoca?



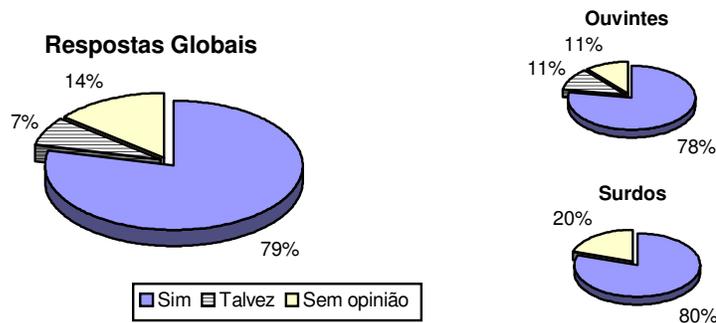
Esta questão, obteve uma resposta em massa para a opção “Poucas Vezes”, registrando unicamente uma única resposta fora da maioria, essa resposta foi “Muitas vezes” dada por um inquirido ouvinte. Temos de ter em linha de conta que os inquiridos não tinham um grau muito elevado de conhecimentos de informática mesmo ao nível do utilizador, o que torna ainda mais difícil o trabalho com objectos tridimensionais.

Exprima de 0 (fácil) a 4 (difícil) como considera a construção de gestos e a respectiva manipulação dos modelos do construtor 3D.



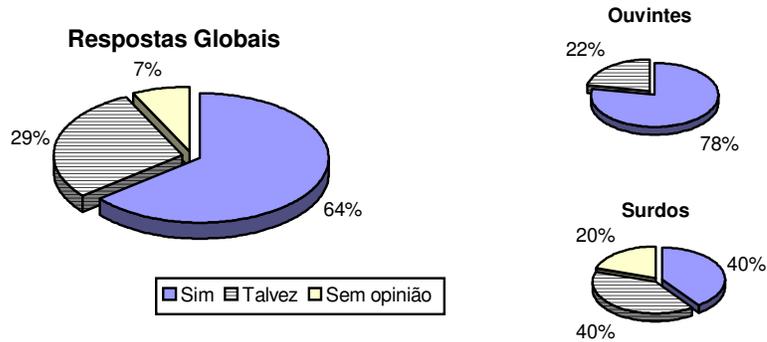
A questão nº 12, teve como resposta maioritária o “3” com 50% no global das respostas. Entre inquiridos surdos e ouvintes, há uma clara tendência para uma maior facilidade no uso do Construtor 3D do lado dos inquiridos surdos.

Considera que há pouca informação do Gestuário entre a comunidade surda o que permite por vezes o aparecimento de vários gestos para a mesma palavra?



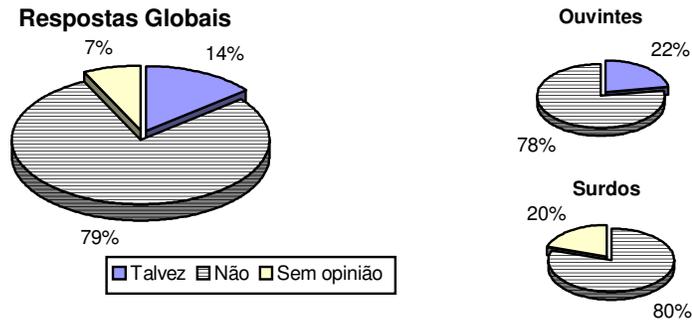
A maioria dos inquiridos optaram pelo “Sim” com 79 % das respostas, 14% não têm opinião sobre esta questão e 7% respondeu “Talvez”. Entre surdos e ouvintes ambos estão sincronizados nas respostas e na aceitação da fraca divulgação do Gestuário entre a comunidade surda.

Acha que o Gestuário transposto na integra para a Internet facilitaria a aprendizagem e popularização da LG em Portugal?



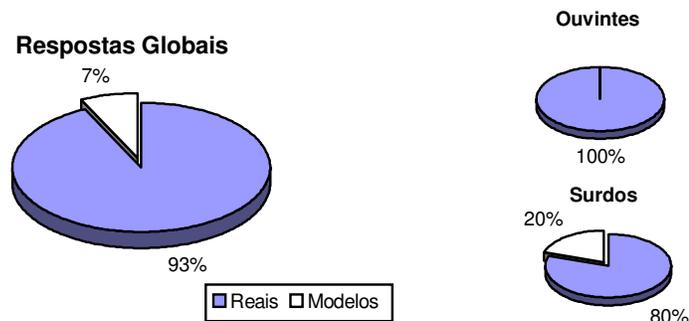
Temos que claramente e no geral o “Sim” ganha com 64% das respostas, mas olhando para o gráficos comparativos entre respostas de inquiridos surdos e ouvintes temos que os ouvintes são os mais crentes na facilidade de aprendizagem de LG recorrendo à transposição do Gestuário na Internet.

Acha que um Gestuário 2D e consequentes vídeos são suficiente para o ensino da LGP?



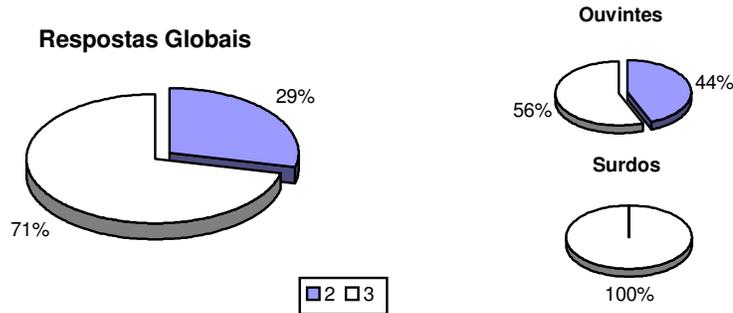
A larga maioria dos inquiridos (79%) acha que estes dois elemento não são suficientes para o ensino de LGP. De salientar ainda, que todos os inquiridos surdos responderam “Não” com exceção de uma resposta sem opinião.

Prefere animações de personagens reais ou animações com modelos tal como são apresentados nas animações vídeo?



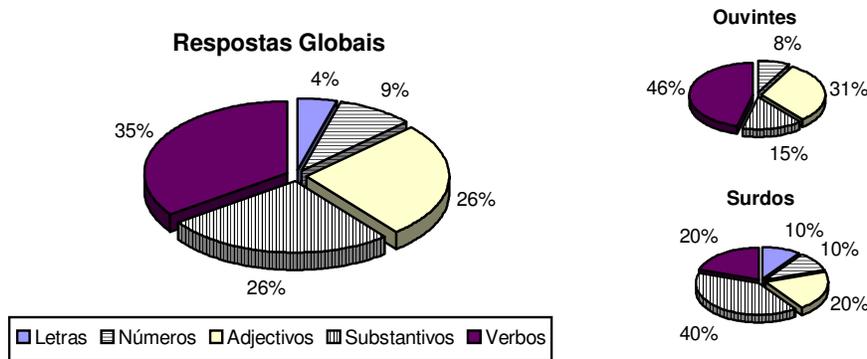
Mais uma vez, as personagens reais, neste caso para as animações vídeo, são as preferidas pelos inquiridos com 93% das respostas. A única resposta a preferir modelos virtuais para as animações vídeo foi dada por um inquirido surdo.

Exprima de 0 (min) a 4 (max) qual acha ser seu grau detalhe e perfeição das animações vídeo apresentadas?



Ambos os inquiridos (surdos e ouvintes) têm maioritariamente o “3” como resposta maioritária a esta questão, mas como se pode verificar, são os inquiridos surdos que consideram uma perfeição superior nas animações apresentadas.

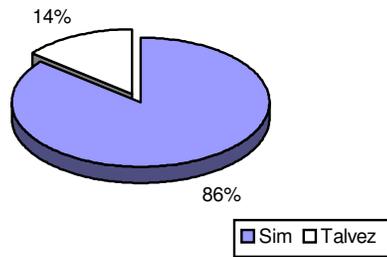
Se tivesse que escolher, o que seria preferencial animar em vídeo?



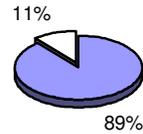
Esta foi a única questão que teve como opção, respostas múltiplas para cada inquirido. Estas, foram muito repartidas, principalmente entre os verbos (35%), substantivos (26%) e adjectivos (26%). Na análise entre inquiridos surdos e ouvintes de notar que enquanto os ouvintes referem os verbos em maioria, os surdos referem os substantivos. As letras são unicamente referidas pelos inquiridos surdos.

Acha que seria benéfico dar continuidade ao trabalho até aqui efectuado?

Respostas Globais



Ouvintes



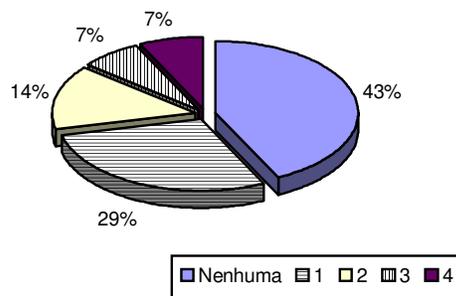
Surdos



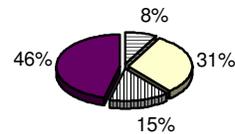
86% do total das respostas decidiram pelo “Sim”, enquanto 14% responderam “Talvez”. Entre ouvintes e surdos não houve diferenciação nas respostas. Assim, temos que ambos acham que o trabalho efectuado até aqui é válido, e que deveria ter continuidade.

Além desta, quantas aplicações informáticas conhece para apoio ou ensino da LGP?

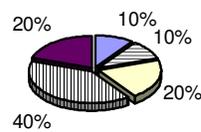
Respostas Globais



Ouvintes



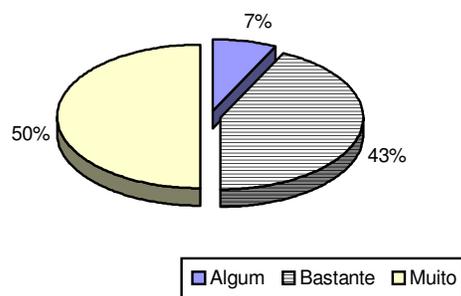
Surdos



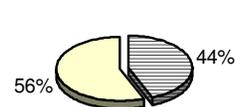
As respostas a esta questão denotam que a maioria dos inquiridos, tanto surdos como ouvintes não conhece nenhuma aplicação de apoio a LGP além desta. De salientar ainda que nenhum dos inquiridos conhece mais de 4 aplicações para este fim.

Acha que ainda há muito caminho a percorrer para que a Informática seja o apoio efectivo às pessoas com dificuldades auditivas?

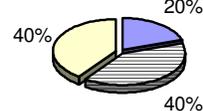
Respostas Globais



Ouvintes

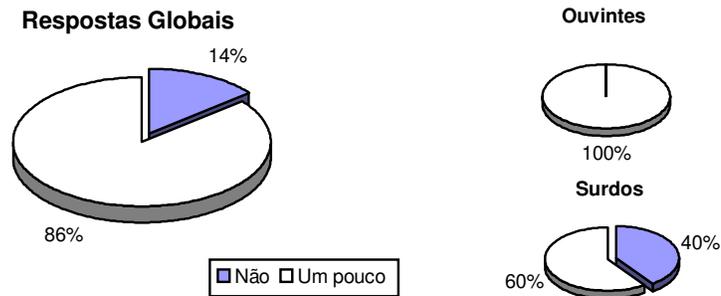


Surdos



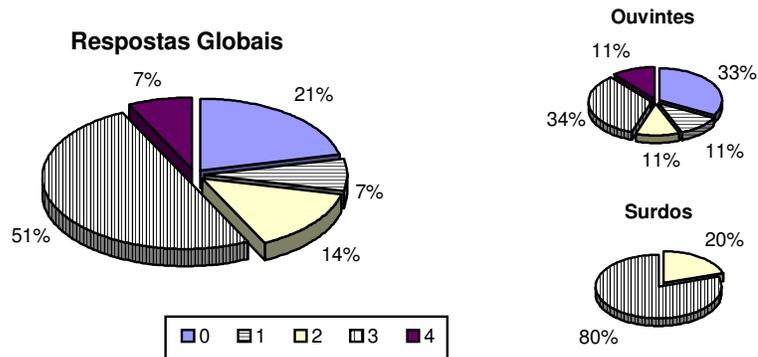
A resposta maioritária é o “Muito” com 50%, seguida de perto o “Bastante” com 43%. Entre surdos e ouvintes nota-se levemente que os inquiridos surdos se sobressaem em concordar que a informática é já hoje um apoio às pessoas com deficiências auditivas.

Acha que a utilização desta ferramenta é de alguma forma lenta ou difícil?



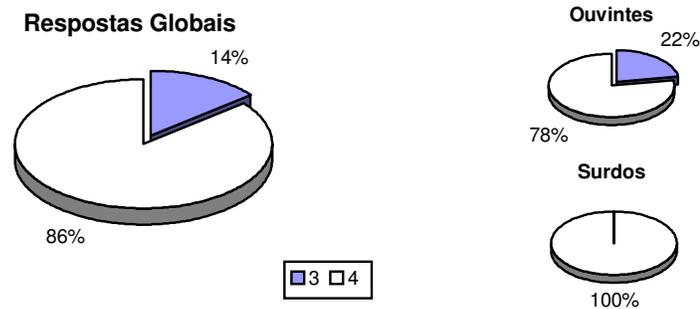
As respostas a esta questão, permitem verificar, que grande parte dos inquiridos se adaptaram com alguma dificuldade à interface da aplicação, sendo que consideram a utilização da ferramenta no geral um pouco difícil e lenta. Na comparação entre respostas de inquiridos surdos e ouvintes, temos que os primeiros se dividem entre o “Um pouco” e o “Não” enquanto que todos os segundos acham a utilização um pouco difícil ou lenta.

Exprima de 0 (min) a 4 (max) e de uma forma geral a utilização que dará a esta aplicação no futuro.



A resposta “3” foi a escolhida por 51% dos inquiridos, na comparação entre respostas de surdos e ouvintes esta questão revela o importante facto de serem os inquiridos surdos que mais contam com esta aplicação no futuro, as respostas destes situam-se somente entre a faixa “2” e “3” sendo que 80% destes optaram pela última.

Exprima de 0 (min) a 4 (max) necessidade para um surdo aprender a LGP?



No global, a larga maioria dos inquiridos (86%) revelaram a pontuação máxima para a necessidade de surdos aprenderem LGP. Entre as respostas dos inquiridos surdos temos a totalidade das respostas nesta mesma escolha.

5.2 Resumo Geral

5.2.1 A Aprendizagem de Língua Gestual Recorrendo à Informática

A maioria dos inquiridos acham que a informática/Internet pode ajudar ao ensino de língua gestual Portuguesa. O grau de conhecimento da LGP por parte dos inquiridos revelou ser elevado, em especial o dos inquiridos surdos. Os inquiridos acham que existe falta no mercado aplicações informáticas que apoiem a LGP. Para aprender a estabelecer uma conversa fluida em LGP os inquiridos surdos acharam que se necessitaria mais de um ano, enquanto que os inquiridos ouvintes aumentaram este tempo para mais de dois anos. De notar, a escassez de conhecimento de aplicações deste género por parte dos inquiridos, sendo mesmo, que a maioria não conhece nenhuma com estes fins. Os inquiridos dividem-se entre o “bastante” e “muito” no respeitante ao caminho por percorrer para que a informática seja o apoio efectivo às pessoas com dificuldades auditivas.

5.2.2 O Abecedário3D

Em resumo e relativamente ao abecedário3D temos que os inquiridos prefeririam imagens reais na apresentação do abecedário3D, como já se tinha verificado em testes efectuados em aplicações multimédia e referidos na tese no 2º Capítulo “*State of the art*”, mas como facilmente se percebe, este método não tem as mesmas condições e potencialidades de apresentação das imagens 3D. No que se refere à utilização do abecedário3D, a ilação resultante é de que o abecedário3D é útil para a aprendizagem e aperfeiçoamento dos gestos correspondentes às letras LGP. Para os inquiridos, o grau de perfeição dos gestos e do modelo apresentado situa-se num grau 3 numa escala de 0 a 4.

5.2.3 O Construtor 3D

Em conclusão no que diz respeito à utilização do Construtor 3D temos que este se apresenta como algo difícil de utilizar e nem sempre os inquiridos conseguiram construir gestos capazes de representar de forma inequívoca letras, números ou palavras em LGP. No entanto, a maior parte dos inquiridos acham que construção de um determinado gesto tal como acontece no Construtor 3D é viável à posterior memorização do mesmo. De ter atenção que o grau de conhecimento de informática não é elevado e o tempo reservado para a aprendizagem e manipulação da ferramenta não foi nalguns dos casos talvez o suficiente. Entre surdos e ouvintes, mais uma vez se confirma que os primeiros inquiridos conseguem melhores desempenhos na utilização da aplicação, talvez pela necessidade e interesse dos seus objectivos.

5.2.4 O Gestuário2D

A simples passagem (digitalização) do Gestuário para a Internet não é considerada por parte dos inquiridos, suficiente para o ensino de LGP, mas consideram maioritariamente os ouvintes que este trabalho facilitaria a sua aprendizagem. Em unanimidade estão inquiridos surdos e ouvintes na escassa utilização do Gestuário por parte da comunidade surda, o que provoca por vezes, o aparecimento em Portugal de mais de um gesto para cada palavra, ou mesmo letra ou número.

5.2.5 As Animações Vídeo Propostas

Em resumo, e no geral para os inquiridos, as animações com base em imagens reais são as preferidas para o ensino de LGP. No global, a perfeição das animações é bastante aceitável, uma vez mais, são os inquiridos surdos a acharem a aplicação mais completa, neste caso, em termos da perfeição das animações. Em relação ao que seria preferencial animar nas animações vídeo temos resultados diferentes nas respostas de surdos e ouvintes. Os surdos dividem as respostas por todas opções, com preferência por substantivos, os ouvintes preferiam ver animados principalmente verbos e depois os adjectivos.

5.2.6 A Aplicação no Geral

No geral, os inquiridos acham algo difícil a utilização da aplicação no seu conjunto, no entanto, a maioria dos inquiridos surdos não acha difícil a sua utilização. Ambos estão de acordo quando se lhes pergunta se acham que seria benéfico continuar com o trabalho apresentado. A utilização desta ferramenta foi de alguma forma um pouco difícil de utilizar por parte dos inquiridos, mas não tanto entre os inquiridos surdos. Metade dos inquiridos afirmam vontade de passar a utilizar regularmente esta ferramenta no futuro, sendo que a totalidade inquiridos surdos pretende utilizar muito frequentemente no futuro esta aplicação prática.

Capítulo 6

Conclusões e Trabalho Futuro

6.1 Conclusões

As questões de acessibilidade na utilização de interfaces nos meios informáticos, por parte de pessoas com certas necessidades especiais, têm vindo a tomar relevância especial de algum tempo a esta parte, com a formação de uma opinião pública cada vez mais consciente do tipo de dificuldades associadas a esta problemática. A comunicação com pessoas com deficiências auditivas pode ser melhorada, oferecendo a uma vasta audiência a oportunidade de aprender língua gestual a uma escala mais alargada. A comunidade linguística em língua gestual Portuguesa necessita de ser estendida, por forma a minimizar a exclusão social de alguns cidadãos, para isso, terá de haver investimento em apoio tecnológico para ensinar o português escrito aos surdos e língua gestual Portuguesa a surdos e também aos ouvintes.

Cada vez mais, intérpretes de língua gestual, estão ao serviço de construtores de aplicações gráficas de apoio à língua gestual, com o objectivo de transmitir os movimentos gestuais a aplicações que intentarão ajudar os anteriores em propósitos educacionais de língua gestual. Aplicações gráficas estas, por norma constituídas com base em humanóides 3D, ou parte deles, capazes de documentar língua gestual na World-Wide-Web. Este desenvolvimento poderá marcar um grande impacto no campo dos “livros de ensino de língua gestual”. O que se pretende principalmente, é conseguir converter texto escrito em gestos correspondentes numa das línguas gestuais, em tempo real, e sem necessidade de intervenção humana.

A componente prática desta dissertação passou pelo desenho, construção e implementação de um sistema de ajuda ao ensino de língua gestual Portuguesa com base numa aplicação gráfica tridimensional composta de “ferramentas” diversas, por forma a se enquadrar a um nível mais abrangente no propósito educacional de língua gestual Portuguesa a todos os sectores etários e linguísticos .

Este sistema é proposto via Internet nos endereços: <http://www.ipb.pt/~leodeus/LGP> ou <http://sim.di.uminho.pt/mestrados/deusdado> e também através de um CD-ROM que se apresenta em anexo.

A principal vertente, como já foi referido, é o ensino de língua gestual a adultos e crianças com dificuldades auditivas e também o ensino à comunidade ouvinte para que as barreiras que os separam possam ser cada vez mais curtas, assim sendo, foram construídos humanóides 3D, ou parte deles, com carácter interactivo e manobrável no caso do construtor3D e abecedário3D. Ou seja, uma interface que permitiu aos utilizadores no caso do construtor3D, definir novas posições e associar a estas uma letra, palavra ou frase por forma a memorizarem os gestos construídos. No caso do abecedário3D, a apresentação dos gestos representativos de todas as letras em língua gestual Portuguesa, permite uma melhor aprendizagem e um aperfeiçoamento superior destes gestos, já que compreende modelos e movimentos bastante perfeitos, bem como a possibilidade de visualizar os movimentos de ângulos diversos.

A biomecânica do tronco superior humano é muito complexa, a representação dos modelos 3D propostos para a aprendizagem de língua gestual recorrem como é lógico unicamente uma a aproximação dos movimentos humanos. Houve no entanto, uma simplificação não uniforme, do tronco superior, já que, como facilmente se percebe, as posições do corpo em língua gestual envolvem mais os movimentos dos membros superiores do que por exemplo, movimentos da espinha dorsal.

Recorreu-se assim à cinemática inversa de modo a criar movimentos de figuras articuladas, modeladas por um conjunto de segmentos rígidos conectados por junções organizados numa estrutura em árvore. A esse conjunto de articulações procurou dar-se-lhes expressões de uma forma simples, ordenada e natural, de modo a que quando houvesse interacção do utilizador com os modelos, os movimentos se assemelhassem com o maior rigor possível, aos movimentos humanos. Procurou-se assim, por um lado, averiguar quais as capacidades de animação de objectos tendo em conta os níveis de interacção entre estes (i.e., verificar como o movimento de um corpo se repercute sobre os corpos a ele ligados), por outro lado investigou-se a expressividade em termos de modelação gráfica de objectos com níveis de detalhe elevado, para se saber, até que ponto um modelo com um grau de detalhe elevado mantém a capacidade de ser animado num desempenho tão próximo quanto possível de uma animação em tempo real.

Foi assim encontrado no Pulse3D uma aplicação Web3D apropriada para a implementação desta aplicação já que conta com a facilidade de importação e modelação de objectos tridimensionais, facilidade de codificação, nível de controlo sobre o programa, portabilidade, desempenho e principalmente a capacidade de organizar um mundo 3D conseguindo resultados bastante satisfatórios oferecidos num ambiente Web. Gráficos 3D na Web em tempo real, capazes de “correrem” num browser ainda é uma algo difícil de otimizar, terá de haver sempre a instalação de um *plug-in* externo e conseguir uma estrutura de modelos e movimentos de tal forma leve, que não provoque frustração na hora de manipular e visualizar as personagens virtuais criadas.

Foi proposto à Associação de Pais para a Educação de Crianças Deficientes Auditivas (APECDA) situada em Braga, apoio e parceria nesta investigação para testar e devolver *feed-back* da aplicação construída, bem como, das necessidade ou falta de

aplicações do género em instituições que apoiem a aprendizagem de língua gestual Portuguesa. Os resultados e conclusões foram retirados sob a forma de um inquérito realizados a surdos e ouvintes utilizadores por algum tempo desta ferramenta. Da análise das respostas e numa vista geral podem-se retirar as seguintes deduções:

- A maioria dos inquiridos acham necessário e estão de acordo numa utilização mais alargada da Informática/Internet no ensino e aprendizagem de língua gestual Portuguesa.
- Existe falta no mercado de aplicações informáticas que de alguma forma apoiem a língua gestual Portuguesa, a maior parte dos inquiridos desconhece mesmo, qualquer aplicação informática de apoio a esta língua.
- Aprender, com o que se tem hoje, a comunicar fluidamente em língua gestual Portuguesa é um processo por muitos considerado aliciante mas moroso, podendo gastar até, mais de dois anos para se conseguir esse objectivo.
- A aplicação testada revelou ser no seu conjunto um pouco difícil de utilizar, principalmente devido aos inquiridos e utilizadores em geral não estarem habituados a utilizar ferramentas 3D.
- Os inquiridos sentiram-se motivados com a aplicação, mas ressaltaram o facto de quando possível preferirem visualizar gestos por imagens reais.
- Consideram útil e viável a utilização do Construtor3D e Abecedário3D para os propósitos que foram construídos, ou seja, a memorização e aperfeiçoamento dos gestos respectivamente.
- Acrescentam ainda, que a aplicação ajuda a normalizar a língua gestual por forma a evitar o aparecimento de vários gestos para a mesma palavra.
- Um Gestuário 2D na Internet e consequentes vídeos são importantes, mas não suficientes no apoio à aprendizagem de língua gestual Portuguesa.
- 86% dos inquiridos consideram útil dar continuidade ao trabalho até aqui desenvolvido.
- De uma forma geral e principalmente do lado dos inquiridos surdos afirmam a vontade de continuar a utilizar esta ferramenta futuramente.

Assim, e paralelamente à utilização de personagens 3D virtuais, seria benéfico também criar outros meios alternativos de apoio a instituições direccionadas ao ensino de língua gestual Portuguesa, tal como meios de comunicações mais alargados e eficientes passando pela criação de sites Web interligados capazes de fomentar o gosto e alargar conhecimentos e informação neste tipo de linguagem.

6.2 Trabalho Futuro

É possível estabelecer algum trabalho futuro baseado nas conclusões retiradas e nos problemas debatidos.

Em primeiro lugar, melhorar a interface da aplicação no seu conjunto, principalmente ao nível do construtor de movimentos, que se revelou ser algo difícil de utilizar por parte dos inquiridos. Construir além de um Abecedário3D que se demonstrou ser de grande utilidade e facilidade no uso, animações 3D interactivas para verbos, substantivos e adjetivos, bem como a possibilidade de guardar os movimentos criados no Construtor3D, utilizando um motor de base de dados, permitindo assim implementar

em tempo real, um tradutor de texto escrito para língua gestual Portuguesa com base nos movimentos guardados.

A separação entre a componente visual (modelo) e a estrutura dinâmica (biomecânica) dos objectos em causa, apresenta-se como um dos objectivos futuros, de modo a poder ter vários modelos (interfaces) para diferentes utilizadores, ganhando assim um leque muito mais alargado de interesse ao nível visual sem aumentar consideravelmente a aplicação e não perdendo qualidades de interacção e movimentação em tempo real.

Bibliografia

- [1] – Associação de Surdos do Porto - Disponível em:
<http://www.asurdosporto.pt/inicial.html> - 7 de Setembro de 2001.
- [2] – Em conversa com o Director da Associação de País para a Educação de Crianças Dificientes Auditivas (APECDA), Braga, Novembro 2001.
- [3] – Gestuário da Língua Gestual Portuguesa, 2ª Edição, 1992
- [4] – Justine Cassell, Catherine Pelachaud, Normand Badler, Mark Steedman, Brett Achorn, Tripp Becker, Brett Douville, Scott Prevost, Matthew Stone, "Animated Conversation: Rule-based Generation of Facial Expression, Gesture & Spoken Intonation for Multiple Conversational Agents", *Proceedings of SIGGRAPH 94 In Computer Graphics Proceedings, Annual Conference Series, 1994, ACM SIGGRAPH*, pp. 413-420.
- [5] – Daniel Thalmann, Hansrudi Noser and Zhiyong Huang, "How to Create a Virtual Life", In *Interactive Computer Animation*, 1996, Springer, pp 263-291.
- [6] – Shan Lu and Seiji Igi, "Towards a Dialog System Based on Recognition and Synthesis of Japanese Sign Language", *Gesture- and Sign-Language Based Communication in Human-Computer Interaction*, 1997, Springer.
- [7] – "Evaluación de soportes Multimedia para educación y supresión de barreras de comunicación con personas sordas", ver em:
<http://tid.es/presencia/boletin7/art004.htm>. Outubro 2001.
- [8] – G. Vanderheiden, "Making software more accesible for people with disabilities" 1995.
- [9] – Repositório VRML em <http://www.sdsc.edu/vrml/>
- [10] – Geitz, S., Hansen, T. and Maher, S. Computer generated 3-Dimensional models of manual alphabet handshapes for the World Wide Web, in *Proceedings of ASSETS '96 (Vancouver, Canada, April 1996)*, ACM Press, 27-31.
<http://holodeck.gsfc.nasa.gov/sign/sign.html>.
- [11] – Su, S.A. and Furuta, R. A logical hand device in virtual environments, in *Virtual Reality Software & Technology: Proceedings of the VRST'94 Conference (Singapore, August 1994)*. World Scientific Publishing Co., Singapore, 33-42.
- [12] – Spence, A.P. and Mason, E.B. *Human Anatomy and Physiology*, 3rd ed. Benjamin Cummings Pub. Co., 1987.

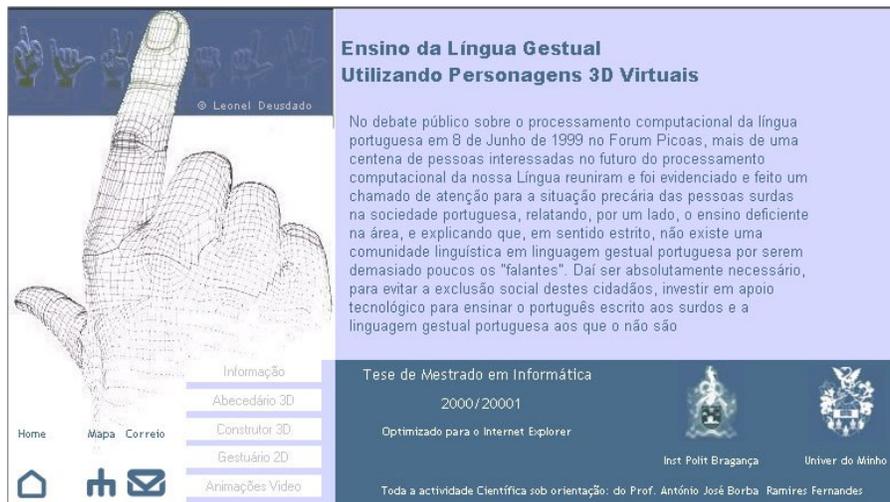
- [13] – Riekehof, L.L. *The Joy of Signing*, 2nd ed. Gospel Publishing House, Springfield, MO, 1987.
- [14] – Christopher Lee, Yangsheng Xu. "Online, Interactive Learning of Gestures for Human/Robot Interfaces." 1996 IEEE International Conference on Robotics and Automation, Minneapolis, MN. vol. 4, pp 2982-2987.
- [15] – "Multimédia and 3D for Deaf People", ver em <http://www.fhs-hagenberg.ac.at/mtd/projekte/FFF/3dSign/english.html>, Outubro 2001.
- [16] – Verlinder, M., Tijsseling, C., Frowein, H.: "A Signing Avatar on the www". 2000.
- [17] – Schulmeister R.: ViSiCAST: Virtual Signing, Capture, Animation, Storage and Transmission , 2000.
- [18] – Bangham, J. A., Cox, S.J., Elliott, R., Glauert, J.R.W, Marshall, I., Rankov, S., Wells, M.: An Overview of ViSiCAST. Presented at IEEE Seminar on "Speech and language processing for disable and elderly people" London, Abril 2000.
- [19] – Elliot, R., Glauert, J.R.W, Kennaway, J.R., Marshall, I.: The Development of language processing support for the ViSiCAST project. Presented at 4th International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2000) Washington, Novembro 2000.
- [20] – Kennaway, J.R.: Synthetic of Deaf Signing Gestures. Presented at International Gesture Workshop 2001, City University, London, Abril 2001.
- [21] – Bangham, J. A., Cox, S.J., Lincoln, M., Marshall, I., Tutt, M., Wells, M.: Signing for the Deaf using Virtual Humans. Presented at IEEE Seminar on "Speech and Language processing for disable and elderly people", London, Abril 2000.
- [22] – Project "Simon"- Disponível em: <http://www.televirtual.com>, Janeiro, 2002.
- [23] – Frank Godenschweger, Thomas Strothotte and Huber Wagener, "Rendering Gestures as Line Drawings", *Gesture- and Sign-Language Based Communication in Human-Computer Interaction*, 1997, Springer.
- [24] – Stefan Schlechtweg and Andreas Raab, "Rendering Line Drawings For Illustrative Purposes", In Thomas Strothotte and Hubert Wagener (eds), *Abstraction in Interactive Computational Visualization: Exploring Complex Information Spaces*, Springer 1997.
- [25] – Rogers, David F. ed "Computer Graphics Techniques", 1990 CG3010.
- [26] – Pinho, M. S. 1997: "Interacção em Ambientes Tridimensionais. 1997.
- [27] – Pinho, M et al: "A User Interface Model for Navigation in Virtual Worlds" in: *Conferência Latino-America de informática*, 1999, Assunción, Paraguai.

- [28] – Open Inventor. Silicon Graphics Inc. Disponível em:
<http://www.sgi.com/Technology/Inventor/> , Setembro 2000.
- [29] – GLEEM. OpenGL Extremely Easy-to-use Manipulators. Disponível em
<http://www.media.mit.edu/~kbrussel/gleem/>, Setembro. 2000.
- [30] – Jianmin Zhao and Norman I. Badler: “Inverse Kinematics Positioning Using Nonlinear Programming for Highly Articulated Figures” University of Pennsylvania Research, Outubro,1994.
- [31] – Nick Lowe, James Strauss, Sam Yeates, Eun Jung Holden, “Auslan Jam: A Graphical Sign Language Display System” University of Western Australia, 2000.
- [32] – Naganand Madhavapeddy and Stuart Ferguson, “ Specialised Constraints for an Inverse Kinematics Animation System Applied to Articulated Figures”, The Queen’s University of Belfast, Ireland, 1998
- [33] – Martin J, Crowley L, “An Appearance-Based Approach to Gesture Recognition”, proc. of the 9th Conf on Image Analysis And Processing, Italy, 1997.
- [34] – Regh J. M., Kanade T., “Visual Tracking of Height DOF Articulated Structures: an Application to Human hand Tracking”, Proc 3rd ECCV, Vol 2 1994, pp. 37-46
- [35] – Trevor Darrell and Alex Pentland. Attention-driven expression and gesture analysis in an interactive environment. MIT Media Lab Perceptual Computing Group Technical Report N^o 312; Massachusetts Institute of Technology; 1995.
- [36] – Ouhaddi H, Horain P., “Modélisation et suivi de la main” (CORESA 99) França, 1999 pp 109-114
- [37] – Martin, Private Communication. Appears in Baecker, “Picture-Driven Animation.” Proceedings of the 1969 Spring Joint Computer Conference. (1969):273-88
- [38] – Brown, C. M. L. Human-Computer Interface Design Guidelines. Ablex, Norwood, NJ, 1988
- [39] – Smith, S. L., e Mosier, J. N. Design Guidelines for Designing User Interface Software. Technical Report MTR-10090, The MITRE Corporation, Bedford, MA 01730, USA, 1986
- [40] – Turner R, Gobbetti E, Balanguier F, Mangili A, Thalmann D, Magnenat-Thalmann N: An Object Oriented Methodology Using Dynamic Variables for Animation and Scientific Visualization Proceedings Computer Graphics International 90 Springer-Verlag: 317-328, 1990

Anexos

Respostas Individuais aos Inquéritos

Conteúdo e Instalação do CDROM



O conteúdo do CDROM inclui a aplicação prática realizada para esta dissertação. É composta de uma introdução à tese com o resumo e objectivos da mesma, um Abecedário3D, um Construtor de movimentos 3D, um Gestuário 2D e Animações Vídeo de alguns gestos de verbos em língua gestual.

A instalação do CDROM é muito fácil:

- O Computador deverá contar com um *browser*, de preferência o Internet Explorer. Deverá ser instalado primeiro o *plug-in* do Pulse3D (Player). Esta instalação poderá ser feita directamente da Internet quando o utilizador entrar no Abecedário3D ou Construtor3D. Ou então através do Ficheiro "Player.exe" que se encontra no CDROM, escolhendo na opção de instalação "Setup Type" o tipo "Custom" e seleccionando somente o Player para instalação.
- Executar o ficheiro "index.htm" no Internet Explorer.