



Cristina Nader Vasconcelos

**Segmentação de vídeo no domínio comprimido baseada na
história da compactação**

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: Bruno Feijó
Co-orientadora: Dilza Szwarcman



Cristina Nader Vasconcelos

Segmentação de vídeo no domínio comprimido baseada na história da compactação

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.

Bruno Feijó
Orientador
PUC-Rio

Dilza Szwarcman
Co-Orientador
Midiarte

Luiz Fernando Gomes Soares
PUC-Rio

Raul Queiroz Feitosa
PUC-Rio

Mônica Costa
Midiarte

Prof. Eugenio Leal
Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 31 de março de 2005

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

Cristina Nader Vasconcelos

Cristina Nader Vasconcelos graduou-se é Bacharel em Informática pela Universidade Federal do Rio de Janeiro (UFRJ). Fez iniciação científica no Instituto de Matemática Pura e Aplicada (IMPA) no projeto Vismed de análise de imagens de Tomografia Computadorizada. Atualmente é membro do laboratório ICAD/VisionLab onde participa como pesquisadora do projeto IndVideo desenvolvido em parceria formada pela PUC-RIO, a FINEP e a empresa Midarte.

Ficha Catalográfica

Vasconcelos, Cristina Nader

Segmentação de vídeo no domínio comprimido baseada na história da compactação / Cristina Nader Vasconcelos ; orientador: Bruno Feijó ; co-orientadora: Dilza Szwarcman. – Rio de Janeiro : PUC, Departamento de Informática, 2005.

97 f. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui referências bibliográficas.

1. Informática – Teses. 2. Segmentação de vídeo. 3. Detecção de tomadas de câmera. 4. Análise de dados MPEG comprimidos. 5. Recuperação de conteúdo. 6. Navegação em vídeo. I. Feijó, Bruno. II. Szwarcman, Dilza. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. IV. Título.

CDD: 004

A Deus.

Aos meus pais, Anacleto e Geni,
por todo amor que dedicam a nossa família.
Eles que adivinham meus pensamentos e dizem,
com apoio incondicional, exatamente o que eu preciso ouvir.

Às minhas irmãs, Ana e Bitá,
minhas inseparáveis companheiras e fontes de minha alegria.

Agradecimentos

À minha família, pelo apoio e incentivo demonstrados em todos os momentos de minha vida.

Aos meus amigos, por me ensinar que a felicidade ocorre na medida em que fazemos os outros felizes

Ao Professor Dr. Bruno Feijó, por ter me orientado com paciência e respeito, sempre solícito em ouvir minhas inquietações.

Aos membros da Midiarte, Mônica, Hélio e em especial a Dilza, por todos os momentos de reflexão e incentivo para sempre fazer o melhor, indispensável para o desenvolvimento deste trabalho.

Aos meus amigos do ICAD e da Midiarte, Pozzer, Raul, Ruben, Lucas, Lauro, Chico, Rodrigo, Esteban, Fábio e Gilliard, Ana, Eduardo, Gustavo e Romildo. Pela amizade que amenizou esta batalha em momentos muito alegres da arte de viver entre amigos. Em especial ao Dr. Pozzer por estar sempre pronto a ajudar e dividir seus conhecimentos.

À CAPES pelo apoio financeiro que proporcionou a realização deste trabalho.

Resumo

Vasconcelos, Cristina. Nader **Segmentação de vídeo no domínio comprimido baseada na história da compactação**. Rio de Janeiro, 2005. 97p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Este trabalho apresenta uma proposta de solução do problema de detecção de tomada de câmera de vídeos MPEG-1 e MPEG-2. A abordagem proposta está baseada na aplicação de diversas heurísticas para eliminação de quadros semelhantes, de forma a extrair um conjunto de quadros que representam os cortes entre tomadas de câmera vizinhas. Essas heurísticas analisam informações no domínio compactado, obtidas diretamente do fluxo de dados codificado dos vídeos, como forma de eliminar o processo de descompressão MPEG e diminuir o volume de dados manipulados durante a análise. A observação dos valores assumidos pelas diversas métricas utilizadas demonstrou a existência de padrões falsos de corte relacionados à história do processo de codificação do vídeo. Por esta razão, as análises das informações codificadas para detecção das tomadas de câmera procuram identificar padrões estabelecidos pelo processo de codificação, considerados assinaturas dos codificadores. Para distinção entre quadros com características de corte, de quadros com características influenciadas pelo codificador, são propostas filtragens para suavizar a influência dessas assinaturas nos valores obtidos pelas métricas de caracterização de similaridade.

Palavras-chave

Segmentação de vídeo; Detecção de tomadas de câmera; Análise de dados MPEG comprimidos; Recuperação de conteúdo; Navegação em vídeo;

Abstract

Vasconcelos, Cristina. Nader **Video segmentation in the compressed domain based on the compression history**. Rio de Janeiro, 2005. 97p. M.Sc. Thesis - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

This work presents a proposal for finding shot cuts in MPEG-1 and MPEG-2 videos. The proposed approach is based on heuristics for eliminating similar frames and thus extracting a set of frames positioned at cut points. These heuristics analyze the compressed data, retrieved from MPEG video streams, without any decompression, thus saving time and space during the shot finding process. The existence of false cut patterns is noticed by studying the data returned by the chosen metrics. In face of such false positives (related to choices made during the history of the video encoding process), the analysis of the compressed data tries to identify patterns in the encoded stream, considered as compressor signatures. To distinguish between cut frames and frames characterized by the encoding process, some filters are proposed in order to alleviate the compressor influence on the similarity metrics results.

Palavras-chave

Segmentação de vídeo; Detecção de tomadas de câmera; Análise de dados MPEG comprimidos; Recuperação de conteúdo; Navegação em vídeo;

Sumário

1	Introdução	13
1.1.	Indexação, Recuperação e Segmentação de Vídeo	13
1.2.	Os Padrões MPEG-1 e MPEG-2	15
1.3.	Escopo da Dissertação	16
1.4.	A Organização da Dissertação	18
2	Segmentação de vídeo	20
2.1.	Tipos de transição	20
2.2.	Abordagens ao problema de detecção de tomada de câmera	22
2.3.	Domínio Comprimido	26
3	Algoritmo Proposto para a Detecção de Cortes de Câmera	37
3.1.	Busca de Padrão de Codificação	38
3.2.	Refinamento por Histograma	42
3.3.	Refinamento por Espaço de Armazenamento	45
3.4.	Probabilidade de corte	49
3.5.	Definição da Probabilidade Ponderada por Canal:	52
3.6.	Refinamento pela Vizinhança no interior das subsequências	53
3.7.	Refinamento por Vizinhança Mínima	55
3.8.	Refinamento dos Candidatos Simples por Canal	56
3.9.	Seleção final	57
4	Resultados	59
4.1.	Critérios e Métricas	60
4.2.	Gráficos de Desempenho	61
4.3.	Etapas de Refinamento da Detecção	64
5	Conclusão	68
5.1.	O Problema do Limiar	68

5.2. Segmentação Baseada na História da Codificação	69
5.3. Trabalhos Futuros	72
6 Referências	75
Apêndice I Os Padrões MPEG-1 e MPEG-2	82
Aspectos históricos dos padrões MPEG-1 e MPEG-2	82
Apêndice II Estrutura hierárquica MPEG-1 e MPEG-2	85
Apêndice III Compressão MPEG	93
Compressão MPEG-1/2	93
Compressão Temporal	94
Compressão Espacial	95

Lista de figuras

Figura 1 – Estrutura Cinemática Hierárquica	15
Figura 2– Tipos de Transição: a.Corte Seco; b. <i>Dissolve</i> ; c. <i>Wipe</i> ; d. <i>Fade-in</i> ; e. <i>Fade-out</i> .	21
Figura 3 – Imagem DC	25
Figura 4 – Possibilidade de posicionamento dos cortes.	29
Figura 5 – Distribuição de uma métrica não ideal para caracterização de corte	38
Figura 6 – Universo da segmentação	41
Figura 7 – Estado da candidatura dos quadros	42
Figura 8 – Macroblocks dependentes	43
Figura 9 – Intervalo de quadros descartados na análise de histogramas	45
Figura 10 – Mudança de estado da candidatura dos quadros (Refinamento por Histograma).	45
Figura 11 – Mudança de estado da candidatura dos quadros (Refinamento por Espaço de armazenamento).	46
Figura 12 – Mudança de estado da candidatura dos quadros (Refinamento por Padrão de Distribuição dos quadros).	48
Figura 13 – Separação em subseqüências	50
Figura 14 – Distribuição da Função de Probabilidade Adaptada – Video 1	51
Figura 15 –Distribuição da Função de Probabilidade Adaptada – Video 2	51
Figura 16 – Distribuição da Função de Probabilidade Adaptada – Video 3	52
Figura 17 – Canal de posição	53
Figura 18 – Possibilidade de corte no interior de uma subseqüência	53
Figura 19 – Mudança de estado da candidatura dos quadros (Refinamento pela Vizinhaça no interior das subseqüências).	55
Figura 20 – Mudança de estado da candidatura dos quadros (Refinamento por Vizinhaça Mínima).	56
Figura 21– Mudança de estado da candidatura dos quadros (Refinamento dos Candidatos Simples por Canal).	57
Figura 22 – Amostras de vídeos do acervo: a. Vídeo Clipe b. Documentário	

ecológico c.Partida de Futebol (preto e branco) d. Partida de Futebol (cores)	
e. Animação	59
Figura 23 - Taxa de acerto X Vizinhança mínima	62
Figura 24 - Taxa de falsos cortes X Vizinhança mínima.	62
Figura 25 – Taxa de acerto X Sensibilidade	63
Figura 26 –Taxa de falsos cortes X Sensibilidade	63
Figura 27 - Desempenho	64
Figura 28 – Distribuição 1:2:3	66
Figura 29 – Distribuição da Função de Probabilidade Adaptada – Vídeo com subseqüência de tamanho 8	67
Figura 30 - <i>Trick Modes</i>	83
Figura 31 – Hierarquia MPEG	85
Figura 32 – Ordem de exibição e ordem de transmissão	86
Figura 33 – Macroblock de referência Forward	89
Figura 34 – Macroblock de referência Backward	90
Figura 35– Macroblock de referência Bidirecional	90
Figura 36 – Distribuição dos blocos no padrão MPEG-2	91
Figura 37 – Camada Block	92
Figura 38 – Codificador MPEG Abstrato	93
Figura 39 - Etapas da compressão espacial	95

Lista de tabelas

Tabela 1

Análise das características esperadas para cortes em quadros B 29

Tabela 2

Análise das características esperadas para cortes em quadros P 30

Tabela 3

Análise das características esperadas para cortes em quadros B 31