

Geração Automática de Olhares Expressivos em Humanos Virtuais

por

Rossana Baptista Queiroz

UNIVERSIDADE DO VALE DO RIO DOS SINOS

ROSSANA BAPTISTA QUEIROZ

**Geração Automática de Olhares
Expressivos em Humanos Virtuais**

Monografia apresentada como requisito
parcial para a obtenção do grau de
Bacharel em Ciência da Computação

Prof. Ms. Leandro Motta Barros
Orientador

São Leopoldo, junho de 2007

“Ora, Àquele que é poderoso para fazer tudo muito mais abundantemente além daquilo que pedimos ou imaginamos, a Esse seja glória para todo o sempre. Amém.”
— TRECHOS DE EFÉSIOS 3:20-21, BÍBLIA

AGRADECIMENTOS

Agradeço, em primeiro lugar, ao grande Artista e Imaginador de todos nós, que nos modelou, animou (comportamentalmente!) e nos renderizou neste mundo com tanta perfeição. Agradeço à minha mãe, por ter me apoiado durante essa jornada e também à “minha irmã gêmea de três anos de diferença”, Janne, com quem sempre pude compartilhar minhas imaginações sem qualquer restrição. Deixo uma homenagem ao meu pai, apesar de não estar presente em vida, mas sempre em minha memória. Apesar de não poder agradecê-lo, por hora, talvez seja a pessoa a qual sinto-me mais grata, pois foi ele que me passou (talvez pelos próprios genes) o gosto por esta ciência fantástica que é a Computação. Tenho certeza de que, assim como ele fazia com meus desenhos e poesias de quando eu era pequena, ele estaria todo “bobo” e orgulhoso com esta monografia pendurada no mural da empresa (se houvesse grampos fortes para segurá-la, claro), mostrando aos seus colegas: “foi minha filhinha quem fez!”... Ah, não posso esquecer, de forma alguma, de agradecer minha vó Joanna, que me emprestou o seu *notebook* tantas vezes e me tirou do sufoco! O resto da minha família agradeço em um lote só. Eles sabem que eu os amo.

Aos meus colegas e companheiros destes últimos sete anos e meio que passei... Foram tantos, que mesmo que eu tente restringir somente para os mais “chegados”, tenho receio de deixar passar alguém. Mas vou tentar: Antonio, Ferronato, Priscila, Tana e Cássio (pessoal do CTPetro). O Antonio merece mais uma citação, não tenho palavras para agradecer o quanto que ele me ensinou, inclusive sobre essas coisas complicadas de viver entre os seres humanos, fazer e ser amigos. Continuando: Ruthiano, Leandro (Coração Gelado e Docinho de Nozes... vem mais depois), Leandro (Hoffmann), Cadu, Fauzi, Root, Vinícius, Mierlo (especial agradecimento ao “*mieurlo* anjo da guarda”...). São meus preciosos. E a jovem-guarda, claro: Renato, Fabiane, Mírian, Marcelo, Solon, Rodrigo (meu “bolsista”), Jezer (meu “assistente de bolsista”), Chips, Bryan... Esse pessoal que entrou depois de mim, mas que eu pude me divertir bastante junto! Meus monitores, que espero que nenhum tenha desistido, assustado, por minha causa. E todos do Laboratório Kronos¹, digo, Cromos! Também aos meus professores, num lote só. Todos foram importantes, alguns me fizeram dormir um pouco mais e outros, bem menos. Falando sério, essa turma não tem preço. Se esqueci de alguém e esse alguém porventura ler este agradecimento, sim, prometo que pagarei uma prenda bem ridícula, como cantar o hino nacional em baleiês²! Também agradeço a todas as “cobaías” que participaram de meus experimentos e à minha família na fé, que são muitos, muitos, muitos... Obrigada!!!

Ainda preciso deixar tributos a três pessoas que foram imprescindíveis para que este

¹Laboratório do Síndrome, em Os Incríveis, Pixar (2004)

²Língua que as baleias falam. Fonte: Procurando Nemo, Pixar (2001)

trabalho existisse: Soraia, minha orientadora oficial na parte um e extraordinária na parte dois. O Leandro (viu, eu disse que tinha mais depois!). . . quem diria que este passaria da condição de um amigo tão admirado para orientador!!! A vida traz surpresas interessantes. Ou seriam *Easter Eggs* colocados pelo grande Imaginador? Prefiro não escrever mais, porque senão vai parecer puxa-saquismo e pegaria mal. O terceiro elemento é o Koray Balci, desenvolvedor do Xface. Eu não o conheço pessoalmente, mas nos comunicamos tanto no último ano que atrevo a chamá-lo de “meu amigo turco-italiano”. Sem a ajuda do Koray e sua Alice, que vocês vão conhecer se lerem o resto desta monografia, este trabalho não teria ficado com uma cara tão bonita. Literalmente.

Para finalizar os agradecimentos a seres humanos, ainda presto homenagem a um cara que eu nem cheguei a conhecer, mas que sou muito, muito grata: o Walt Disney. Sim, ele mesmo. Sou grata pelo fato de que seu trabalho pioneiro me proporciona hoje o que é mais do que diversão para mim. Desenho animado é minha grande paixão. Juntando com a computação, deu nisso que vemos nos cinemas, jogos e cada vez mais em aplicações diversas hoje: a Animação Computadorizada; simplesmente um “visualizador” tridimensional de nossos sonhos, um *player* da nossa imaginação, que todos podem ver. E pra finalizar de verdade esta seção, agradeço a todos os meus seres imaginários (elefantes rimantes, fellowsheeps, pterodákeos, lemingues persas, tartarugas-com-sorvete-na-cabeça. . .) Alguns deles já se materializaram em poemas, desenhos e até em Computação Gráfica. Se não fosse sua presença constante em minha mente imaginante, é provável (o bastante) que eu já tivesse enlouquecido — como um elefante.

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	7
LISTA DE FIGURAS	8
LISTA DE TABELAS	9
RESUMO	10
ABSTRACT	11
1 INTRODUÇÃO	12
2 FUNDAMENTOS	14
2.1 Animação Comportamental	14
2.2 Animação Facial	15
2.2.1 Padrão MPEG-4 de Animação Facial	15
2.3 Os Movimentos dos Olhos	18
2.3.1 As Sacadas Visuais	20
2.4 O Comportamento do Olhar	20
2.4.1 Padrões de Olhar em Processos Cognitivos	22
2.4.2 Troca de Turnos na Conversação	22
2.4.3 Olhar e Expressividade	23
2.5 Emoção e Estado Afetivo	24
2.5.1 As Seis Emoções Básicas de Ekman	26
2.5.2 Modelos Emocionais	27
3 ESTADO DA ARTE	29
3.1 Principais Abordagens	29
3.2 Modelos de Animação de Olhos	30
3.3 Contextualização do Trabalho no Estado-da-Arte	34
4 MODELO	35
4.1 Da Arte dos Animadores a um Modelo Comportamental	35
4.1.1 Metodologia	36
4.1.2 Comportamentos dos Olhos e sua Descrição	39
4.1.3 Modelo Padrão	42
4.1.4 Associação dos Movimentos da Cabeça, Pálpebras e Olhos	44
4.1.5 Arquitura do Modelo	44

5	PROTÓTIPO	46
5.1	Linguagem de Descrição de Olhares	46
5.2	A Base de Dados Comportamental	46
5.3	Gerador de Olhares Expressivos	47
5.4	Modelo Padrão	47
5.5	Plataforma de Animação Facial	48
5.5.1	Xface	50
5.5.2	Os Olhos de Alice	51
5.5.3	Integração do Xface no OSG	52
5.5.4	<i>Alice's Eyes Adventures in OSGLand</i>	52
6	RESULTADOS	54
6.1	Impacto das Sacadas em Conjunto com Expressões Faciais	54
6.2	Comportamentos Gerados no Protótipo	57
6.3	Avaliação dos Comportamentos com Usuários	58
7	CONSIDERAÇÕES FINAIS	65
7.1	Trabalhos Futuros	66
7.1.1	Metodologia de Coleta de Dados	66
7.1.2	Consolidação da Base de Dados Comportamental	67
7.1.3	Inclusão de Novos Comportamentos	67
7.1.4	Melhorias nas Expressões Faciais	68
7.1.5	Integração com um Sistema de Sincronização de Fala	68
7.1.6	Integração com um Modelo Emocional	68
7.1.7	Melhorias no Protótipo e Testes com Outros Modelos de <i>Talking Heads</i>	68
APÊNDICE A	LISTA DOS FAPS	69
APÊNDICE B	SINTAXE DA GDL	73
B.1	<i>Storyboard</i> de Ações	73
B.2	Comportamentos do Olhar	73
B.3	Descrição das Expressões	74
APÊNDICE C	SCRIPTS DAS ANIMAÇÕES AVALIADAS	77
C.1	Animações "A"	77
C.2	Animações "B"	78
C.3	Animações "C"	79
C.4	Animações "D"	79
REFERÊNCIAS		81

LISTA DE ABREVIATURAS E SIGLAS

AU	<i>Action Unit</i>
CG	Computação Gráfica
ECA	<i>Embodied Conversational Agent</i>
FACS	<i>Facial Action Code System</i>
FAP	<i>Facial Animation Parameter</i>
FAPU	<i>Facial Animation Parameter Units</i>
FP	<i>Feature Point</i>
GDL	<i>Gaze Description Language</i>
MPEG	<i>Moving Picture Experts Group</i>
OSG	<i>Open Scene Graph</i>

LISTA DE FIGURAS

Figura 2.1:	Pontos de controle do padrão MPEG-4.	16
Figura 2.2:	Descrição dos MPEG-4 FAPU.	17
Figura 2.3:	Exemplo de um arquivo FAP, mostrando a sua estrutura.	17
Figura 2.4:	Corte mostrando a estrutura do olho humano.	18
Figura 2.5:	Cinemática típica da cabeça/olhos/olhar.	19
Figura 2.6:	Representação de termos emocionais.	26
Figura 3.1:	Modelo de Lee <i>et al.</i>	31
Figura 3.2:	Modelo de Poggi <i>et al.</i>	32
Figura 3.3:	Modelo de Gu e Badler	33
Figura 4.1:	Diagrama da arquitetura do modelo.	45
Figura 5.1:	Diagrama de classes do módulo Gerador de Olhares Expressivos.	48
Figura 5.2:	Diagrama de classes do módulo Modelo Padrão.	49
Figura 5.3:	Ambiente experimental montado para os testes.	50
Figura 5.4:	Olhos de Alice: original e modificados.	51
Figura 5.5:	Problemas na animação dos olhos de Alice.	52
Figura 5.6:	<i>Screenshot</i> da execução de uma animação utilizando a interface provida pelo OSGXface	53
Figura 5.7:	<i>Screenshot</i> do protótipo desenvolvido.	53
Figura 6.1:	<i>Snapshots</i> da simulação dos Diálogo 1.	56
Figura 6.2:	Resultados da avaliação subjetiva para os diálogos 1 e 2	56
Figura 6.3:	<i>Snapshots</i> de animação rodando o Modelo Padrão.	57
Figura 6.4:	<i>Snapshots</i> de animação rodando com olhares direcionados (anotados), sem expressão facial.	57
Figura 6.5:	<i>Snapshots</i> de animação rodando com expressão anotada.	58
Figura 6.6:	<i>Snapshots</i> das animações com os comportamentos propostos	58
Figura 6.7:	Resultados da avaliação subjetiva da Parte 1 do questionário.	62
Figura 6.8:	Resultados da avaliação subjetiva para da Parte 1 do questionário: comparativo.	63
Figura 6.9:	Resultados da avaliação subjetiva para as quatro questões da Parte 2 do questionário.	63
Figura 6.10:	Resultados da avaliação subjetiva para as quatro questões da Parte 3 do questionário.	64

LISTA DE TABELAS

2.1	Comportamentos do olhar relacionados a atividades cognitivas.	22
2.2	Comportamentos do olhar relacionados a troca de turnos em uma conversação	23
2.3	Comportamentos do olhar relacionados com algumas expressões emocionais.	23
2.4	Emoções e significados conforme o modelo OCC, por Elliot. . .	27
4.1	Parâmetros observados nas cenas selecionadas.	37
4.2	Exemplo de parte da tabela de observação do filme Era do Gelo	38
4.3	Distribuição das direções das sacadas.	43
6.1	Diálogo simulado: informação anotada no arquivo de configuração.	55
6.2	Informações técnicas sobre os trechos dos filmes.	58
A.1	Definições dos FAPs.	69
A.2	Grupos dos FAPs.	72

RESUMO

Os olhos desempenham um importante papel na comunicação entre os seres humanos. A movimentação dos olhos expressa emoções e regula o fluxo de conversação. Em vista disso, considera-se fundamental que humanos e outros seres virtuais em aplicações como agentes conversacionais, jogos e filmes, apresentem aspectos convincentes em seu olhar. Percebe-se, no entanto, que para aplicações que requerem a geração automática de movimentos faciais, como os agentes conversacionais, a movimentação dos olhos ainda é muitas vezes desprovida de significado, principalmente quando relacionada à expressividade do agente. Este trabalho propõe um modelo de geração automática de olhares expressivos, levando em consideração os comportamentos dos olhos em diferentes estados afetivos. Para a coleta de dados referentes ao olhar e expressividade, foram feitas observações em filmes de Computação Gráfica. Esses dados serviram de base para a descrição das expressões no modelo proposto. Foi construído um protótipo com a implementação do modelo e realizados alguns testes com voluntários, a fim de se observar o impacto do comportamento do olhar em algumas expressões emocionais. Os resultados obtidos mostram que o modelo é capaz de gerar movimentações dos olhos coerentes com os estados afetivos do personagem virtual.

Palavras-chave: Animação comportamental, animação de olhos, expressividade.

Automatic Generation of Expressive Gaze in Virtual Humans

ABSTRACT

Eyes play an important role in communication among people. The eye motion expresses emotions and regulates the flow of conversation. Therefore, we consider fundamental that virtual humans or other characters, in applications such as Embodied Conversational Agents (ECAs), games and movies, present convincing aspects in their gaze. However, we perceive that in applications that require automatic generation of the facial movements, such as ECAs, the eye motion does not have meaning, in many times, mainly when gaze is related to agent's expressiveness. This work proposes a model for automatic generation of expressive gaze, considering the eye behavior in different affective states. To collect data related to gaze and expressiveness, we made observations in Computer Graphics movies. These data were used as basis for expressions description in the proposed model. We also implemented a prototype and some tests was performed with users, in order to observe the impact of eye behavior in some emotional expressions. The results we obtained show that the model is capable to generate eye motion coherent with the affective states of virtual character.

Keywords: behavioral animation, eye animation, expressiveness.

1 INTRODUÇÃO

“Os olhos são a janela da alma e o espelho do mundo.”
—LEONARDO DA VINCI

A citação acima, atribuída ao grande inventor e artista Leonardo Da Vinci, traz em si por completo a justificativa para o desenvolvimento deste trabalho. Os olhos representam um canal de comunicação que envia muitos sinais e reage aos vários estímulos que recebe. O movimento dos olhos possui um papel fundamental na comunicação não-verbal entre os seres humanos, pois revela muito mais do que o que é expresso em um diálogo. A movimentação dos olhos expressa emoções e regula o fluxo de conversação. Além do mais, a direção dos olhos indica o foco atencional do indivíduo. Diante disso, nota-se a importância de modelos animados de humanos e de outros seres virtuais apresentarem comportamentos convincentes em seu modo de olhar, de maneira a expressar corretamente os sinais desejados.

Sabe-se, nos dias de hoje, que a geração de comportamentos de olhar em ECAs (*Embodied Conversational Agents*) mostra-se problemático no sentido de se produzir resultados que expressem convincentemente seu estado interno e suas intenções de comunicação. Vários trabalhos [CASSELL; TORRES; PREVOST, 1997; COHEN; COLBURN; DRUCKER, 2000; FUKAYAMA et al., 2002; LEE; BADLER; BADLER, 2002] estudam maneiras de gerar comportamentos de olhar coerentes com aspectos conversacionais e alguns outros [POGGI; PELACHAUD; ROSIS, 2000; GU; BADLER, 2006] incluem também aspectos cognitivos e perceptivos. Poucos trabalhos, como o de Poggi *et al.* [POGGI; PELACHAUD; ROSIS, 2000], procuram relacionar o comportamento do olhar e o estado afetivo do agente. Seu enfoque, no entanto, fica na expressão facial (sobrancelhas, pálpebras, rugas) e na direção do olhar durante determinado estado afetivo, e não na maneira como os olhos se movimentam.

Em contrapartida, observa-se em bons filmes de Computação Gráfica (CG), que a movimentação dos olhos dos personagens é extremamente convincente, expressando claramente o que os artistas querem mostrar. Neste tipo de animação, cada sinal é artisticamente produzido pelo animador para conduzir o significado desejado ao público. Esta não é uma tarefa trivial¹ e os animadores ajustam manualmente cada olhar até alcançarem o resultado desejado.

Nesse contexto, este trabalho busca explorar os principais elementos que produzem a expressividade e interatividade do olhar humano e introduzir essa abordagem comportamental a um modelo de animação de olhos já existente [LEE; BADLER; BADLER, 2002],

¹O *Making of* de A Era do Gelo, Blue Sky (2002) mostra o esforço dos animadores para fazer com que seus personagens alcançassem as expressões e olhares extremamente convincentes.

para a geração automática e coerente do movimento dos olhos em aplicações interativas, como ECAs e Jogos.

Mais especificamente, são objetivos deste trabalho:

- Mapear olhares expressivos às principais emoções e sinais comunicativos que os seres humanos emitem. Os dados para esse estudo foram coletados na literatura e na análise de filmes de CG;
- Propor um modelo de animação de olhos que integre a abordagem comportamental estudada aos movimentos realísticos de olhos providos pelo modelo estatístico de Lee *et al.* [LEE; BADLER; BADLER, 2002];
- Implementar um protótipo para a visualização do modelo. Esse protótipo utiliza as ferramentas de animação e visualização 3D de faces sintéticas Xface [BALCI, 2004] no ambiente gráfico Open Scene Graph (OSG) [BURNS; OSFIELD, 2007], o que o torna uma ferramenta portátil e extensível para diversas plataformas e aplicações;
- Realizar testes com usuários e avaliar qualitativamente os resultados obtidos.

De forma diferente dos trabalhos estudados, este trabalho explora o comportamento do olhar não somente a nível de expressão facial e direção do olhar, mas procura modelar diferentes tipos de movimentos dos olhos como comportamentos, que ocorrem de acordo com o estado interno do personagem, assim como em outros sinais comunicativos. São também levados em consideração alguns detalhes que contribuem para a geração de olhares mais convincentes, como a associação dos movimentos dos olhos, pálpebras e cabeça.

O modelo proposto é baseado principalmente na hipótese de que filmes de CG podem prover uma boa parametrização da movimentação dos olhos em função do estado afetivo do personagem. Isto poderia, obviamente, ser observado em humanos reais; entretanto, acredita-se que os filmes de computação gráfica sejam mais úteis no sentido de que a linguagem utilizada é considerada mais adequada para as aplicações de interesse. Para fazer isso, seguiu-se uma metodologia inspirada na proposta do trabalho de Lance *et al.* [LANCE; STACY; KOIZUMI, 2004], trabalho cujo objetivo assemelha-se ao deste. Entretanto, a metodologia seguida difere consideravelmente da metodologia de Lance, pois a intenção neste trabalho não é gerar movimentos de olhos precisos através de medições cuidadosas, mas identificar comportamentos e codificá-los em um sistema computacional de maneira a melhorar a empatia dos personagens. Assim, este estudo fornece uma base para a melhoria na qualidade da expressividade de personagens com comportamentos gerados automaticamente, sem a intervenção dos artistas.

Esta monografia está organizada da seguinte forma: o próximo capítulo apresenta alguns fundamentos sobre a movimentação dos olhos nos contextos físico, comportamental e computacional. Em seguida, é apresentado o estado-da-arte da animação de olhos, com enfoque nos trabalhos que estão diretamente relacionados a este. O Capítulo 4 apresenta a metodologia adotada para a coleta de dados e o modelo proposto. Detalhes sobre a implementação do protótipo que valida o modelo são apresentados no Capítulo 5 e os resultados obtidos, no Capítulo 6. Por fim, as contribuições deste trabalho são apresentadas e discutidas no Capítulo 7.

2 FUNDAMENTOS

“The different branches of Arithmetic - Ambition, Distraction, Uglification, and Derision.”

— LEWIS CARROLL, em *“Alice’s Adventures in Wonderland”*

Neste capítulo são apresentados alguns conceitos de animação comportamental e animação facial, a fim de contextualizar o trabalho nas áreas da Computação Gráfica em que está inserido. Em seguida, são apresentadas informações sobre a movimentação de olhos humanos, sob o ponto de vista biológico e comportamental (psicológico e social), relevantes para a construção de um modelo sintético de animação.

2.1 Animação Comportamental

De maneira geral, em *softwares* de animação, os animadores precisam descrever detalhadamente cada movimento de seus personagens, tendo no máximo a ajuda de algum interpolador entre os quadros-chave (*keyframes*). Não basta apenas indicar ao personagem para que ele aja “feliz” ou que “vá para lugar X”, mas é necessário especificar cada movimento que expressa a felicidade no ator ou cada passo e trejeitos de seu percurso até X. Na animação comportamental, a abordagem utilizada é diferente ao dessas ferramentas de animação convencionais: procura-se modelar o *comportamento* do personagem, e não somente suas propriedades físicas. Neste caso, o foco principal são suas ações e os estímulos para estas, além de sua movimentação em particular. Os comportamentos podem incluir desde atividades de um simples trajeto planejado até interações emocionais complexas entre os personagens [REYNOLDS, 1987].

Segundo Reynolds, na animação comportamental, um personagem autônomo determina suas próprias ações, pelo menos até um certo nível. Isso dá ao personagem alguma habilidade de improvisar, e libera o animador de precisar especificar cada detalhe da movimentação de cada personagem. De acordo com Musse [MUSSE, 2001], o principal objetivo da animação comportamental é prover “autonomia” aos atores virtuais, de maneira que eles possam “evoluir” num determinado ambiente, sem a intervenção do usuário. Trata-se de uma área que utiliza com frequência técnicas de Inteligência Artificial [THÓRISSON et al., 2004], desde algoritmos clássicos e sistemas de regras até modelos avançados do campo de *Vida Artificial* [TERZOPOULOS, 1999; MAES, 1995] e Computação Afetiva [PICARD, 1997].

Este trabalho contextualiza-se na área de animação comportamental pelo fato de que o modelo construído provê um conjunto de ações (olhares) que são descritas em alto nível

(por exemplo, “olhar para a direita”, ou “olhar triste”) e são executadas automaticamente, sem a necessidade de especificar detalhes sobre cada movimento. Apesar do modelo não fornecer o estado interno do agente, foram necessários estudos relativos a síntese de emoções em agentes virtuais [ELLIOTT, 1992; PICARD, 1997] para a inclusão do conjunto de ações que o modelo contempla.

2.2 Animação Facial

Pesquisas em animação facial são desenvolvidas desde a década de 70 e têm alcançado vários resultados, incluindo a definição de técnicas para a execução de expressões faciais em modelos de faces sintéticas. Estudos relevantes de relações entre os músculos e expressões faciais são vistos no sistema FACS (*Facial Action Code System*), desenvolvido por Ekman e Friesen [EKMAN; FRIESEN, 1978]. Este modelo codifica as expressões faciais humanas, através da determinação dos conjuntos de músculos responsáveis pelas expressões. O sistema FACS define unidades de ação (*Action Units* ou AUs), representadas pela contração e relaxamento dos músculos (conjuntos de músculos responsáveis pelas mudanças faciais). O sistema passou por algumas revisões, tendo uma nova edição lançada em 2002 [EKMAN; FRIESEN; HAGER, 2002] e continua sendo a melhor base para a parametrização de expressões em baixo-nível. Porém, do ponto de vista do animador, a base parametrização fornecida é um tanto abstrata, segundo [PARADISO, 2002].

Em 1999, foi lançado o padrão MPEG-4¹ de Animação Facial, que procura padronizar os parâmetros de deformação facial para a comunicação e integração de conteúdos de multimídia [OSTERMANN, 1998; PARADISO, 2002]. O desenvolvimento do modelo de animação facial MPEG-4, com a introdução dos FAPs (*Facial Animation Parameters*), permite que os desenvolvedores de aplicações entrem com dados para a animação (expressões, visemas e movimentos de músculos específicos da face) e então a interpretem e tratem as deformações do modelo da sua maneira. Entretanto, ainda assim a utilização das diretivas MPEG-4 pode ser considerada como de baixo nível para os animadores, como mostrado na subseção seguinte.

No contexto de animação facial, a busca por padrões de comportamento é bastante vasta, como será visto no decorrer deste trabalho. A animação comportamental em faces é uma abordagem que se justifica facilmente, pelo fato da face ser o centro da comunicação humana. Por esse motivo, ela exerce atividade intensa e cada movimento expressa as emoções e percepções sensoriais, que são baseadas diretamente em padrões comportamentais do indivíduo.

No que diz respeito à síntese da animação, este trabalho segue o padrão MPEG-4 de animação facial. Este padrão trabalha com algumas definições importantes dos parâmetros, que serão vistas a seguir.

2.2.1 Padrão MPEG-4 de Animação Facial

O padrão MPEG-4 de Animação Facial descreve os passos para a criação de um agente comunicativo através da definição dos vários parâmetros necessários de uma maneira padronizada. Pode-se distinguir duas fases principais para a criação de um agente comunicativo: *i*) a escolha de pontos de controle no modelo 3D que definem as regiões de deformação da face (*feature points*); e *ii*) a geração e interpretação dos parâmetros que vão modificar esses pontos de maneira a criar as animações. MPEG-4 abstrai estes dois

¹Moving Picture Experts Group

passos de uma maneira padronizada, dando liberdade aos desenvolvedores de aplicações para focar em seu campo de trabalho.

Para a criação de uma face padronizada, MPEG-4 define 84 pontos de controle (*feature points* ou FPs) localizados na cabeça do modelo. Esses pontos são usados para a definição dos parâmetros de animação, assim como para calibrar os modelos quando há alguma troca da plataforma de visualização. A Figura 2.1 mostra o conjunto de FPs.

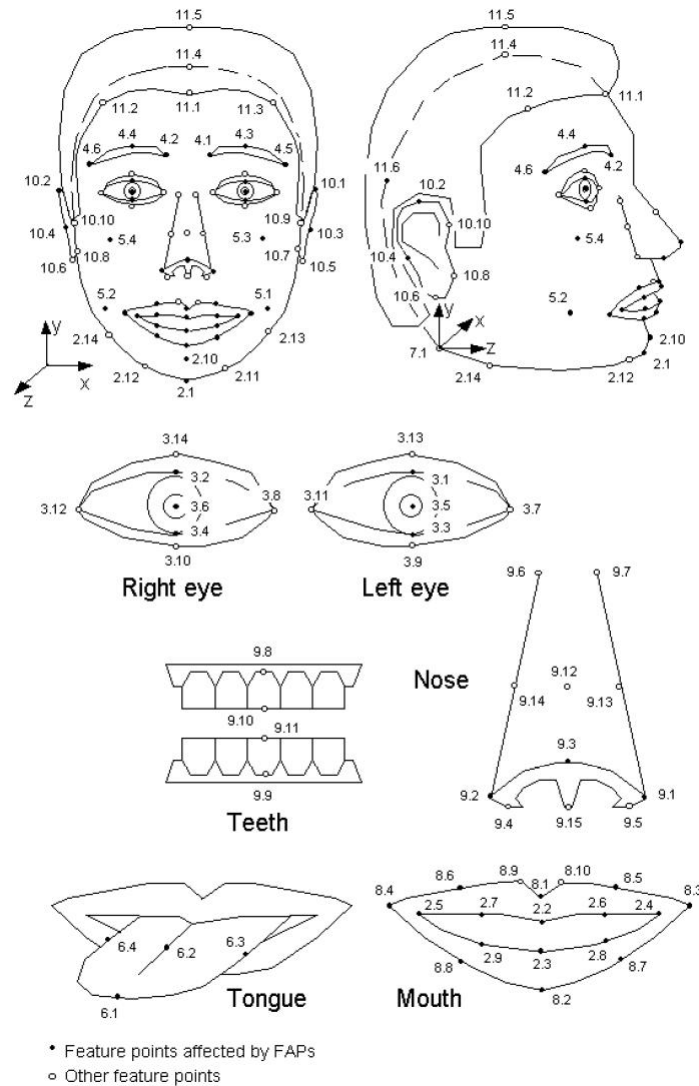


Figura 2.1: Pontos de controle do padrão MPEG-4 [BALCI, 2004].

Para a animação da face, são definidos 68 parâmetros de animação facial (*Facial Animation Parameters* ou FAPs). Os dois primeiros são parâmetros que representam, em alto nível, visemas e expressões faciais. Visemas são a parte visual dos fonemas na fala e as expressões faciais consistem nas 6 emoções básicas de Ekman [EKMAN, 1992] (raiva, alegria, tristeza, surpresa, desgosto e medo) como protótipos. Utilizando-se apenas estes dois FAPs em um modelo de face já é possível obter resultados satisfatórios, junto com a interpolação linear entre cada um desses protótipos [BALCI, 2004]. Entretanto, para uma melhor qualidade nas animações, recomenda-se utilizar os FAPs restantes, que tratam de regiões específicas da face, como o “canto do lábio esquerdo” ou “parte média da sobrancelha direita”, por exemplo. Cada FAP atua sobre um FP e define as defor-

2.3 Os Movimentos dos Olhos

Os olhos humanos movimentam-se intensamente pelo fato que, ao contrário das câmeras de vídeo convencionais, que amostram uniformemente o mundo visual, eles somente concentram as informações visuais detalhadas em uma pequena região central do campo de visão, chamada *fovea*. Por este motivo, o ser humano é dotado de olhos extremamente móveis que podem quase que instantaneamente apontar a *fovea* para qualquer local do mundo visual [ITTI; DHAVALÉ; PIGHIN, 2004].

A Figura 2.4 apresenta, em um corte do globo ocular humano², os principais elementos que o formam. Para os fins deste trabalho, é importante apenas a identificação da íris, pupila e a região da fóvea, já que é por causa dela que é necessário “apontar” os olhos para o alvo de interesse.

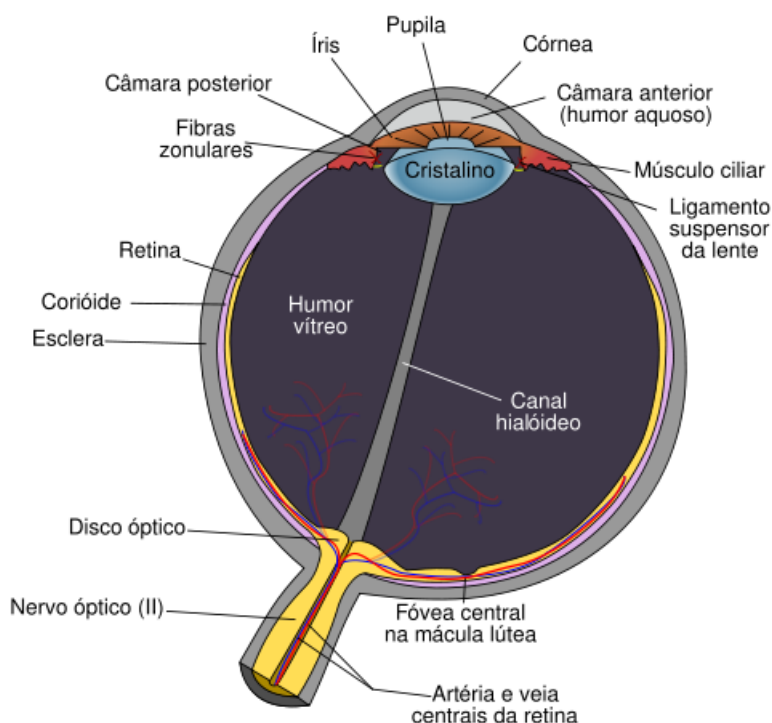


Figura 2.4: Corte mostrando a estrutura do olho humano. Origem: Wikipédia, a enciclopédia livre.

O trabalho de Lee [LEE, 2002] indentifica quatro tipos de movimentos de olhos:

- **Movimentos sacádicos:** a região *foveal*, que tem a abertura de aproximadamente 2° do campo de visão, é especializada para uma visão de alta precisão e colorida. Para ver um objeto claramente, os movimentos de deslocamento de olhar direcionam o olho para o alvo. Esses movimentos são os chamados sacadas visuais, ou simplesmente sacadas. São movimentos muito rápidos, com controle voluntário. As sacadas visuais serão detalhadas adiante, pois constituem o tipo de movimento ocular que será simulado no modelo.

²Disponível em http://pt.wikipedia.org/wiki/Olho_humano

- **Movimentos de retenção do olhar:** Pelo fato que o olhar é a soma da posição da cabeça com a posição do olho, estes movimentos dos olhos compensam os movimentos da cabeça e corpo, a fim de aumentar a estabilidade da imagem retinal. Esses movimentos compreendem os *reflexos optocinéticos*, que são dirigidos pelo movimento da imagem retinal (conhecido como fluxo ótico) e os *reflexos vestibulo-oculares*, que são dirigidos pelos órgãos de equilíbrio no ouvido interno.
- **Movimentos de fixação:** Mesmo quando fixos em um objeto parado, os olhos fazem contínuos micro-movimentos de três tipos: lentas trações, rápidos tremores de pequena amplitude e micro-sacadas que recuperam o olhar quando a pequena tração tiver movido o olho para muito longe do alvo. Quando o objeto-alvo realiza um pequeno deslocamento, os dois olhos devem convergir para o alvo. Isso se chama movimento de convergência.
- **Movimentos de perseguição:** Ocorrem quando os olhos seguem um objeto em movimento, voluntária ou involuntariamente. Este movimento é um pouco diferente das sacadas: são mais suaves, mais lentos e com menor intervalo entre os movimentos, em reação ao deslocamento do objeto.

Visto que a cabeça tem uma massa significativamente maior que a dos olhos, a dinâmica da cabeça é muito mais ociosa que a dinâmica dos olhos. De acordo com [CARPENTER, 1988], quando um sujeito voluntariamente rotaciona bastante a cabeça e os olhos para um alvo no plano horizontal, o movimento dos olhos consiste em uma sacada inicial seguindo a direção do movimento da cabeça, seguido de um retorno devagar para o centro orbital. Durante a aquisição do alvo, a velocidade da cabeça é normalmente correlacionada com a amplitude do deslocamento do alvo visual.

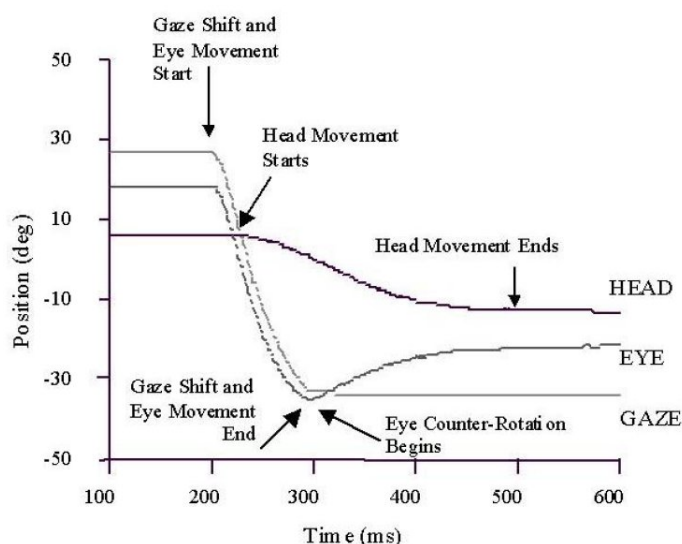


Figura 2.5: Cinemática típica da cabeça/olhos/olhar [TERZOPOULOS; LEE, 2004].

Segundo [TERZOPOULOS; LEE, 2004], o *olhar* é a direção do olho no espaço. Entretanto, neste trabalho, o termo *olhar* é usado de maneira mais abrangente, referindo-se à movimentação dos olhos como um todo. A cinemática típica da cabeça/olhos/olhar (direção) está representada na Figura 2.5. As posições da cabeça, olhos e olhar são plotadas

como funções do tempo durante um deslocamento de olhar de 60° , composto dos movimentos coordenados de olhos e cabeça. A cabeça contribui aproximadamente com 10° da mudança total do deslocamento do olho durante este movimento. Os 50° restantes são executados pelo movimento sacádico do olho.

2.3.1 As Sacadas Visuais

Sacadas são movimentos rápidos de ambos os olhos de uma posição de olhar para outra [LEIGH; ZEE, 1991]. São os únicos movimentos de olhos que podem ser executados prontamente, conscientemente e voluntariamente por seres humanos. Há poucas convenções usadas na literatura sobre movimentos de olhos que descrevem as sacadas. O trabalho de Lee [LEE; BADLER; BADLER, 2002] relaciona quatro características principais que podem ser medidas em uma sacada:

- **Magnitude:** a magnitude de uma sacada é a medida, em graus, da extensão do movimento sacádico. Trata-se do ângulo através do qual o globo ocular rotaciona quando ele muda de uma posição para outra.
- **Direção:** a direção da sacada define o seu eixo 2D de rotação. Por exemplo, uma sacada com magnitude 10° e direção de 45° é equivalente a um globo ocular rotacionando 10° na direção direita-cima.
- **Duração:** é o tempo que o movimento leva em sua execução, da mudança de uma posição para outra.
- **Intervalo inter-sacádico:** é a quantidade de tempo que transcorre entre o término de uma sacada e o início do próxima (tempo de fixação do olhar).

As métricas (características espaço-temporais) das sacadas já foram bem estudadas. Um movimento normal de sacada começa com uma alta aceleração inicial ($30.000^\circ/s^2$) e termina com uma rápida desaceleração. A velocidade de pico para grandes sacadas está entre 400 e 600 $^\circ/s$. O tempo de reação sacádico é entre 180 e 220 ms, em média. O intervalo inter-sacádico mínimo varia de 50 a 100ms.

A duração e velocidade de uma sacada são funções de sua magnitude. Para sacadas entre 5° e 50° , a duração tem uma taxa de aumento constante com a magnitude, que pode ser aproximada por uma função linear (a Subseção 4.1.3 do Capítulo 4 fornece as fórmulas de cálculo desses parâmetros). Sacadas com grandes deslocamentos sempre incluem uma rotação de cabeça, sob condições naturais, conforme mostrado na figura 2.5. No entanto, as sacadas raramente têm uma magnitude maior que 15° . De acordo com a bibliografia [CHOPRA-KHULLAR; BADLER, 1999], deslocamentos horizontais maiores que 25° ou deslocamentos verticais maiores que 10° produzem movimentos combinados com a cabeça. As pálpebras também acompanham os olhos durante os deslocamentos.

2.4 O Comportamento do Olhar

O trabalho de Lee [LEE, 2002] apresenta quatro funções para o olhar (cada uma seguida de exemplos ilustrativos):

- **Envio de sinais sociais:**
 - Quem está falando tende a enfatizar palavras, frases ou expressões inteiras.

- Quem está ouvindo tende a sinalizar a atenção ou o interesse em algum ponto particular da conversa. Em caso de reprovação, ele pode ainda evitar ou desviar o olhar.

- **Abertura de um canal para recebimento de informações:**

- Quem está falando irá olhar para os ouvintes, durante pausas na fala, para julgar como suas palavras estão sendo recebidas e se os ouvintes demonstram interesse que ele continue.
- Os ouvintes continuamente monitoram as expressões e direção do olhar da pessoa que está falando.

- **Regular o fluxo da conversação:**

- Quando está parando de falar, normalmente a pessoa lança um olhar para o ouvinte, indicando o fim de seu turno de conversa.
- Os participantes da conversa podem olhar para um dos ouvintes para sugerir que ele seja o próximo a falar.

- **Expressar emoções:**

- As pessoas tendem a olhar para baixo quando estão tristes, tímidas ou envergonhadas.
- Quando uma pessoa fica surpresa, as pálpebras se mantêm bem abertas por algum tempo.

Alguns dos fatores que interferem no comportamento do olhar durante uma conversação são [COHEN; COLBURN; DRUCKER, 2000]:

- **Idade:** mais longo (tempo de fixação do olhar) em adultos do que em crianças;
- **Sexo:** mais longo em mulheres que em homens;
- **Distância entre os conversantes:** olhar mútuo é mais longo quanto mais próximos os conversantes estão;
- **Cultura:** por exemplo, a postura do olhar dos povos ocidentais e orientais, de acordo com crenças e regras disciplinares.

Em relação a tarefas, as pesquisas categorizam os padrões de movimento dos olhos em quatro diferentes tipos [LEE, 2002]:

- **Olhar espontâneo:** o olhar espontâneo é a percepção de uma cena sem uma tarefa específica em mente. O padrão do olhar durante um olhar espontâneo é direcionado às partes da imagem que contenham mais informação.
- **Olhar relevante à tarefa:** este tipo de olhar resulta da observação de uma cena com uma tarefa específica em mente. O padrão do olhar em uma cena particular é correlacionado à questão que está sendo analisada.
- **Orientação do pensamento:** é o padrão de olhar de quando a pessoa não está dando atenção para o que seus olhos estão olhando em virtude dos seus pensamentos (a Subseção 2.4.1 apresenta uma relação destes padrões de olhar e atividades).

- **Olhar intencional de manipulação:** resulta do movimento dos olhos para uma localização específica ou de uma maneira específica, com a intenção de manipular alguém através da sua movimentação do olhar (como por exemplo, a sinalização do fim do turno de uma conversa, como mostra a Subseção 2.4.2).

2.4.1 Padrões de Olhar em Processos Cognitivos

A programação neurolingüística afirma que a direção do movimento do olho é uma reflexão de atividades cognitivas. Nesta teoria, a hipótese é de que existe uma associação entre a posição do olho com diferente tipo de de processamento de informações. A Tabela 2.1 relaciona esses tipos de padrão. [GU; BADLER, 2006; BROOKER, 2007]

Tabela 2.1: Comportamentos do olhar relacionados a atividades cognitivas [GU; BADLER, 2006; BROOKER, 2007].

Processamento de Informação	Movimento dos Olhos
Construção visual: ocorre quando tenta-se visualizar mentalmente um evento que nunca foi visto antes	deslocamento direita-cima
Lembrança visual: ocorre quando um evento já visto é lembrado	deslocamento esquerda-cima
Construção de áudio: ocorre na elaboração de uma nova frase	deslocamento para direita
Lembrança de áudio: ocorre quando se está lembrando de algo que foi dito	deslocamento para a esquerda
Cinestético: solucionando, ou avaliando as sensações do corpo (relacionado a sentimentos e emocional)	deslocamento direita-baixo
Diálogo Interno: quando há conversação interna	deslocamento esquerda-baixo
Visualização interna: quando há a visualização interna de um evento real ou imaginário	olhar sem um foco específico, em um ponto fixo no espaço

2.4.2 Troca de Turnos na Conversação

Comportamentos de olhar sinalizam e monitoram a inicialização, manutenção e término de mensagens comunicativas. Contatos visuais ocorrem quando os participantes querem sinalizar o desejo de interagir com os outros. Tipicamente, o olhar mútuo (o encontro do olhar entre duas pessoas) é curto, de aproximadamente 1 segundo. O olhar “diminui” quando há a desaprovação do contato social (diminui a frequência em que o participante que desaprova olha para o outro) [GU; BADLER, 2006].

Uma conversação ocorre em turnos. O participante pode encontrar-se em dois estados mutuamente exclusivos: ouvinte ou interlocutor³. O olhar provê sinais para regular o fluxo da comunicação, como já escrito anteriormente [GU; BADLER, 2006]. A tabela 2.2 mostra como comportamentos de olhar atuam para manter e controlar os turnos de conversação entre múltiplos participantes:

³Neste trabalho, o termo interlocutor é utilizado para designar quem está falando

Tabela 2.2: Comportamentos do olhar relacionados a troca de turnos em uma conversação [GU; BADLER, 2006].

Estado	Sinal	Comportamento do Olhar
Interlocutor	Ceder o turno	Olhar direcionado para um ouvinte específico
Interlocutor	Reinvindicar o fim do turno	Desviar o olhar dos ouvintes
Interlocutor	Continuar com o turno	Olhar direcionado para os ouvintes em geral
Interlocutor	Nenhum sinal	Olhar não direcionado (<i>away</i>)
Ouvinte	Retornar o sinal	Olhar direcionado para o interlocutor
Ouvinte	Reinvindicar o turno	Procurar contato com o olhar do interlocutor
Ouvinte	Fim de turno	Desviar do contato de olhar do interlocutor
Ouvinte	Reinvindicar o fim do turno	Olhar direcionado para os outros ouvintes para prevenir que eles falem
Ouvinte	Sem resposta	Aleatório

Em diálogos, ao final de um pronunciamento, o interlocutor dá uma olhada de relance, para o ouvinte para ceder o turno. Esta indicação de olhar repete-se até que o ouvinte assuma o papel de interlocutor. [GU; BADLER, 2006]

2.4.3 Olhar e Expressividade

Existem poucos trabalhos que relacionam o olhar com expressividade e emoções, como será visto no próximo capítulo. O trabalho de Gu e Badler [GU; BADLER, 2006] fornece uma tabela com alguns comportamentos encontrados na literatura sobre a movimentação dos olhos em algumas expressões faciais relacionados com emoções. A Tabela 2.3 reproduz estes dados, que estão descritos de forma bastante subjetiva:

Tabela 2.3: Comportamentos do olhar relacionados com algumas expressões emocionais [GU; BADLER, 2006].

Tipo	Descrição do Comportamento	Movimento dos Olhos
“Risada”	Riso submisso: apreensão em volta dos olhos	Olhar abatido e diminuição do contato visual
“Risada”	Sorriso: relaxado, dentes juntos, lábios um pouco afastados	Olhar positivo e seguro
“Risada”	Risada: dentes frequentemente afastados, parcialmente cobertos pelos lábios	Olhar levantado e sem foco, grande abertura dos olhos

Surpresa	Repentina abertura dos olhos, seguido por um misto de emoções: satisfação, raiva, estado de choque	Olhar fixo e para cima, no misto de emoções
Medo	Semelhante à Surpresa	Desvio do olhar
Raiva	Olhos bem abertos e fixos	Fixação
Interesse	Olhos bem abertos (se o objeto está próximo) ou com as pálpebras fechando um pouco os olhos, ajustando-se ao foco (em grandes distâncias), fixos no objeto	Fixação, examinando com olhadas com grande intervalo de fixação
Desprezo	Olhos um pouco fechados, rugas entre os olhos, fixados em objetos insignificantes	Olhares para os lados
Desgosto	Pálpebras superiores podem estar parcialmente fechadas, ou um pouco abertas em um lado	Desvio de olhar

Neste trabalho, é utilizado o termo *olhar expressivo* para denotar a junção da expressão facial (posição da cabeça, sobrancelhas e pálpebras) com o comportamento do olhar (movimentação do globo ocular).

2.5 Emoção e Estado Afetivo

Para desenvolver uma face animada sintética com o comportamento baseado em emoções, é necessário: *i*) ter um modelo emocional que determine a emoção que o personagem está sentindo no momento e *ii*) fazer a associação das emoções às expressões faciais correspondentes, determinando os parâmetros de animação facial envolvidos. Este trabalho considera apenas o segundo aspecto, ou seja, o mapeamento de emoções para as expressões do olhar. Para isso, parte-se do princípio de que este trabalho deverá ser integrado a um sistema que forneça as informações sobre a emoção do personagem a cada instante de tempo. Entretanto, para a idealização do modelo, foi necessária uma pesquisa sobre o conceito de emoção, sua classificação e utilização em modelos emocionais de computação, a fim de selecionar as emoções (ou estados afetivos) que podem ser fornecidos como entrada para a geração de olhares expressivos.

Existem diversas linhas de pensamento e conceitos para emoções. O termo *emoção* popularmente é usado para muitos fenômenos de ordem afetiva que, na literatura, conceitualmente não constituem emoções. Para os fins científicos, como o que é objeto de exploração deste trabalho, o termo genérico mais adequado é *estado afetivo* [JACQUES; VICCARI, 2005]. Esse é um termo mais abrangente, o qual inclui outros estados além das emoções, como por exemplo, o humor.

De acordo com Scherer [SCHERER, 2000], emoção é um episódio relativamente breve de resposta sincronizada para a maioria ou todos os seres vivos, para a avaliação de um evento interno ou externo, como sendo de maior importância. Alguns exemplos de emoções são: raiva, tristeza, alegria, medo, desespero e vergonha. Ortony, Clore e Collins

[ORTONY; CLORE; COLLINS, 1988] propõem uma definição similar, porém mais precisa, para emoções. De acordo com eles, emoções são reações com valência (avaliação positiva ou negativa) a eventos, agentes ou objetos, cuja natureza particular é determinada pela maneira que a situação disparadora é construída. De acordo com esta definição, surpresa não é uma emoção, pois ela não tem uma valência [JACQUES; VICCARI, 2005].

Um outro tipo de estado afetivo é o humor. Ele é um estado afetivo difuso, que consiste em mudanças no sentimento subjetivo, possui baixa intensidade, mas longa duração, e não tem causa aparente [SCHERER, 2000]. Alguns exemplos de humor são: irritação, depressão, chateação e nervosismo. Fridja [FRIJDA, 1994] considera que o humor difere das emoções mais fortemente em não possuir um objeto intencional (não é direcionado a um objeto). Sua causa é tipicamente conceitual ou avaliativa (por exemplo, “as coisas não estão indo muito bem”). Como pode ser observado, cientistas tentam diferenciar emoções de outros estados afetivos; entretanto, não há uma definição única e clara para emoção [JACQUES; VICCARI, 2005]. Existe, no entanto, um consenso acerca de três vertentes que constituem a expressão emocional, segundo [MORGADO, 2006]:

- **Resposta fisiológica:** corresponde a alterações diversas no organismo, produzindo sensações como aceleração do ritmo cardíaco ou tremores;
- **Expressão comportamental:** corresponde a alterações na atividade motora e na expressão corporal, como alteração da expressão facial ou do tom de voz;
- **Experiência subjetiva:** corresponde à percepção consciente da experiência emocional, que pode ser caracterizada através de termos emocionais, como alegria, raiva ou medo.

Uma das formas de caracterização da emoção é através da identificação de *dimensões* de classificação emocional. A principal dimensão de classificação emocional está relacionada com a experiência (subjetiva) de prazer ou de dor, agradável ou desagradável ou, como é designada normalmente na psicologia, valência (atribuição de valor) positiva ou negativa, a que corresponde a noção de afeto (positivo ou negativo) [MORGADO, 2006].

Ainda há outras dimensões de classificação emocional que têm sido propostas, como é o caso de excitação/depressão, ou tensão/relaxamento. Vários estudos têm sido feitos a fim de identificar e relacionar dimensões relevantes. Esses estudos já identificaram, de modo consistente, duas dimensões: valência (positiva/negativa) e excitação/depressão, mais recentemente descrita como ativação/passividade. Com essas duas dimensões, é possível relacionar um conjunto de *termos emocionais*, localizados em regiões específicas do espaço por elas formado, ilustrado na Figura 2.6 [MORGADO, 2006].

É importante notar que o fato de ser possível fazer uma projeção de termos emocionais numa representação, como a ilustrada na Figura 2.6, é sempre uma simplificação. De fato, o número enorme de termos e expressões emocionais existentes em diferentes línguas e culturas indica claramente a existência de uma diferenciação muito mais sutil e, possivelmente, mais adequada à caracterização dos fenômenos emocionais. [MORGADO, 2006]

Neste trabalho, utiliza-se, algumas vezes, o termo emoção de forma abrangente. Entretanto, vistas as definições acima, percebe-se que este trabalho lida com estados afetivos que podem não ser considerados emoções em alguns modelos. Por isso, optou-se dar preferência ao termo estado afetivo, ou ainda, estado interno do agente.

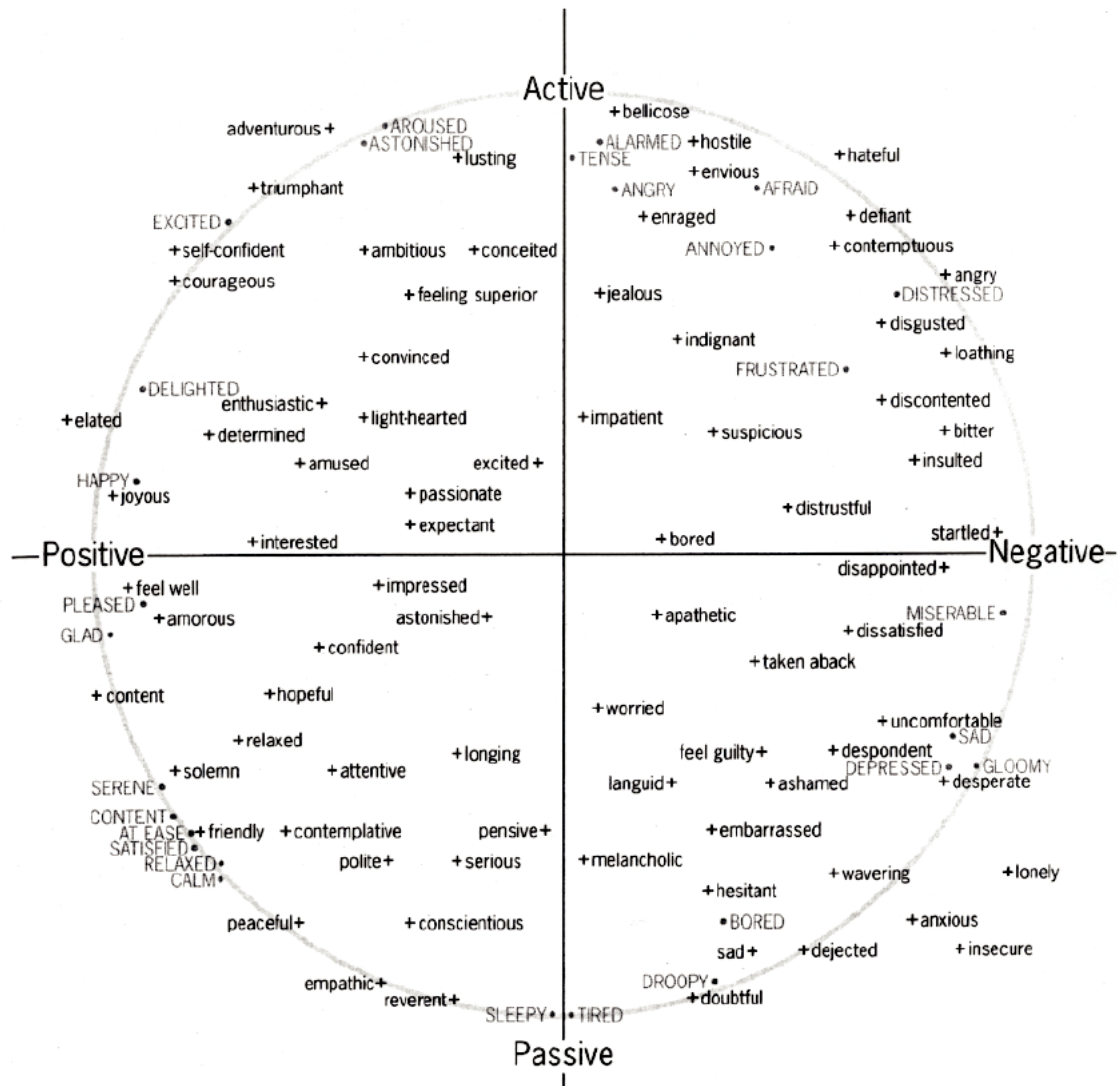


Figura 2.6: Representação bidimensional de termos emocionais - dimensão vertical: ativação (ativo/passivo); dimensão horizontal: valência (positiva/negativa) [SCHERER, 2001].

2.5.1 As Seis Emoções Básicas de Ekman

O pesquisador Paul Ekman, um dos criadores do sistema FACS [EKMAN; FRIESEN, 1978], defende em muitos de seus trabalhos, como [EKMAN, 1992] e anteriores, um conjunto de emoções que são chamadas de *emoções básicas*. Sua hipótese é de que existem sinais-padrão emitidos pelos humanos em sua fisiologia (expressões faciais) que são reconhecidos universalmente, independente de cultura. Esse conjunto de emoções básicas sofreu algumas adições em trabalhos posteriores [EKMAN, 1993, 1999a,b], inclusive em resposta a críticas sofridas, como a de Ortony e Turner em [ORTONY; TURNER, 1990] (“O que há de básico nas emoções básicas?”). Entretanto, até hoje, muitos trabalhos (como o próprio padrão MPEG-4), que relacionam expressões faciais e emoção, utilizam como base 6 emoções, chamadas de *as 6 emoções básicas de Ekman*. São elas: *alegria, surpresa, tristeza, desgosto, medo e raiva*. Essas emoções claramente podem ser usadas como um *framework* de expressões faciais, como o padrão MPEG-4 sugere. Segundo Ekman [EKMAN, 1992], cada emoção básica não constitui apenas um estado afetivo, mas

uma *família* de estados relacionados. Por exemplo, na família “raiva”, pode-se relacionar mais de 60 expressões de raiva.

2.5.2 Modelos Emocionais

Existem muitos estudos de modelos emocionais, com diferentes níveis de alcance e complexidade. Este trabalho baseou-se, em sua metodologia, principalmente nos conceitos encontrados no trabalho de Elliot [ELLIOTT, 1992], cujo modelo emocional é baseado no modelo de Ortony, Clore e Collins (OCC) [ORTONY; CLORE; COLLINS, 1988], sob o qual se baseiam vários projetos e modelos computacionais de emoção, como pode ser visto em [PICARD, 1997]. A tabela do trabalho de Elliot está transcrita (traduzida) na Tabela 2.4 e foi a principal base utilizada no processo de identificação de emoções nos filmes de CG, descrito no Capítulo 4.

Tabela 2.4: Emoções e significados conforme o modelo OCC, por Elliot [ELLIOTT, 1992].

Grupo	Especificação	Nome e tipo de emoção
Bem-estar	avaliação de uma situação como um evento	alegria: satisfação sobre um evento angústia: insatisfação sobre um evento
Destino dos outros	valor pressuposto de uma situação como sendo um evento que afeta outro	feliz-por: satisfação sobre um evento desejável para o outro satisfação maliciosa (<i>gloating</i>): satisfação sobre um evento indesejável para o outro ressentimento: insatisfação sobre um evento desejável para o outro pesar, “sentir muito por” (<i>sorry-for</i>): insatisfação sobre um evento indesejável para o outro
Baseado em perspectivas	avaliação de uma situação como um evento sobre o qual se tem perspectivas	esperança: satisfação sobre um evento com perspectivas desejáveis medo: insatisfação sobre um evento com perspectivas indesejáveis ciúmes: ressentimento sobre uma meta desejada de maneira mutualmente exclusiva inveja: ressentimento sobre uma meta desejada de maneira não exclusiva
Confirmação	avaliação de uma situação como confirmação ou des	satisfação: satisfação sobre um evento desejável confirmado

	confirmação de uma expectativa	<p>alívio: satisfação sobre um evento não desejável desconfirmado</p> <p>medos-confirmados: insatisfação sobre um evento indesejável confirmado</p> <p>desapontamento: insatisfação sobre um evento desejável desconfirmado</p>
Atribuição	avaliação de uma situação como sendo uma ação importante de alguém	<p>orgulho: aprovação de seu próprio ato</p> <p>admiração: aprovação do ato de outro</p> <p>vergonha: desaprovação do seu próprio ato</p> <p>reprovação, censura: desaprovação do ato de outro</p>
Atração	avaliação de uma situação como contendo um objeto atrativo ou não-atrativo	<p>gostar de (<i>liking</i>): achar um objeto atraente</p> <p>desgostar de (<i>disliking</i>): achar um objeto não atraente</p>
Bem-estar/ Atribuição	emoções compostas	<p>gratidão: admiração + alegria</p> <p>raiva: reprovação + angústia</p> <p>gratificação: orgulho + alegria</p> <p>remorso: vergonha + angústia</p>
Atração/Atribuição	composição de extensões de emoções	<p>amor: admiração + gostar de</p> <p>ódio: reprovação + desgostar de</p>

3 ESTADO DA ARTE

*“Qualquer idéia que te agrada
Por isso mesmo... é tua
O autor nada mais fez que vestir a verdade
Que dentro de ti se achava inteiramente nua”*

— MARIO QUINTANA

Sabe-se que os olhos desempenham um papel fundamental no processo de comunicação não-verbal entre os seres humanos. Existem muitos estudos de animação facial relacionados com estados afetivos e expressividade, mas seu enfoque principal geralmente não está nos olhos. Percebe-se, analisando a bibliografia, que a relação entre movimento de olhos e estados afetivos foi pouco explorada até o momento.

Este capítulo apresenta os principais problemas abordados e uma visão geral sobre os principais modelos computacionais de animação de olhos encontrados na bibliografia. Esses modelos baseiam-se, principalmente, nas funções comunicativas do olhar, como a sinalização de troca de turno em um diálogo. Ao final do capítulo, é feita a relação dos trabalhos que serviram de base para o desenvolvimento do modelo proposto neste trabalho.

3.1 Principais Abordagens

Analisando a bibliografia, foram encontradas diferentes pesquisas sobre animação de olhos. Em geral, essas pesquisas têm como forte base estudos psicológicos, como [KENDON, 1967; ARGYLE; COOK, 1976], que estudam o comportamento do olhar na interação social. Argyle e Cook são responsáveis por um estudo psicológico dos padrões de olhar e olhar mútuo (*Gaze and Mutual Gaze*) [ARGYLE; COOK, 1976]. Este livro não aborda apenas aspectos fisiológicos, mas principalmente comportamentais, buscando padrões nos olhares trocados entre as pessoas. Trata-se de uma obra bastante respeitada, que serviu como base para os modelos de simulação de olhar em grande parte dos trabalhos estudados, tais como [COHEN; COLBURN; DRUCKER, 2000; CASSELL; TORRES; PREVOST, 1997; VERTEGAAL et al., 2001; LEE; BADLER; BADLER, 2002; LANCE; STACY; KOIZUMI, 2004]. Entretanto, os dados fornecidos por artigos psicológicos necessitam ser, de alguma forma, traduzidos para um sistema computacional. Para isso, os pesquisadores adotaram diferentes técnicas, de acordo com o problema que resolveram abordar.

Vários trabalhos exploram o comportamento do olhar em sistemas de conversação face-a-face com os usuários, como ECAs. Alguns estudam padrões de olhar em um di-

álogo, como o olhar mútuo e o desvio de olhar durante os turnos de conversação [CASSELL; TORRES; PREVOST, 1997; COHEN; COLBURN; DRUCKER, 2000; FUKAYAMA et al., 2002]. Segundo [VERTEGAAL et al., 2001], o olhar pode ser usado para reconhecer quem vai assumir o discurso quando na troca de turno. Modelos como [COHEN; COLBURN; DRUCKER, 2000; FUKAYAMA et al., 2002; LEE; BADLER; BADLER, 2002; VINAYAGAMOORTHY et al., 2004] levam em consideração se o participante da conversa encontra-se no estado interlocutor ou ouvinte, para a determinação de parâmetros que geram olhares diferenciados. Verificou-se em [LEE; BADLER; BADLER, 2002], através da análise de dados de um rastreador de olhos (*eye tracker*), que esses dois estados apresentam diferentes distribuições de probabilidade no intervalo entre os movimentos oculares.

Existem também abordagens que buscam determinar o foco atencional em agentes imersos em um mundo virtual [ITTI; DHAVALÉ; PIGHIN, 2004] ou em aplicações de realidade virtual [COURTY; BRETON; PELÉ, 2003], provendo ao agente capacidade de percepção do mundo através do uso técnicas de visão artificial. O recente trabalho de Gu e Badler [GU; BADLER, 2006] procura incluir noções de *engajamento* e *distração* nos agentes em um ambiente dinâmico, onde movimentos periféricos podem afetar o comportamento do olhar.

O trabalho de Lance *et al.* [LANCE; STACY; KOIZUMI, 2004] procura fornecer um modelo de olhares expressivos para humanos virtuais. A abordagem do estudo deles é um tanto diferente dos trabalhos citados até agora, uma vez que os dados de padrões de olhar são extraídos de filmes de Computação Gráfica. Porém, esse trabalho ainda encontra-se nas etapas iniciais. A ênfase do artigo está na metodologia e em alguns dados coletados, sendo que os resultados ainda são preliminares (tabelas com algumas correlações entre os parâmetros coletados). Nenhum modelo de geração de olhares é apresentado, uma vez que, segundo os autores, este seria construído a partir dos dados coletados. Foi pesquisado um pouco mais a fundo a respeito dos pesquisadores desse estudo, a fim de verificar se houve continuidade nessa pesquisa, mas não foi encontrada nenhuma outra publicação deles sobre essa abordagem.

A abordagem mais completa encontrada, levando em consideração a expressividade do olhar foi encontrada no trabalho de Poggi *et al.* [POGGI; PELACHAUD; ROSIS, 2000] [POGGI; PELACHAUD, 2002]. Este trabalho procura relacionar expressões do olhar com o significado da fala do agente. De acordo com a semântica do texto, são identificadas as intenções a serem sinalizadas pelo comportamento do olhar do agente, seguindo uma taxonomia proposta pelas pesquisadoras. Alguns detalhes desse trabalho, assim como dos outros citados, são apresentados na próxima seção.

3.2 Modelos de Animação de Olhos

Para simular os movimentos oculares, diferentes modelos foram construídos utilizando diversas abordagens, contemplando mais ou menos itens de animação, de acordo com o enfoque do problema dado pelos pesquisadores.

Cassel [CASSELL; TORRES; PREVOST, 1997] descreve um modelo de animação de olhos baseado na estrutura de conversação entre duas pessoas. O modelo leva em consideração dois aspectos: *i*) a troca de turnos durante uma conversa, onde os dois participantes intercalam suas falas e *ii*) a estrutura do discurso, com a análise do conteúdo proposicional. Foram feitos experimentos em busca de padrões de olhar durante a ocorrência dos elementos da estrutura de informação identificados em um discurso. Com base

nesses experimentos e em algum conhecimento prévio a respeito de regras do olhar, o artigo apresenta um algoritmo para a animação dos olhos, que foi implementado no agente comunicativo Gandalf [THÓRISSON, 1997].

O modelo proposto por Cohen *et al.* [COHEN; COLBURN; DRUCKER, 2000] foi baseado principalmente na psicologia. É um modelo estocástico, representado por uma máquina de estados hierárquica. O contexto onde os humanos virtuais estão inseridos é o de agente comunicativo (conversa com o usuário), conversa com outro avatar ou entre vários avatares (em uma videoconferência).

Lee *et al.* [LEE; BADLER; BADLER, 2002] propõe um modelo estatístico construído com base na coleta de dados reais, via rastreamento do globo ocular (*eye tracking*). Os dados coletados servem para a geração automática de sacadas visuais, cujos parâmetros considerados foram apresentados no Capítulo 2, a saber: *i*) magnitude, *ii*) direção *iii*) duração e *iv*) intervalo inter-sacádico. Através da análise dos dados capturados, foram encontrados diferentes padrões dos parâmetros observados para os estados ouvinte e interlocutor, expressos por diferentes distribuições de probabilidade de ocorrência. O modelo apresenta-se bastante robusto mediante a avaliação feita, onde é comparado com um agente de olhos parados e um com movimentos aleatórios dos olhos. A Figura 3.1 ilustra a performance do humano virtual utilizado por Lee em suas animações, que seguem o padrão MPEG-4.

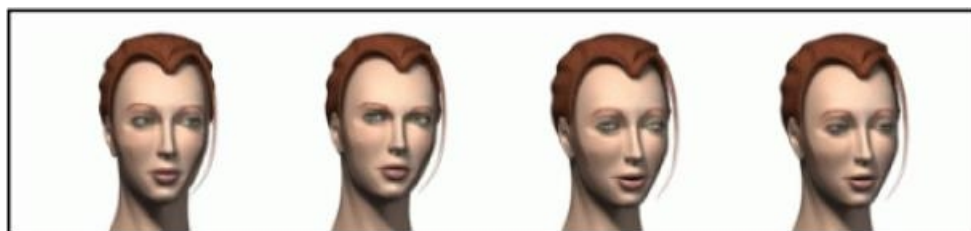


Figura 3.1: *Snapshots* do humano virtual utilizado para as animações no modelo de Lee [LEE; BADLER; BADLER, 2002].

O modelo de Lee [LEE; BADLER; BADLER, 2002] é estendido por Vinayagamoorthy *et al.* [VINAYAGAMOORTHY *et al.*, 2004], que constrói um modelo comportamental usado para simular um comportamento de olhar realístico e outras animações corporais, que fazem parte do comportamento não-verbal, para avatares representando participantes de um ambiente imersivo compartilhado.

Fukayama *et al.* [FUKAYAMA *et al.*, 2002] propõem um modelo dirigido a agentes comunicativos. Para isso, eles procuram padrões de olhar que causem impacto ao usuário. O modelo é implementado com uma cadeia de Markov de 2 estados e os parâmetros de olhar por eles considerados são três: *i*) quantidade de olhar (porcentagem do tempo total que o agente olha para o usuário na interação), *ii*) duração média do olhar e *iii*) pontos de olhar distante (*away*, pontos que não estão de encontro com o olhar do usuário), retirados de estudos psicológicos. Apresentam, então, uma avaliação das impressões causadas por padrões de olhar criados pela alteração dos parâmetros de olhar, através de hipóteses considerando fatores de o “quão amigável” pareceu o olhar, ou a impressão de o “quão dominante” pareceu.

Poggi, Pelachaud e Rosis tratam de uma abordagem que procura gerar faces expressivas [POGGI; PELACHAUD, 2000] a partir da identificação da semântica do conteúdo proposicional da conversa do agente. Elas também desenvolveram um trabalho considerando especificamente o comportamento do olhar [POGGI; PELACHAUD; RO-

SIS, 2000] [POGGI; PELACHAUD, 2002]. Esses trabalhos fazem parte de um contexto maior, um grupo de pesquisa que há alguns anos desenvolve um agente comunicativo, chamado Greta [ROSIS et al., 2003]. Os olhares incluem dois diferentes aspectos, sob os quais é possível montar um “alfabeto do olhar” [POGGI; PELACHAUD; ROSIS, 2000], relacionando-os:

- **Significado:** informação que o agente pretende transmitir;
- **Sinal:** expressão não-verbal do significado. O sinal é constituído do conjunto de feições físicas e do comportamento dinâmico dos olhos no olhar: suas ações musculares e seu estado fisiológico.

Seguindo uma taxonomia própria, baseada na literatura, seu modelo classifica os “significados” de acordo com as *informações do mundo* e as *informações da mente* do agente, que são: *i*) crenças, *ii*) metas e *iii*) estado afetivo. Os sinais foram atribuídos para cada significado (parâmetros para cada expressão) empiricamente ou com base na literatura. O trabalho de Pelachaud em [PELACHAUD; BILVI, 2003] propõe, ainda, um modelo estatístico associado às informações comunicativas do seu trabalho anterior, para preencher “lacunas” quando não há informações relevantes para o agente. A Figura 3.2 mostra a performance do agente virtual no tipo de olhar classificado como “olhos performativos”, incluído no tipo de significado “metas” em uma expressão que sugere o termo “Eu imploro...” [POGGI; PELACHAUD, 2002].



Figura 3.2: Olhar “performativo” gerado pelo modelo de Poggi, Pelachaud e Rosis [POGGI; PELACHAUD, 2002].

No trabalho de Itti *et al.* [ITTI; DHAVALÉ; PIGHIN, 2004], os autores têm um estudo voltado ao *foco atencional* de um indivíduo. O modelo é baseado no sistema neurobiológico de macacos. São utilizadas técnicas de processamento de imagens e Redes Neurais Artificiais para a determinação do foco atencional do indivíduo, ou seja, para onde ele direciona o olhar. Nesse artigo, o ambiente não é mais o de conversa um-para-um ou de videoconferência, como visto em outros trabalhos. O humano virtual tem capacidade de reação dos olhos a uma variedade de vídeos clipes, *outdoors* reais e saídas de consoles de jogos. A idéia é que os avatares possam estar imersos em um mundo virtual, com diversos elementos que podem ou não ser percebidos (através de visão artificial) e, de acordo com sua focalização, seja gerada uma animação coerente e realista de seus olhares.

Uma proposta mais simples de detectar um foco de atenção, com a finalidade de aplicação em um sistema de realidade virtual, é apresentada no trabalho de Courty [COURTY; BRETON; PELÉ, 2003]. Nesse trabalho, o foco atencional do humano virtual é fornecido por um sistema de detecção de faces em tempo real, que captura as imagens através de uma *web cam*.

Terzopoulos [TERZOPOULOS; LEE, 2004] traz uma abordagem bastante completa sobre animação comportamental facial. A parte de movimentação dos olhos é diretamente sincronizada com os movimentos da cabeça. Eles baseiam-se em observações e dados coletados na fisiologia humana [CARPENTER, 1988; FREEDMAN; SPARKS, 2000]. Foi implementado um comportamento de cabeça e olhos que leva em consideração os fenômenos reportados na literatura. O esquema usa funções exponenciais com diferentes constantes de tempo para a cabeça e os olhos, para aproximar as curvas cinemáticas observadas empiricamente, mostradas na Figura 2.5, no Capítulo 2. O modelo, que suporta os movimentos de retenção e deslocamento de olhar, implementa a camada de controle de movimentação cabeça-olhos da face sintética. Para dar ainda um ar um pouco mais natural à cabeça, os ângulos de rotação foram perturbados com ruído usando a técnica de Perlin (*Perlin Noise*) [PERLIN, 1997].

Peters *et al.* [PETERS *et al.*, 2005] propõe um modelo de atenção e interesse usando o comportamento do olhar em agentes comunicativos. Segundo o artigo, quando o transmissor produz seus sinais comunicativos, para que a interação prossiga, o receptor precisa passar pelos seguintes passos: *atenção, percepção, compreensão, reação interna, decisão e geração*. O agente conversacional deve ser capaz de começar, manter e terminar uma conversação. O modelo proposto procura contemplar essas habilidades, a fim de estabelecer a noção de engajamento em uma conversa dos pontos de vista do ouvinte e do interlocutor.

O modelo de Gu e Badler [GU; BADLER, 2006] busca padrões para o desenvolvimento de um modelo que simule os olhares de agentes conversacionais num ambiente de conversação, considerando não apenas o engajamento na conversa, mas também a distração que ocorre em um ambiente dinâmico, em que movimentos periféricos podem influenciar no comportamento do olhar. O modelo apresentado no artigo contempla, além da percepção, regras referentes a trocas de turno em conversação. Mas o intuito dos autores é que o modelo contemple também comportamentos referentes ao estado interno dos agentes, como atividades cognitivas e fenômenos emocionais, como os que foram listados nas tabelas 2.3 e 2.2, no capítulo anterior. A Figura 3.3 ilustra alguns comportamentos de uma agente de acordo com diferentes eventos (movimentação de outro agente) ocorridos no ambiente em que está inserida.



Figura 3.3: Agentes inseridos em um ambiente, dotados de percepção e com diferentes reações, pelo modelo de Gu e Badler [GU; BADLER, 2006].

3.3 Contextualização do Trabalho no Estado-da-Arte

Este trabalho é baseado principalmente em três trabalhos. O primeiro é o trabalho de Lee *et al.* [LEE; BADLER; BADLER, 2002], que provê um modelo estocástico para a geração de sacadas. Os resultados deste trabalho chamam atenção por serem muito realísticos, mas não consideram, no entanto, estados afetivos. Por este motivo, utilizou-se o modelo de Lee como uma *engine* de sacadas, fornecedora dos parâmetros para os movimentos do globo ocular, ora de acordo com as funções probabilísticas do modelo de Lee, ora de acordo com as funções do modelo proposto, para determinados comportamentos relacionados ao estado afetivo do agente.

O segundo trabalho no qual este trabalho é baseado é o de Poggi *et al.* [POGGI; PELACHAUD; ROSIS, 2000; POGGI; PELACHAUD, 2002]. Desse trabalho, foram aproveitados os conceitos de “alfabeto do olhar”, isto é, o mapeamento entre os significados e os sinais dos olhos, através de parâmetros específicos. Os parâmetros referentes a expressão do olhar do modelo proposto foram fortemente baseados nos parâmetros apresentados nesse trabalho, como mostrado no Capítulo 4. Seguindo a taxonomia proposta por Poggi, este trabalho teria o seu enfoque de estudo enquadrado na classificação “olhar afetivo”, que faz parte das “informações da mente” do agente. Entretanto, como o modelo proposto não está vinculado a nenhum modelo emocional ou comunicativo, esta classificação não se apresenta claramente, como um módulo específico. Tanto que é possível fornecer como entrada ao modelo dados referentes às “informações mundo”, como posições para a geração de um olhar direcionado.

Finalmente, o terceiro trabalho, inspirador da metodologia aplicada para a obtenção de dados comportamentais sob o qual o modelo é construído, é o trabalho de Lance *et al.* [LANCE; STACY; KOIZUMI, 2004]. Esse é o trabalho cujo objetivo mais se assemelha ao do presente trabalho. O trabalho de Lance aborda a problemática da falta de empatia e expressividade no olhar dos ECAs e propõe uma metodologia de coleta de dados a partir da observação dos filmes de CG. Ele apresenta duas razões para basear o estudo em filmes de CG:

1. A literatura não provê dados suficientes relacionados em como os deslocamentos dos olhos variam de acordo com o estado interno do personagem;
2. Em filmes de CG, os deslocamentos de olhar são explicitamente trabalhados para serem entendidos pela audiência. Em outras palavras, a linguagem utilizada nos filmes de CG parece ser a mais adequada para o uso em ECAs e outras aplicações que requeiram a geração automática de olhares.

Porém, é necessário salientar que a metodologia seguida neste trabalho difere-se da metodologia proposta por Lance, que optou por fazer medições (não automatizadas) precisas de itens como ângulo e direção do olhar, ângulos da cabeça e pescoço e tempo de duração do olhar. O presente trabalho tem como principal enfoque, em sua metodologia, identificar e descrever comportamentos de olhares salientes em filmes de CG (como apresentado no Capítulo 4), mas não se preocupa em medições precisas dos parâmetros, no atual estágio de desenvolvimento (que abrange esta monografia).

4 MODELO

“A arte não reproduz o que vemos. Ela nos faz ver.”

— PAUL KLEE

Este capítulo apresenta o modelo proposto. Inicialmente, é apresentado o escopo do problema estudado e a metodologia proposta para a obtenção de dados comportamentais. Em seguida, são descritos os comportamentos considerados no modelo e a linguagem de descrição utilizada para realizar a tradução dos dados coletados para um sistema computacional. Logo após, são apresentados detalhes do modelo de Lee [LEE; BADLER; BADLER, 2002] e como é feita a associação dos olhares com movimentos de cabeça e pálpebras. Por fim, é fornecida a visão geral da arquitetura proposta, com a descrição de todos os módulos do modelo.

4.1 Da Arte dos Animadores a um Modelo Comportamental

Este trabalho investiga padrões de comportamento dos olhos e sua relação com estados afetivos. Como apresentado no Capítulo 2, foram encontradas na bibliografia algumas informações sobre comportamentos do olhar relacionados com atividades cognitivas e alguns estados emocionais (tabelas 2.1 e 2.3). Entretanto, esses comportamentos são descritos somente em termos de direção do olhar ou ainda “olhar ou não para” (*away/toward*) um determinado alvo. Não há informação, por exemplo, sobre como estes deslocamentos de olhar são realizados (intervalos de magnitudes, velocidade, intervalos inter-sacádicos). Artigos psicológicos fornecem informações sobre o olhar no contexto de interação social [KENDON, 1967; ARGYLE; COOK, 1976], mas não diretamente relacionadas com estados afetivos.

Por outro lado, observa-se em bons filmes de CG que o comportamento dos olhos dos personagens é muito convincente. Pode-se facilmente entender seu estado interno, mesmo em situações mais complexas (por exemplo, quando o personagem está mentindo ou dissimulando). Os animadores conseguem reproduzir os sentimentos desejados nos personagens através de comportamentos que algumas vezes podem parecer exagerados, mas que são bons o suficiente para produzir a receptividade esperada nos expectadores. De fato, os artistas podem trabalhar nas animações até elas “parecerem boas”, num processo cíclico. Por este motivo, entende-se que propor métodos de geração automática desses comportamentos é um desafio. Sabe-se que, para produzir bons resultados, é necessário o artista ter uma boa habilidade de abstração dos comportamentos humanos. E este é o ponto que se mostra mais atraente para este estudo: verificar o que se pode aprender com

os artistas de CG sobre animação de olhos, de maneira a aplicar este conhecimento em um modelo computacional.

Neste contexto, a construção deste modelo baseou-se principalmente em três idéias:

1. O modelo estatístico de Lee [LEE; BADLER; BADLER, 2002], para a geração de sacadas realistas;
2. A idéia do “alfabeto do olhar”, do trabalho de Poggi [POGGI; PELACHAUD; ROSIS, 2000], para o mapeamento de sinais e significados do estado interno do agente, e;
3. A idéia da extração de dados de filmes de CG, do trabalho de Lance *et al.* [LANCE; STACY; KOIZUMI, 2004].

As próximas subseções explicam com detalhes como estes itens foram inseridos na construção do modelo, cuja arquitetura completa é apresentada na Subseção 4.1.5.

4.1.1 Metodologia

Esta seção descreve a metodologia adotada para a observação dos comportamentos e coleta dos parâmetros de interesse. O objetivo dessa metodologia é a identificação de comportamentos de olhar que refletem o estado interno do personagem e sua associação com expressões faciais.

Os parâmetros observados estão listados na Tabela 4.1. A escolha destes parâmetros foi baseada a lista de parâmetros apresentados no trabalho de Poggi *et al.* [POGGI; PELACHAUD; ROSIS, 2000], com algumas modificações pertinentes ao foco deste trabalho:

1. Observação da frequência das piscadas: de maneira empírica, a frequência das piscadas é anotada (*lenta, padrão* ou *rápida*), com o objetivo de no futuro se verificar se há alguma alteração com o estado afetivo observado. O valor “*padrão*” deve ser usado se nada de especial na frequência das piscadas foi notado;
2. Observação de parâmetros sacádicos: alcance das magnitudes e intervalos intersacádicos (frequência), ou informação de um olhar específico direcionado (como ocorre no caso da Tabela 2.1. Algumas considerações: *i*) o valor “*padrão*” para magnitude, direção e frequência significa que nada em especial foi observado e o olhar pode ser gerado pelo modelo estatístico; e o valor “*nenhuma*” para a magnitude significa o olho localizado na sua posição central;
3. Este trabalho não considera parâmetros relacionados a “*rugos*”, como proposto no trabalho de Poggi *et al.* [POGGI; PELACHAUD; ROSIS, 2000];
4. O campo “*observações extras*” pode ser usado se for encontrado alguma característica muito marcante no comportamento que não possa ser anotada nos parâmetros padrão, para estudos posteriores;
5. Neste modelo, como será descrito posteriormente, a cabeça segue os deslocamentos de olhar com “*grandes*” magnitudes, de acordo com a literatura. Mas em algumas expressões, a movimentação da cabeça pode ser diferenciada, não seguindo essas regras. Para esses casos, julgou-se necessário observar os parâmetros mostrados na Tabela 4.1. O valor “*padrão*” deve ser especificado se nenhum comportamento específico foi notado.

Tabela 4.1: Parâmetros observados nas cenas selecionadas.

Sobrancelhas: direita/esquerda	
parte interna	para cima / centro / para baixo
parte média	para cima / centro / para baixo
parte externa	para cima / centro / para baixo
Pálpebras: direita/esquerda	
superior	padrão / levantada / abaixada
inferior	padrão / levantada / abaixada
Comportamento das Piscadas	
frequência	padrão / rápida / lenta
Comportamento dos Olhos	
umidade	seco / úmido / lágrimas
dilatação da pupila	padrão / dilatada / contraída
magnitude	padrão / grande / pequena / nenhuma
direção	padrão / esquerda / esquerda-cima / cima / direita-cima / direita / direita-baixo / baixo / esquerda-baixo
frequência	padrão / rápida / lenta / fixação
observações “extras”	descrição livre
Cabeça	
inclinação vertical	padrão / para cima / para baixo
virar para	padrão / esquerda / direita
inclinação horizontal	padrão / esquerda / direita

Nas cenas observadas, é necessário também identificar o estado afetivo do personagem, de maneira a mapear a expressão facial e o comportamento do olhar. Identificar emoções não é uma tarefa trivial [EKMAN, 1993; ORTONY; TURNER, 1990; COHN, 2006]. A base para a identificação de emoções foi a tabela do trabalho de Elliot [ELLIOTT, 1992], transcrita na Tabela 2.4 do Capítulo 2.

A metodologia proposta consiste nos seguintes passos, apresentada como um relatório do que foi realizado neste trabalho:

1. Inicialmente, foram escolhidos filmes cuja animação de olhos dos personagens foi considerada convincente. Optou-se por observar filmes que seguem três linguagens diferentes, de maneira a “capturar” uma diversidade de comportamentos:
 - Humanos cartuns (filme escolhido: Os Incríveis¹)
 - Humanos realistas (filme escolhido: Final Fantasy: the Spirits Within²)
 - Não-humanos cartuns (filmes escolhidos: Era do Gelo 1 e 2³)
2. Os filmes foram assistidos inteiramente, sendo anotadas as cenas que foram consideradas de interesse para a pesquisa (as que apresentam expressões e comportamentos de olhos destacáveis), identificando-se o estado afetivo do personagem, se

¹Pixar (2004)

²Chris Lee Productions (2001)

³Blue Sky (2002 e 2006)

pertinente. Neste estágio, para cada filme foi criada uma tabela⁴, tal como a Tabela 4.2 ilustra.

Tabela 4.2: Exemplo de parte da tabela de observação do filme Era do Gelo

Início	Breve descrição da cena	Tipo de Emoção
00:01:35	Quando o Scrat percebe que causou uma fissura no gelo. Surpreso ele abre bem seus olhos.	Surpresa
00:02:00	Quando o Scrat percebe que causou um desmoronamento no gelo	Medos-confirmados
00:02:58 00:03:20	Scrat olha com desejo para a noz...	<i>Gostar de</i> Amor
00:04:...	Bebê anta contrariada, quando diz: “Tá bom!”	Aborrecimento
00:05:...	<i>Stress</i> no trânsito dos animais que estão migrando. Antas e Manny.	Raiva
00:06:..	Sid, quando percebe que foi abandonado pelos seus parentes	Ressentimento
00:08:...	Manny: “É, eu tava brincando...”, horrorizado quando o Sid provocou os rinocerontes	Medo
00:13:...	Manny se mostra irônico e impaciente quanto aos questionamentos de Sid	Impaciência Reprovação
00:16:...	Mãe do bebê, hesitante se deveria pular no rio com o bebê para se salvarem dos tigres	Hesitação Medo
...

3. As anotações foram analisadas. Essa análise consiste em três passos:

- Identificar (dar nomes a) os comportamentos encontrados⁵;
- Descrevê-los e parametrizá-los (Subseção 4.1.2 apresenta os comportamentos que foram encontrados);
- Buscar, nos trechos anotados, os estados afetivos que apresentaram esses comportamentos; selecionar trechos mais relevantes.

4. Para cada trecho de filme selecionado:

- Descrevê-lo e parametrizá-lo de acordo com a Tabela 4.1;
- Traduzir essas informações para a linguagem de descrição de comportamentos de olhar (descrita na próxima subseção).

Ao término desses passos, é necessário incluir os dados coletados (comportamentos de olhar e expressões) em um modelo computacional que reproduza os parâmetros do olhar em uma face animada. A próxima subseção detalha como foi construído o modelo, baseado nas informações coletadas.

⁴Estes dados, anotados de maneira informal e subjetiva, estão disponíveis em <http://www.fellovsheep.wait4.org/2007-1/TCC/Stuff>

⁵Dessa análise subjetiva, definiu-se os quatro comportamentos propostos na Subseção 4.1.2

4.1.2 Comportamentos dos Olhos e sua Descrição

De maneira a codificar os parâmetros em um modelo de animação facial, foi desenvolvida uma simples linguagem chamada *Gaze Description Language* (GDL). Esta linguagem possui dois propósitos: *i*) Descrever os parâmetros observados, conforme a Tabela 4.1; e *ii*) prover a entrada para o módulo *Gerador de Olhares Expressivos* do modelo (descrito na Subseção 4.1.5), descrevendo uma ou mais *ações de olhar*, em seqüência. Estas seqüências de ações podem ser provenientes de uma aplicação, como por exemplo o sistema de um ECA, ou manualmente editadas em um arquivo de *script*.

Existem dois tipos de *scripts* GDL:

1. **Arquivos de Descrição de Expressões de Olhares:** no modelo proposto, cuja arquitetura é apresentada na Subseção 4.1.5, cada estado afetivo capturado das anotações das cenas de filmes é codificado em um arquivo de descrição, que contém a informação coletada sobre as sobrancelhas, pálpebras, olhos e cabeça. Todos os arquivos de descrição, como o do exemplo abaixo, constituem o módulo *Base de Dados Comportamental*, que também é explicado na Subseção 4.1.5. Os valores numéricos dos parâmetros, vistos no exemplo abaixo descrevem a intensidade do valor do parâmetro, codificado para o sistema FAP. Os valores zero e um representam os limites inferior e superior dos valores FAP para cada parâmetro, definidos no módulo de geração dos parâmetros de animação do modelo.

```
Animation =
{
  -- Eyebrows Info
  eyebrows = {
    inner = {
      { "l", "up", 1.0 },
      { "r", "up", 1.0 },
    },
    medial = {
      { "r", "down", 0.5 },
      { "l", "down", 0.5 },
    },
    outer = {
      { "l", "down", 0.3 },
      { "r", "down", 0.3 },
    },
  },

  -- Eyelids Info
  eyelids = {
    upper = { "lr", "lowered", 0.4 },
    lower = { "lr", "raised", 0.0 },
    blink = "quick",
  },

  -- Eyeballs Info
  eyeballs = {
    humidity = "wet",
    pupil = "default",

    -- Eye behavior
    behavior = {"distress", 0.9},
  },
}
```



```

-- Head Info
head = {
    pitch = 0.5,
    yaw = 0.0,
    roll = 0.0,
},

-- Mouth Info
mouth = "none",
}

```

2. **Arquivos de “Storyboard” de Olhares:** são os arquivos de descrição das ações, provendo uma interface de alto nível para o modelo de geração dos olhares. Nesses arquivos, descreve-se uma ou mais ações parametrizadas. Existem dois tipos de ações que podem ser definidas nesses *scripts*: *i*) estados afetivos que já estão definidos em arquivos de descrição de olhares na Base de Dados Comportamental e o tempo (em quadros) da animação a ser gerada; e *ii*) comportamentos de olhares implementados no modelo, parâmetros do comportamento e tempo da animação. Se a ação for um comportamento de olhar, a expressão da face se mantém neutra. O exemplo abaixo ilustra um arquivo simples de *storyboard* de olhares. É importante observar, ainda, que os comportamentos de olhares (descritos em detalhes a seguir) são os mesmos que fazem parte dos Arquivos de Descrição de Expressões de Olhares. Cada animação é automaticamente gerada com o tempo indicado, de acordo com a expressão e/ou o comportamento de olhar anotado.

```

Storyboard = {
    {"default", "talking", 3000},
    {"sad", 1000},
    {"default", "listening", 1000},
    {"lookTo", "up", 17.0, 200},
    {"internalDialogue", 50},
}

```

O modelo proposto implementa diferentes comportamentos de olhares. Os nomes dos comportamentos foram atribuídos em inglês, visando diretamente a implementação, por isso a apresentação dos comportamentos segue relacionando os termos nas duas línguas. Para gerar o “olhar padrão”, ou seja, um olhar não relacionado a estados internos do agente, optou-se em utilizar o modelo de Lee [LEE; BADLER; BADLER, 2002]. Têm-se, então, dois comportamentos-padrão:

1. **Ouvinte Padrão (*Default Listening*):** modelo estatístico de Lee para o modo ouvinte;
2. **Interlocutor Padrão (*Default Talking*):** modelo estatístico de Lee para o modo interlocutor.

Para olhares especificamente direcionados, foi implementada uma ação chamada **Olhar para (*Look To*)**, cujos parâmetros são a magnitude e a direção do deslocamento do olhar e o tempo total da animação. A duração do deslocamento do olhar (sacada) é calculada pela fórmula do modelo estatístico de Lee (apresentada na próxima subseção).

Da bibliografia [GU; BADLER, 2006; BROOKER, 2007], foram implementados os seguintes comportamentos (Tabela 2.1 do Capítulo 2):

1. **Construção Visual (*Visual Construct*)**: olhar com direção direita-cima, seguido de fixação;
2. **Lembrança Visual (*Visual Recall*)**: olhar com direção esquerda-cima, seguido de fixação;
3. **Construção de Áudio (*Audio Construct*)**: olhar para a direita, seguido de fixação;
4. **Lembrança de Áudio (*Audio Recall*)**: olhar para a esquerda, seguido de fixação;
5. **Cinestético (*Kinesthetic*)**: olhar com a direção direita-baixo, seguido de fixação;
6. **Diálogo Interno (*Internal Dialogue*)**: olhar com direção esquerda-baixo, seguido de fixação.

A duração do deslocamento é dada pelo modelo de Lee. A magnitude da sacada e o tempo de duração total do olhar devem ser fornecidos como parâmetros.

Baseado nas observações dos filmes, foram identificados e implementados os seguintes comportamentos:

1. **Concentração (*Concentration*)**: os deslocamentos de olhares são menores e menos freqüentes. Estados afetivos que requeiram um olhar fixo podem ser descritos com este comportamento, tais como: raiva, *gostar de* e surpresa. Este comportamento ainda pode ser usado para denotar atenção (olhar mútuo) por um tempo determinado. Neste comportamento, é utilizado, inicialmente, o modelo de Lee para gerar os parâmetros sacadas. Porém, os parâmetros magnitude e intervalo inter-sacádico são alterados posteriormente, de acordo com a parametrização fornecida, da seguinte forma:
 - Magnitude: as magnitudes podem ser limitadas em um intervalo de 0° a 5° , especificando-se valores de zero a um. Zero faz com que os olhos fiquem estáticos (0°). O valor um faz com que o valor da magnitude gerado sofra uma normalização de limite de 5° ⁶.
 - Intervalo inter-sacádico: valores especificados no intervalo de zero a um geram o aumento de 0% a 100% no valor de intervalo inter-sacádico fornecido pelo modelo de Lee.
2. **Desconforto (*Discomfort*)**: os deslocamentos do olhar são mais freqüentes, com grandes (espiadas, desvios de olhar) ou pequenas (sensação de descontrole) magnitudes. Estados afetivos que requeiram uma sensação de perda de controle ou sentimentos confusos/escondidos como comoção, aborrecimento, nervosismo, gratidão e medo podem ser descritos utilizando esse comportamento. O comportamento foi parametrizado em função da magnitude e intervalo inter-sacádico, da seguinte forma:
 - Magnitude: as magnitudes podem sofrer uma diminuição de 0% (mantém o intervalo normal) a 99,95% do seu valor, utilizando-se valores no intervalo de zero a um para indicar a intensidade;

⁶As constantes definidas para as alterações dos parâmetros foram determinadas empiricamente, através da análise subjetiva dos filmes

- Intervalo inter-sacádico: o intervalo inter-sacádico pode ser decrementado de 0% a 99,95% do seu valor, de forma análoga a da magnitude.
3. **Angústia (*Distress*)**: a direção do olhar tende a ser para baixo. Estados afetivos como tristeza e vergonha podem ser descritos utilizando este comportamento. O comportamento foi parametrizado em função da direção, aumentando a probabilidade de ocorrência das direções esquerda-baixo, baixo e direita-baixo para 26,67% (totalizando 80% de probabilidade de ocorrência, enquanto o restante das direções dividem igualmente os 20% restantes), e utilizando-se o modelo de Lee para a geração dos outros parâmetros da sacada. Foi implementado também, para complementar este comportamento, um outro comportamento chamado ***SlideDown***, que gera uma seqüência de olhares para baixo, direita-baixo, baixo, esquerda-baixo e baixo, o que cria uma animação inspirada na observação dos filmes, que não poderia ser gerada diretamente com o modelo de Lee. Este comportamento “complementar” pode ser ativado no comportamento *distress*, aleatoriamente (sorteio no início da geração dos quadros, com 50% de probabilidade de ocorrência), caso o tempo de animação especificado por parâmetros suporte o número de quadros desse (que é fixo em 300 quadros). Este comportamento também pode ser especificado isoladamente no arquivo de descrição de olhar ou de *storyboard*, desde que seja especificado corretamente 300 quadros de duração.
 4. **Ironia (*Irony*)**: a direção do olhar tende a ser para cima. Estados afetivos como impaciência e reprovação podem ser descritos utilizando este comportamento. O comportamento foi parametrizado em função da direção, aumentando a probabilidade de ocorrência das direções esquerda-cima, cima e direita-cima para 26,67%, e utilizando-se o modelo de Lee para a geração dos outros parâmetros da sacada. De maneira similar ao comportamento *distress*, foi implementado um comportamento complementar, chamado ***TurnEyes***. Esse comportamento gera uma seqüência de olhares para esquerda-cima, cima, direita-cima e direita, gerando uma animação de “revirar os olhos”, observado também nos filmes. Essa animação possui o tempo fixo de 40 quadros.

Para todos esses comportamentos, é necessário especificar o tempo da animação a ser gerada, em quadros, no *script* GDL. Pode-se utilizá-los sem expressão facial (diretamente no arquivo de *storyboard*) ou com expressões anotadas em arquivos de descrição de olhares.

4.1.3 Modelo Padrão

Como já descrito neste trabalho, o modelo estatístico apresentado por Lee em [LEE; BADLER; BADLER, 2002] foi implementado no modelo proposto, sendo utilizado como um motor de geração de sacadas. Nesta subseção são apresentados os passos e as fórmulas utilizadas para a geração dos parâmetros.

Inicialmente, é gerado um número randômico P entre 0 e 15, correspondendo a frequência de ocorrência de determinada magnitude. Então a magnitude da sacada pode ser obtida pela Equação 4.1,

$$A = -6,9 \log(P/15,7) \quad (4.1)$$

onde A é a magnitude da sacada, em graus. A magnitude máxima limitada no modelo de Lee é de 27,5° para o modo interlocutor e 22,5° para o modo ouvinte. A direção da

sacada é determinada baseada em oito direções discretizadas, correspondentes aos termos utilizados até o momento, segundo a probabilidade de ocorrência dada na Tabela 4.3.

Tabela 4.3: Distribuição das direções das sacadas.

Direção	Ângulo correspondente	Percentual (%)
direita	0°	15,54
direita-cima	45°	6,46
cima	90°	17,69
esquerda-cima	135°	7,44
esquerda	180°	16,80
esquerda-baixo	225°	7,89
baixo	270°	20,38
direita-baixo	315°	7,19

A duração e velocidade de um sacada são funções de sua magnitude. Para sacadas entre 5° e 50°, a duração tem aproximadamente uma taxa de aumento constante com a magnitude e pode ser aproximado pela função linear da Equação 4.2,

$$D = D_0 + dA \quad (4.2)$$

onde D e A são a duração e a magnitude do movimento do olho, respectivamente. A reta d representa o aumento da velocidade, em duração por grau. Ela varia, nos dados reais de 2 a 2.7ms/°. D_0 é o tempo de intersecção, ou alcance da sacada, que varia, nos dados reais, de 20 a 30ms. Para o modelo de Lee, estes valores são fixos: $d = 2,4 \text{ ms/}^\circ$ e $D_0 = 25 \text{ ms}$.

Para a determinação da velocidade da sacada, utilizou-se a fórmula proposta no trabalho de Vinayagamoorthy *et al.* [VINAYAGAMOORTHY *et al.*, 2004] (Equação 4.3), onde os autores propõem uma equação diferente da apresentada por Lee, evitando a geração de valores negativos.

$$V = 14e^{[-(\pi/4)(X-3)^2]} \quad (4.3)$$

onde V é a velocidade instantânea (°/quadro). X é relacionado com o tempo em quadros da duração da sacada. No modelo de Lee, X é o valor normalizado de cada quadro da duração para 6 quadros (representando 6 pontos da curva gerada pelos dados coletados), a fim de gerar uma aproximação da curva observada em seus experimentos. No modo interlocutor, Lee define equações diferentes para o olhar mútuo (magnitude de 0°) e *away* (outras magnitudes), apresentadas nas equações 4.4 e 4.5, respectivamente.

$$Y = 2,5524 \times 10^{-4}X^2 - 0,1763X + 32,2815 \quad (4.4)$$

$$Y = 1,8798 \times 10^{-5}X^4 - 0,0034X^3 + 0,2262X^2 - 6,7021X + 78,831 \quad (4.5)$$

No modo ouvinte, o intervalo inter-sacádico é obtido através da geração de um número aleatório seguindo uma distribuição gaussiana, com as médias e desvios-padrão de valores 237,5 e 47,1 para o olhar mútuo, e 13,0 e 7,1 para o *away*.

4.1.4 Associação dos Movimentos da Cabeça, Pálpebras e Olhos

Para aumentar o realismo na animação, o modelo implementa uma associação entre os movimentos da cabeça e olhos e das pálpebras e olhos. Para a associação cabeça-olhos, utilizou-se, como regra, a informação encontrada na literatura e apresentada no Capítulo 2. Esta regra estabelece que a cabeça segue os olhos em deslocamentos horizontais com magnitude maior que 25° e verticais maiores que 10° . Optou-se em utilizar esta regra para os comportamentos *Default*, *Look To*, *Visual Construct*, *Visual Recall*, *Audio Construct*, *Audio Recall*, *Kinesthetic* e *Internal Dialogue*. Para o comportamento *Discomfort*, optou-se em não movimentar a cabeça, a fim de gerar a sensação de descontrole só dos olhos, observada dos filmes. O comportamento *Concentration* gera sacadas com, no máximo, 5° de magnitude, portanto esta regra nunca é aplicada. Já nos comportamentos *Irony* e *Distress*, onde observou-se nos filmes que a cabeça mexe-se com mais frequência, faz-se a associação independente do valor da magnitude. A associação dos valores FAP para os parâmetros da cabeça foi feita de forma empírica, seguindo a relação apresentada na Equação 4.6,

$$fap_value = magnitude \times m/27,5 \quad (4.6)$$

onde m (um multiplicador da intensidade) foi fixado em 0,15 para os comportamentos que seguem a regra de associação da literatura e o *Irony*, e 0,5 para o comportamento *Distress*. A constante 27,5 é o valor máximo de magnitude gerado pelo modelo de Lee.

No modelo proposto, as pálpebras acompanham o globo ocular em olhares para baixo, esquerda-baixo, direita-baixo, cima, esquerda-cima e direita-cima. Para olhares com a direção esquerda-cima, cima e direita-cima, o parâmetro FAP da pálpebra inferior recebe o valor segundo a Equação 4.7. Olhares com a direção esquerda-baixo, baixo e direita-baixo geram valores para os FAPs da pálpebra superior segundo a Equação 4.8,

$$fap_value = magnitude \times 0,5/27,5 \quad (4.7)$$

$$fap_value = magnitude \times 0,7/27,5 \quad (4.8)$$

onde 27,5 é o valor máximo de magnitude gerado pelo modelo de Lee e os valores 0,5 e 0,7 (multiplicadores de intensidade) são constantes que foram definidas empiricamente.

Para as piscadas, que não foram alvo de estudo aprofundado no período de desenvolvimento do trabalho, implementou-se a diferenciação de 3 frequências, que ocorrem em intervalos fixos: padrão (85 quadros), lento (105 quadros) e rápido (25 quadros).

4.1.5 Arquitura do Modelo

A arquitetura do modelo proposto está representada na Figura 4.1. O modelo é composto dos seguintes módulos:

1. **Storyboard de Olhares:** é a entrada para o modelo, em forma de Arquivos de *Storyboard* ou chamadas diretas às funções do módulo *Gerador de Olhares Expressivos*;
2. **Base de Dados Comportamental:** é a coleção de expressões de olhares, representando estados afetivos, que foram coletados e descritos em arquivos GDL. De acordo com as ações descritas no *Storyboard* de Olhares, este módulo provê os parâmetros de animação para as expressões;

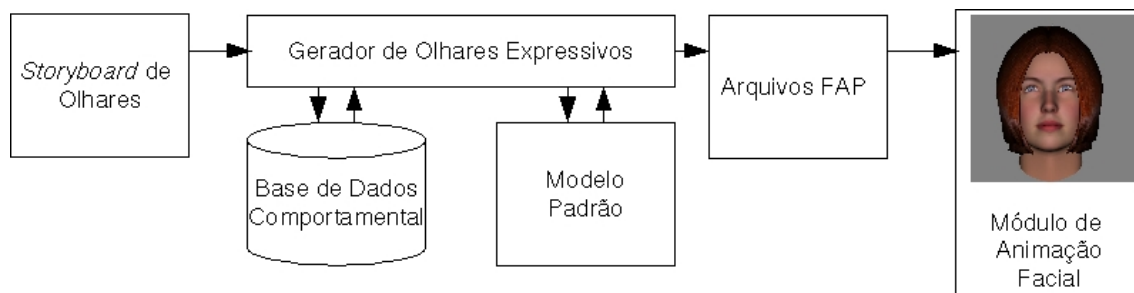


Figura 4.1: Diagrama da arquitetura do modelo.

3. **Gerador de Olhares Expressivos:** este módulo interpreta as ações descritas no *Storyboard* de Ações, carrega os arquivos GDL requeridos da Base de Dados Comportamental e gera arquivos FAP de animação. Este módulo se comunica com o *Modelo Padrão* para gerar os parâmetros das sacadas;
4. **Modelo Padrão:** este módulo compreende a implementação do modelo de Lee, conforme o descrito na Subseção 4.1.3 e alguns outros métodos auxiliares para gerar apenas parâmetros isolados, quando o comportamento requer alterações na magnitude, direção ou intervalo inter-sacádico. A duração das sacadas sempre é calculada usando o modelo de Lee;
5. **Módulo de Animação Facial:** trata-se da plataforma de visualização, que segue o padrão MPEG-4 de Animação Facial. Este módulo foi construído utilizando-se o Xface [BALCI, 2004, 2007], que é um conjunto de ferramentas de código aberto para animação facial e pesquisas com *talking heads*. A função deste módulo é, basicamente, rodar as animações FAP produzidas pelo modelo.

Explicações da implementação do modelo no protótipo construído são fornecidas no Capítulo 5.

5 PROTÓTIPO

“It sounded an excellent plan, no doubt, and very neatly and simply arranged; the only difficulty was that she had not the smallest idea how to set about it;(...)”

— LEWIS CARROLL, em *“Alice’s Adventures in Wonderland”*

Este capítulo fornece um relatório técnico da implementação do modelo descrito no capítulo anterior em um protótipo, chamado *Alice’s Eyes Adventures in OSGLand*¹. Este protótipo foi desenvolvido em ambiente Linux, na linguagem C++ [STROUSTRUP, 2000]. A linguagem Lua [IERUSALIMSCHY, 2003] também foi utilizada, para os arquivos de descrição GDL. As seções seguintes apresentam os aspectos mais relevantes da implementação deste protótipo.

5.1 Linguagem de Descrição de Olhares

A linguagem de descrição de olhares, GDL, é construída em cima da linguagem Lua, para sua rápida interpretação e comunicação com os módulos em C++, feita através da ferramenta Diluculum [BARROS, 2007]. Um *script* GDL é um arquivo em Lua, seguindo uma sintaxe específica. O propósito da GDL é similar ao da APML [CAROLIS et al., 2004], que serve como uma interface de alto nível para a geração de expressões faciais utilizando o padrão MPEG-4 de Animação Facial. A GDL, no entanto, desenvolvida para este trabalho, por enquanto compreende apenas os parâmetros faciais que foram considerados para a geração de olhares e umas poucas ações de boca, predefinidas e não parametrizáveis. Porém, pela sua estruturação, ela pode ser facilmente estendida (como trabalhos futuros, apresenta-se algumas sugestões de melhorias e acréscimos, no Capítulo 7).

Como já descrito no Capítulo 4, os arquivos GDL podem ser de dois tipos: de *storyboard* de ações ou de descrição de expressões. O Apêndice B apresenta a sintaxe completa da linguagem.

5.2 A Base de Dados Comportamental

A Base de Dados Comportamental compreende a descrição dos parâmetros faciais observados para diferentes estados afetivos. Ela, por exemplo, contém o arquivo `sad.gdl`,

¹Alice é a humana virtual utilizada, que foi introduzida, neste protótipo, no ambiente OSG.

com as descrições do olhar em GDL para esse estado afetivo. Se, em um arquivo de *storyboard*, for especificado para “agir triste por 300 quadros”, o módulo gerador de olhares irá buscar os parâmetros em *sad.gdl* e gerar a animação de acordo com esses, com a duração de 300 quadros.

5.3 Gerador de Olhares Expressivos

O Gerador de Olhares Expressivos é o módulo que traduz as informações descritas em GDL para arquivos de animação FAP. Ele possui duas classes principais:

- **Storyboard**: o construtor desta classe recebe como parâmetro o nome de um arquivo de *storyboard*, o carrega e interpreta, gerando o arquivo FAP de animação com a seqüência de ações lidas. Para cada ação, utiliza-se os métodos da classe *GDL2FAP*, descrita a seguir.
- **GDL2FAP**: esta classe contém os métodos que fazem a tradução dos parâmetros encontrados nos arquivos de descrição de olhares ou do comportamento requerido e gera os FAP *streams* (linhas com os valores dos FAPs) para cada frame da animação, de acordo com o número de quadros especificado para cada ação do arquivo de *storyboard*. O cálculo dos valores FAP de cada parâmetro do rosto atingido pela animação em cada frame está implementado em classes separadas, correspondentes aos parâmetros observados: *Eyebrows*, *Eyelids*, *Eyeballs* e *Head*. Objetos da classe *Eyeballs* comunicam-se com o Modelo Padrão, para a geração dos parâmetros das sacadas, e também com os objetos das classes *Eyelids* e *Head* para movimentos associados. Implementou-se, também, a classe *Mouth*, que provê algumas animações de boca, a fim de se gerar animações de rosto inteiro mais coerentes. Como a boca não é o foco de estudo deste trabalho, foram criados alguns comportamentos-chave, como sorrisos e movimentação dos lábios (simulando a fala), que sempre ocorrem da mesma maneira (não são parametrizáveis).

A Figura 5.1 apresenta o diagrama de classes do módulo implementado. Os parâmetros referentes aos olhos *umidade e dilatação das pupilas*, apesar de poderem ser descritos nos *scripts* GDL, não foram implementados no protótipo. Estes itens são colocados, posteriormente, como trabalhos futuros.

5.4 Modelo Padrão

O Modelo Padrão é o módulo que implementa o modelo estatístico de geração de sacadas, conforme descrito no Capítulo 4. Este módulo também possui duas classes principais:

- **EyesAliveModel**: classe que contém a implementação das fórmulas apresentadas na Subseção 4.1.3 do Capítulo 4.
- **SaccadeGen**: interface para a geração de sacadas, que podem ser geradas de forma completamente automática (todos os parâmetros são gerados pelo modelo estatístico) ou com alguns parâmetros especificados (como magnitude, direção e intervalo inter-sacádico), a fim de gerar as animações com comportamentos diferenciados.

A Figura 5.2 apresenta o diagrama de classes do módulo implementado.

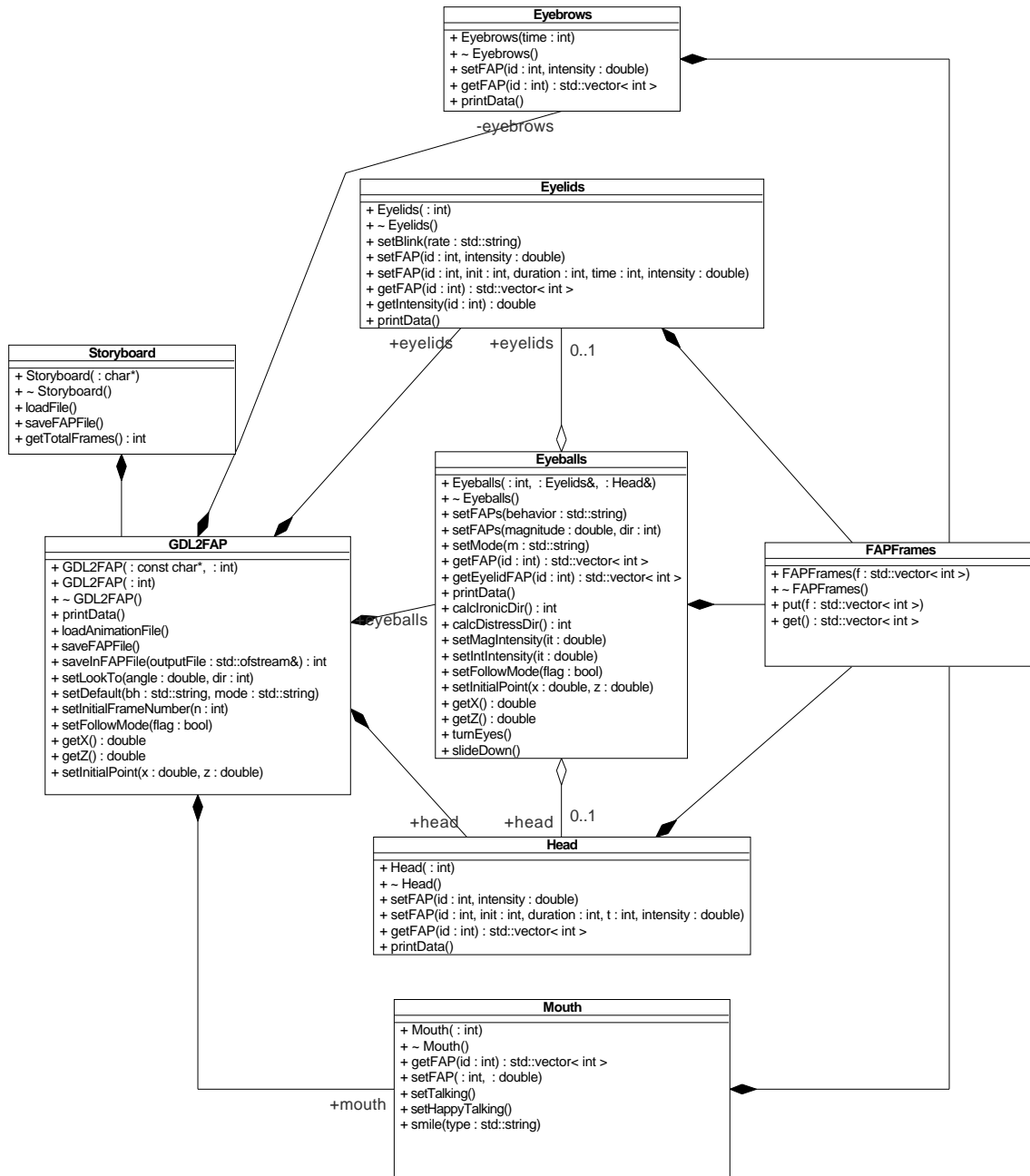


Figura 5.1: Diagrama de classes do módulo Gerador de Olhares Expressivos.

Este módulo é independente e pode ser usado em qualquer outra aplicação que necessite de um modelo de geração automática de sacadas, não necessariamente seguindo o padrão MPEG-4.

5.5 Plataforma de Animação Facial

A escolha da plataforma de animação facial foi, sem dúvida, a etapa mais desafiadora na construção do protótipo. Foi, também, o fator determinante para que se optasse desenvolver o modelo seguindo o padrão MPEG-4 de Animação Facial.

Inicialmente, em uma etapa anterior ao desenvolvimento do modelo proposto, começou-se o estudo da animação de olhos utilizando-se somente o Modelo Padrão em conjunto

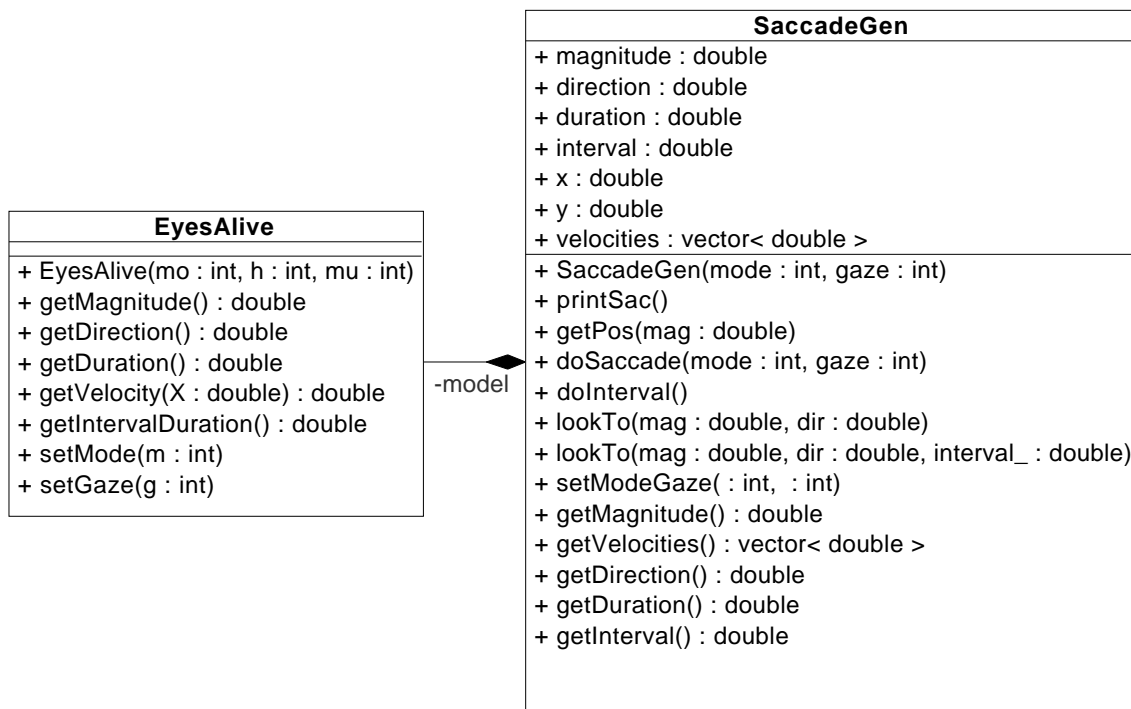


Figura 5.2: Diagrama de classes do módulo Modelo Padrão.

com expressões faciais. Para a visualização, utilizou-se o sistema de *talking head ResponsiveFace*, de Ken Perlin [PERLIN, 1997], desenvolvido na linguagem Java [Sun Microsystems, Inc., 2007]. Foi construído, em parceria com um grupo de pesquisa do Rio de Janeiro², um protótipo que permite a visualização da região dos olhos do agente comunicativo. Na interface do protótipo, é possível selecionar 8 estados afetivos (a saber, surpresa, alegria, arrogância, raiva, desapontamento, aborrecimento e susto), e 3 modos de animação dos olhos: *i*) estático (sem movimentação), *ii*) movimentação gerada aleatoriamente, através do Ruído de Perlin (*Perlin Noise* [PERLIN, 1997]) e *iii*) movimentação gerada pelo modelo de Lee [LEE; BADLER; BADLER, 2002]. A Figura 5.3 ilustra a interface do protótipo.

Este protótipo foi utilizado para experimentos iniciais do estudo sobre animação de olhos e expressividade, em que procurou-se verificar o impacto do comportamento sacádico provido pelo modelo de Lee em conjunto com expressões faciais. Entretanto, para os fins específicos deste trabalho, a plataforma acima não se mostrou adequada. Os fatores que mais influenciaram a busca por uma outra plataforma foram:

- O modelo da face do agente é bastante simplificado (poucos polígonos), especificado diretamente no código da aplicação (*hardcoded*), o que limita o uso de outros modelos e extensibilidade para outras aplicações;
- Apesar do rosto ser definido como uma malha tridimensional, os olhos são bidimensionais (círculos para íris e pupila, não contendo globo ocular), o que limitou bastante a qualidade dos resultados (a saída do modelo de Lee teve de ser adaptada);
- Plataforma incompatível com as aplicações desenvolvidas no grupo de pesquisa

²Projeto Virtus, financiado pelo CNPq em cooperação com PUC-Rio/UFRJ e IMPA

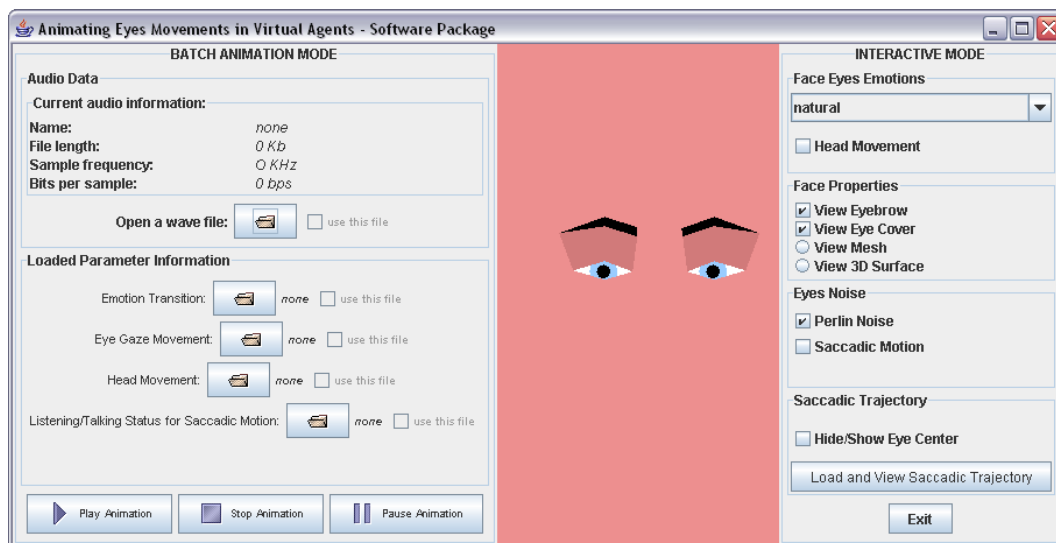


Figura 5.3: Ambiente experimental montado para os testes, na primeira fase do trabalho.

do Laboratório Cromos³, o que impossibilita a sua futura integração aos demais projetos.

Por estes motivos, optou-se mudar a plataforma de visualização. Adotou-se as bibliotecas de animação facial do conjunto de ferramentas Xface [BALCI, 2007], discutido em mais detalhes na próxima subseção.

5.5.1 Xface

O Xface é um conjunto de ferramentas de código aberto para o desenvolvimento de *talking heads*, desenvolvido na linguagem C++ e utilizando a API (*Application Programming Interface*) OpenGL [WOO; DAVIS; SHERIDAN, 1999] para a renderização. Para a animação facial, o Xface provê suporte a animação por quadros-chave (*keyframes*) ou conforme o padrão MPEG-4. O Xface é formado por quatro módulos (ou ferramentas):

- **Xface Core:** é a biblioteca principal, que fornece as funções de animação e renderização da face;
- **XfaceEd:** é um editor que permite associar os pontos de uma malha tridimensional de rosto aos FPs, FAPU e FAPs, necessários para a animação seguindo o padrão MPEG-4, gerando os arquivos com a descrição dessas associações, que precisam ser interpretadas por um *player* de animações (o próprio editor já possui um, embutido em sua interface);
- **XfacePlayer:** é uma aplicação simples, que lê os arquivos do modelo facial padronizado (um arquivo com a descrição da malha, dos FAPU e associação dos FPs aos FAPs, e um arquivo contendo a malha do modelo) e roda animações contidas em arquivos FAP especificados pelo usuário. Este *player* pode ser controlado remotamente;

³O Laboratório Cromos faz parte do Programa Interdisciplinar de Pós-Graduação em Computação Aplicada (PIPICA), Unisinos

- *XfaceClient*: possibilita o controle remoto para o XfacePlayer, via rede.

Para o desenvolvimento do protótipo, o XfaceEd foi utilizado inicialmente para a preparação da face utilizada (Alice, que será descrita na próxima subseção) e para rodar as primeiras animações geradas pelo módulo Gerador de Olhares Expressivos. O módulo de visualização do protótipo (Módulo de Animação Facial), no entanto, utiliza o XFace Core integrado ao ambiente *Open Scene Graph* [BURNS; OSFIELD, 2007], descrito na Subseção 5.5.3.

5.5.2 Os Olhos de Alice

O Xface vem com um modelo-exemplo de *talking head* chamado Alice, junto com alguns exemplos de arquivos de animação FAP. Esse modelo já está com os parâmetros MPEG-4 de Animação Facial ajustados. Entretanto, encontrou-se dois graves problemas nos olhos de Alice, referentes à modelagem e animação do modelo: *i*) os olhos não eram esferas, e sim, semi-esferas que apenas cobriam a parte visível (abertura do olho) da malha; e *ii*) a rotina de animação do Xface Core, referente aos FAPs dos olhos, continha erros e não funcionava adequadamente.

Foi necessário, então, modelar olhos esféricos e incorporar ao modelo de Alice, criando as devidas associações de FPs e FAPs. Os olhos foram modelados na ferramenta de modelagem Wings3D [GUSTAVSSON; CEDERLUND et al., 2007] e incorporados à Alice utilizando-se o XfaceEd. A Figura 5.4 ilustra a modificação feita no modelo, feita com permissão do desenvolvedor do Xface.

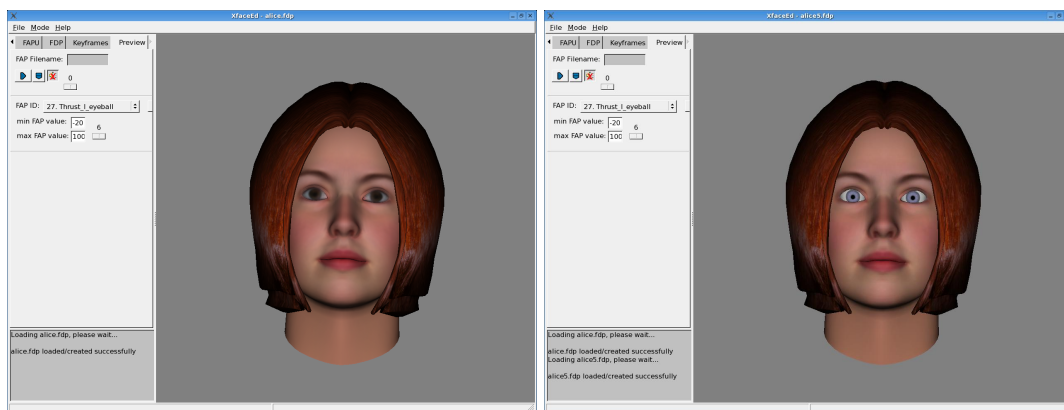


Figura 5.4: As duas imagens mostram o ambiente do XfaceEd exibindo, respectivamente, a Alice original (com olhos semi-esféricos) e a Alice com os olhos esféricos, desenvolvidos para os fins deste trabalho.

A segunda etapa foi estudar o código do Xface Core e corrigir os erros referentes a animação dos olhos. A Figura 5.5 mostra o que ocorre com a animação dos olhos no pacote original do Xface.

O erro foi descoberto, com a análise do código e a ajuda do desenvolvedor do Xface. Consistia em dois problemas: *i*) O *pivot* de rotação da malha dos olhos estava centrado na cabeça, o que gerava as rotações relativas ao ponto central da malha; e *ii*) problemas com as matrizes de transformação implementadas no Xface. A correção consistiu na modificação de alguns métodos em duas classes, localizadas no módulo XEngine do Xface Core: *Drawable* e *RendererGL*. As correções foram enviadas como um *patch*⁴

⁴Disponível em <http://www.fellowsheep.wait4.org/2007-1/TCC/Stuff>

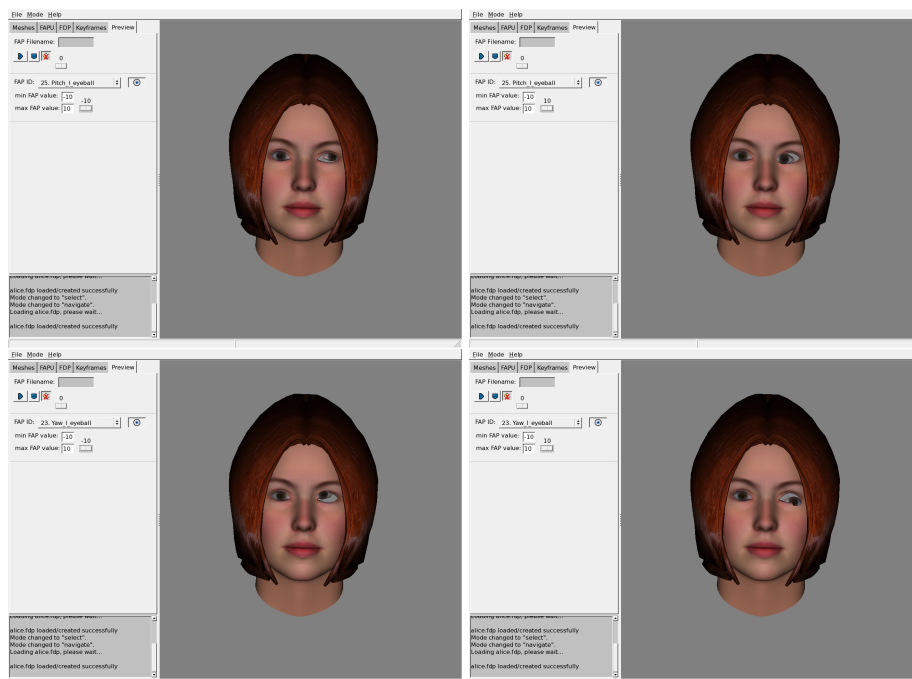


Figura 5.5: Imagens que ilustram os problemas encontrados nas rotinas de animação do Xface Core, deslocando erroneamente o globo ocular de Alice.

ao desenvolvedor do Xface, que pretende incorporá-las na sua próxima versão. Figuras que mostram animações com o pacote corrigido são apresentadas no Capítulo 6.

5.5.3 Integração do Xface no OSG

Para a integração do Xface no protótipo, optou-se em utilizar o XFace Core integrado ao ambiente OSG. As rotinas de integração foram desenvolvidas em conjunto com uma equipe do Laboratório Cromos, para uma aplicação que utiliza o modelo proposto neste trabalho⁵. Dessa integração, resultou uma biblioteca chamada OSGXface (Figura 5.6). O Módulo de Animação Facial do modelo proposto, que permite a visualização dos resultados obtidos, utiliza esta interface.

5.5.4 Alice's Eyes Adventures in OSGLand

O protótipo, que inclui a implementação de todos os módulos do modelo em uma só aplicação, possui a aparência ilustrada na Figura 5.7.

Este protótipo permite a geração instantânea de animações contidas na Base de Dados Comportamental, ou outras predefinidas em *scripts* de *storyboard* fornecidos pelo usuário. Trata-se de uma interface simples, desenvolvida com a finalidade de apresentar os resultados desenvolvidos por este trabalho, mas que serve como um *framework* para o desenvolvimento de aplicações que incluam o modelo proposto neste trabalho, como por exemplo, um ECA.

⁵Aplicação em andamento, não pertencente ao escopo deste trabalho



Figura 5.6: *Screenshot* da execução de uma animação utilizando a interface provida pelo OSGXface



Figura 5.7: *Screenshot* do protótipo desenvolvido, no seu estágio atual.

6 RESULTADOS

“Insults are effective only where emotion is present.”
— SPOCK, *“Who Mourns for Adonais?”*

Inicialmente, este capítulo apresenta os resultados obtidos na primeira etapa deste trabalho, quando estudou-se o impacto da movimentação provida pelo modelo de Lee em conjunto com expressões faciais. Logo a seguir, são apresentados resultados obtidos pelo protótipo desenvolvido. A fim de validar qualitativamente o resultado visual gerado pelos comportamentos de olhar propostos no modelo, realizou-se uma pesquisa com usuários, que é apresentada no fim deste capítulo.

6.1 Impacto das Sacadas em Conjunto com Expressões Faciais

Os primeiro esforço referente ao estudo da animação de olhos e sua relação com expressividade foi a integração do movimento sacádico de Lee a um ambiente simplificado no *ResponsiveFace*, como descrito na Seção 5.5 do Capítulo 5. A fim de se verificar o impacto do comportamento sacádico em conjunto com as expressões faciais, realizou-se um estudo comparativo [RODRIGUES et al., 2007]. Neste estudo, comparou-se os seguintes estados:

- Modelo sem movimentação do globo ocular (olhos estáticos);
- Modelo com movimentação gerada apenas pelo ruído de Perlin;
- Modelo com movimentação gerada pelo modelo estatístico de Lee, integrada com o ruído de Perlin (utilizado para a geração das piscadas).

Para a avaliação dos resultados, foram reproduzidos no *ResponsiveFace* dois diálogos de filmes de Computação Gráfica conhecidos. O agente representava um dos participantes do diálogo. As emoções foram anotadas em arquivos de configuração e executadas junto com o áudio desses diálogos.

A Tabela 6.1 mostra o esquema do arquivo de configuração de um dos diálogos reproduzidos, a saber, do filme *Shrek 2*¹. Neste diálogo, o personagem desempenhou o papel da Princesa Fiona.

¹Dreamworks (2005)

Tabela 6.1: Diálogo simulado: informação anotada no arquivo de configuração.

Diálogo Simulado			
Turno	Fala do Personagem	Papel da Fiona	Emoção da Fiona
Fiona	Very nice, Shrek	Talking	Angry
Shrek	What? I told you coming here was a bad idea.	Listening	Natural
Fiona	You could've at least tried to get along with my father.	Talking	Angry
Shrek	I don't think I was going to get Daddy's blessing, even if I did want it.	Listening	Natural
Fiona	Do you think it might be nice if somebody asked me what I wanted?	Talking	Half Angry
Shrek	Sure. Do you want me to pack for you?	Listening	Natural
Fiona	You're unbelievable!	Talking	Surprised
Fiona	You're behaving like a...	Talking	Half angry
Fiona	(Silence)	Silence	Disappointed (2 segundos)
Shrek	Go on! Say it!	Listening	Natural
Fiona	Like an ogre!	Talking	Angry
Shrek	Here's a news flash for you! Whether your parents like it or not... I am an ogre!	Listening	Natural
Shrek	ARGHHHHHHHHHHHHHHHHHHHH!	Listening	Frightened
Shrek	And guess what, Princess? That's not about to change.	Listening	Half Angry
Fiona	(Silence)	Silence	Natural (11 segundos)
Fiona	I've made changes for you, Shrek. Think about that.	Talking	Disappointed

A Figura 6.1 mostra alguns *snapshots* da simulação do diálogo. O segundo diálogo representado foi retirado do filme *Ice Age 2*² e o personagem desempenhou o papel da mamute fêmea Ellie.

Com o objetivo de avaliar as emoções representadas pelos olhos dos personagens virtuais, foi feita uma avaliação subjetiva com 11 voluntários. Neste experimento, foi mostrado aos participantes:

1. Os trechos dos filmes de CG com o diálogo desejado;
2. A reprodução dos diálogos no ResponsiveFace sem movimentação do globo ocular;
3. A reprodução dos diálogos no ResponsiveFace com movimentação gerada apenas pelo ruído de Perlin;
4. A reprodução dos diálogos no ResponsiveFace com movimentação gerada pelo modelo estatístico de Lee.

²Blue Sky (2006)

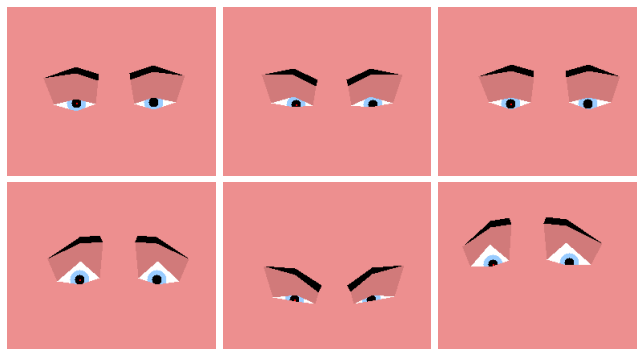


Figura 6.1: *Snapshots* da simulação. Os *snapshots* se referem, respectivamente, à primeira e segunda fala da Fiona, segunda fala do Shrek, quarta e quinta fala da Fiona e sexta fala do Shrek, como detalhado na Tabela 6.1.

Durante a exibição dos vídeos, os participantes preencheram o seguinte questionário, baseado no trabalho de validação de Lee *et al.* [LEE; BADLER; BADLER, 2002]:

- Q1: O personagem na tela pareceu interessado (5) ou indiferente (1) para você?
- Q2: O personagem pareceu engajado na conversação (5) ou distraído (1)?
- Q3: A personalidade do personagem pareceu amigável (5) ou não (1)?
- Q4: O rosto do personagem pareceu ter um olhar "vivo"(5) ou estático, sem vida (1)?
- Q5: O olhar do personagem pareceu emotivo (5) ou sem emoção (1)?
- Q6: No geral, que nota você daria ao desempenho dos personagens frente aos diálogos? (5) Bom (1) Ruim

A Figura 6.2 mostra o resultado da avaliação pelos 11 participantes do experimento.

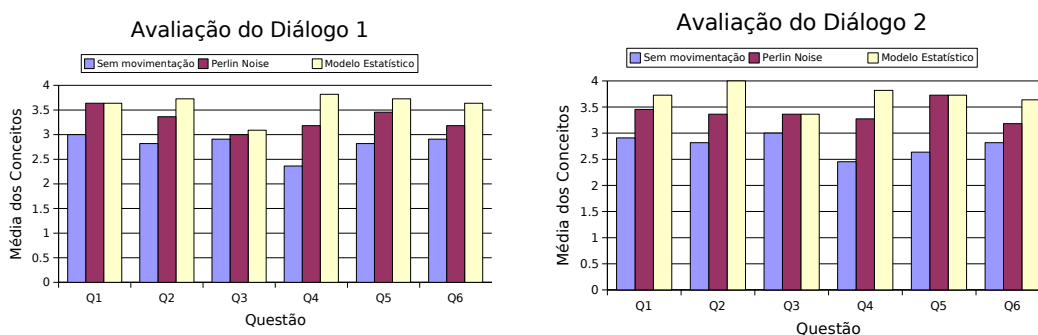


Figura 6.2: Resultados da avaliação subjetiva para o Diálogo 1 (*Shrek 2*) e Diálogo 2 (*Ice Age 2*). O gráfico mostra a pontuação média para cada questão.

Analisando os gráficos, observou-se que para todas as questões, em ambos os filmes, a média da pontuação para o movimento sacádico nunca foi pior do que a dos outros casos. Percebe-se, também, que a ausência de movimentos oculares obteve menor pontuação na avaliação da performance do personagem, como se esperava.

Para as questões 2, 4 e 6, identificou-se que o movimento sacádico teve impacto positivo na performance do personagem, apresentando sempre resultados melhores que os olhos estáticos e que os olhos movimentados apenas com o ruído de Perlin. Isso não ocorre sempre nas questões 1, 3 e 5, onde a pontuação média do modelo estatístico foi muito semelhante à da animação com ruído de Perlin. Quanto ao “parecer interessado na conversa” (questão 1) e as questões relacionadas com emoção (3 e 5), considera-se que o motivo pelo qual o modelo de Lee obteve a avaliação semelhante à do Perlin Noise deve-se ao seguinte fato: pela natureza do modelo, o personagem lança olhares para cima, para baixo ou para os lados, que não possui nenhuma associação com a expressão facial do agente e a semântica do diálogo, o que pode dar uma impressão de “distração” ou incoerência com o engajamento da conversa.

Os resultados da pesquisa foram satisfatórios. Entretanto, percebe-se que a plataforma de animação do ResponsiveFace é um tanto limitada para os estudos de movimentação ocular. Entende-se que essas limitações, apresentadas na Seção 5.5 (Capítulo 5), comprometem a apresentação dos resultados do modelo implementado. Apesar da visualização não constituir o cerne do trabalho, acredita-se que a escolha de uma boa plataforma de animação é essencial para que se consiga avaliar efetivamente os resultados do modelo.

6.2 Comportamentos Gerados no Protótipo

Os comportamentos apresentados no Capítulo 4 foram implementados no protótipo e vários arquivos de *storyboard* e de descrição de olhares foram editados para a geração de animações com esses comportamentos. As figuras 6.3, 6.4 e 6.5 apresentam alguns *snapshots* de animações geradas com os comportamentos *Default*, *Look To* (sem anotações de expressões) e com algumas expressões e comportamentos anotados em arquivos de descrição GDL³.

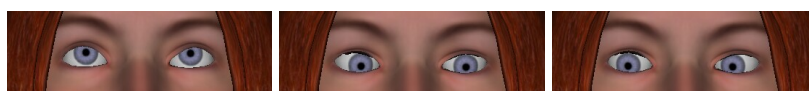


Figura 6.3: *Snapshots* de animação rodando o Modelo Padrão.



Figura 6.4: *Snapshots* de animação rodando com olhares direcionados (anotados), sem expressão facial.

Seguindo a metodologia apresentada na Seção 4.1.1 (Capítulo 4), foram selecionados quatro trechos de filmes. Desses quatro trechos, cada um representando um dos comportamentos propostos no modelo, foram gerados vídeos⁴ a fim de se avaliar o impacto dos comportamentos implementados em usuários, como será descrito na próxima seção.

³Muitos vídeos com animações diversas estão disponíveis em www.fellowsheep.wait4.org/2007-1/TCC/Videos

⁴Disponíveis em <http://www.fellowsheep.wait4.org/2007-1/TCC/quiz>



Figura 6.5: *Snapshots* de animação rodando com expressão anotada em arquivos de descrição GDL (comportamento dos olhos definido na anotação também).

A Tabela 6.2 e a Figura 6.6 identificam os quatro trechos selecionados para os testes e mostram alguns *snapshots* das animações geradas.

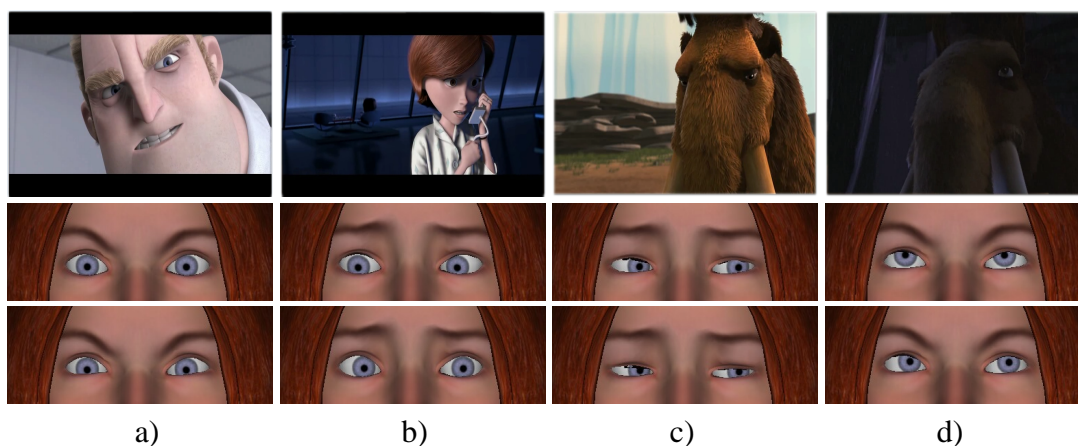


Figura 6.6: *Snapshots* das animações com os comportamentos: a) *Concentration* b) *Discomfort* c) *Distress* e d) *Irony*

Tabela 6.2: Informações técnicas sobre os trechos dos filmes cujos dados foram coletados para a geração das animações ilustradas na Figura 6.6.

Rótulo	Filme	Início	Fim	Estado Afetivo	Comportamento dos Olhos
a)	Os Incríveis	00:29:23	00:29:25	Raiva	<i>Concentration</i>
b)	Os Incríveis	01:00:49	01:00:52	Medos-confirmados	<i>Discomfort</i>
c)	Era do Gelo 2	01:17:27	01:27:34	Vergonha	<i>Distress</i>
d)	Era do Gelo	00:13:05	00:13:06	Arrogância	<i>Irony</i>

6.3 Avaliação dos Comportamentos com Usuários

A fim de se avaliar os resultados obtidos pelo modelo, optou-se em fazer uma nova avaliação subjetiva com usuários. Esta avaliação foi realizada através um questionário, via internet⁵, com o intuito de se conseguir mais participantes do que na avaliação anterior, além de proporcionar a possibilidade de se obter opiniões de um público mais variado.

Para esta avaliação, foram disponibilizados 12 vídeos para *download*, sendo que não foi indicado aos voluntários que método de animação de olhos estava sendo utilizado em cada animação:

- **a.avi**: Vídeo contendo a cena inspiradora das animações:

⁵Disponível em <http://www.fellowsheep.wait4.org/2007-1/TCC/quiz>

- **a1.avi**: Animação gerada com o modelo de Lee
- **a2.avi**: Animação gerada com o comportamento *Concentration*
- **b.avi**: Vídeo contendo a cena inspiradora das animações:
 - **b1.avi**: Animação gerada com o comportamento *Discomfort*
 - **b2.avi**: Animação gerada com o modelo de Lee
- **c.avi**: Vídeo contendo a cena inspiradora das animações:
 - **c1.avi**: Animação gerada com o comportamento *Distress*
 - **c2.avi**: Animação gerada com o modelo de Lee
- **d.avi**: Vídeo contendo a cena inspiradora das animações:
 - **d1.avi**: Animação gerada com o modelo de Lee
 - **d2.avi**: Animação gerada com o comportamento *Irony*

O questionário foi dividido em 4 partes, com diferentes objetivos de avaliação:

Parte 1 Avaliação das animações geradas: todas as animações geradas (pelo modelo de Lee e pelos comportamentos propostos) receberam uma nota de 1 a 5, significando:

1 (muito fraco) | 2 (fraco) | 3 (médio) | 4 (bom) | 5 (muito bom)

Nesta etapa, o usuário respondeu as 8 questões seguintes:

- Q1: Dê uma nota de 1 a 5 para o “desempenho” da Alice no video a1.
- Q2: Dê uma nota de 1 a 5 para o “desempenho” da Alice no video a2.
- Q3: Dê uma nota de 1 a 5 para o “desempenho” da Alice no video b1.
- Q4: Dê uma nota de 1 a 5 para o “desempenho” da Alice no video b2.
- Q5: Dê uma nota de 1 a 5 para o “desempenho” da Alice no video c1.
- Q6: Dê uma nota de 1 a 5 para o “desempenho” da Alice no video c2.
- Q7: Dê uma nota de 1 a 5 para o “desempenho” da Alice no video d1.
- Q8: Dê uma nota de 1 a 5 para o “desempenho” da Alice no video d2.

Parte 2 Avaliação da identificação de estados afetivos e relação com os comportamentos de olhar. Nesta etapa, os usuários respondiam as seguintes questões, indicando qual vídeo se identificava mais com o comportamento (ou nenhum, caso não tivessem achado relação):

- Q1: Em alguma das animações a Alice pareceu mais concentrada, olhando fixo para você? Qual?
- Q2: Em alguma das animações a Alice pareceu mais apreensiva, desconfortável? Qual?
- Q3: Em alguma das animações a Alice pareceu não estar bem, meio “tristonha” ou envergonhada? Qual?

Q4: Em alguma das animações a Alice pareceu estar impaciente, ou falando ironicamente? Qual?

Parte 3 Questões mais gerais, a fim de se avaliar as impressões de “vida”, atenção e expressão das emoções do personagem:

Q1: A Alice pareceu ter alguma “vida” (nota 5) ou você achou ela extremamente apática (nota 1)?

Q2: Ela parecia ter algum interesse em estar “falando com/ouvindo você” (nota 5)? Ou parecia estar falando “para as paredes” (nota 1)?

Q3: Ela parecia expressar alguma emoção (nota 5) ou parecia não ter sentimentos (nota 1) enquanto falava/ouvía?

Parte 4 Questões dissertativas, a fim de avaliar aspectos do modelo e animação como um todo:

Q1: Alguma animação pareceu incoerente, no sentido da expressão do que a Alice parecia estar sentindo? (Em outras palavras, teve algumas animações que você realmente não gostou ou sentiu desconforto, no sentido visual?) Qual(is)? Por quê?

Q2: Alguma animação chamou mais sua atenção, no sentido de que ela estava demonstrando claramente seus sentimentos? (Em outras palavras, teve alguma animação que você realmente gostou ou sentiu-se “envolvido”, no sentido visual?) Qual(is)? Por quê?

Q3: A Alice é um protótipo de cabeça falante. Se você fosse melhorar algo na movimentação dela (cabeça, olhos, sobrancelhas, boca, nariz, testa, bochechas), o que você faria primeiro? (Só não vale cabelo).

Vinte e uma pessoas responderam o questionário. Os resultados da Parte 1 do questionário estão expressos na Figura 6.7, mostrando a porcentagem dos conceitos atribuídos pelos usuários a cada um dos vídeos gerados e na Figura 6.8, que faz um comparativo das médias da pontuação obtida pelo Modelo Padrão e pelos comportamentos propostos neste trabalho.

Analisando os gráficos, percebe-se que todos os conjuntos de animações (A, B, C e D) obtiveram pontuação média bastante próximos. De um modo geral, os conceitos atribuídos foram bastante positivos (“médio” a “muito bom”), exceto o conjunto D, que recebeu uma porcentagem significativa de conceito “fraco”. Acredita-se que uma das razões disso está no que mais de um usuário expôs nas questões dissertativas⁶: a animação d, que contém a cena inspiradora, é muito curta (2s, aproximadamente) e um tanto escura. Por isso, as pessoas relataram ter dificuldades em entender o que estava sendo representado.

Percebe-se também que a animação c1, que foi gerada utilizando o comportamento *Distress* foi avaliada como sendo a mais convincente do conjunto de animações apresentadas. Isso se confirma nas questões dissertativas, onde c1 foi colocada por mais de um usuário como a “mais expressiva”. Os voluntários mostraram-se bastante controversos quanto às animações b1 e d2 (comportamentos *Discomfort* e *Irony*). Como se percebe nos gráficos de sua pontuação e no gráfico comparativo entre a média das pontuações entre

⁶As respostas dadas pelos usuários estão disponíveis para consulta, em <http://www.fellowsheep.wait4.org/2007-1/TCC/quiz/dissertativas.html>

o Modelo Padrão e os comportamentos, essas animações obtiveram conceito mais baixo que as geradas com o Modelo Padrão (b2 e d1). Analisando os gráficos em conjunto com as respostas das questões dissertativas, percebeu-se que algumas pessoas estranharam o comportamento dos olhos nessas duas animações (frequência rápida em b1 e o “revirar” dos olhos em d2). Entretanto, observou-se nas questões da Parte 2 do questionário, que são mostradas a seguir, que na hora de atribuir uma das animações aos estados afetivos correspondentes, essas duas animações foram as mais apontadas nos itens referentes aos estados afetivos que elas representam.

A Figura 6.9 mostra os resultados obtidos nas quatro questões da Parte 2 do questionário, que consiste na avaliação da identificação dos estados afetivos pelos usuários nos vídeos observados. Analisando os gráficos da Parte 2, percebe-se que as animações a2, b1, c1, e d2 (geradas com os comportamentos *Concentration*, *Discomfort*, *Distress* e *Irony*, respectivamente), foram as mais indicadas como representativas dos estados afetivos questionados. Estes resultados refletem positivamente, indicando que houve coerência não só nas expressões faciais, mas que o comportamento do olhar contribuiu na identificação dos estados afetivos.

A Figura 6.10 apresenta os resultados obtidos nas três questões da Parte 3 do questionário, que pede ao usuário para avaliar alguns aspectos gerais da animação. Os aspectos avaliados foram o quanto a Alice parece ter “vida” nas animações (Q1), o quanto ela parece estar atenta ao usuário (Q2) e o quanto ela parece expressar emoções (Q3). Para todos estes itens, nota-se que os conceitos atribuídos pelos usuários também foram positivos (“médio” a “muito bom”), em pelo menos 90% das avaliações em cada questão. Em relação à expressividade, os resultados mostram que Alice pareceu, aos voluntários, convincente na expressão de suas emoções nas animações (95.24% das respostas foram “bom” e “muito bom” e o restante “médio”).

As questões dissertativas foram elaboradas de forma que os voluntários pudessem expressar suas opiniões livremente, e mostraram aspectos muito importantes que sugerem melhorias tanto no modelo 3D da Alice como nas animações. Os itens abaixo apresentam algumas dessas considerações, que estão mais relacionadas ao escopo deste trabalho e que merecem destaque:

- Em relação aos olhos, sugere-se a dilatação e contração da pupila;
- O movimento da cabeça às vezes se apresenta de forma muito brusca, sugere-se melhorias;
- Sugere-se também a exploração de rugas na testa, para acentuar a expressividade;
- Sugere-se a melhoria da movimentação das sobrancelhas, que permanecem estáticas em cada expressão facial.

Abaixo apresenta-se outras sugestões, que apesar de não estarem ligadas diretamente ao escopo deste trabalho, são relevantes para a construção de um sistema de animação facial robusto ou ainda uma *talking head* mais convincente:

- Sugere-se melhorias na movimentação da boca (incluindo os dentes, que não acompanham os lábios) e bochechas;
- Um voluntário observou que, em sua opinião, o cabelo tapando as orelhas sugere a “pré-disposição de não ouvir”.

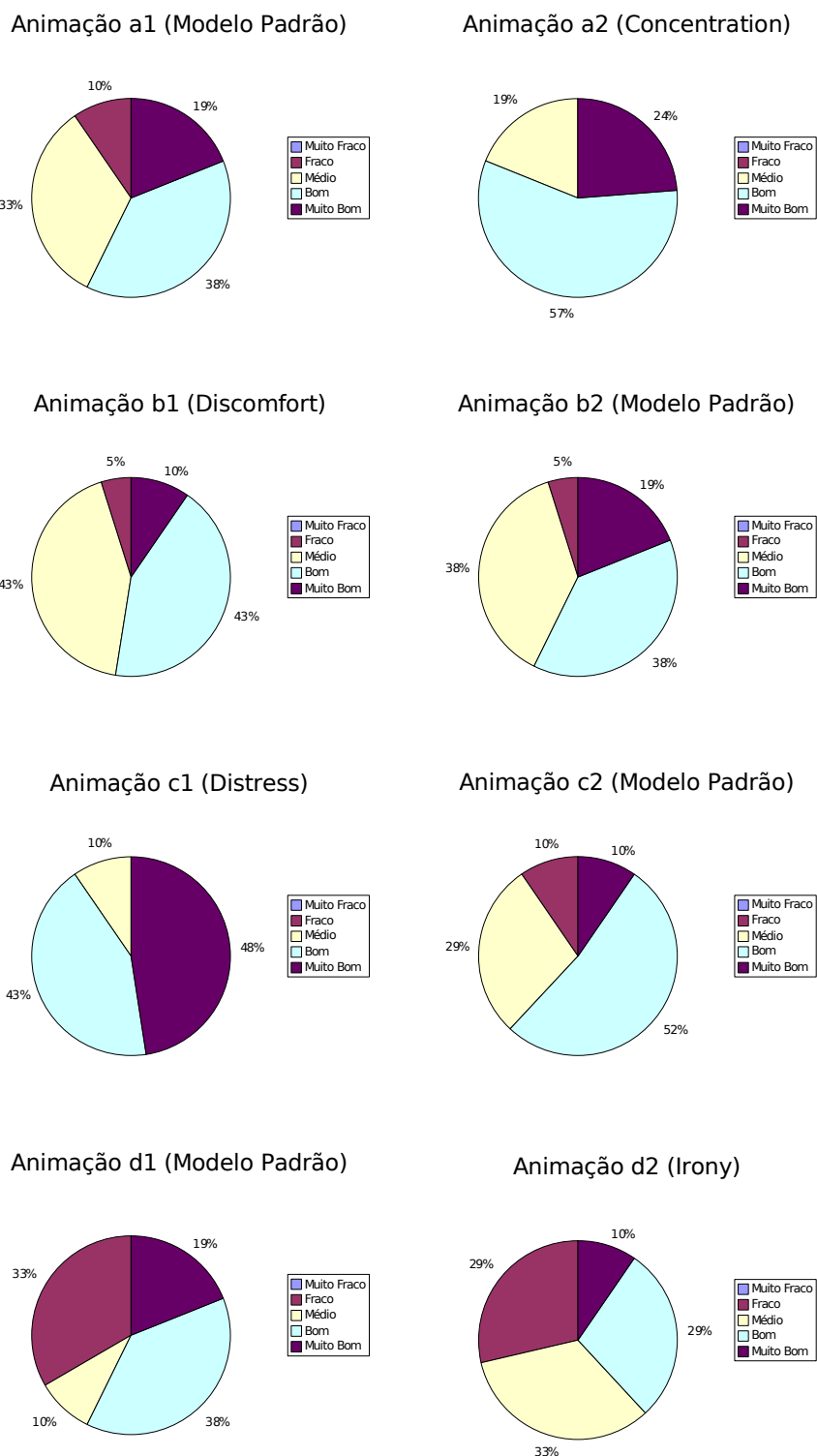


Figura 6.7: Resultados da avaliação subjetiva da Parte 1 do questionário (questões Q1 a Q8): o gráfico mostra a porcentagem dos conceitos atribuídos pelos usuários às oito animações geradas.

Comparativo das Médias dos Conceitos

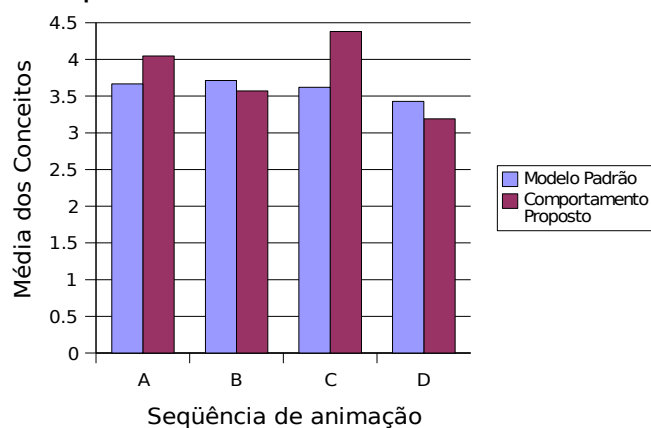


Figura 6.8: Resultados da avaliação subjetiva para da Parte 1 do questionário: o gráfico mostra a pontuação média para cada questão, fazendo um comparativo entre os modelos utilizados.

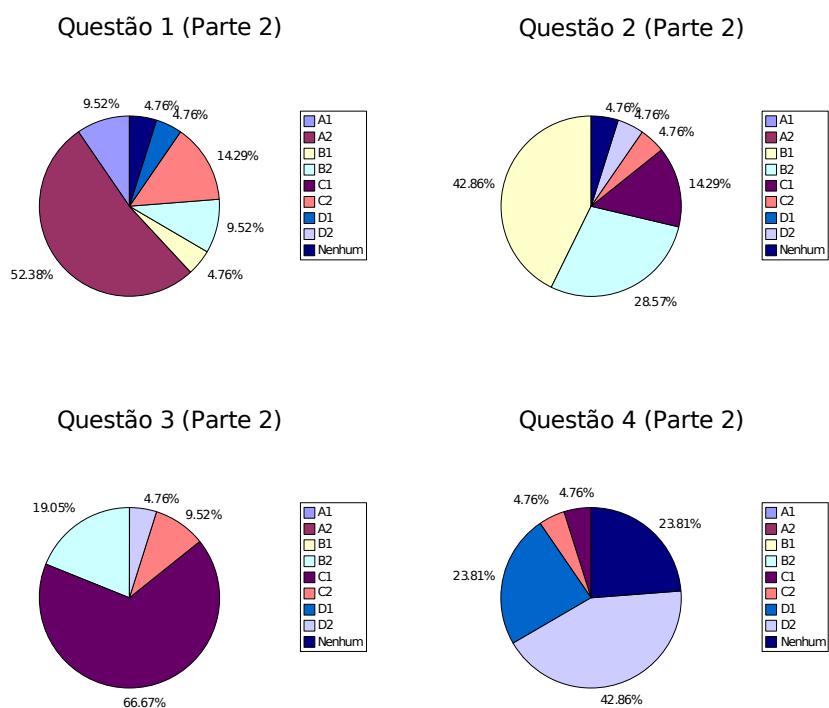


Figura 6.9: Resultados da avaliação subjetiva para as quatro questões da Parte 2 do questionário: o gráfico mostra a porcentagem das respostas dadas pelos usuários para cada questão.

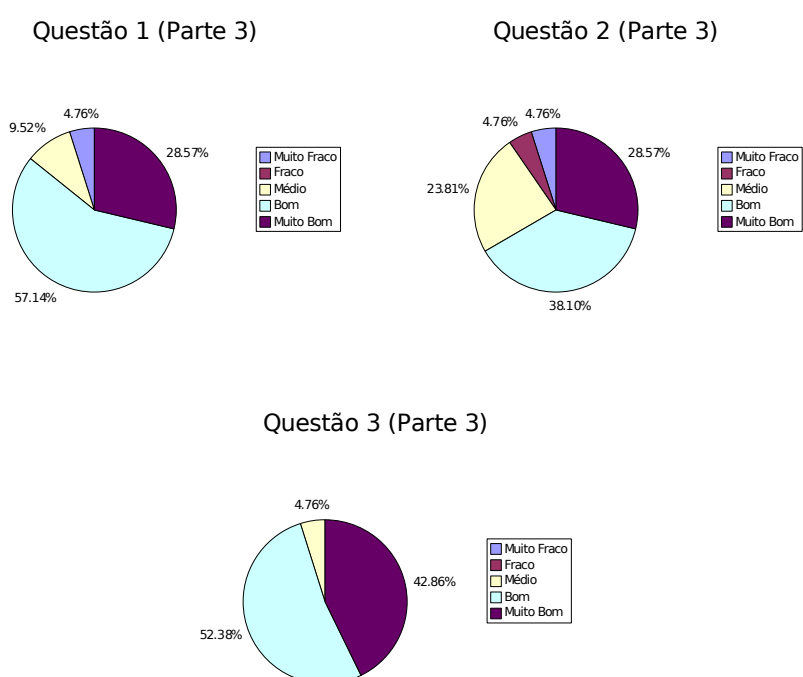


Figura 6.10: Resultados da avaliação subjetiva para as quatro questões da Parte 3 do questionário: o gráfico mostra a porcentagem das respostas dadas pelos usuários para cada questão.

7 CONSIDERAÇÕES FINAIS

“No fim dá tudo certo, ou tão errado, que funciona.”

— LUIS FERNANDO VERISSIMO

Este trabalho apresentou um modelo de geração automática de olhares expressivos, aplicado em uma plataforma de visualização que segue o padrão MPEG-4 de animação facial. Para o desenvolvimento desse modelo, foram feitos estudos e coleta de dados necessários para se estabelecer algumas relações entre estados afetivos, expressividade e comportamentos de olhar.

Os primeiros esforços feitos neste trabalho resultaram na inclusão do comportamento sacádico de Lee junto com expressões faciais e uma avaliação do impacto desse resultado [RODRIGUES et al., 2007]. Esta junção melhora o realismo da animação, porém o modelo estatístico pode gerar, pela sua natureza, olhares incoerentes com a expressão emocional. Por exemplo, algumas emoções requerem um olhar fixo para o parceiro de conversa, enquanto outros produzem olhadas de relance ou o desvio do olhar. De fato, as distribuições de probabilidade obtidas pelos experimentos de Lee não levaram em consideração estados emocionais, gerando apenas um comportamento “padrão” para o modo interlocutor e ouvinte. Por este motivo, o modelo construído propõe alguns comportamentos de olhar diferenciados, para serem aplicados em conjunto com as expressões faciais de maneira a tornar as animações mais coerentes.

As principais contribuições deste trabalho são a metodologia proposta para a coleta de dados e o modelo proposto, que inclui diferentes comportamentos de olhares, o que constitui em um *framework* de animação automática de olhos que pode ser usado em diversas aplicações interativas, como ECAs e Jogos. É importante frisar que o modelo de animação de olhos proposto não é dependente do uso do padrão MPEG-4 de Animação Facial, podendo ser utilizado em diferentes plataformas. Além disso, os dados coletados para a Base de Dados comportamental, descritos em alto-nível através da GDL, podem ser utilizados por outras aplicações que requeiram faces expressivas, também independente da plataforma de animação. Pode-se considerar ainda, como contribuições válidas em nível mais técnico, o protótipo da linguagem de descrição de olhares (GDL), que pode ser facilmente estendida, as melhorias no Xface e o protótipo de visualização, que permite a exibição de faces animadas no ambiente OSG.

Através da inclusão de comportamentos de olhar inspirados na observação de filmes de CG, acrescentou-se um diferencial em relação aos modelos de animação de olhos vistos na bibliografia. Entende-se, pelos resultados da avaliação com usuários, que esses comportamentos precisam ser aplicados no “contexto certo”, isto é, aliado a expressões faciais correspondentes ao estado afetivo do personagem. Por exemplo, notou-se, pelas

notas atribuídas às animações do grupo D (que representavam arrogância, num tom de ironia), em conjunto com algumas críticas nas questões dissertativas, que o fato de a animação inspiradora (cena do filme) d ter sido muito curta e escura, não permitiu a compreensão de qual estado afetivo o personagem representava. Por este motivo, algumas pessoas acharam “estranho” o revirar dos olhos da Alice na animação d2. Outras pessoas criticaram a animação b1, que foi gerada com o comportamento *Discomfort*, uma vez que os olhos de Alice parecem estar um tanto “descontrolados”. De fato, os comportamentos *Irony* e *Discomfort*, em que os olhos comportam-se de forma mais diferenciada que o comportamento “padrão” dos olhos, foram as que provocaram a atribuição dos conceitos mais baixos nas animações. Entretanto, na segunda parte do questionário, onde perguntou-se especificamente sobre em qual animação Alice parecia ter um determinado estado afetivo, os voluntários apontaram essas duas animações (b1 e d2) como as mais coerentes com os estados afetivos que elas realmente deveriam estar representando. Portanto, conclui-se que deve-se atentar para que o usuário compreenda bem o contexto da animação (o que está sendo avaliado nela), para que as movimentações não pareçam “estranhas” ou “incoerentes”.

No geral, percebeu-se que, para este tipo de avaliação, a qualidade da interface de visualização é muito importante: as pessoas, em geral, expressaram sua aprovação ao visual de Alice, deixando também algumas críticas construtivas, mostradas no final da Seção 6.3 do Capítulo 6. Percebeu-se que a problemática da movimentação dos olhos, na percepção de um público geral, é mais sutil do que as expressões faciais em si. Entretanto, acredita-se que, de forma alguma esse fato desmereça os estudos realizados. Os resultados referentes à Parte 2 do questionário mostram que os voluntários indicaram justamente as animações com os comportamentos de olhos diferenciados propostos como sendo as mais expressivas. Isso confirma a contribuição dos movimentos dos olhos aliados à expressividade.

7.1 Trabalhos Futuros

Esta seção apresenta alguns tópicos que podem ser explorados futuramente, a fim de melhorar o modelo proposto e o protótipo desenvolvido.

7.1.1 Metodologia de Coleta de Dados

Durante a aplicação da metodologia, identificou-se alguns problemas, que poderiam ser alvos de novas pesquisas e considerados em trabalhos futuros:

- A identificação dos estados afetivos dos personagens nos trechos dos filmes é complexa: em cenas dinâmicas é comum ocorrerem emoções mistas e olhares relacionados à tarefa, além da influência de características individuais dos personagens (personalidade, idade, sexo, cultura, etc.). Estes fatores não pertencem ao escopo definido para o trabalho, não sendo contemplados no modelo proposto;
- Medir e quantificar parâmetros em imagens de filmes não é uma tarefa simples. Optou-se, neste trabalho, em descrever os parâmetros em alto-nível. A intensidade dos parâmetros de animação no modelo proposto, portanto, continuou sendo definida empiricamente, como em outros trabalhos apresentados ([POGGI; PELACHAUD; ROSIS, 2000], por exemplo), exceto a dos parâmetros gerados pelo Modelo Padrão. Pensou-se, como alternativa, em explorar o uso de técnicas de processamento de imagens para se realizar algumas medições, mas de antemão entendeu-se que esta seria

uma tarefa bastante complexa, dado que os personagens são extremamente dinâmicos e sua morfologia corpórea variada (diferentes formas e tamanhos de cabeça e olhos, por exemplo);

- Em relação aos filmes observados, que foram escolhidos de forma a contemplar três diferentes representações de personagens (humanos realísticos, humanos cartuns e não-humanos cartuns), observou-se que, para os fins deste trabalho, os filmes com personagens cartuns foram mais úteis, no sentido de que, nessa linguagem, a obviedade (abstração) nos movimentos contribuíram para uma identificação mais clara de comportamentos de olhar. No filme com humanos realistas observado, supõe-se que a preocupação de manter as feições do rosto fiéis a de humanos reais, fez com que muitas expressões faciais parecessem um tanto “fracas” (poucas alterações nos músculos). Além disso, os olhos dos personagens são pequenos (proporções humanas), o que em cenas onde o enfoque não é um rosto, dificulta a observação. Em vista da aparência realística de Alice, precisou-se tomar cuidado na hora de calibrar os parâmetros dos comportamentos propostos. Com essas observações, sugere-se testes com outras faces e estudos de aperfeiçoamento e extensão do modelo, para verificar se a calibragem dos valores FAP necessita alterações de acordo com a linguagem de representação do personagem.

7.1.2 Consolidação da Base de Dados Comportamental

A Base de Dados Comportamental constitui o diferencial do trabalho realizado, onde estão mapeadas as associações entre expressividade e comportamentos de olhar que o modelo contempla. Entretanto, fornecer uma rica base de dados para o modelo é uma tarefa que exige tempo, tanto para a observação das cenas e determinação em alto-nível dos parâmetros, quanto para o ajuste das intensidades de cada um. Portanto, entende-se que o modelo se mostrará mais rico e robusto no momento em que a Base de Dados Comportamental fornecer entrada para uma grande variedade de estados afetivos. Assim, o modelo poderia ser incorporado em diversas ferramentas, reduzindo ou eliminando a necessidade dos desenvolvedores editarem *scripts* com a descrição dos comportamentos desejados, focando-se apenas em sua aplicação.

7.1.3 Inclusão de Novos Comportamentos

Neste trabalho, foram propostos quatro comportamentos diferentes, inspirados na observação subjetiva de filmes. A inclusão de novos comportamentos, sem dúvidas, enriqueceria o modelo. Pensa-se na possibilidade de observar os fatores propostos por [FUKAYAMA et al., 2002], relativos a “quantidade de olhar” (tempo em que um participante olha para o outro durante uma interação) nos filmes, relacionando com estados afetivos.

Além disso, este trabalho poderia ser estendido de forma que as sobrancelhas e pálpebras também pudessem ser descritos em mais alto nível. Outro item que certamente trará melhoras na qualidade das animações será contemplar a dilatação e diminuição da pupila (suportado pelo padrão MPEG-4, mas que foi inviável de implementar devido a restrições no modelo do rosto utilizado). Já o item observado umidade dos olhos, que corresponde ao brilho do olhar, necessitaria de estudos mais avançados e inclusão de técnicas de renderização.

7.1.4 Melhorias nas Expressões Faciais

Uma vez que o foco deste trabalho foram os olhos (globo ocular), percebeu-se (e as avaliações comprovam isso) que a parte das expressões, principalmente envolvendo as sobrancelhas, necessita de um refino. Especificamente sobre as sobrancelhas, encontrou-se algumas dificuldades de representação das expressões devido a algumas limitações do modelo facial, provavelmente no mapeamento dos FPs e FAPs. Também as rotinas de deformação de face do Xface poderiam ser estudadas, mas isso foge do escopo deste trabalho, especificamente. As pálpebras também possuem alguns *bugs* na parte do mapeamento FPs – FAPs, sendo perceptíveis alguns artefatos nas imagens, dependendo da animação. Sabe-se, no entanto, que o mapeamento dos pontos de controle aos parâmetros de animação é uma tarefa manual; no caso de Alice, aproveitou-se esta parte pronta, vinda do pacote do Xface.

A fim de tornar as animações das sobrancelhas mais dinâmicas nas expressões, pensa-se na possibilidade de acrescentar o ruído de Perlin [PERLIN, 1997], que era utilizado no primeiro protótipo para dar um aspecto de dinamismo aleatório à cabeça e na geração das piscadas.

7.1.5 Integração com um Sistema de Sincronização de Fala

Rostos humanos sintéticos, em especial, necessitam de animações convincentes da boca, uma vez que sua movimentação desempenha um papel primordial como veículo da comunicação verbal. Além disso, são fundamentais na expressividade do rosto.

O trabalho desenvolvido deixa claro que o foco de estudo eram os olhos. Entretanto, com o desenvolvimento da GDL, percebeu-se que, dado a facilidade de extensão da mesma, a futura inclusão de ações de boca é perfeitamente possível de se implementar. Pode-se observar, nos vídeos disponíveis em <http://www.fellowsheep.wait4.org/2007-1/TCC/Videos> que algumas funções simples de boca foram implementadas (sorrisos e uma simulação simples de fala), com o intuito de complementar a animação do rosto. Pensa-se em, como trabalhos futuros, incluir no protótipo algum modelo que forneça movimentações sincronizadas de fala e expressão, como, por exemplo, o de [RODRIGUES; FEIJÓ; VELHO, 2003].

7.1.6 Integração com um Modelo Emocional

A construção do modelo proposto neste trabalho sempre levou em consideração a fácil integração com algum modelo emocional, que fornecesse os estados afetivos do agente, ou seja, as “ações” de entrada para o modelo proposto. A interface com um modelo de emoção pode ser facilmente implementada: basta traduzir o estado afetivo fornecido pelo modelo emocional em um *script* de *storyboard* GDL (considerando que a Base de Dados Comportamental contenha a descrição de todos os estados afetivos contemplados pelo modelo emocional).

7.1.7 Melhorias no Protótipo e Testes com Outros Modelos de *Talking Heads*

O protótipo desenvolvido pode ser facilmente estendido para suportar outros modelos de *talking heads*. Isso possibilitaria a avaliação do modelo em outras cabeças, diferentes de Alice, a fim de se verificar a robustez desse. Além disso, a implementação do protótipo poderia ser melhorada a fim de que se disponibilizasse um *framework* de animação facial para o ambiente OSG.

APÊNDICE A LISTA DOS FAPS

Este apêndice apresenta a definição de todos os FAPs do padrão MPEG-4 de Animação Facial. A Tabela A.1 mostra a relação de todos eles, sua função e agrupamento, por área do rosto. A seguir, a Tabela A.2 mostra os grupos ao qual os FAPs pertencem. O trabalho focou-se nos FAPs dos grupos 3, 4 e 7.

Tabela A.1: Definições dos FAPs. Fonte: [LEE, 2002]

#	Nome do FAP	Descrição do FAP	(U)ni ou (B)idirecional	Movimento Positivo	Grupo
1	viseme	um conjunto de valores para representação dos visemas	n/a	n/a	1
2	expression	um conjunto de valores para a representação de expressões faciais	n/a	n/a	1
3	open_jaw	deslocamento vertical da mandíbula	U	para baixo	2
4	lower_t_midlip	deslocamento vertical da parte média do lábio superior interno	B	para baixo	2
5	raise_b_midlip	deslocamento vertical da parte média do lábio inferior interno	B	para cima	2
6	stretch_l_cornerlip	deslocamento horizontal do canto do lábio esquerdo interno	B	esquerda	2
7	stretch_r_cornerlip	deslocamento horizontal do canto do lábio direito interno	B	direita	2
8	lower_t_lip_lm	deslocamento vertical do ponto central entre o canto esquerdo e a parte média do lábio superior interno	B	para baixo	2
9	lower_t_lip_rm	deslocamento vertical do ponto central entre o canto direito e a parte média do lábio superior interno	B	para baixo	2
10	raise_b_lip_lm	deslocamento vertical do ponto central entre o canto esquerdo e a parte média do lábio inferior interno	B	para cima	2

11	raise_b_lip_rm	deslocamento vertical do ponto central entre o canto direito e a parte média do lábio inferior interno	B	para cima	2
12	raise_l_cornerlip	deslocamento vertical do canto do lábio esquerdo interno	B	para cima	2
13	raise_r_cornerlip	deslocamento vertical do canto do lábio direito interno	B	para cima	2
14	thrust_jaw	deslocamento em profundidade da mandíbula	U	para frente	2
15	shift_jaw	deslocamento lado a lado da mandíbula	B	direita	2
16	push_b_lip	deslocamento em profundidade da parte média do lábio inferior interno	B	para frente	2
17	push_t_lip	deslocamento em profundidade da parte média do lábio superior interno	B	para frente	2
18	depress_chin	movimento de compressão para cima do queixo (como na expressão da tristeza)	B	para cima	2
19	close_t_l_eyelid	deslocamento vertical da pálpebra superior esquerda	B	para baixo	3
20	close_t_r_eyelid	deslocamento vertical da pálpebra superior direita	B	para baixo	3
21	close_b_l_eyelid	deslocamento vertical da pálpebra inferior esquerda	B	para cima	3
22	close_b_r_eyelid	deslocamento vertical da pálpebra inferior direita	B	para cima	3
23	yaw_l_eyeball	rotação horizontal do olho esquerdo	B	esquerda	3
24	yaw_r_eyeball	rotação horizontal do olho direito	B	para cima	3
25	pitch_l_eyeball	rotação vertical do olho esquerdo	B	para baixo	3
26	pitch_r_eyeball	rotação vertical do olho direito	B	para baixo	3
27	thrust_l_eyeball	deslocamento em profundidade do olho esquerdo	B	para frente	3
28	thrust_r_eyeball	deslocamento em profundidade do olho direito	B	para frente	3
29	dilate_l_pupil	dilatação da pupila esquerda	B	crescente	3
30	dilate_r_pupil	dilatação da pupila direita	B	crescente	3
31	raise_l_i_eyebrow	deslocamento vertical da sobrancelha interna esquerda	B	para cima	4
32	raise_r_i_eyebrow	deslocamento vertical da sobrancelha interna direita	B	para cima	4
33	raise_l_m_eyebrow	deslocamento vertical da sobrancelha média esquerda	B	para cima	4
34	raise_r_m_eyebrow	deslocamento vertical da sobrancelha média direita	B	para cima	4
35	raise_l_o_eyebrow	deslocamento vertical da sobrancelha externa esquerda	B	para cima	4

36	raise_r_o_eyebrow	deslocamento vertical da sobrancelha externa direita	B	para cima	4
37	squeeze_l_eyebrow	deslocamento horizontal da sobrancelha esquerda	B	direita	4
38	squeeze_r_eyebrow	deslocamento horizontal da sobrancelha direita	B	esquerda	4
39	puff_l_cheek	deslocamento horizontal da bochecha esquerda	B	esquerda	5
40	puff_r_cheek	deslocamento horizontal da bochecha direita	B	direita	5
41	lift_l_cheek	deslocamento vertical da bochecha esquerda	U	para cima	5
42	lift_r_cheek	deslocamento vertical da bochecha direita	U	para cima	5
43	shift_tongue_tip	deslocamento horizontal da ponta da língua	B	direita	6
44	raise_tongue_tip	deslocamento vertical da ponta da língua	B	para cima	6
45	thrust_tongue_tip	deslocamento em profundidade da ponta da língua	B	para frente	6
46	raise_tongue	deslocamento vertical da língua	B	para cima	6
47	tongue_roll	enrolar a língua em forma de “U”	U	côncavo para cima	6
48	head_pitch	rotação vertical da cabeça	B	para baixo	7
49	head_yaw	rotação horizontal da cabeça	B	esquerda	7
50	head_roll	rotação da cabeça em direção aos ombros (inclinado para os lados)	B	direita	7
51	lower_t_midlip_o	deslocamento vertical da parte média do lábio superior externo	B	para baixo	8
52	raise_b_midlip_o	deslocamento vertical da parte média do lábio inferior externo	B	para cima	8
53	stretch_l_cornerlip_o	deslocamento horizontal do canto do lábio esquerdo externo	B	esquerda	8
54	stretch_r_cornerlip_o	deslocamento horizontal do canto do lábio direito externo	B	direita	8
55	lower_t_lip_lm_o	deslocamento vertical do ponto central entre o canto esquerdo e a parte média do lábio superior externo	B	para baixo	8
56	lower_t_lip_rm_o	deslocamento vertical do ponto central entre o canto direito e a parte média do lábio superior externo	B	para baixo	8
57	raise_b_lip_rm_o	deslocamento vertical do ponto central entre o canto esquerdo e a parte média do lábio inferior externo	B	para cima	8

58	raise_b_lip_rm_o	deslocamento vertical do ponto central entre o canto esquerdo e a parte média do lábio inferior externo	B	para cima	8
59	raise_l_cornerlip_o	deslocamento vertical do canto esquerdo externo do lábio	B	para cima	8
60	raise_r_cornerlip_o	deslocamento vertical do canto direito externo do lábio	B	para cima	8
61	stretch_l_nose	deslocamento horizontal do lado esquerdo do nariz	B	esquerda	9
62	stretch_r_nose	deslocamento horizontal do lado direito do nariz	B	esquerda	9
63	raise_nose	deslocamento vertical da ponta do nariz	B	para cima	9
64	band_nose	deslocamento horizontal da ponta do nariz	B	direita	9
65	raise_l_ear	deslocamento vertical da orelha esquerda	B	para cima	10
66	raise_r_ear	deslocamento vertical da orelha direita	B	para cima	10
67	pull_l_ear	deslocamento horizontal da orelha esquerda	B	esquerda	10
68	pull_r_ear	deslocamento horizontal da orelha direita	B	direita	10

Tabela A.2: Grupos dos FAPs. Fonte: [LEE, 2002]

Grupo	Número de FAPs
1: visemas e expressões	2
2: mandíbula, queixo, lábio superior e inferior interno, cantos dos lábios interno e parte central dos lábios	16
3: olhos, pupilas e pálpebras	12
4: sobrancelhas	8
5: bochechas	4
6: língua	5
7: cabeça	3
8: posições externas dos lábios	10
9: nariz	4
10: orelhas	4

APÊNDICE B SINTAXE DA GDL

Esta seção apresenta a sintaxe da linguagem de descrição de parâmetros criada para o trabalho. A GDL ainda é um protótipo de linguagem. Trata-se, na verdade, de *scripts* em Lua seguindo a sintaxe apresentada a seguir.

B.1 Storyboard de Ações

Os arquivos de *storyboard* de ações são muito simples. Seguem a seguinte sintaxe:

```

<Sequência de Ações> ::= ‘Storyboard = {’
    ( <Estado Afetivo> ‘,’
      | <Comportamento> ‘,’
    )+
    ‘}’

```

O “Estado Afetivo” deve existir na Base de Dados comportamental, como um arquivo de descrição de expressão. Trata-se do nome do estado afetivo desejado, como por exemplo “tristeza” ou “alegria”, desde que exista na Base de Dados comportamental os arquivos de descrição *tristeza.gdl* e *alegria.gdl*. Além disso, deve-se indicar o tempo da animação com o estado afetivo desejado, em número de frames.

```

<Estado Afetivo> ::= ‘{’ <Nome> ‘,’ <Tempo> ‘}’
<Nome> ::= ‘”’ <Caracter diferente de “aspas”>+ ‘”’
<Tempo> ::= (0...9)+

```

B.2 Comportamentos do Olhar

O “Comportamento” refere-se diretamente ao tipo de movimentação de olhos que o personagem vai ter. Quando indicado diretamente no arquivo de *storyboard*, o restante da face se mantém neutra, apenas com o comportamento de olhar indicado. Os comportamentos reconhecidos pela GDL são os descritos no Capítulo 4. O esquema a seguir mostra a sintaxe (nome do comportamento e parâmetros) de cada um deles. É importante notar que os diferentes comportamentos possuem, em sua maioria, parâmetros diferenciados. Este, inclusive, é o motivo principal pelo qual se julgou necessária esta documentação.

$\langle \text{Comportamento} \rangle ::= \{ \langle \text{Nome do Comportamento e seus Parâmetros} \rangle \}$
 $\langle \text{Nome do Comportamento e seus Parâmetros} \rangle ::= \text{"default"} \text{' , ' } \langle \text{Modo} \rangle \text{' , ' }$
 $\quad | \text{"static"} \text{' , ' }$
 $\quad | \text{""} \langle \text{Cognitivo} \rangle \text{" " ' , ' } \langle \text{Magnitude} \rangle \text{' , ' }$
 $\quad | \text{"lookTo"} \text{' , ' } \langle \text{Direção} \rangle \text{' , ' } \langle \text{Magnitude} \rangle \text{' , ' }$
 $\quad | \text{"concentration"} \text{' , ' } \langle \text{Alcance das Magnitudes} \rangle \text{' , ' }$
 $\quad | \text{"discomfort"} \text{' , ' } \langle \text{Frequência} \rangle \text{' , ' } \langle \text{Alcance das Magnitudes} \rangle \text{' , ' }$
 $\quad | \text{"distress"} \text{' , ' }$
 $\quad | \text{"irony"} \text{' , ' }$
 $\quad \langle \text{Tempo} \rangle$

 $\langle \text{Cognitivo} \rangle ::= \text{'visualConstruct'} | \text{'visualRecall'} | \text{'audioConstruct'} |$
 $\quad \text{'audioRecall'} | \text{'kinesthetic'} | \text{'internaDialogue'}$

 $\langle \text{Modo} \rangle ::= (\text{"listening"} | \text{"talking"})$

 $\langle \text{Magnitude} \rangle ::= (0.0 \dots 27.5)$

 $\langle \text{Direção} \rangle ::= \text{""} (\text{'up'} | \text{'down'} | \text{'left'} | \text{'right'} | \text{'lup'} | \text{'ldown'} | \text{'rup'} |$
 $\quad \text{'rdown'}) \text{" "}$

 $\langle \text{Frequência} \rangle ::= (0.0 \dots 1.0)$

 $\langle \text{Alcance das Magnitudes} \rangle ::= (0.0 \dots 1.0)$

Esses comportamentos também podem estar anotados, seguindo a mesma sintaxe, em um arquivo de descrição das expressões.

B.3 Descrição das Expressões

O arquivo de descrição das expressões segue a seguinte sintaxe:

$\langle \text{Seqüência de Ações} \rangle ::= \text{'Animation = \{}$
 $\quad \langle \text{Descrição das Sobrancelhas} \rangle$
 $\quad \langle \text{Descrição das Pálpebras} \rangle$
 $\quad \langle \text{Descrição dos Olhos} \rangle$
 $\quad \langle \text{Descrição da Cabeça} \rangle$
 $\quad \langle \text{Descrição da Boca} \rangle$
 $\quad \text{'\}'}$

A “Descrição das Sobrancelhas” fornece os parâmetros para cada sobrancelha, seguindo a seguinte sintaxe:

$\langle \text{Descrição das Sobrancelhas} \rangle ::= \text{'eyebrows = \{}$
 $\quad \langle \text{Parte Interna das Sobrancelhas} \rangle$
 $\quad \langle \text{Parte Média das Sobrancelhas} \rangle$
 $\quad \langle \text{Parte Externa das Sobrancelhas} \rangle$
 $\quad \text{'\}'}$

$\langle \text{Parte Interna das Sobrancelhas} \rangle ::= \text{'inner = \{}$
 $\langle \text{Sobrancelhas} \rangle$
 $\text{'},'$

$\langle \text{Parte Média das Sobrancelhas} \rangle ::= \text{'medial = \{}$
 $\langle \text{Sobrancelhas} \rangle$
 $\text{'},'$

$\langle \text{Parte Externa das Sobrancelhas} \rangle ::= \text{'outer = \{}$
 $\langle \text{Sobrancelhas} \rangle$
 $\text{'},'$

$\langle \text{Sobrancelhas} \rangle ::= \text{'\{ "1" ', \langle \text{Movimento da Sobrancelha} \rangle ', \langle \text{Intensidade} \rangle '\},'$
 $\text{'\{ "r" ', \langle \text{Movimento da Sobrancelha} \rangle ', \langle \text{Intensidade} \rangle '\},'$
 $\text{'\{ "1r" ', \langle \text{Movimento da Sobrancelha} \rangle ', \langle \text{Intensidade} \rangle '\},'$

$\langle \text{Movimento da Sobrancelha} \rangle ::= \text{'" ('up' | 'down') "'}$

$\langle \text{Intensidade} \rangle ::= (0.0 \dots 1.0)$

O parâmetro “Intensidade” regula o quanto a parte da sobrancelha vai subir ou descer; o valor 0.0 deixa inalterado e 1.0 é o valor máximo do FAP definido no módulo Gerador de Olhares Expressivos. De forma semelhante, segue-se a descrição das pálpebras:

$\langle \text{Descrição das Pálpebras} \rangle ::= \text{'eyelids = \{}$
 $\langle \text{Pálpebra Superior} \rangle$
 $\langle \text{Pálpebra Inferior} \rangle$
 $\langle \text{Piscadas} \rangle$
 $\text{'},'$

$\langle \text{Pálpebra Superior} \rangle ::= \text{'upper = \{}$
 $\langle \text{Pálpebras} \rangle$
 $\text{'},'$

$\langle \text{Pálpebra Inferior} \rangle ::= \text{'lower = \{}$
 $\langle \text{Pálpebras} \rangle$
 $\text{'},'$

$\langle \text{Pálpebras} \rangle ::= \text{'\{ "1" ', \langle \text{Movimento da Pálpebra} \rangle ', \langle \text{Intensidade} \rangle '\},'$
 $\text{'\{ "r" ', \langle \text{Movimento da Pálpebra} \rangle ', \langle \text{Intensidade} \rangle '\},'$
 $\text{'\{ "1r" ', \langle \text{Movimento da Pálpebra} \rangle ', \langle \text{Intensidade} \rangle '\},'$

$\langle \text{Movimento da Pálpebra} \rangle ::= \text{'" ('raised' | 'lowered') "'}$

$\langle \text{Piscadas} \rangle ::= \text{'blink = " ('default' | 'slow' | 'quick') "'},'$

Para a descrição dos olhos (globo ocular), utiliza-se os comportamentos apresentados na seção anterior. Os parâmetros referentes às pupilas e umidade do globo ocular já estão

presentes, para fins de se ter na Base de Dados Comportamental os dados referentes a essas características na descrição dos estados afetivos. Entretanto, como especificado no trabalho, não foram implementadas tais funções no protótipo.

```

<Descrição dos Olhos> ::= 'eyeballs = {'
    <Umidade>
    <Pupila>
    <Comportamento dos Olhos>
    '},'

```

```

<Umidade> ::= 'humidity = "' ('dry' | 'wet' | 'tears') '" , '

```

```

<Pupila> ::= 'pupil = "' ('default' | 'dilated' | 'narrowed') '" , '

```

```

<Comportamento dos Olhos> ::= 'behavior = ' <Comportamento> ' , '

```

O “Comportamento” segue a mesma sintaxe dos arquivos de *storyboard*. A descrição da cabeça é um pouco mais simples: apenas a intensidade do movimento (de acordo com os limiares dos valores FAP determinados no módulo Gerador de Olhares Expressivos) é indicada. Este é um item da GDL que ainda deverá passar por reformas.

```

<Descrição da Cabeça> ::= 'head = {'
    'yaw = ' <Intensidade Pos/Neg> ' , '
    'pitch = ' <Intensidade Pos/Neg> ' , '
    'roll = ' <Intensidade Pos/Neg> ' , ' ' , '

```

```

<Intensidade Pos/Neg> ::= (-1...1)

```

Por fim, como descrito no trabalho, umas poucas funções pré-definidas de boca (não parametrizáveis) foram implementadas. São elas:

```

<Descrição da Boca> ::= 'mouth = ' "' ('none' | 'talk' | 'happyTalk' |
    'closedSmile' | 'smile' | 'laugh' | 'sulk') '" , '

```

Exemplos de arquivos de descrição são mostrados no Apêndice C, que apresenta os quatro *scripts* utilizados para gerar as animações avaliadas na Seção 6.3 do Capítulo 6.

APÊNDICE C SCRIPTS DAS ANIMAÇÕES AVALIADAS

Este apêndice apresenta os arquivos GDL utilizados para a geração das animações utilizadas na validação da Seção 6.3 do Capítulo 6. São apresentados 4 *scripts*, cada qual com as informações das seqüências A, B, C e D de animação. Cada um dos *scripts* contém comentada uma linha, referente ao uso do comportamento *Default*, que foi utilizado para os fins de comparação apresentados no Capítulo 6. Ou seja: cada um desses *scripts* gerou as duas animações de cada seqüência, apenas comentando-se uma das linhas da descrição do comportamento de olhos.

C.1 Animações “A”

```

Animation =
{
  -- Eyebrows Info
  eyebrows = {
    inner = {
      { "l", "down", 0.7 },
      { "r", "down", 0.7 },
    },
    medial = {
      { "r", "down", 0.5 },
      { "l", "down", 0.5 },
    },
    outer = {
      { "l", "up", 0.5 },
      { "r", "up", 0.5 },
    },
  },

  -- Eyelids Info
  eyelids = {
    upper = { "lr", "raised", 0.0 },
    lower = { "lr", "raised", 0.5 },
    blink = "slow",
  },

  -- Eyeballs Info
  eyeballs = {
    humidity = "wet",
    pupil = "default",

    -- For animation a2
    behavior = {"concentration", 0.01},
  },
}

```

```

        -- For animation a1
        --behavior = {"default", "listening"},
    },

-- Head Info
head = {
    pitch = -0.05,
    yaw = 0.0,
    roll = 0.0,
},
mouth = "none",
}

```

C.2 Animações “B”

```

Animation =
{
    -- Eyebrows Info
    eyebrows = {
        inner = {
            { "l", "up", 0.6 },
            { "r", "up", 1.0 },
        },
        medial = {
            { "r", "down", 0.5 },
            { "l", "down", 0.5 },
        },
        outer = {
            { "l", "down", 0.3 },
            { "r", "down", 0.3 },
        },
    },

    -- Eyelids Info
    eyelids = {
        upper = { "lr", "lowered", 0.0 },
        lower = { "lr", "raised", 0.0 },
        blink = "slow",
    },

    -- Eyeballs Info
    eyeballs = {
        humidity = "wet",
        pupil = "default",

        -- For animation b1
        behavior = {"discomfort", 0.9995, 0.7},
        -- For animation b2
        behavior = {"default", "listening"},
    },

-- Head Info
head = {
    pitch = 0.0,
    yaw = 0.0,
    roll = 0.0,
},
}

```

```

    mouth = "none",
}

```

C.3 Animações “C”

```

Animation =
{
  -- Eyebrows Info
  eyebrows = {
    inner = {
      { "l", "up", 0.9 },
      { "r", "up", 0.9 },
    },
    medial = {
      { "r", "down", 0.5 },
      { "l", "down", 0.5 },
    },
    outer = {
      { "l", "down", 0.5 },
      { "r", "down", 0.5 },
    },
  },

  -- Eyelids Info
  eyelids = {
    upper = { "lr", "raised", 0.0 },
    lower = { "lr", "raised", 0.2 },
    blink = "default",
  },

  -- Eyeballs Info
  eyeballs = {
    humidity = "wet",
    pupil = "default",

    -- For animation c1
    behavior = {"distress", 0.7},
    -- For animation c2
    behavior = {"default", "talking"},
  },

  -- Head Info
  head = {
    pitch = -0.05,
    yaw = 0.0,
    roll = 0.0,
  },

  mouth = "talking",
}

```

C.4 Animações “D”

```

Animation =
{

```



```
-- Eyebrows Info
eyebrows = {
  inner = {
    { "l", "down", 0.5 },
    { "r", "down", 0.5 },
  },
  medial = {
    { "r", "down", 0.5 },
    { "l", "down", 0.5 },
  },
  outer = {
    { "l", "down", 0.5 },
    { "r", "down", 0.5 },
  },
},

-- Eyelids Info
eyelids = {
  upper = { "lr", "raised", 0.1 },
  lower = { "lr", "lowered", 0.3 },
  blink = "default",
},

-- Eyeballs Info
eyeballs = {
  humidity = "dry",
  pupil = "default",

  -- For animation d2
  behavior = {"irony", 0.5},
  -- For animation d1
  -- behavior = {"default", "talking"},
},

-- Head Info
head = {
  pitch = 0.3,
  yaw = 0.0,
  roll = 0.0,
},

mouth = "talking",
}
```

REFERÊNCIAS

- ARGYLE, M.; COOK, M. **Gaze and Mutual Gaze**. London: Cambridge University Press, 1976.
- BALCI, K. Xface: mpeg-4 based open source toolkit for 3d facial animation. In: AVI '04: PROCEEDINGS OF THE WORKING CONFERENCE ON ADVANCED VISUAL INTERFACES, 2004, New York, NY, USA. **Anais...** ACM Press, 2004. p.399–402.
- BALCI, K. **XFace**. Acesso em 09 de junho de 2007, Disponível em <http://xface.itc.it/>.
- BARROS, L. M. **Diluculum — Making Lua and C++ live happily together**. Acesso em 09 de junho de 2007, Disponível em <http://www.stackedboxes.org/Projects/Diluculum/>.
- BROOKER, D. **Body Language – Eyes**. Acesso em 28 de abril, 2007, Available in <http://www.hash.com/users/threechikens/eyel.html>.
- BURNS, D.; OSFIELD, R. **Open Scene Graph**. Acesso em 09 de junho de 2007, Disponível em <http://www.openscenegraph.org/>.
- CAROLIS, B. D.; PELACHAUD, C.; POGGI, I.; STEEDMAN, M. APML, a Mark-up language for Believable Behavior Generation. **H. Prendinger, ed., Life-like Characters.**, [S.l.], p.65–85, 2004.
- CARPENTER, R. **Movements of the Eyes**. 2.ed. London: Pion, 1988.
- CASELL, J.; TORRES, O.; PREVOST, S. Modeling Gaze Behavior as a Function of Discourse Structure. In: FIRST INTERNATIONAL WORKSHOP ON HUMAN-COMPUTER CONVERSATIONS, 1997, Bellagio, Italy. **Anais...** [S.l.: s.n.], 1997.
- CASELL, J.; VILHJÁLMSOHN, H. H.; BICKMORE, T. BEAT: the behavior expression animation toolkit. In: SIGGRAPH, 2001, New York, EUA. **Anais...** ACM Press, 2001. p.477 – 486.
- CHOPRA-KHULLAR, S.; BADLER, N. I. Where to look? Automating attending behaviors of virtual human characters. In: AGENTS '99: PROCEEDINGS OF THE THIRD ANNUAL CONFERENCE ON AUTONOMOUS AGENTS, 1999, New York, NY, USA. **Anais...** ACM Press, 1999. p.16–23.
- COHEN, M. F.; COLBURN, R. A.; DRUCKER, S. M. The role of eye gaze in avatar mediated conversational interfaces. In: TECHNICAL REPORT MSR-TR-2000-81, 2000. **Anais...** Microsoft Corporation, 2000.
- COHN, J. F. Foundations of human computing: facial expression and emotion. In: ICMI '06: PROCEEDINGS OF THE 8TH INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 2006, New York, NY, USA. **Anais...** ACM Press, 2006. p.233–238.
- COURTY, N.; BRETON, G.; PELÉ, D. Embodied in a Look: bridging the gap between humans and avatars. In: IVA, 2003. **Anais...** [S.l.: s.n.], 2003. p.111–118.
- EKMAN, P. An Argument for Basic Emotions. In: DALGLEISH; POWER, M. (Ed.). **Cognition and Emotion**. [S.l.]: John Wiley & Sons, 1992. p.169–200.

- EKMAN, P. Facial Expression and Emotion. **American Psychologist**, [S.l.], v.46, n.9, p.384–392, April 1993.
- EKMAN, P. Basic Emotions. In: DALGLEISH; POWER, M. (Ed.). **Handbook of Cognition and Emotion**. [S.l.]: John Wiley & Sons, 1999.
- EKMAN, P. Facial Expressions. In: DALGLEISH; POWER, M. (Ed.). **Handbook of Cognition and Emotion**. [S.l.]: John Wiley & Sons, 1999.
- EKMAN, P.; FRIESEN, W. **Facial Action Code System**. Palo Alto, CA: Consulting Psychologists Press, Inc., 1978.
- EKMAN, P.; FRIESEN, W. V.; HAGER, J. C. **Facial Action Coding System – The Manual**. [S.l.]: A Human Face, 2002.
- ELLIOTT, C. D. **The affective reasoner**: a process model of emotions in a multi-agent system. 1992. Tese (Doutorado em Ciência da Computação) — , Evanston, IL, USA.
- FREEDMAN, E. G.; SPARKS, D. L. Coordination of the eyes and head: movement kinematics. **Experimental Brain Research** **131**, [S.l.], p.22 – 32, 2000.
- FRIJDA, N. Varieties of affect: emotions and episodes, moods, and sentiments. In: EKMAN P.; DAVIDSON, R. J. (Ed.). **The Nature of Emotion**. [S.l.]: Oxford University Press, 1994. p.59–67.
- FUKAYAMA, A.; OHNO, T.; MUKAWA, N.; SAWAKI, M.; HAGITA, N. Messages embedded in gaze of interface agents — impression management with agent’s gaze. In: CHI ’02: PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2002, New York, NY, USA. **Anais...** ACM Press, 2002. p.41–48.
- GU, E.; BADLER, N. I. Visual Attention and Eye Gaze During Multiparty Conversations with Distractions. In: IVA, 2006. **Anais...** [S.l.: s.n.], 2006. p.193–204.
- GUSTAVSSON, B.; CEDERLUND, J. et al. **Wings 3D**. Acesso em 09 de junho de 2007, Disponível em <http://www.wings3d.com/>.
- IERUSALIMSCHY, R. **Programming in Lua**. Rio de Janeiro, Brazil: Lua.org, 2003. 288p.
- ITTI, L.; DHAVALÉ, N.; PIGHIN, F. Realistic avatar eye and head animation using a neurobiological model of visual attention. **SPIE Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation VI**, [S.l.], v.5200, p.64–78, 2004.
- JACQUES, P. A.; VICCARI, R. Estado da Arte em Ambientes Inteligentes de Aprendizagem que Consideram a Afetividade do Aluno. **Revista Informática na Educação: Teoria & Prática**, [S.l.], v.8, n.1, 2005.
- KENDON, A. Some Functions of Gaze-direction in Social Interaction. **Acta Psychologica**, [S.l.], v.26, p.22–63, 1967.
- LANCE, B.; STACY, M.; KOIZUMI, D. **Towards Expressive Gaze Manner in Embodied Virtual Agents**. 2004.
- LEE, S. P. **Facial Animation System with Realistic Eye Movement Based on a Cognitive Model for Virtual Agents**. 2002. Tese (Doutorado em Ciência da Computação) — University of Pennsylvania.
- LEE, S. P.; BADLER, J. B.; BADLER, N. I. Eyes Alive. In: SIGGRAPH ’02: PROCEEDINGS OF THE 29TH ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 2002, New York, NY, USA. **Anais...** ACM Press, 2002. p.637–644.
- LEIGH, R.; ZEE, D. **The Neurology of Eye Movements**. 2.ed. [S.l.]: FA Davis, 1991.
- MAES, P. Artificial life meets entertainment: lifelike autonomous agents. **Communications of the ACM**, [S.l.], v.38, p.108 – 114, 1995.

- MORGADO, L. **Integração de Emoção e Raciocínio em Agentes Inteligentes**. 2006. Tese (Doutorado em Ciência da Computação) — Departamento de Informática, University de Lisboa. DI/FCUL TR-06-2.
- MUSSE, S. R. Behavioral Modeling of Virtual Human Actors. **Revista de Informática Teórica e Aplicada - RITA**, [S.l.], v.8, n.2, p.125–141, 2001.
- ORTONY, A.; CLORE, G.; COLLINS, A. **The cognitive structure of emotions**. [S.l.]: UK: Cambridge University Press, 1988.
- ORTONY, A.; TURNER, T. What's basic about basic emotions? **Psychological Review**, [S.l.], v.97, p.315–331, 1990.
- OSTERMANN, J. Animation of Synthetic Faces in MPEG-4. In: CA '98: PROCEEDINGS OF THE COMPUTER ANIMATION, 1998, Washington, DC, USA. **Anais...** IEEE Computer Society, 1998. p.49.
- PARADISO, A. An algebra for combining MPEG-4 compliant facial animations. In: INT. WORKSHOP LIFE LIKE ANIMATED AGENTS: TOOLS, AFFECTIVE FUNCTIONS, AND APPLICATIONS, 2002. **Proceedings...** H. Prendinger [Online], 2002.
- PELACHAUD, C.; BILVI, M. Modelling Gaze Behaviour for Conversational Agents. In: IVA 2003, 2003. **Proceedings...** Springer LINA Series, 2003. p.93–100.
- PERLIN, K. Layered Compositing of Facial Expression. In: ACM SIGGRAPH - TECHNICAL SKETCH, 1997. **Anais...** [S.l.: s.n.], 1997.
- PETERS, C.; PELACHAUD, C.; BEVACQUA, E.; MANCINI, M.; POGGI, I. A model of attention and interest using Gaze behavior. , London, p.229–240, 2005.
- PICARD, R. W. **Affective computing**. [S.l.]: MIT Press, 1997.
- POGGI, I.; PELACHAUD, C. Emotional meaning and expression in animated faces., New York, NY, USA, p.182–195, 2000.
- POGGI, I.; PELACHAUD, C. Signals and meanings of gaze in animated faces. In: LANGUAGE, VISION AND MUSIC, 2002. **Anais...** Mc Kevitt: Paul: Seán Ó Nualláin and Conn Mulvihill, 2002. p.133–144.
- POGGI, I.; PELACHAUD, C.; ROSIS, F. de. Eye Communication in a Conversational 3D Synthetic Agent. **AI Communications**, [S.l.], v.13, n.3, p.169–182, 2000.
- REYNOLDS, C. W. Flocks, herds and schools: a distributed behavioral model. In: SIGGRAPH '87: PROCEEDINGS OF THE 14TH ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1987, New York, NY, USA. **Anais...** ACM Press, 1987. p.25–34.
- RODRIGUES, P. S. L.; FEIJÓ, B.; VELHO, L. Expressive Talking Heads: uma ferramenta com fala e expressão facial sincronizadas para o desenvolvimento de aplicações interativas. In: IX SIMPÓSIO BRASILEIRO DE SISTEMAS MULTIMÍDIA E WEB - WEBMÍDIA 2003, 2003. **Anais...** [S.l.: s.n.], 2003.
- RODRIGUES, P. S.; QUEIROZ, R. B.; BARROS, L.; FEIJÓ, B.; VELHO, L.; MUSSE, S. R. Automatically Generating Eye Motion in Virtual Agents. In: IX SYMPOSIUM ON VIRTUAL AND AUGMENTED REALITY (SVR2007), 2007. **Anais...** [S.l.: s.n.], 2007. p.84–91.
- RODIS, F. de; PELACHAUD, C.; POGGI, I.; CAROFIGLIO, V.; CAROLIS, B. D. From Greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. **Int. J. Hum.-Comput. Stud.**, Duluth, MN, USA, v.59, n.1-2, p.81–118, 2003.
- SCHERER, K. Psychological models of emotion. In: BOROD, J. (Ed.). **The neuropsychology of emotion**. [S.l.]: Oxford University Press, 2000. p.137–162.
- SCHERER, K. Construction of the Social World: emotion. In: HEWSTONE; M., S. W. (Ed.). **Introduction to Social Psychology**. [S.l.]: Blackwell Publishers, 2001.

STROUSTRUP, B. **The C++ Programming Language**. Special Third.ed. Reading, USA: Addison-Wesley, 2000. 911p.

Sun Microsystems, Inc. **Java**. Acesso em 09 de junho de 2007, Disponível em <http://java.sun.com/>.

TERZOPOULOS, D. Artificial life for computer graphics. **Commun. ACM**, New York, NY, USA, v.42, n.8, p.32–42, 1999.

TERZOPOULOS, D.; LEE, Y. Behavioral Animation of Faces: parallel, distributed, and real time. In: FACIAL MODELING AND ANIMATION, ACM SIGGRAPH 2004 COURSE NOTES, 2004. **Anais...** [S.l.: s.n.], 2004. p.119 – 128.

THÓRISSON, K. R. Gandalf: an embodied humanoid capable of real-time multimodal dialogue with people. In: FIRST ACM INTERNATIONAL CONFERENCE ON AUTONOMOUS AGENTS, 1997, California. **Anais...** [S.l.: s.n.], 1997.

THÓRISSON, K. R.; PENNOCK, C.; LIST, T.; DIPIRRO, J. Artificial intelligence in computer graphics: a constructionist approach. **SIGGRAPH Comput. Graph.**, New York, NY, USA, v.38, n.1, p.26–30, 2004.

VERTEGAAL, R.; SLAGTER, R.; VEER, G. van der; NIJHOLT, A. Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In: CHI '01: PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2001, New York, NY, USA. **Anais...** ACM Press, 2001. p.301–308.

VINAYAGAMOORTHY, V.; GARAU, M.; STEED, A.; SLATER, M. An Eye Gaze Model for Dyadic Interaction in an Immersive Virtual Environment: practice and experience. In: COMPUTER GRAPHICS FORUM, 2004. **Anais...** [S.l.: s.n.], 2004. v.23.

WOO, M.; DAVIS; SHERIDAN, M. B. **OpenGL Programming Guide**: the official guide to learning opengl, version 1.2. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1999.