

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Suzana Mesquita de Borba Maranhão

**Ajuste Elástico em Tempo de Exibição
para Fluxos de Áudio Comprimido**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Informática da PUC-Rio.

Orientador: Luiz Fernando Gomes Soares

Rio de Janeiro, abril de 2006.



Suzana Mesquita de Borba Maranhão

Ajuste Elástico em Tempo de Exibição para Fluxos de Áudio Comprimido

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Informática da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Luiz Fernando Gomes Soares

Orientador

Departamento de Informática - PUC-Rio

Geber Lisboa Ramalho

Centro de Informática - UFPE

Sérgio Colcher

Departamento de Informática - PUC-Rio

Rogério Ferreira Rodrigues

Departamento de Informática - PUC-Rio

José Eugênio Leal

Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 03 de abril de 2006.

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, da autora e do orientador.

Suzana Mesquita de Borba Maranhão

Graduou-se em Ciência da Computação pela UFPE (Universidade Federal de Pernambuco) em 2003. Foi bolsista da CAPES em áreas de Probabilidade e Engenharia de Software durante a graduação.

Ficha Catalográfica

Maranhão, Suzana Mesquita de Borba

Ajuste Elástico em Tempo de Exibição para Fluxos de Áudio Comprimido / Suzana Mesquita de Borba Maranhão; orientador: Luiz Fernando Gomes Soares – Rio de Janeiro: PUC, Departamento de Informática, 2006.

136 f. ; 29,7 cm

Incluí referências bibliográficas.

1. Informática – Teses. 2. Ajuste elástico. 3. Áudio Comprimido (MP3, MP2, AAC, AC-3). 4. MPEG-2 Sistemas. 5. Sincronização Intermídia. 6. Apresentação Hipermídia. 7. Formatador Hipermídia. 8. Formatador Hyperprop. I. Soares, Luiz Fernando Gomes. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Este trabalho é dedicado

Aos meus amados pais, que mesmo longe, sempre estiveram por perto. Grandes educadores, eles me ensinaram muito mais do que pode ser encontrado em livros ou expresso em palavras.

A Marcio, essa pessoa tão especial por quem me apaixonei no decorrer deste mestrado. Um poço de loucuras, carinhos e histórias para contar, esse mineirinho me faz muito feliz.

Agradecimentos

Agradeço a Luiz Fernando, exemplo de orientador, cientista e ser humano. Pessoa que trabalha de modo brilhante, contagiante e, ao mesmo tempo, com extrema simplicidade. Sua personalidade forte e transparente é melhor descrita nas palavras de Chang Yü, extraídas do livro A Arte da Guerra: “Bons generais são ao mesmo tempo amados e temidos”.

Em especial agradeço também ao professor Geber Ramalho do Centro de Informática da UFPE. Geber é admirado há muitos anos não só por mim, mas por todos que lotam sua sala de aula para qualquer disciplina que ele se disponha a ser professor. Foi Geber que me orientou para fazer o mestrado na PUC e agora me deu a honra de participar da banca de avaliação do meu trabalho. Também a Sérgio Colcher, o conhecido por tudo saber, que também se dispôs a contribuir na avaliação deste trabalho.

Como não poderia esquecer, agradeço à professora Marcília Campos, também da UFPE, por ter me introduzido na vida acadêmica e ter sempre feito tudo que foi possível, e às vezes o impossível, para incentivar essa minha vocação.

Agradeço também às pessoas do TeleMídia, que estão sempre alegres, solícitas e prontas para o próximo desafio. A energia dessas pessoas torna o laboratório um lugar realmente especial de se trabalhar. Em especial, obrigada a Rogério, que foi meu guia em vários pontos do trabalho, fundamental em vários momentos. Também não posso deixar de ressaltar os professores e funcionários da PUC-Rio, que fazem da instituição um centro de excelência de ensino no Brasil e no mundo.

Por fim, agradeço ainda ao CNPq, a PUC-Rio e ao Laboratório TeleMídia pelo apoio, inclusive financeiro, indispensável à realização deste trabalho.

Resumo

Maranhão, Suzana Mesquita de Borba. **Ajuste Elástico em Tempo de Exibição para Fluxos de Áudio Comprimido**. Rio de Janeiro, 2006. 136p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Ajuste elástico é a técnica de modificar a duração de exibição de um objeto de mídia. Este trabalho propõe algoritmos de ajuste elástico que funcionam em tempo de exibição para fluxos de áudio comprimidos, com aplicabilidade principal em cenários onde não existe interferência do usuário como, por exemplo, a manutenção da consistência temporal de um documento hipermídia. Os algoritmos de ajuste de áudio são propostos, desenvolvidos e integrados a algoritmos de ajuste de vídeo. Além disso, um algoritmo de resincronização é proposto, e implementado, para preservar o sincronismo intermídia durante a realização de ajuste em fluxos de sistemas MPEG-2. A união dos algoritmos implementados com o desenvolvimento da integração a exibidores de conteúdo deu origem à ferramenta de ajuste, capaz de ser facilmente utilizada por aplicações que precisam solicitar ajuste elástico em fluxos de mídia comprimidos. Como exemplo, a ferramenta de ajuste foi integrada a ferramentas de exibição de um formatador hipermídia. Os algoritmos propostos podem ajustar a duração de áudio em até 10% mantendo a qualidade da percepção do áudio dentro de limites aceitáveis, conforme demonstrado através de medidas de qualidade que também são apresentadas neste trabalho.

Palavras-chave

Ajuste elástico; Áudio Comprimido (MP3, MP2, AAC, AC-3); MPEG-2 Sistemas; Sincronização Intermídia; Apresentação Hipermídia; Formatador Hipermídia; Formatador Hyperprop.

Abstract

Maranhão, Suzana Mesquita de Borba. **On-the-fly Timescale for Compressed Audio Streams**. Rio de Janeiro, 2006. 136p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Timescale is a technique used to modify media-object playing duration. This work proposes timescale algorithms, to be performed on-the-fly, for compressed audio streams. The focus is situations where there is no user interference as, for example, the temporal consistency maintenance of a hypermedia document. The algorithms are proposed, developed and integrated to video timescale algorithms. Furthermore, a resynchronization algorithm is also proposed, and implemented, in order to maintain inter-media synchronization in MPEG-2 system streams. The timescale tool is composed by the aforementioned algorithms and their integration to content rendering tools. The proposed solution can be easily used in applications that need timescale compressed media streams. As an example, the adjustment tool was integrated to a hypermedia formatter. The proposed algorithms can adjust audio durations up to 10%, maintaining the audio perceptual quality within acceptable limits, as observed through quality measurements also described in this work.

Keywords

Timescale; Compressed Audio (MP3, MP2, AAC, AC-3); MPEG-2 Systems; Inter-media Synchronization; Hypermedia Presentation; Hypermedia Formatter; HyperProp Formatter.

Índice

1	Introdução	16
1.1.	Motivação	16
1.2.	Objetivos	19
1.3.	Organização do restante do documento	19
2	Trabalhos Relacionados	21
2.1.	Categorização de algoritmos de ajuste para áudio	21
2.2.	Ferramentas que realizam ajuste elástico	25
2.3.	Comparação dos trabalhos relacionados com o proposto	32
3	Ajuste Elástico de Áudio no Fluxo Comprimido	35
3.1.	Análise Resumida do Ajuste Elástico Proposto: Requisitos do Algoritmo	35
3.2.	Modelagem do Fluxo de Áudio Com Compressão	37
3.3.	Mecanismo de Ajuste Elástico de Áudio com Compressão	39
3.4.	Ajuste Elástico em Fluxos de Áudio Suportados	45
3.5.	Ajuste Elástico em Fluxos de Sistemas	54
4	Ferramenta de Ajuste Elástico	66
4.1.	APIs de ajuste elástico	67
4.2.	Visão Geral da Implementação da Ferramenta de Ajuste Elástico	70
4.3.	Implementação do Algoritmo de Ajuste Elástico de Áudio	75
4.4.	Implementação do Algoritmo de Elástico em Fluxos de Sistemas	81
4.5.	Mudanças no Algoritmo de Ajuste Elástico de Vídeo para Integração	88
5	Integração da Ferramenta de Ajuste com Exibidores de Conteúdo	90
5.1.	Mecanismo de Comunicação utilizado com Ajuste Elástico em Tempo de Compilação	91

5.2. Mecanismo de Comunicação utilizado com Ajuste Elástico em Tempo de Execução	92
5.3. Considerações Adicionais	94
6 Utilização da Ferramenta de Ajuste em Sistemas HiperMídia	95
6.1. Necessidades de Adaptabilidade de Sistemas HiperMídia	95
6.2. Utilização da Ferramenta de Ajuste em Sistema HiperMídia	98
6.3. Integração da Ferramenta de Ajuste com o Sistema HyperProp	100
7 Resultados Obtidos	104
7.1. Medidas de Qualidade Subjetiva	105
7.2. Medidas de Qualidade Objetivas	110
7.3. Análise da Implementação	113
8 Considerações Finais	117
9 Referências	120
10 Apêndice A: Exemplos de Código para Solicitar Ajuste Elástico	126
11 Apêndice B: Aplicação de envio e recepção de fluxo de mídia	129
12 Apêndice C: Exemplo da Integração com Formatador HyperProp	130
13 Apêndice D: Aplicação de Teste da Ferramenta de Ajuste	135

Lista de figuras

Figura 1 - Mudança da taxa de exibição de um sinal.	22
Figura 2 - Exemplo de cortes de quadros de um sinal.	22
Figura 3 - Algoritmo <i>Ajuste Regular</i> .	23
Figura 4 - Ajuste elástico com reprodução dos quadros intercalados.	24
Figura 5 - Ajuste elástico de quadros do sinal no domínio da frequência.	25
Figura 6 - Interface de aplicação de ajuste elástico em arquivos de áudio no <i>Sound Forge</i> .	26
Figura 7 - Interface de aplicação de ajuste elástico em arquivos de áudio no <i>Window Media Player 10</i> .	27
Figura 8 - Interface do <i>software Amazing Slow Downer</i> .	28
Figura 9 - Interface do <i>Enounce 2xAV</i> .	28
Figura 10 - Interface de controle do <i>585 Time Scaling Processor</i> .	29
Figura 11 - Estrutura de funcionamento do <i>FastMPEG</i> .	30
Figura 12 - Decodificador MPEG-4 - fonte: (Wolters & Kjörling, 2003).	31
Figura 13 - Realização do ajuste em fluxo comprimido, utilizando decodificação prévia e nova codificação.	35
Figura 14 - Realização do ajuste em fluxo comprimido, utilizando decodificação prévia.	36
Figura 15 - Realização do ajuste diretamente no fluxo comprimido.	36
Figura 16 - Representação dos quadros que compõem um fluxo de áudio com unidade lógica de dados igual ao campo de dados.	38
Figura 17 - Representação dos quadros que compõem um fluxo de áudio.	38
Figura 18 - Representação dos quadros que compõem um fluxo de áudio com bits de PAD no quadro 1 e 3.	39
Figura 19 - Representação dos quadros no primeiro passo para remover o quadro 3.	40
Figura 20 - Representação dos quadros após remoção do quadro 3.	40

Figura 21 - Representação dos quadros no primeiro passo para duplicar o quadro 2.	41
Figura 22 - Representação dos quadros após duplicação do quadro 2.	41
Figura 23 - Representação dos quadros após transferência do PAD para quadro 4.	42
Figura 24 – Reposicionamento de âncoras em um fluxo de mídia devido ao ajuste.	44
Figura 25 - Cálculo do tamanho de um quadro MP2 ou MP3.	46
Figura 26 - Cálculo do tamanho de um quadro MP1.	46
Figura 27 - Codificador e decodificador MPEG-1 e MPEG-2 áudio MC - fonte: (Noll, 1999).	49
Figura 28 - Tipos de codificação para diferentes taxas de bits em MPEG-4 natural - fonte: (ISO, 2001b).	52
Figura 29 - Quadro do formato AC-3.	53
Figura 30 - Fluxo de sistemas com fluxos elementares de áudio e vídeo uniformemente distribuídos.	55
Figura 31 - Fluxo de sistemas com fluxos elementares de áudio e vídeo com distribuição não-uniforme.	55
Figura 32 – Conjunto de algoritmos para realizar ajuste elástico em fluxos de sistemas.	56
Figura 33 - Transformação de um fluxo de sistemas em um fluxo elementar para aplicar algoritmo de ajuste elástico.	57
Figura 34 - Representação de divisão dos fluxos elementares em trechos para recálculo de fator de ajuste durante aplicação de ajuste elástico.	59
Figura 35 - Cálculo do fator de ajuste a ser aplicado no trecho $i+1$ dos fluxos elementares componentes da mídia original.	59
Figura 36 - Estrutura de fluxo MPEG-2 - fonte: (Cavendish, 2005).	61
Figura 37 - Representação de parte de um fluxo de sistemas, com PACKs, PACKETs e bits de fluxos elementares.	62
Figura 38 - Cálculo do SCR em fluxos MPEG-2.	63
Figura 39 - Cálculo do SCR base em fluxos MPEG-2.	63
Figura 40 - Cálculo do SCR extension em fluxos MPEG-2.	63
Figura 41 - Intervalo de variação do <i>system_clock_frequency</i> .	63

Figura 42 - Cálculo do PTS em fluxos MPEG-2.	64
Figura 43 - Cálculo do DTS em fluxos MPEG-2.	64
Figura 44 - Contexto de utilização da ferramenta de ajuste elástico.	66
Figura 45 - Implementação das APIs de ajuste elástico.	69
Figura 46 - Detalhes da ferramenta de ajuste elástico.	71
Figura 47 - Realização das APIs de ajuste elástico pela ferramenta de ajuste.	72
Figura 48 - Classes do subsistema de entrada.	73
Figura 49 - Classes do subsistema de entrada.	74
Figura 50 - Visão geral da implementação do algoritmo de ajuste de áudio.	75
Figura 51 - Principais classes do formato genérico de áudio.	76
Figura 52 - Principais classes do formato MP3.	77
Figura 53 - Controladores do ajuste elástico de áudio.	77
Figura 54 - Montagem (incluindo leitura), processamento e desmontagem (incluindo escrita) de dados de áudio.	78
Figura 55 - <i>Assemblers</i> e <i>disassembler</i> de objetos de formatos de áudio.	80
Figura 56 - Principais classes do subsistema de processamento de áudio.	80
Figura 57 - Visão geral da implementação do algoritmo de ajuste em fluxos de sistemas.	81
Figura 58 - Algoritmo de ajuste de dados audiovisuais.	82
Figura 59 - Tempo de vida das <i>threads</i> responsáveis pelo ajuste elástico.	82
Figura 60 - Classes do formato de MPEG-2 sistemas.	84
Figura 61 - Classes dos subsistemas de montagem e desmontagem.	84
Figura 62 - Classes dos subsistemas de multiplexação e demultiplexação.	85
Figura 63 - Controladores do algoritmo de ajuste em fluxos de sistemas.	86
Figura 64 - Classes para realizar ajuste em fluxo elementar de áudio.	87
Figura 65 - Classe para adaptar serviços oferecidos pelo subsistema de ajuste de vídeo à ferramenta de ajuste elástico.	89
Figura 66 - Diagrama das classes dos exibidores de conteúdo.	91

Figura 67 - Diagrama das classes da integração via programação com JMF.	92
Figura 68 - Integração de um formatador hipermídia com uma ferramenta de exibição.	97
Figura 69 - Integração de uma ferramenta de exibição com ajuste elástico com um formatador hipermídia.	99
Figura 70 - Escala de votação utilizada para avaliar cada áudio no teste subjetivo.	106
Figura 71 - MOS obtido para cada tipo de áudio com fator de ajuste 0.90 e 0.95 com coeficiente de confiança 95%.	108
Figura 72 - MOS obtido para cada tipo de áudio com fator de ajuste 1.05 e 1.10 com coeficiente de confiança 95%.	108
Figura 73 - Cálculo da duração de um quadro do fluxo elementar.	110
Figura 74 - Cálculo do tempo total para ajustar fluxo de sistemas.	114
Figura 75 - Exibição de uma mídia com aplicação de ajuste elástico recebidas em tempo de exibição.	129
Figura 76 - Interface gráfica da aplicação de teste.	135

Lista de tabelas

Tabela 1 - Comparação entre camadas do MPEG-1 - fonte: (Noll, 2000; Soares, 2005).	45
Tabela 2 - Duração para exibir as amostras associadas a um quadro (em milisegundos).	47
Tabela 3 - Descrição dos métodos da API para realizar ajuste elástico em tempo de compilação.	67
Tabela 4 - Descrição dos métodos da API para realizar ajuste elástico em tempo de execução.	68
Tabela 5 - Descrição dos métodos de configuração da ferramenta de ajuste elástico.	70
Tabela 6 - Mapeamento dos métodos da API FF com os da API de ajuste elástico em tempo de compilação.	100
Tabela 7 - Mapeamento dos métodos da API FF com os da API de ajuste elástico em tempo de execução.	102
Tabela 8 - Tipos de áudio utilizados nos teste de qualidade.	105
Tabela 9 - Categorias de tipos de áudio recomendadas para testes subjetivos, definidas pela ITU-T - fonte:(ITU, 1998).	105
Tabela 10 - Interpretação de cada nível de qualidade.	107
Tabela 11 - MOS obtido para cada tipo de áudio e fator de ajuste.	107
Tabela 12 - Precisão do processamento de arquivos MP3 (em milisegundos).	112
Tabela 13 - Medida de não-linearidade do processamento de arquivos MP3.	112
Tabela 14 - Efeitos do aumento do retardo em conversa telefônica - fonte: (Soares, 2005).	116
Tabela 15 - Trechos de código para solicitar a geração e exibição de um arquivo ajustado.	127
Tabela 16 - Trechos de código para solicitar a exibição de um fluxo sendo ajustado.	128

Tabela 17 - Hiperdocumento escrito em NCL contendo especificação que precisa da aplicação de ajuste elástico para correta exibição. 134