

# Expressive Talking Heads

Defesa de Proposta de Dissertação



Por: Paula Salgado Lucena

Orientadores: Luiz Velho e Marcelo Gattass

19-fevereiro-2002





# Estrutura da Apresentação

- Introdução
- Aspectos Conceituais
- Trabalhos Relacionados
- O *Expressive Talking Heads*
- Considerações Finais
- Apresentação do Sistema



# Introdução

- Motivação

- ✍ Por que animar a face humana?

- Desafio

- ✍ Comunicação entre um **personagem virtual** e o usuário através da **fala expressiva** em um sistema multimídia em **tempo real**.



# Aspectos Conceituais

## Palavras-Chaves:

*fonema* *difone* *trifone* *visema* *lip sync* *talking head*

## Fala

- Em sistemas *talking head* a fala está diretamente ligada ao áudio que é reproduzido junto com a animação facial.

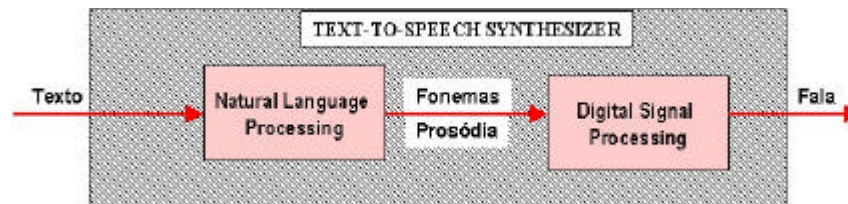
### Abordagens:

- voz capturada
- voz sintetizada



Voz **capturada**: gravar a fala a ser reproduzida.

Voz **sintetizada**: áudio gerado através de um sistema *text-to-speech*.



# Aspectos Conceituais

## Face

- ✍ Abordagens para modelagem e animação da face humana.
- ✍ Modelagem ✍ duas abordagens:
  - ✍ Face definida através de imagens capturadas.
  - ✍ Face definida através de um modelo geométrico.

### Face **Capturada**:

- imagem bidimensional ou tridimensional
- animação feita através da técnica de metamorfose (*morphing*)

### Face **Sintetizada**:

- modelo geométrico pode ser bidimensional ou tridimensional
- adição de textura produzindo um efeito de maior detalhe na face humana
- animação sobre os músculos da expressão facial

### Estilo:

**realista**: semelhança com fisionomia humana



**caricatural**: presença de fatores de distorção ou exagero

# Aspectos Conceituais

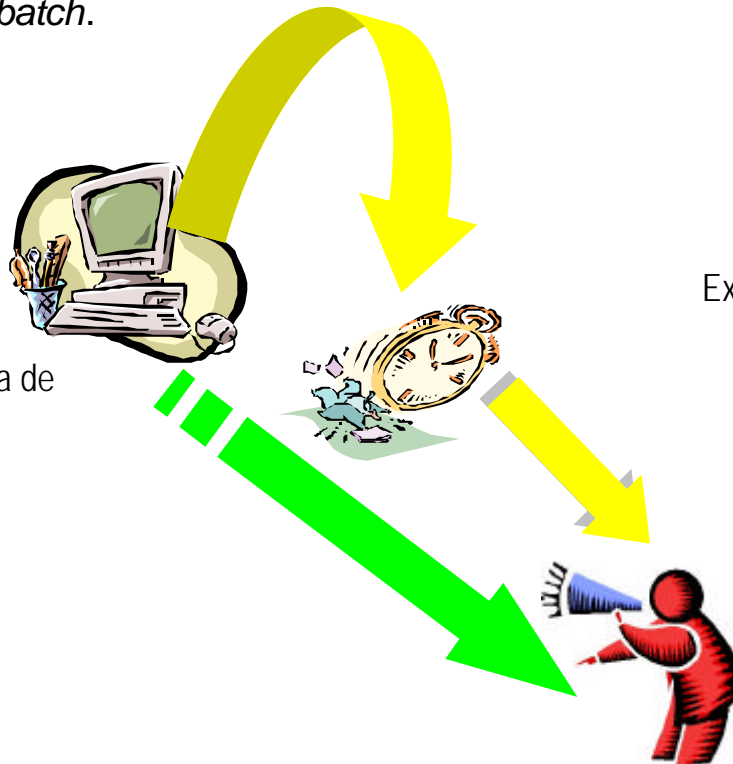
## Formas de Execução

✍ São consideradas duas abordagens para a forma de execução de uma aplicação *talking head*:

- ✍ Execução em tempo real.
- ✍ Execução em *batch*.

### Execução em **Tempo Real**:

- interação com o usuário
- sistemas *talking head*: a entrada de dados permanece ocorrendo em paralelo com a animação facial.

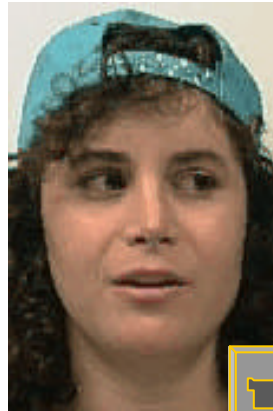


### Execução em **Batch**:

- abordagem passiva
- sistemas *talking head*: produção de um vídeo da animação e posterior apresentação.

# Trabalhos Relacionados

## 🌐 Vídeo-Rewrite



### Definição:

Sistema que faz uso de uma sequência de vídeo existente para automaticamente criar um novo vídeo com o mesmo contexto mas com uma nova trilha sonora.

### Classificação:

- ✍ Fala: **capturada**
- ✍ Face: **capturada**
- ✍ Forma de Execução: **em batch**

## 🌐 MikeTalk

### Definição:

Sistema de síntese de fala texto -visual que converte um texto de entrada em um fluxo de fala audiovisual.

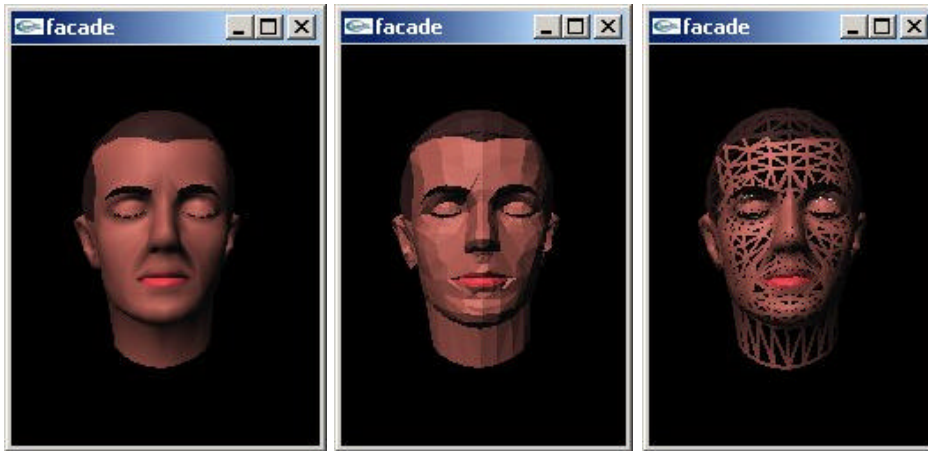
### Classificação:

- ✍ Fala: **sintetizada**
- ✍ Face: **capturada**
- ✍ Forma de Execução: **em batch**



# Trabalhos Relacionados

## Facade



### Definição:

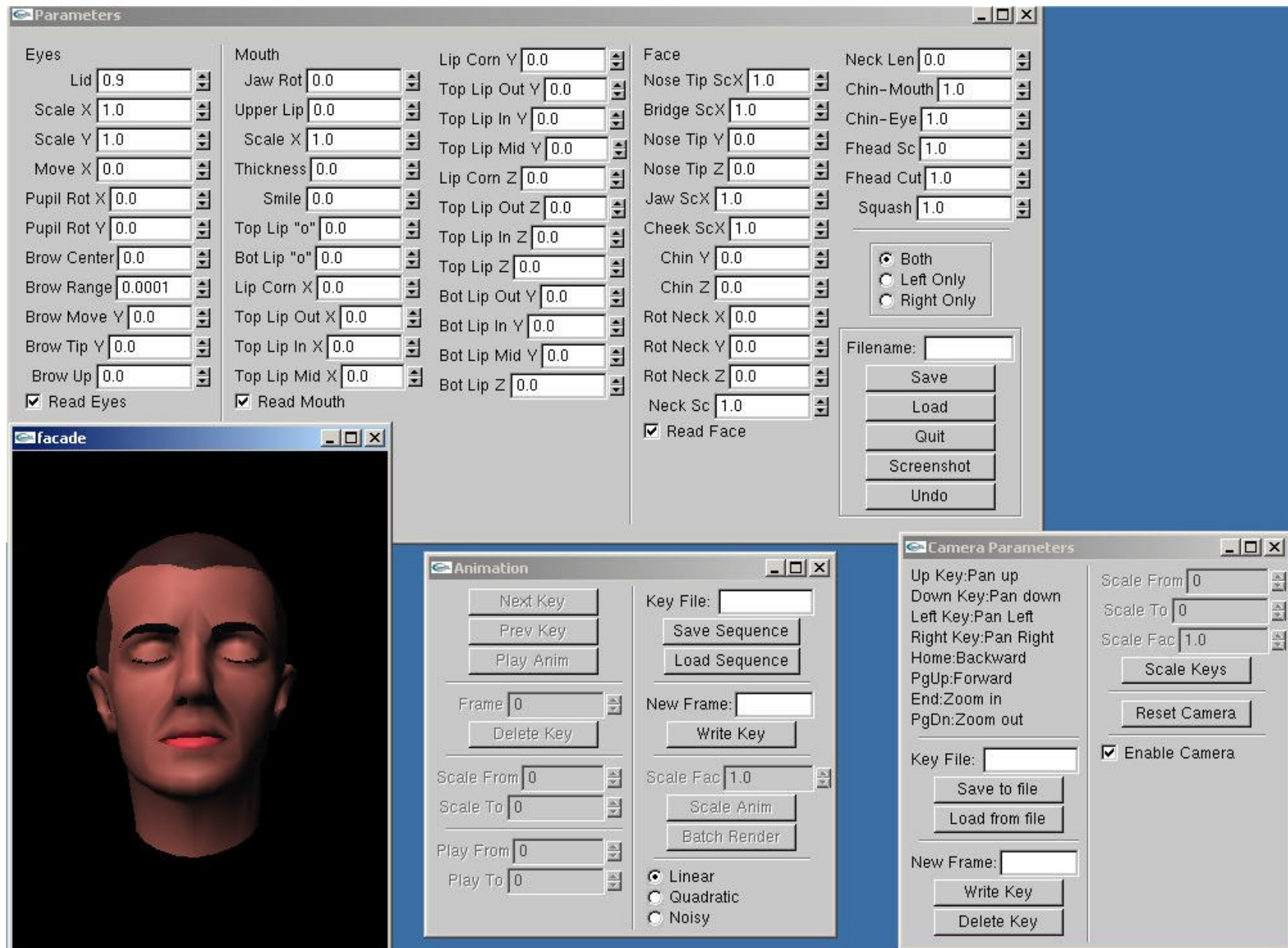
Sistema de animação facial parametrizado que faz uso de uma topologia fixa de face poligonal. Esse sistema pode ser utilizado para criar e animar tipos faciais como também expressões faciais.

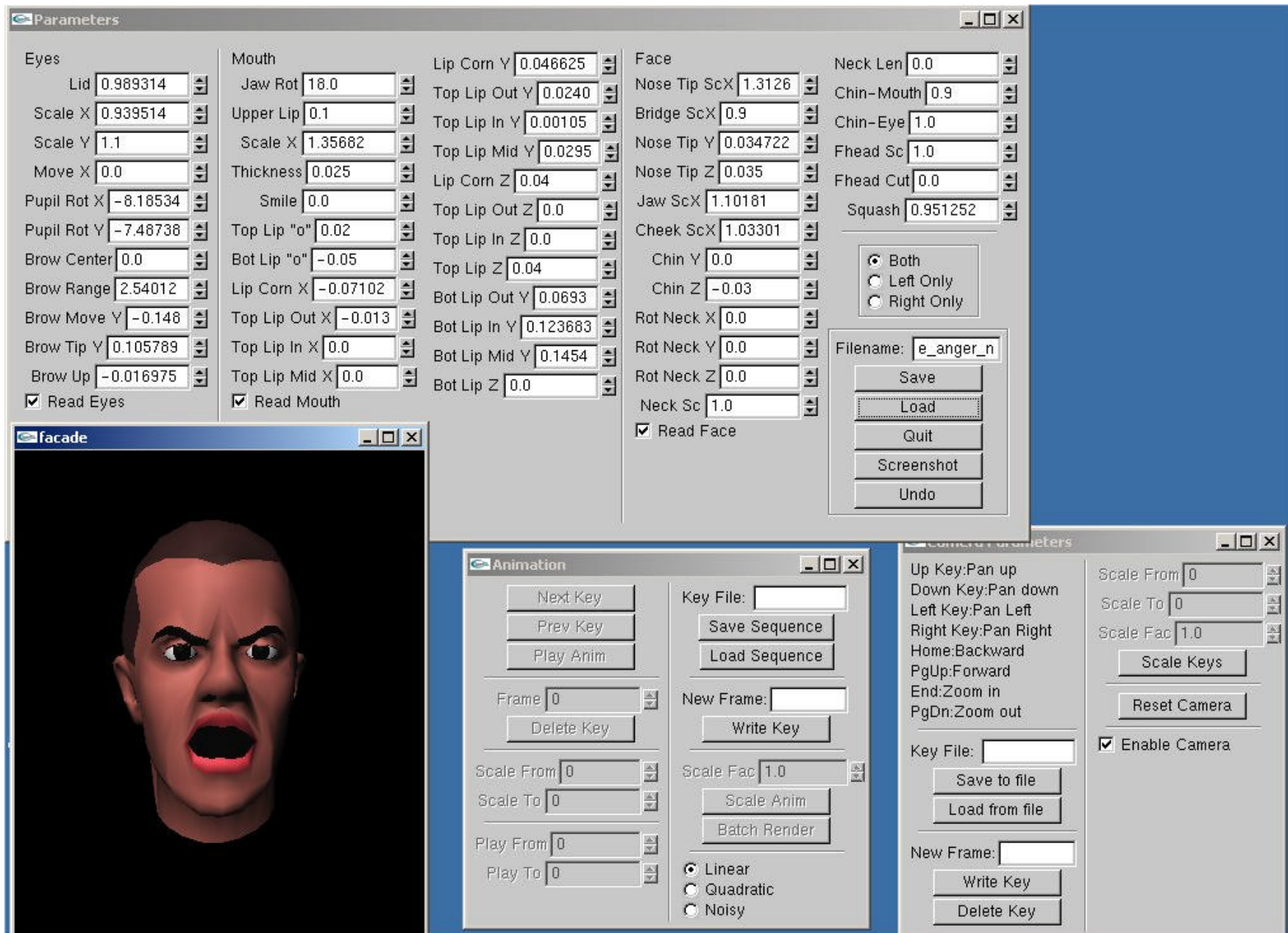
### Classificação:

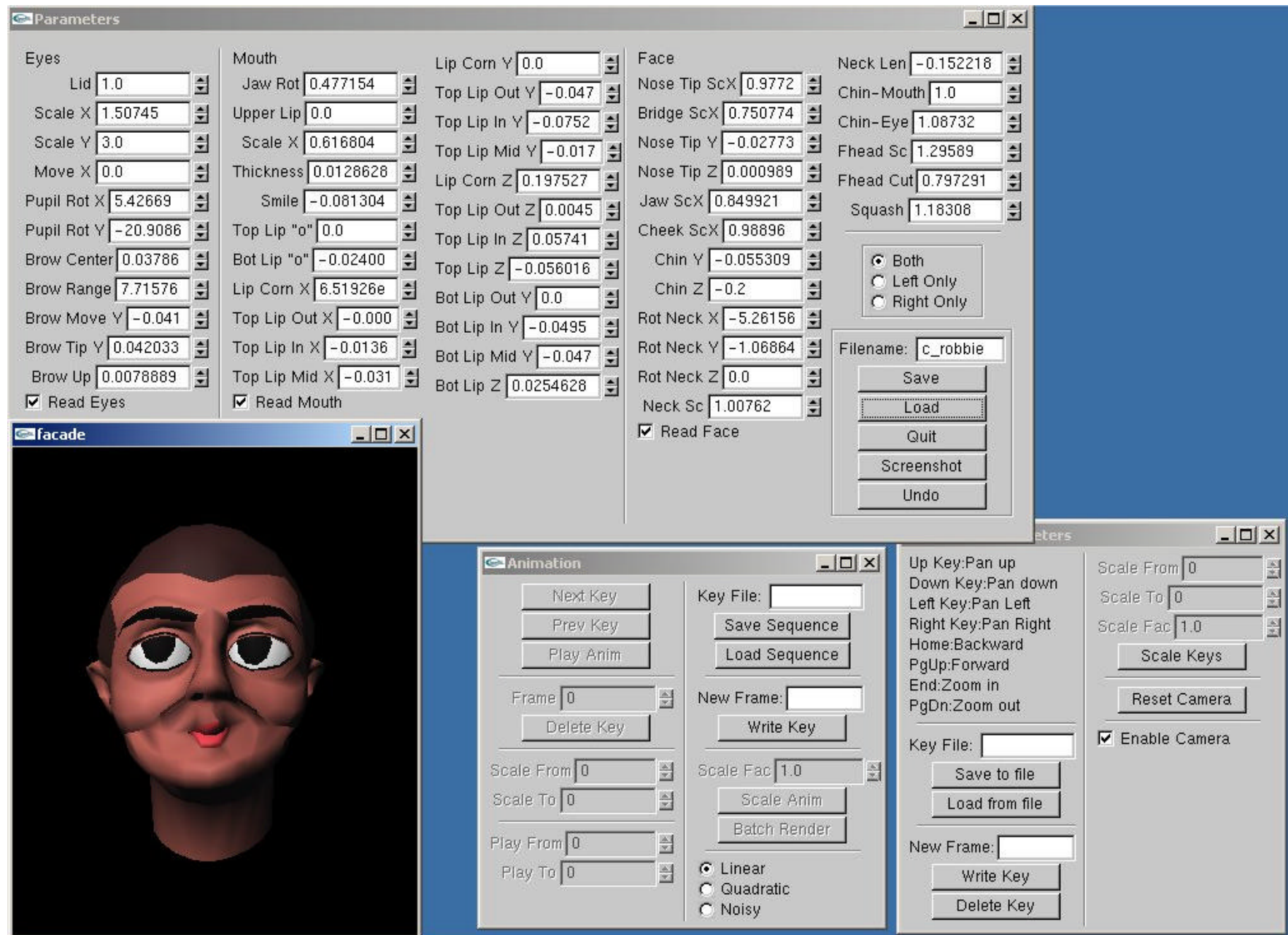
- Fala: **capturada**
- Face: **modelo geométrico**
- Forma de Execução: **em batch**

## FaceWorks





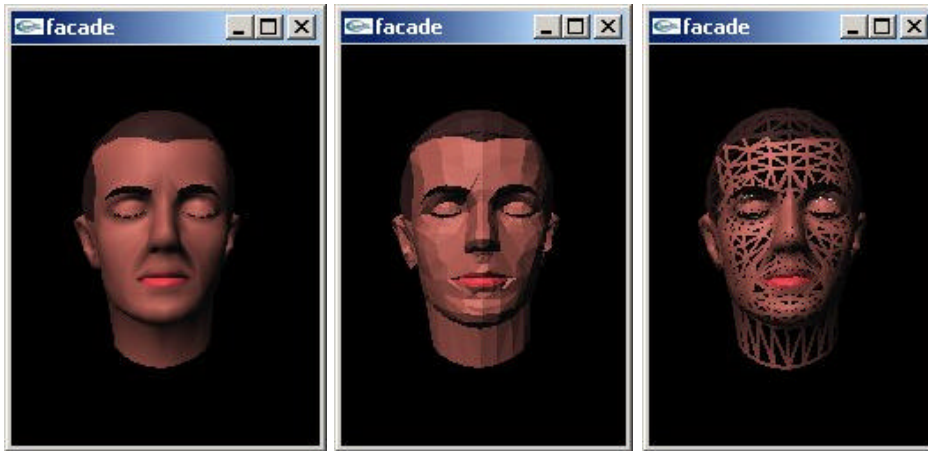






# Trabalhos Relacionados

## Facade



### Definição:

Sistema de animação facial parametrizado que faz uso de uma topologia fixa de face poligonal. Esse sistema pode ser utilizado para criar e animar tipos faciais como também expressões faciais.

### Classificação:

- Fala: **capturada**
- Face: **modelo geométrico**
- Forma de Execução: **em batch**

## FaceWorks

### Definição:

Sistema multimídia de animação facial que faz uso de um modelo sintético 3D. O FaceWorks simplifica o desenvolvimento da animação facial através da automatização da sincronização do áudio, movimentos dos lábios, expressões e movimentos da cabeça.

### Classificação:

- Fala: **capturada**
- Face: **modelo geométrico podendo fazer uso de textura**
- Forma de Execução: **em batch**





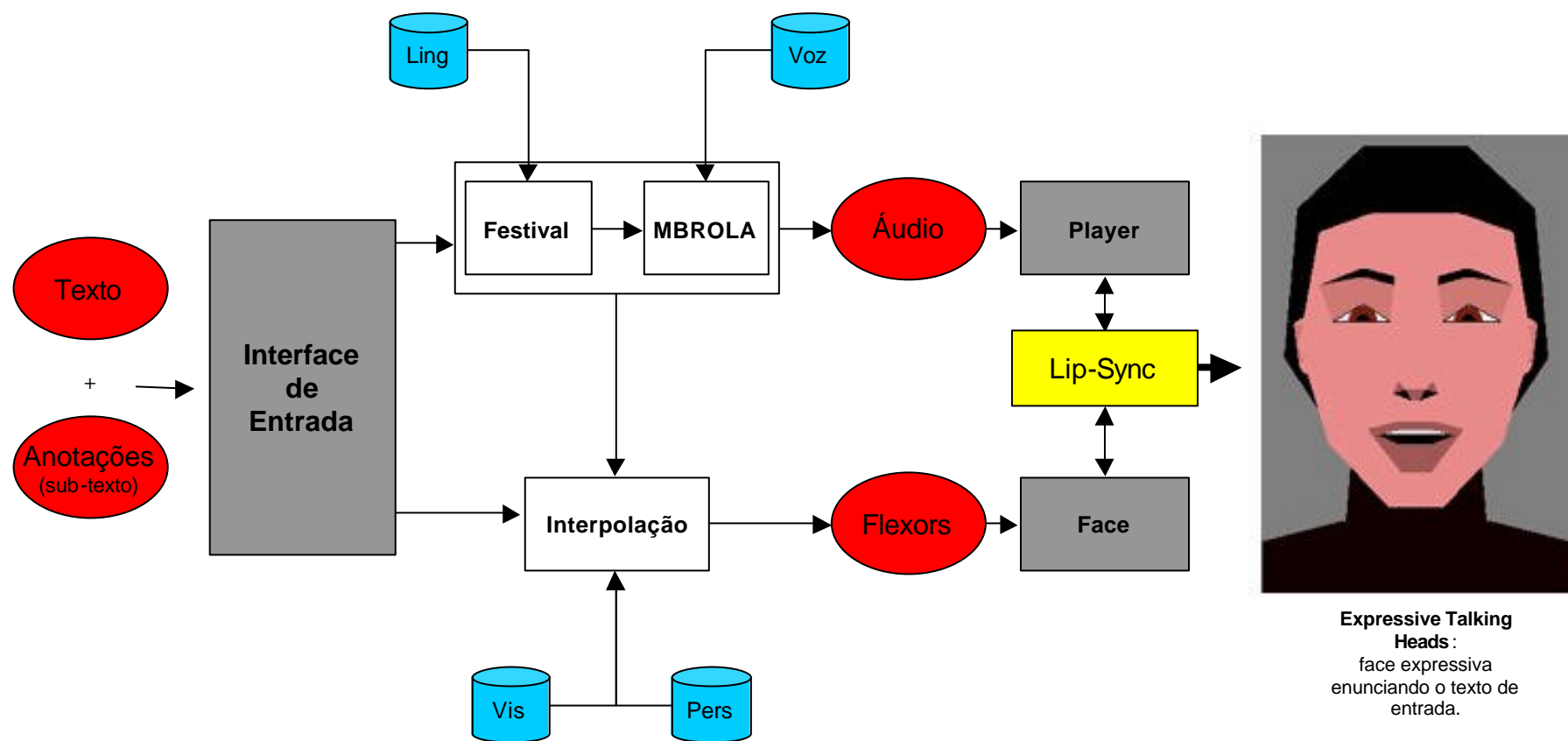
## ○ *Expressive Talking Heads*

- Sistema *talking head* de animação facial capaz de receber um texto e um *sub-texto* de entrada e gerar como saída, em tempo real, uma animação de um personagem virtual enunciando o texto da entrada com o áudio e a expressão facial sincronizados.

(sub-texto: anotações sobre o estado de ânimo do personagem)

- Módulos:
  - ✍ Síntese da entrada
  - ✍ Modelagem da face
  - ✍ Animação da face
  - ✍ *Lip sync*

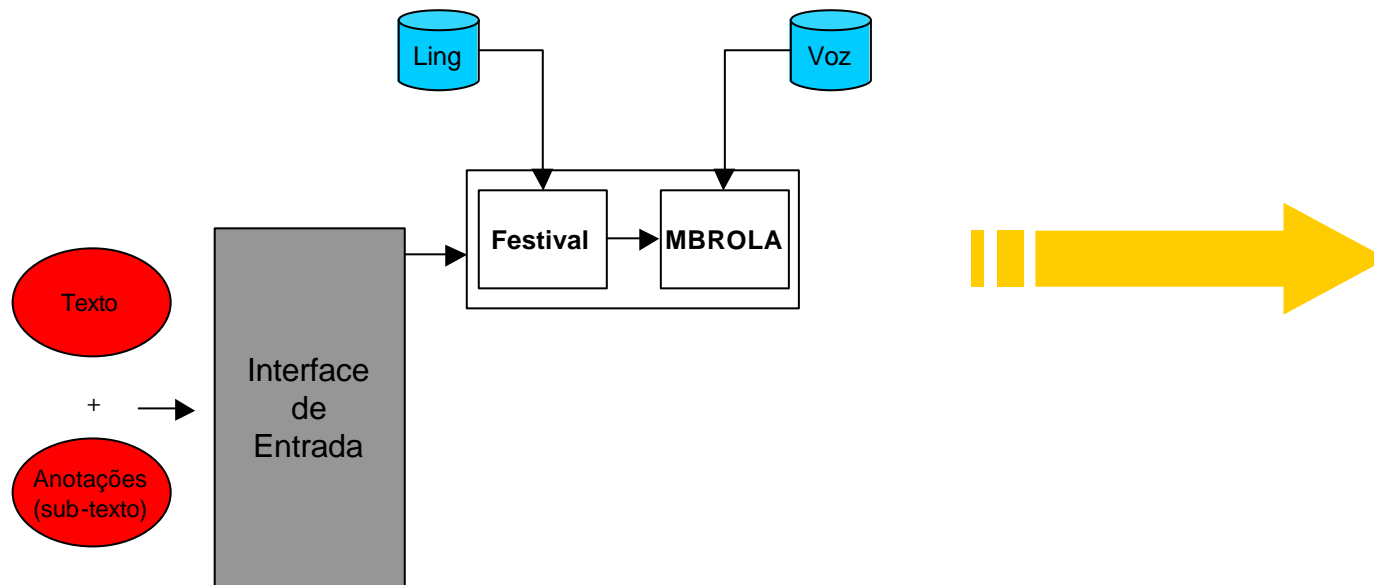
# ETHs - Visão Geral



**Expressive Talking Heads :**  
face expressiva  
enunciando o texto de  
entrada.

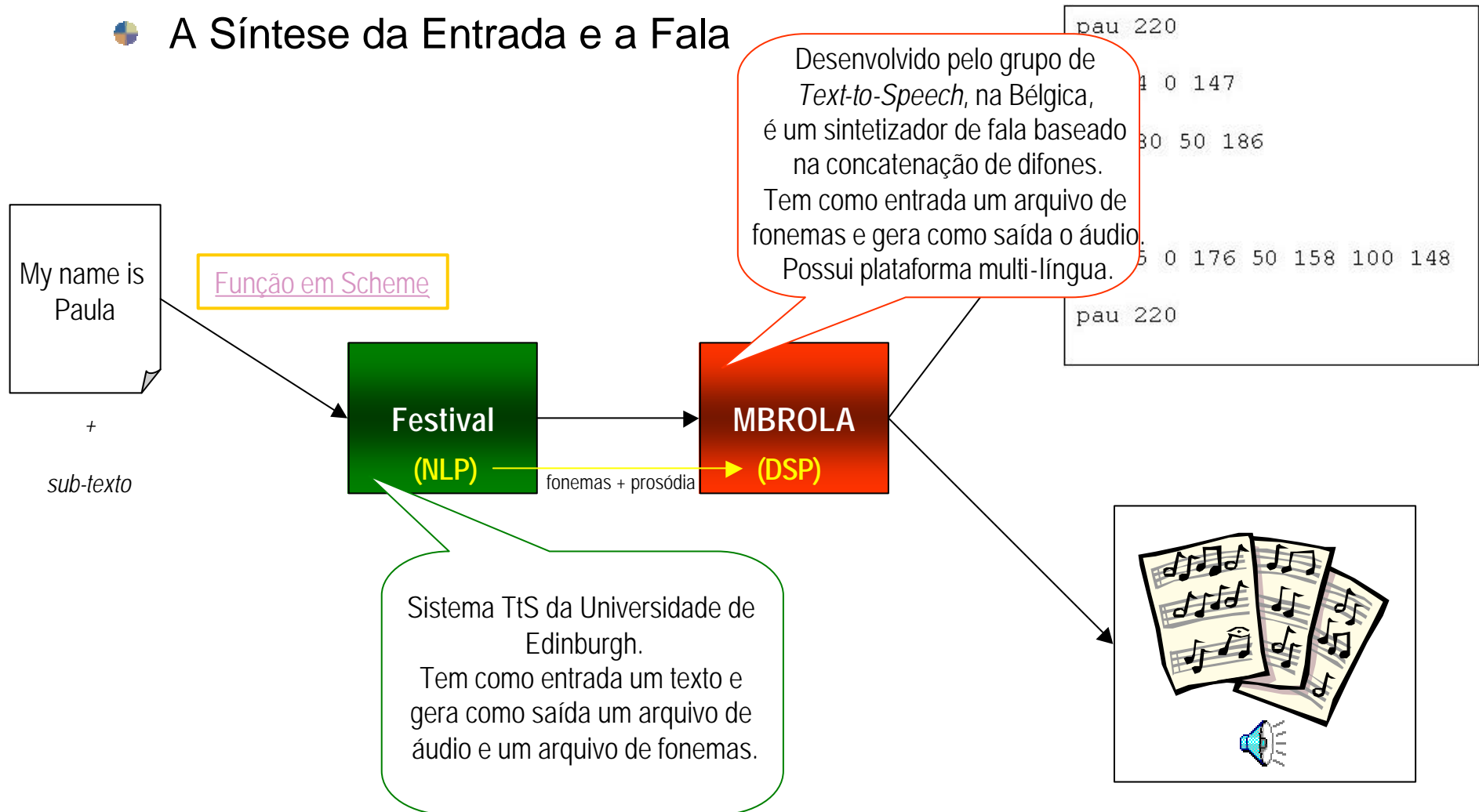
# ETHs - Aspectos de Implementação

## ● A Síntese da Entrada e a Fala



# ETHs - Aspectos de Implementação

## A Síntese da Entrada e a Fala





# ETHs - Aspectos de Implementação

## A Modelagem da Face

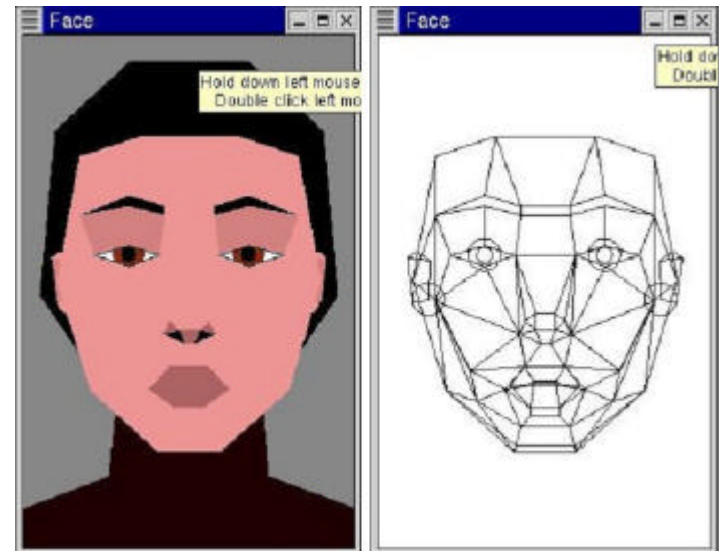
- ✍ Malha poligonal resultado do projeto Responsive Face, desenvolvido por Ken Perlin.
- ✍ Sistema interativo que trabalha com um número mínimo de elementos de expressão facial.
- ✍ Com o agrupamento dos vértices da malha foram formados músculos que são usados para a animação facial.

Músculo da animação labial:

- ✍ *sayAh*
- ✍ *sayOo*
- ✍ *smile*
- ✍ *sneer*

(valor entre  $-1$  e  $+1$  é aplicado a cada músculo para informar quanto cada músculo vai contrair ou relaxar)

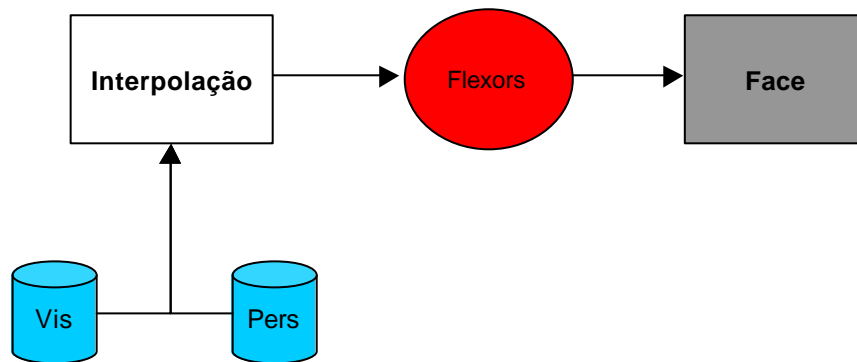
- ✍ Modelagem simples
- ✍ Mínimos controles, mas **bastante expressividade**.



Expressive Talking Heads

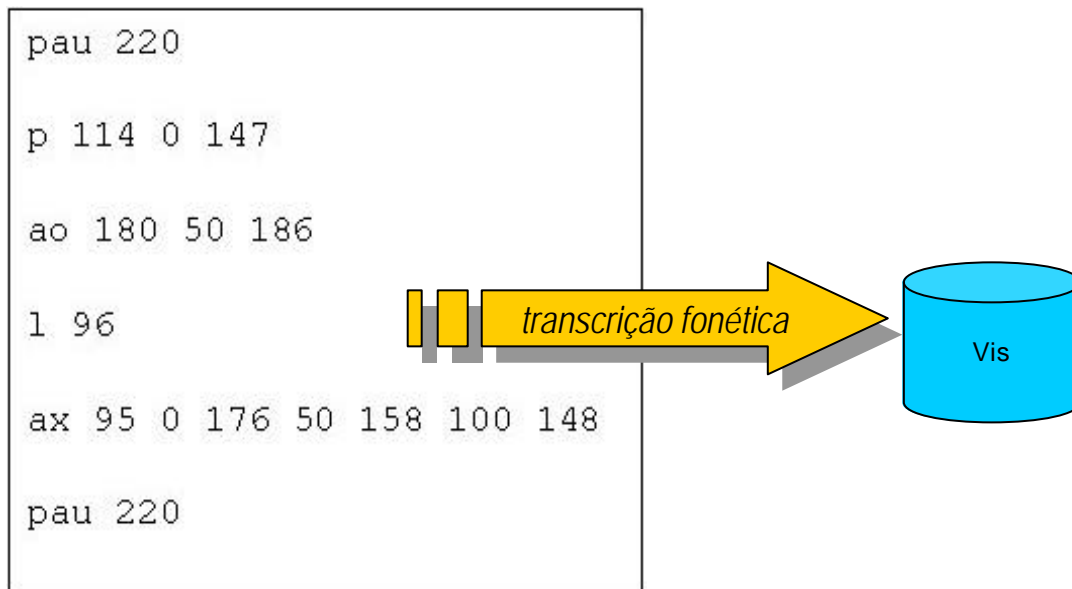
# ETHs – Aspectos de Implementação

## ● A Animação da Face



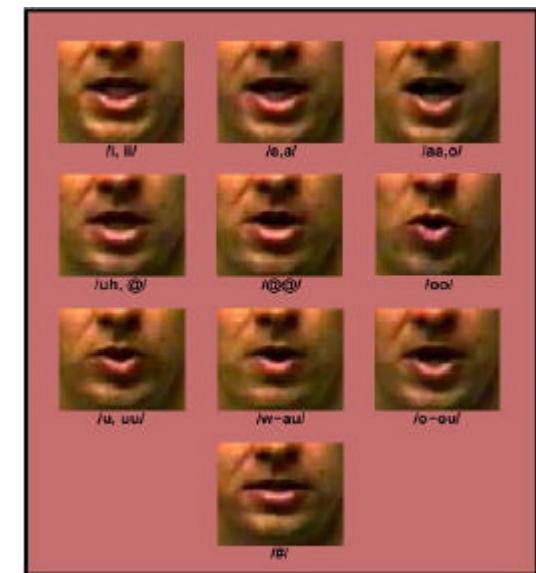
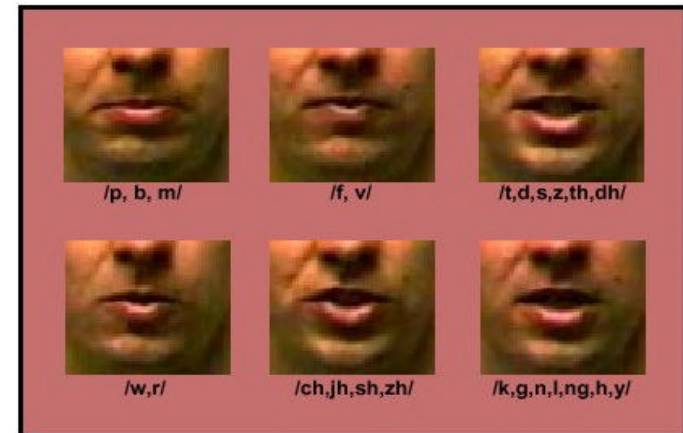
# ETHs – Aspectos de Implementação

## A Animação da Face



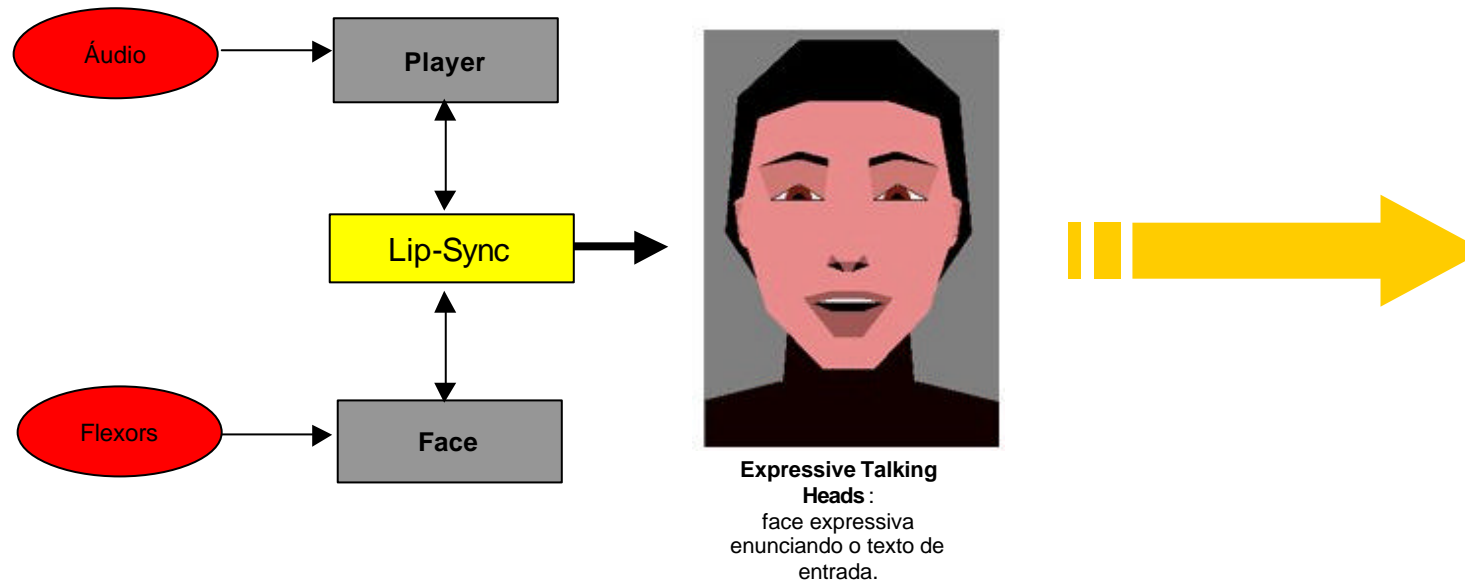
### A Animação:

- mecanismo da fala ? trabalho contínuo dos músculos da face
- transição de visemas ? interpolação linear ? *lip sync*



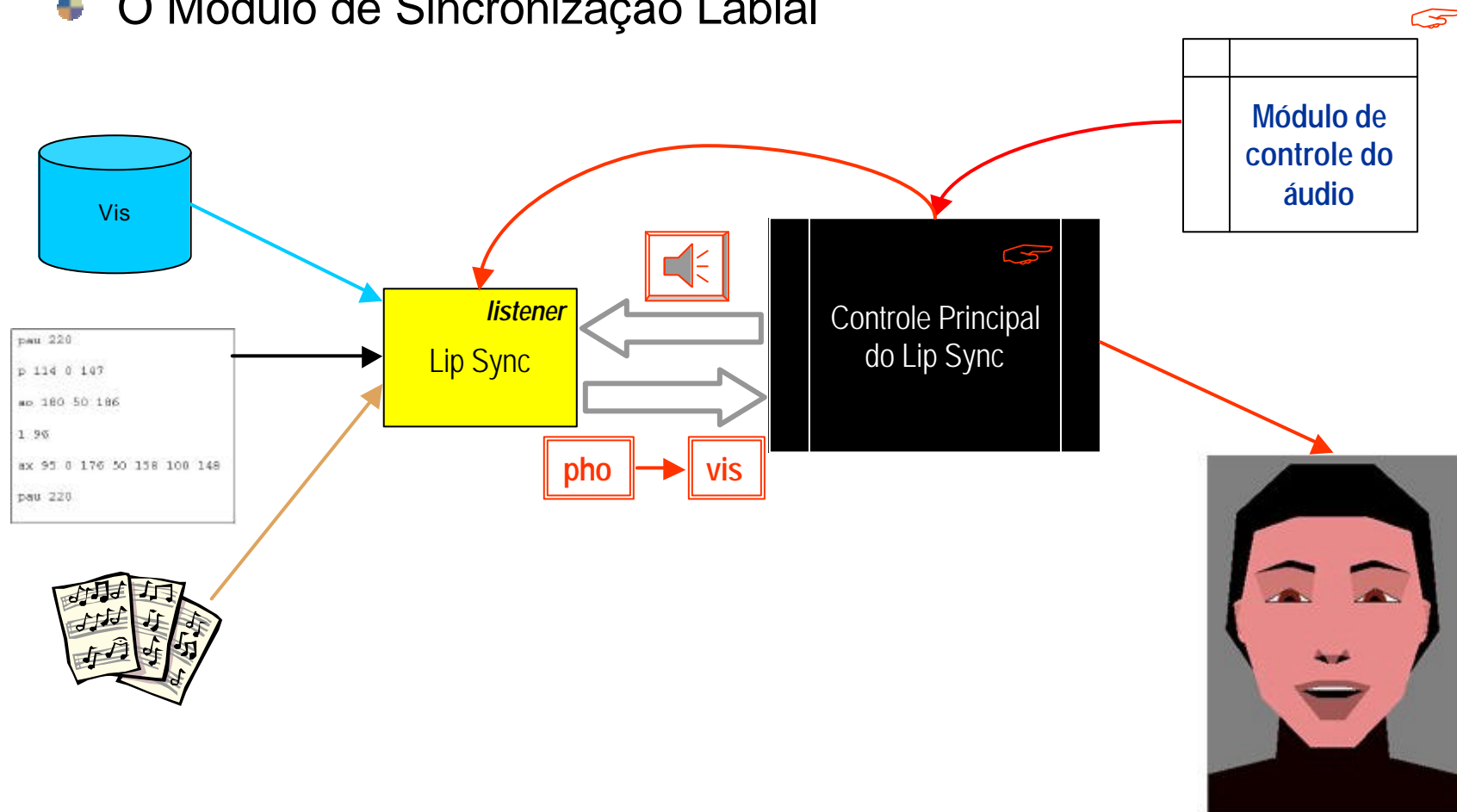
# ETHs – Aspectos de Implementação

## ● O Módulo de *Lip Sync*



# ETHs – Aspectos de Implementação

## ● O Módulo de Sincronização Labial





# ETHs - Finalizações



## Estado Atual

- ✍ Versão do *Expressive Talking Heads* enunciando o texto de entrada sem o atributo de expressividade do personagem.
- ✍ Pesquisa e início da implementação de “anotações(marcações) para expressividade”.



## Melhorias



## Próximos Passos

- ✍ Uso de uma linguagem de marcação para expressividade.
- ✍ Tratamento de ruído de baixa frequência.
- ✍ Implementação do *Expressive Talking Heads* como *applet*.



# Contribuições Esperadas

- Implementação de um sistema *talking head* que, diferente de todos os trabalhos analisados na literatura, efetua a animação facial em tempo real.
  - (tempo real = dados contendo a fala do personagem são interativamente fornecidos pelo usuário, enquanto a animação ocorre)
- Aplicação WWW
- Classificação:
  - ✍ Fala: **sintetizada**
  - ✍ Face: **modelo geométrico**
  - ✍ Forma de Execução: **tempo real**



# Estrutura da Dissertação

1. Introdução
2. Conceituação
3. Trabalhos Relacionados
4. Fala e Som
5. Face e Expressividade
6. Animação
7. Expressive Talking Heads
8. Aspectos de Implementação no “Expressive Talking Heads”
9. Aplicações do “Expressive Talking Heads”
10. Conclusões e Trabalhos Futuros





## Cronograma para Conclusão (\*)

Atividade	Período (ano base 2002)
Defesa da Proposta de Dissertação	fevereiro
Definição e implementação de anotações para controle da expressividade.	fevereiro
Tratamento de ruído de baixa frequência no movimento da face.	março
Geração da versão do “Expressive Talking Heads” como <i>applet</i> .	março e abril
Escrita do documento de Dissertação de Mestrado.	janeiro a abril
Defesa da Dissertação de Mestrado	maio

\* : Cronograma sujeito à alteração.

# Expressive Talking Heads

Defesa de Proposta de Dissertação



**Apresentação do Sistema**

