

## 7

### Conclusão

Esta tese apresenta um sistema de animação facial com suporte à fala e à geração de expressões a partir de emoções básicas e derivadas. Para atingir o seu objetivo, o sistema foi desenvolvido buscando aproximar-se de uma parametrização tida como ideal para os sistemas de animação facial, descrita no primeiro capítulo deste documento.

A partir do estudo dos principais modelos de emoção e dos modelos de representação para fenômenos afetivos, como personalidade e humor, desenvolvido no Capítulo 2, este trabalho optou por ter como base o modelo proposto por Plutchik (Plut62)(Plut80).

Considerando as oito emoções puras (também chamadas de básicas) definidas por Plutchik (Plut62)(Plut80), as emoções derivadas são consequência da combinação de duas emoções básicas não opostas. O modelo proposto nesta tese para combinação e geração das emoções derivadas é denominado *VeeM* (*Virtual emotion-to-expression Model*) estando o mesmo detalhadamente descrito no Capítulo 3. Para validação deste modelo é apresentada uma representação matemática definida sobre o espaço canônico no  $R^4$ . Uma abordagem para incorporação de fenômenos afetivos como personalidade, humor, papel da personagem durante a conversação (ouvinte ou falante), contexto e ambiente físico de simulação para a personagem, movimentação de elementos não-verbais como olhos e cabeça, entre outros, é então apresentada.

Para atingir os requisitos de animação facial colocados no Capítulo 1, o sistema foi construído seguindo o padrão MPEG-4 para animação facial. Um estudo sobre esse padrão é apresentado no Capítulo 4, salientando os requisitos para se ter uma malha facial poligonal conforme com esse padrão.

Objetivando visualizar as emoções puras e derivadas originadas a partir do *VeeM*, foi implementada uma ferramenta que representa as emoções através de expressões faciais. Considerando ainda o aspecto do sistema oferecer suporte, idealmente, a um processamento de fala, a ferramenta incorpora um tratamento para áudios capturados que contenham a fala da personagem. A ferramenta cuida então para que haja um sincronismo temporal entre a fala e as expressões faciais que são visualizadas durante a animação facial. O tra-

balho tem por base a implementação de uma face compatível com o padrão MPEG-4 importada do projeto de código aberto *Xface* (Balc04).

A ferramenta desenvolvida, chamada *DynaFeX*, é um ambiente para autoria e apresentação de animações faciais. O Capítulo 5 detalha os aspectos de implementação da ferramenta, descrevendo seus módulos, os possíveis arquivos de entrada, o processo de reconhecimento da fala, o mapeamento dos fonemas para visemas, o mapeamento de visemas para FAPs (representação baixo-nível do padrão MPEG-4), o processo de criação de novas emoções e de refinamento de emoções previamente definidas, o mapeamento das expressões faciais para FAPs, o mecanismo de combinação de FAPs, e os detalhes de execução da animação facial onde fala e expressões são sincronizadas.

É possível destacar contribuições importantes que foram alcançadas com este trabalho de doutorado. A próxima seção dedica-se a enumerar essas contribuições, sendo seguida de uma seção que destaca os trabalhos que podem dar continuidade à pesquisa desta tese.

## 7.1

### Contribuições da Tese

É possível destacar como resultado desta pesquisa as seguintes contribuições:

- Estudo dos principais modelos de emoção e dos aspectos emocionais existentes em um ser humano. Esse estudo é desenvolvido com o objetivo de avaliar qual é um bom modelo para ser utilizado (implementado) em uma ferramenta de síntese e edição de animações faciais.
- Revisão e generalização do círculo emocional de Plutchik, dando origem a um modelo flexível (*VeeM*) para trabalhar com emoções básicas e derivadas.
- Representação matemática do *VeeM* através de um hipercubo emocional definido no espaço canônico do  $R^4$ . Foram selecionadas oito emoções básicas, distribuídas em quatro eixos cartesianos, sendo cada eixo delimitado por duas emoções puras opostas (eixos bipolares). Emoções puras, pertencentes a diferentes eixos, podem ser combinadas para geração das emoções derivadas.
- Incorporação de fenômenos afetivos ao *VeeM*. A personalidade de cada indivíduo é modelada como uma deformação da personalidade “neutra”. Já fenômenos como humor, papel do personagem durante a conversação e outros aspectos que influenciam na dinâmica dos movimentos faciais (por exemplo, as expressões não-verbais), são modelados como funções

monótonas crescentes que quando aplicadas à emoção geram a mesma emoção com alguns parâmetros modificados.

- Estudo e utilização do padrão MPEG-4 com ênfase nos requisitos necessários para definição de um modelo poligonal paramétrico de face. A padronização MPEG-4 para animação facial favorece uma interoperabilidade entre diferentes sistemas faciais. O uso de uma modelagem de face baseada no padrão MPEG-4 também é traduzido em uma maior eficiência para transmissão das animações faciais, uma vez que se torna possível codificar e decodificar apenas o arquivo *fap* descrevendo a animação e o arquivo *fdp* descrevendo a estrutura de definição do modelo facial.
- Implementação da ferramenta *DynaFeX* que incorpora o *VeeM*. Com a ferramenta é possível desenvolver animações faciais com sincronização entre as expressões faciais e a fala da personagem.
- Definição dos FAPs para representação dos visemas e das expressões faciais geradas a partir do *VeeM* (construção das emoções puras e derivadas). O padrão MPEG-4 define apenas um intervalo de valores para os FAPs das seis expressões universais.
- Utilização do módulo *Baldisync* e a especialização de *scripts* para incorporação do reconhecedor de fala na ferramenta *DynaFeX*.
- Incorporação e adaptação da malha poligonal da ferramenta *Xface* para a *DynaFeX*. A malha facial é definida segundo o padrão MPEG-4 para animação facial. Originalmente, a *Xface* trabalha apenas com as seis expressões universais, a *DynaFeX* estende esse conjunto de possíveis expressões faciais a serem visualizadas.
- Desenvolvimento de um estudo para o movimento dos olhos (Rodr07). Os resultados desse estudo são consequência de um trabalho cooperativo entre a PUC-Rio, o IMPA e a Unisinos.
- Aplicação de personagens virtuais 2D/3D com fala e expressões faciais sincronizadas em outros sistemas. Um primeiro estudo de caso desenvolvido incorpora a personagem como apresentadora de programas multimídia interativos, onde a fala da personagem encontra-se sincronizada com outros objetos de mídia como vídeos, imagens, textos etc. (Rod04a) (Rod04b). O segundo estudo de caso incorpora a personagem em um ambiente de *storytelling*. A partir da criação interativa de histórias, a personagem atua como narradora, onde fala e emoções estão sincronizadas com o enredo apresentado (Rodr06).

## 7.2

### Trabalhos Futuros

Um primeiro trabalho futuro que pode ser investigado diz respeito à combinação de FAPs. Um conjunto de testes qualitativos para escolha do peso ideal para contribuição do visema quando os FAPs da região da boca conflitam com a expressão facial da emoção pode ser um projeto interessante de pesquisa.

Outra possibilidade é explorar o uso de mais de dois eixos de emoção do *VeeM* para gerar novas emoções derivadas. Evidentemente, é necessário especificar o vetor da expressão facial ( $\vec{f_{ev}}$ ) para cada nova emoção pura e adaptar as regras de combinação dos FAPs.

Ainda relacionado ao *VeeM*, percebe-se a necessidade de que seja feita uma investigação mais detalhada e com testes para as funções monótonas crescentes que podem ser utilizadas para modelar os fenômenos afetivos. Nessa mesma linha, se faz necessário aplicar (implementar) a idéia de distorção proposta nesta tese sobre o espaço de emoções para tratar a personalidade na animação. Todas as extensões e testes de implementação poderiam ser feitos com base na ferramenta *DynaFeX*.

Em termos de extensão das funcionalidades da ferramenta, pode ser interessante tratar como entrada outros formatos de áudio para a fala da personagem, exigindo, evidentemente, adaptações no processo de reconhecimento da fala para a geração dos arquivos de estrutura fonética.

Outro trabalho futuro é o desenvolvimento de uma linguagem de marcação alto-nível que incorpore, em um único arquivo, todos os atributos emocionais que podem ser utilizados para especificar o comportamento da personagem durante a animação facial na *DynaFeX*. O interessante dessa linguagem seria encapsular todos os arquivos que atualmente são utilizados na *DynaFeX* e, idealmente, que essa linguagem seguisse a proposta do W3C para linguagens de marcação de emoção (W3C07).

Ainda em termos de implementação, um trabalho futuro bastante interessante é a construção de um sistema de inferência de emoções. Esse sistema de inferência poderia ser desenvolvido como uma fase de pré-processamento na ferramenta *DynaFeX*. O sistema receberia como entrada informações relacionadas à personalidade, humor, contexto físico, eventos temporais e, automaticamente, geraria (inferiria) uma curva emocional para a personagem. Essa curva equivaleria à informação descrita no “arquivo *timeline* das emoções”, que atualmente a ferramenta recebe como entrada.

Em termos de modelagem da face, uma idéia interessante é implementar uma malha adaptativa que funcione como um modelo intermediário entre a face MPEG-4 e uma face qualquer. A malha adaptativa teria a responsabilidade de

saber mapear pontos de controle da nova face nos pontos de controle (grupos de FAPs) do modelo MPEG-4. Através dessa malha adaptativa, *DynaFeX* poderia receber diferentes modelos de face que seriam adaptados para uma malha do padrão MPEG-4.

Um trabalho futuro semelhante ao trabalho anterior, é propor um sistema de reconstrução como nos trabalhos (Blan03) (Blan04). Uma vantagem sobre esses trabalhos de reconstrução seria trabalhar com uma malha facial estruturada e poder incorporar expressões faciais dinâmicas relacionadas à fala.

Outro trabalho futuro é focar em testes que validem a interoperabilidade do padrão MPEG-4. Seria bastante interessante ter um conjunto de diferentes malhas faciais especificadas segundo o padrão incorporadas na *DynaFeX* e permitir que, interativamente, o usuário escolha em qual malha ele quer visualizar as expressões faciais geradas pelo *VeeM*. Ainda ligado a testes sobre o MPEG-4, seria interessante estender os testes para transmissão dos arquivos *faps* através de um codificador.

Voltando à incorporação de novas funcionalidades na ferramenta *DynaFeX*, um trabalho futuro é incorporar o trabalho de movimento dos olhos (Rodr07) na *DynaFeX*. Outro trabalho seria estudar e incorporar o movimento da cabeça, de forma independente do movimento dos olhos.

Por fim, a ferramenta *DynaFeX* atualmente é uma implementação que não é leve e não favorece a portabilidade (entre diferentes sistemas operacionais), o que dificulta a sua integração com outras aplicações. Um trabalho para que essa limitação seja contornada e os resultados obtidos nesta pesquisa possam ser integrados a outras aplicações, como foi feito com o “*Expressive Talking Heads*” e as aplicações multimídia interativas e os sistemas de *storytelling*, seria uma importante contribuição para difundir o uso da *DynaFeX*.