

# Estimação de Movimento baseada em Algoritmo Genético

Rafael Silveira Xavier e Robson C. Lins

**Resumo**—Este artigo apresenta uma técnica baseada em algoritmo genético para estimação de movimento em seqüências de vídeo e realiza uma comparação com algoritmos convencionais de estimação de movimento. Os resultados experimentais mostram que esta técnica apresenta um desempenho satisfatório.

**Palavras-Chave**—Compressão de vídeo, estimação de movimento, block-matching, algoritmo genético.

**Abstract**—This paper presents one technique based on genetic algorithm for motion estimation in video sequences and compares it with conventional motion estimation algorithms. The experimental results show that this technique presents a satisfactory performance.

**Keywords**—Video compression, motion estimation, block-matching, genetic algorithm.

## I. INTRODUÇÃO

As principais diferenças entre quadros consecutivos, numa seqüência de vídeo, são causadas por movimentação de câmera, do objeto ou por variações de forma que aparentam ser movimentos locais. Essas variações podem ser estudadas dentro de um modelo: compensação de movimento. Entretanto este modelo precisa de parâmetros que são estimados no codificador: estimação de movimento. Em outras palavras, estes geradores de movimento podem ser compensados pela detecção do deslocamento de *pixels* ou regiões de *pixels* entre quadros consecutivos e mensurando suas diferenças.

Técnicas baseadas em estimação e compensação de movimentos são usados em vários sistemas de compressão de vídeo, como por exemplo o MPEG-4[1]. Atualmente, esta área de estudo está crescendo, especialmente pelo surgimento da TV digital[4] que vem fomentando a proposta de novos padrões e técnicas.

## II. ESTIMAÇÃO DE MOVIMENTO POR *block-matching*

Nas técnicas do tipo *block-matching* os quadros da seqüência de vídeo são analisados aos pares a fim de detectar os movimentos de um quadro para o outro. Esses quadros são denominados de quadro de referência e quadro alvo, na maioria dos casos o quadro alvo é o corrente e o de referência é o quadro anterior[2]. O quadro atual (ou alvo) é segmentado em vários blocos  $M \times N$ , em que cada bloco será comparado, dentro de uma janela de busca própria na imagem anterior, com outros blocos a fim de encontrar o melhor *matching* (casamento). A janela de busca tem uma dimensão de  $(2dm + M) \times (2dm + N)$ , em que  $dm$  representa o deslocamento máximo, em *pixels*, em relação à coordenada vertical e horizontal do bloco.

O deslocamento entre o bloco alvo e o bloco de referência é chamado de vetor de movimento[2]. Como, por conveniência,

Rafael Silveira Xavier e Robson C. Lins, Curso de Ciência da Computação, Universidade Católica de Pernambuco, Recife, Brasil, E-mails:rsixavier@gmail.com, rcl@dei.unicap.br.

a origem do bloco é o seu canto superior esquerdo  $(k, l)$ , logo se pode dizer que o vetor de movimento é a diferença entre a coordenada  $(k, l)$  do bloco alvo e do bloco de referência. Para encontrar a melhor correspondência entre os blocos pode-se usar a diferença média absoluta, ou seja, MAD (*Mean Absolute Difference*):

$$MAD(x, y) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |C(k+i, l+j) - R(k+x+i, l+y+j)| \quad (1)$$

em que  $C(k+i, l+j)$  representa o nível de luminância do *pixel* na imagem atual,  $R(k+x+i, l+y+j)$  é o nível de luminância na imagem anterior,  $(k, l)$  indica a coordenada do canto superior esquerdo do bloco da imagem atual e  $(x, y)$  é o deslocamento em *pixels* a ser pesquisado na imagem anterior com relação à posição explorada na imagem atual.

A meta está em encontrar um vetor de movimento  $(u, v)$  caracterizado por um MAD(x,y) mínimo, ou seja,  $(u, v) = [(x, y)|(x, y)min, x \in [-dm, dm], y \in [-dm, dm]]$ .

Os algoritmos de estimação de movimento por *block-matching* se diferem apenas no modo de realizar a pesquisa pelo melhor casamento. O *full-search*(FS)[2] que realiza uma busca exaustiva, o 3SS[2] que limita os números de pontos de pesquisa na área de busca e o NPDS[5] que busca a partir das distorções parciais, são exemplos de algoritmos convencionais de *block-matching*.

## III. ALGORITMO GENÉTICO

Inicialmente descrito por Jonh Holland em Michigan, EUA, o algoritmo genético (AG) utiliza busca estocástica e operadores de variação, baseados numa versão simplificada da evolução darwiniana e da genética, para a resolução de problemas[3].

O algoritmo genético pertence a um ramo da inteligência computacional denominado: computação evolutiva.

## IV. ESTIMAÇÃO DE MOVIMENTO BASEADA EM AG

Para o mapeamento do problema dentro do modelo de algoritmo genético utiliza-se a seguinte relação:

1. **população inicial**: formada por indivíduos gerados aleatoriamente;
2. **indivíduo(solução candidata)**: par ordenado  $(x, y)$  que representa um bloco na imagem de referência;
3. **fitness(qualidade da solução)**: baseado no MAD entre o indivíduo e o bloco alvo;
4. **seleção dos pais**: feita utilizando o método da roleta[3];
5.  **cruzamento**: o filho é formado por uma coordenada de cada pai escolhida aleatoriamente;

- 6. **mutação:** é somado um valor arbitrário(dentro de um intervalo) aleatoriamente a uma das coordenadas do filho;
- 7. **critério de parada:** baseado no número de gerações.

V. RESULTADOS E DISCUSSÃO

As simulações foram feitas com o uso dos componentes de luminância para as primeiras setenta imagens das seqüências Claire e Tênis. A primeira seqüência envolve movimentos lentos e suaves. Já a segunda contém movimentos rápidos e bruscos. Portanto, foram analisados tipos diferenciados de movimentação. O tamanho utilizado em cada imagem das seqüências foi 240 × 240, no formato PGM (*Portable Grey Map*). E o tamanho do bloco utilizado foi 8 × 8 por pertencer aos tamanhos padrões do MPEG-4 (8 × 8, 8 × 16, 16 × 16, 16 × 32, 32 × 32).

A métrica utilizada para avaliar a qualidade da imagem estimada foi o MSE (*Mean Square Error*):

$$MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2, \quad (2)$$

em que  $M \times N$  representa a dimensão das imagens,  $\hat{f}(x, y)$  representa a imagem estimada e  $f(x, y)$  a imagem original.

As Figuras 1 e 2 apresentam valores de MSE entre a imagem original e a estimada para 70 imagens(quadros) das seqüências Claire e Tennis, respectivamente. Para a maioria dos quadros da seqüência Tênis, o AG possui um desempenho superior ao 3SS, em relação ao MSE. Os algoritmos 3SS e AG, para a seqüência Claire, possuem um desempenho similar, no que diz respeito ao MSE. A figura 2 mostra que, dentre os métodos em consideração, o de pior desempenho em termos de MSE é o NPDS. Mas convém salientar que o NPDS é levado a efeito no sentido de reduzir a métrica SAE (Sum Absolute Error)[5], e não MSE. Dentre os algoritmos em consideração, o NPDS em geral resulta no melhor quadro estimado no tocante à qualidade subjetiva (percepção visual).

A tabela I mostra o valor médio de MSE para as 70 imagens de cada seqüência em consideração. Para a seqüência Claire, O AG e o 3SS apresentaram desempenhos próximos. Para a Tênis, o melhor desempenho foi apresentado pelo AG. É válido observar que o algoritmo genético apresentou um valor médio de MSE inferior em mais de 16% do 3SS.

TABELA I  
MÉDIA DO MSE PARA AS 70 IMAGENS DE CADA SEQÜÊNCIA.

| Algoritmo | Claire | Tênis  |
|-----------|--------|--------|
| FS        | 5,74   | 87,85  |
| 3SS       | 6,02   | 141,21 |
| NPDS      | 10,46  | 567,06 |
| AG        | 6,17   | 117,34 |

VI. CONCLUSÕES

Este artigo apresentou uma abordagem baseada em algoritmo genético para a estimação de movimento, um dos grandes desafios da compressão de vídeo. No tocante aos valores de MSE entre as imagens original e estimada, os resultados da simulação mostraram que o AG, em relação a

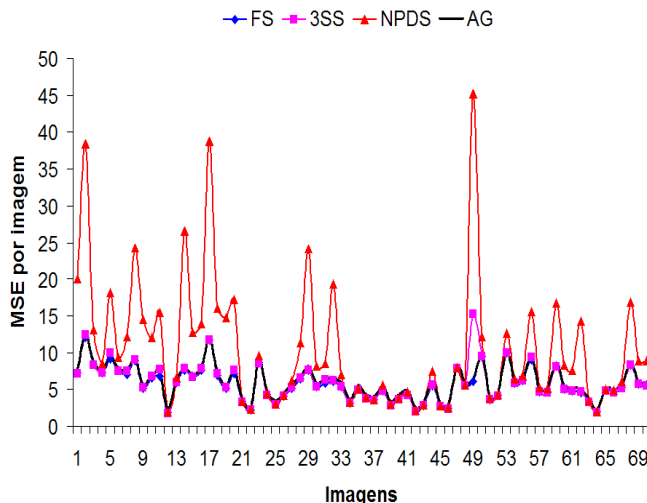


Fig. 1. Gráfico que mostra o MSE entre a imagem original e estimada para as 70 imagens da seqüência Claire.

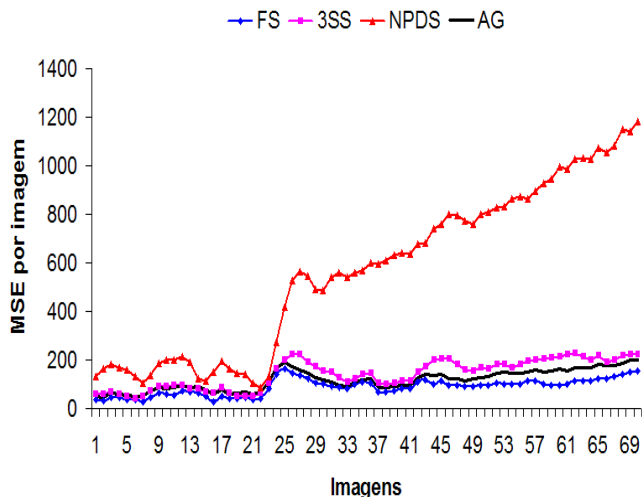


Fig. 2. Gráfico que mostra o MSE entre a imagem original e estimada para as 70 imagens da seqüência Tênis.

seqüência Tênis, apresentou um valor médio de MSE inferior ao valor médio de MSE obtido pelo 3SS. Para a seqüência Claire o AG e o 3SS obtiveram valores similares de MSE.

AGRADECIMENTOS

Os autores agradecem ao PIBIC/UNICAP, pelo suporte financeiro dado à pesquisa.

REFERÊNCIAS

- [1] T. Sikora, "The MPEG-4 video standard verification model", *IEEE Transaction Circuits System Video Technology*, v.7, n.1, p.19-31, 1997.
- [2] Ze-Nian Li and Mark S. Drew, *Fundamentals of multimedia*. Upper Saddle River: Prentice Hall, 2004.
- [3] A.E. Eiben and J.E. Smith, *Introduction to Evolutionary Computing*. Springer, 2003.
- [4] Iain E.G. Richardson, *H.264 and MPEG-4 Video Compression. Video Coding for Next-generation Multimedia*. Wiley, 2005.
- [5] L.M. Po and C.K. Cheung, "Normalized partial distortion search algorithm for block motion estimation", *IEEE Transactions on Circuits and System for Video Technology*, Hong Kong, vol. 10, n. 3, p. 417- 422, 2000.