

# ESTIMAÇÃO DE MOVIMENTO COM USO DE ALGORITMO GENÉTICO MODIFICADO

ROBSON C. LINS\*, RAFAEL S. XAVIER\*, FRANCISCO MADEIRO\*, TIAGO A. E. FERREIRA\*

\*Universidade Católica de Pernambuco  
Recife, PE, Brasil

Emails: rcl@dei.unicap.br, rsixavier@gmail.com, madeiro@dei.unicap.br,  
taef@dei.unicap.br

**Abstract**— This paper presents a modified genetic algorithm applied to motion estimation in video sequences. When compared to the conventional genetic algorithm, the modified version leads to a higher convergence speed. Simulation results are presented for comparing the modified genetic algorithm to classical motion estimation algorithms.

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

**Resumo**— Este artigo apresenta um algoritmo genético modificado aplicado à estimação de movimento em seqüências de vídeo. Quando comparada ao algoritmo genético convencional, a versão modificada leva a uma maior velocidade de convergência. São apresentados resultados de simulações que permitem comparar o desempenho do algoritmo genético modificado com o desempenho de algoritmos clássicos de estimação de movimento.

**Keywords**— Compressão de vídeo, estimação de movimento, casamento de blocos, algoritmo genético.

## 1 Introdução

O principal objetivo das técnicas de compressão de sinais é reduzir o número de bits necessários para representá-los adequadamente, visando minimizar os requisitos de largura de faixa e de capacidade de armazenamento. Apesar de alguns sistemas não apresentarem grandes limitações de largura de faixa, como é o caso das redes de comunicações por fibra ótica, e embora a evolução tecnológica esteja continuamente contribuindo para o surgimento de memórias com grandes capacidades de armazenamento, a compressão de sinais desempenha um papel importante, devido a uma série de fatores, como, por exemplo (Jayant and Noll, 1984; Gersho and Gray, 1992):

- nas redes digitais de serviços integrados (ISDN, *integrated services digital networks*), as técnicas de compressão permitem uma integração eficiente de sinais e dados;
- a enorme utilização dos sistemas multimídia tem levado ao aumento da demanda no tocante ao armazenamento de voz, música, imagens, vídeo e dados em forma comprimida;
- um maior número de canais de comunicação pode ser multiplexado em sistemas de faixa larga, por meio do uso de técnicas de compressão para reduzir os requisitos de largura de faixa de cada sinal a ser multiplexado;
- em telefonia móvel celular, a largura de faixa é severamente limitada, o que tem motivado muitos estudos em compressão de sinais;
- nos sistemas de reconhecimento de fala e de locutor e nos sistemas de resposta vocal, vocabulários maiores podem ser armazenados por meio da redução dos requisitos de memória necessários para cada padrão de voz.

As técnicas de compressão de sinais desempenham um papel importante em aplicações que necessitam minimização dos requisitos de largura de faixa (para o propósito de transmissão) e/ou de capacidade de memória (para o propósito de armazenamento), tais como: sistemas multimídia, redes digitais de serviços integrados, videoconferência, sistemas de resposta vocal, correio de voz (*voicemail*), difusão de música, facsímile de alta resolução, televisão de alta definição (HDTV, *high definition television*), telefonia móvel, sistemas de armazenamento de imagens médicas e de impressões digitais e transmissão de imagens de sensoriamento remoto obtidas por satélites.

Este trabalho tem como foco a compressão de vídeo. Um dos grandes desafios em se tratando de codificação de vídeo é a estimação de movimento, responsável por grande parte do custo computacional da codificação e essencial para a qualidade do vídeo comprimido. Neste trabalho, um algoritmo genético modificado é aplicado à estimação de movimento. O desempenho do método é comparado ao apresentado por algoritmos clássicos de estimação de movimento.

## 2 Estimação de movimento por *block-matching*

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. Este modelo precisa de parâmetros que têm de ser 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 (Sikora, 1997). Atualmente, esta área de estudo está crescendo, especialmente pelo surgimento da TV digital (Richardson, 2005; Wiegand et al., 2003), que vem fomentando a proposta de novos padrões e técnicas (Chen et al., 2004).

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 (Li and Drew, 2004). 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 (Li and Drew, 2004). Como, por conveniência, 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, em que,  $x$  e  $y \in [-dm, dm]$ .

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) (Li and Drew, 2004), que realiza uma busca exaustiva, o 3SS (Li and Drew, 2004), que limita os números de pontos de pesquisa na área de busca, e o NPDS (Po and Cheung, 2000) que busca a partir das distorções parciais, são exemplos de algoritmos convencionais de *block-matching*.

### 3 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 (Eiben and Smith, 2003).

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

#### 4 Estimação de movimento baseada em Algoritmo Genético

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 (Eiben and Smith, 2003);
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.

#### 5 Algoritmo Genético Modificado

O Algoritmo Genético Modificado (AGM), como descrito na Figura 1, utiliza operadores de cruzamento e mutação modificados com o intuito de acelerar a convergência para uma possível solução, baseado no trabalho de Leung et al (Leung et al., 2003). São aplicados quatro diferentes operadores de cruzamento, gerando quatro novos cromossomos:  $\mathbf{C}_1$ ,  $\mathbf{C}_2$ ,  $\mathbf{C}_3$  e  $\mathbf{C}_4$ , definidos por,

$$\mathbf{C}_1 = [C_1^1 \ C_2^1 \ \dots \ C_{nVar}^1] = \frac{\mathbf{p}_1 + \mathbf{p}_2}{2}, \quad (2)$$

$$\mathbf{C}_2 = [C_1^2 \ C_2^2 \ \dots \ C_{nVar}^2] = \mathbf{p}_{max}(1-w) + \max(\mathbf{p}_1, \mathbf{p}_2), \quad (3)$$

$$\mathbf{C}_3 = [C_1^3 \ C_2^3 \ \dots \ C_{nVar}^3] = \mathbf{p}_{min}(1-w) + \min(\mathbf{p}_1, \mathbf{p}_2), \quad (4)$$

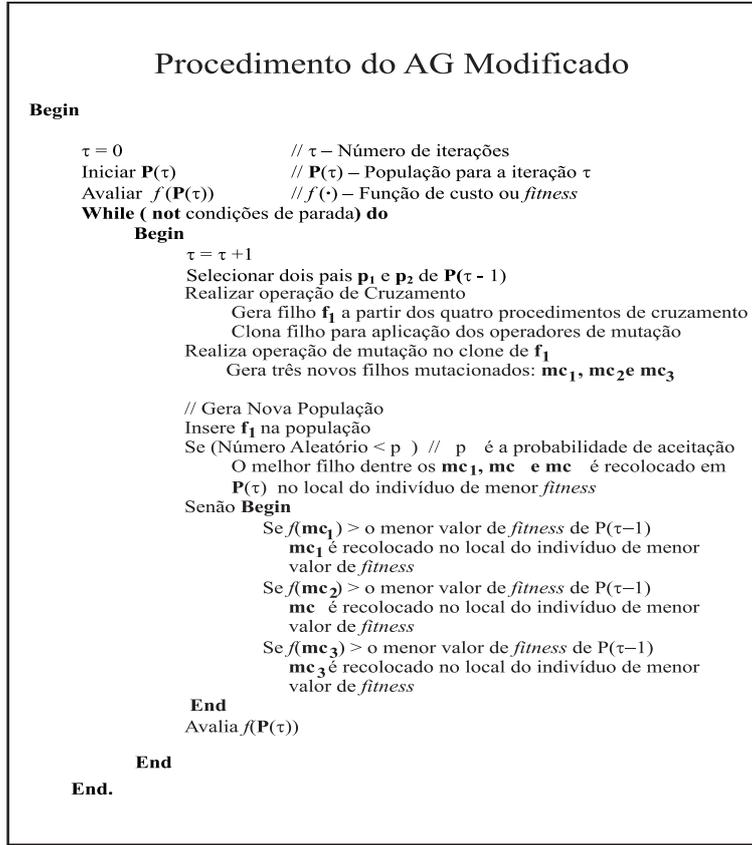


Figura 1: Procedimento do Algoritmo Genético modificado utilizado nos experimentos.

e

$$C_4 = \frac{[C_1^4 C_2^4 \dots C_{nVar}^4]}{(\mathbf{p}_{max} + \mathbf{p}_{min})(1 - w) + (\mathbf{p}_1 + \mathbf{p}_2)w}, \quad (5)$$

em que  $\mathbf{p}_1$  e  $\mathbf{p}_2$  são os cromossomos pais escolhidos com uma chance proporcional aos seus valores de *fitness*,  $\mathbf{p}_{max} = [p_{max}^1 p_{max}^2 \dots p_{max}^{nVar}]$  e  $\mathbf{p}_{min} = [p_{min}^1 p_{min}^2 \dots p_{min}^{nVar}]$  são os valores máximos e mínimos dos genes. O peso  $w \in [0,1]$  é definido pelo usuário: quanto maior seu valor, maior será a contribuição direta dos pais. E, por fim, as funções  $max(\mathbf{p}_1, \mathbf{p}_2)$  e  $min(\mathbf{p}_1, \mathbf{p}_2)$  irão escolher o genes de maior e menor valor respectivamente. As operações de cruzamento dadas pela Equação (2) e pela Equação (5) tendem a varrer a região central do domínio de busca (se  $w \approx 1$  na Equação (5)), e as Equações (3) e (4) tendem a mover a busca para os limites máximo e mínimo do domínio respectivamente.

Dentre os quatro cromossomos gerados, aquele que detiver o maior valor de *fitness* será eleito como o filho gerado pelo processo de cruzamento.

Depois da seleção do filho, este é clonado. Seu clone irá sofrer um processo de mutação, e três novos filhos serão criados. Sendo o clone do filho eleito no processo de cruzamento dado por

$C_g$  ( $g = 1, 2, 3, 4$ ), gera-se as mutações da forma:

$$\mathbf{MC}_j = [C_1^g C_2^g \dots C_{nVar}^g] + [\gamma_1 \Delta MC_1 \gamma_2 \Delta MC_2 \dots \gamma_{nVar} \Delta MC_{nVar}], \quad (6)$$

em que  $\gamma_i$  ( $i = 1, 2, \dots, nVar$ ) assume o valor 0 ou o valor 1, e  $\Delta MC_i$  ( $i = 1, 2, \dots, nVar$ ) são números aleatórios gerados de tal forma a satisfazerem a condição:  $p_{min}^i \leq C_i^g + \Delta MC_i \leq p_{max}^i$  e  $nVar$  é a quantidade de genes por cromossomo.

A primeira mutação ( $j = 1 \Rightarrow \mathbf{MC}_1$ ) é realizada com apenas um único  $\gamma_i$  igual a 1 (escolhido aleatoriamente), e os demais iguais a zero. A segunda mutação ( $j = 2 \Rightarrow \mathbf{MC}_2$ ) decorre com uma quantidade aleatória de  $\gamma_i$ 's (escolhidos ao acaso) iguais a 1 e os demais iguais a zero. E, por fim, a terceira mutação ( $j = 3 \Rightarrow \mathbf{MC}_3$ ) ocorre com todos os  $\gamma_i$ 's iguais a 1.

Cada um dos filhos mutacionados serão aceitos se: 1. o seu *fitness* for superior ao *fitness* do pior indivíduo da população; ou 2. por sorteio, segundo uma probabilidade de aceitação de mutação ( $P_{Mut}$ ).

## 6 Resultados

As simulações foram feitas com o uso dos componentes de luminância para as primeiras setenta

imagens das seqüências Claire e Tennis, apresentadas nas Figuras 2 e 3, respectivamente. 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  $240 \times 240$ , no formato PGM (*Portable Grey Map*). O tamanho do bloco utilizado foi  $8 \times 8$ , um dos tamanhos padrões do MPEG-4 (que também utiliza blocos  $8 \times 16$ ,  $16 \times 16$ ,  $16 \times 32$  e  $32 \times 32$ ).



Figura 2: Quadro de número 70 da seqüência Claire original.



Figura 3: Quadro de número 70 da seqüência Tennis original.

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

$$MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x,y) - f(x,y)]^2, \quad (7)$$

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 4, 5 e 6 apresentam os valores de *fitness* em função do número de gerações para o algoritmo genético original (AG) e para o algoritmo genético modificado (AGM) aplicado à seqüência Claire. Observa-se que o desempenho de AGM é superior ao apresentado pelo AG. De fato, com um menor número de gerações, o AGM “converge” mais rapidamente que AG. Além disso, os

valores finais de *fitness* são praticamente os mesmos para ambos os algoritmos. A superioridade de AGM sobre AG no que diz respeito à “convergência” também pode ser observada nas Figuras 7, 8 e 9, concernentes à seqüência Tennis.

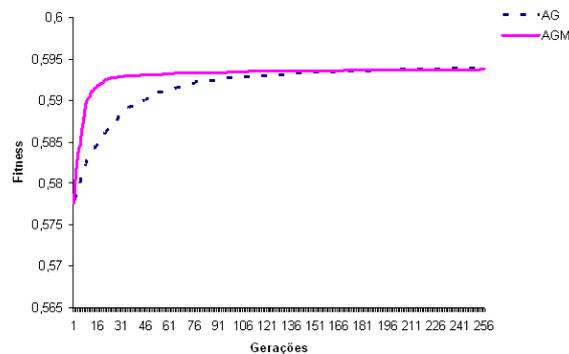


Figura 4: Valor de *fitness* em função do número de gerações para Claire. Estimativa concernente ao quadro 1 em relação ao 2.

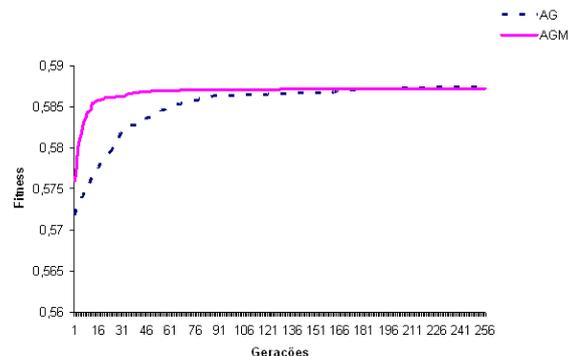


Figura 5: Valor de *fitness* em função do número de gerações para Claire. Estimativa concernente ao quadro 2 em relação ao 3.

As Figuras 10 e 11 apresentam valores de MSE entre a imagem original e a estimada para 70 imagens (quadros) das seqüências Claire e Tennis, respectivamente. Considerando a seqüência Tennis, para a maioria das imagens o algoritmo AGM apresenta uma superioridade em relação a AG, levando a menores valores de MSE entre as imagens original e processada. Para a maioria dos quadros de Tennis, AG e AGM apresentam um desempenho superior ao apresentado por 3SS, no que diz respeito a MSE. A Figura 11 mostra que, dentre os métodos em consideração, o de pior desempenho em termos de MSE é o NPDS. Convém salientar, entretanto, que o algoritmo NPDS é levado a efeito no sentido de reduzir a métrica SAE (*Sum Absolute Error*), e não MSE. Dentre os algoritmos em consideração, o NPDS em geral leva ao melhor quadro estimado no tocante à qualidade subjetiva (via inspeção visual). É possível observar na Figura 10 que, para a seqüência Claire, os algoritmos 3SS, AG e AGM possuem um desempenho similar no que diz respeito a MSE.

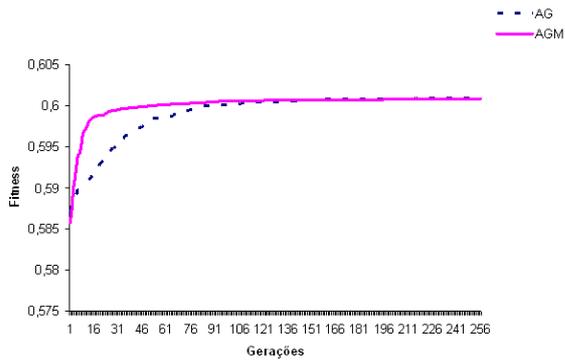


Figura 6: Valor de *fitness* em função do número de gerações para Claire. Estimação concernente ao quadro 3 em relação ao 4.

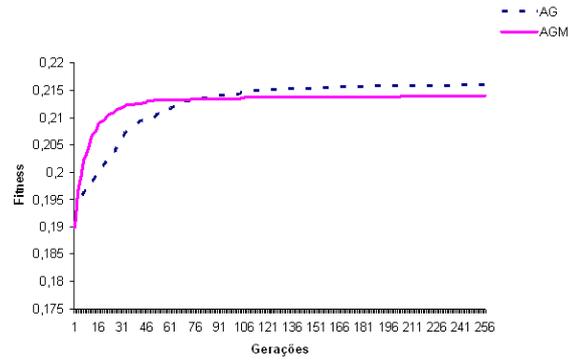


Figura 9: Valor de *fitness* em função do número de gerações para Tennis. Estimação concernente ao quadro 3 em relação ao 4.

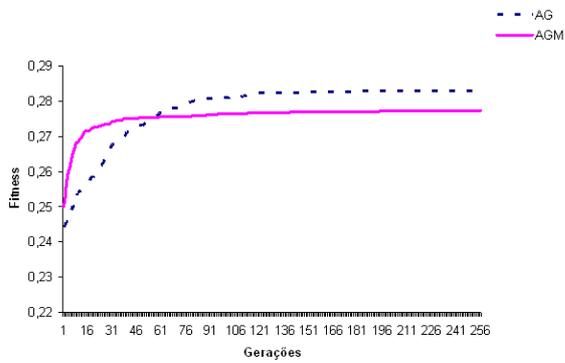


Figura 7: Valor de *fitness* em função do número de gerações para Tennis. Estimação concernente ao quadro 1 em relação ao 2.

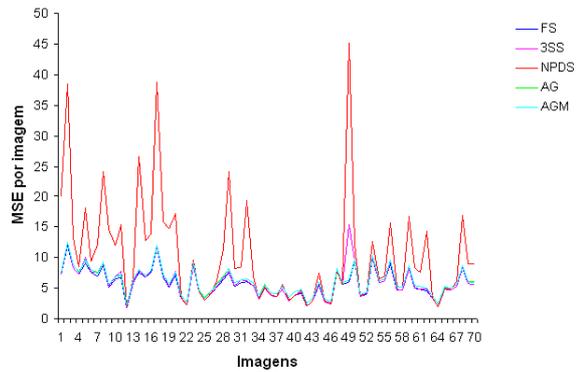


Figura 10: Valor de MSE entre a imagem original e a estimada para 70 imagens (quadros) da seqüência Claire.

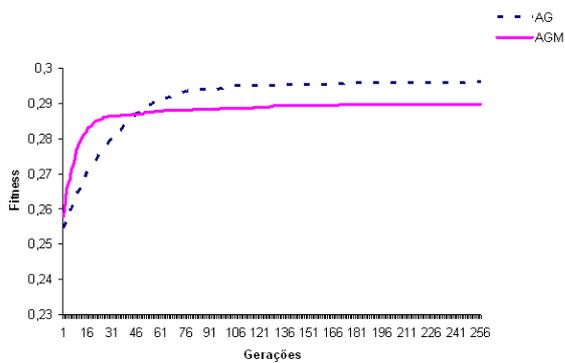


Figura 8: Valor de *fitness* em função do número de gerações para Tennis. Estimação concernente ao quadro 2 em relação ao 3.

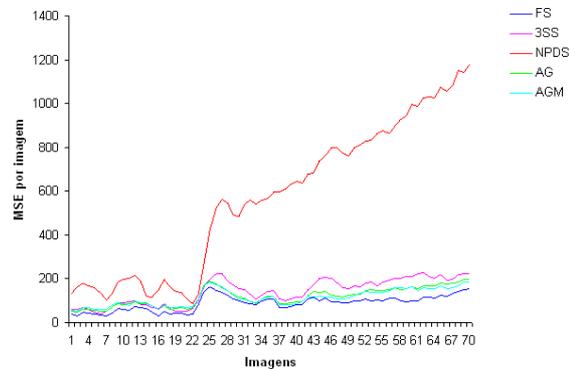


Figura 11: Valor de MSE entre a imagem original e a estimada para 70 imagens (quadros) da seqüência Tennis.

A Tabela 1 mostra o valor médio de MSE para as 70 imagens de cada seqüência em consideração. Para Claire, os algoritmo 3SS, AG e AGM apresentaram um desempenho similar. Para a seqüência Tennis, o melhor desempenho foi apresentado pelo algoritmo GM, seguido de AG e 3SS. É possível observar que o algoritmo AGM apresentou um valor médio de MSE inferior em mais de 20% ao 3SS.

Tabela 1: Média do MSE para as 70 imagens (quadros) de cada seqüência.

Algoritmo	Claire	Tennis
FS	5,74	87,85
3SS	6,02	141,21
NPDS	10,46	567,06
AG	6,17	117,35
AGM	6,16	109,48

## 7 Conclusões

Este artigo apresentou um algoritmo genético modificado, aplicado à estimação de movimento, a qual representa um dos grandes desafios em se tratando de compressão de vídeo. A modificação introduzida no algoritmo genético convencional permite acelerar a velocidade de convergência do algoritmo. No que diz respeito aos valores de erro médio quadrático (MSE, *mean square error*) entre as imagens (quadros de uma seqüência de vídeo) original e estimada, os resultados de simulação mostraram que o algoritmo genético modificado (AGM) apresenta desempenho em geral próximo e por vezes ligeiramente superior ao algoritmo genético convencional (AG). Ambos algoritmos genéticos, AG e AGM, apresentaram uma superioridade em relação ao algoritmo 3SS: levaram a imagens estimadas com MSE (em relação à imagem original) inferior ao apresentado pelas imagens estimadas com uso de 3SS. Simulações realizadas com a seqüência Tennis mostraram que, para as 70 primeiras imagens da seqüência, o valor médio de MSE para as imagens estimadas com uso do algoritmo AGM foi inferior em mais de 20% ao valor médio de MSE para as imagens estimadas por meio do algoritmo 3SS.

## Referências

- Chen, M.-J., Chiang, Y.-Y., Li, H.-J. and Chi, M.-C. (2004). Efficient multi-frame motion estimation algorithms for mpeg-4 avc/jvt/h.264, *Proceedings of the 2004 International Symposium on Circuits and Systems*, pp. 737–740.
- Eiben, A. and Smith, J. (2003). Introduction to evolutionary computing, *Springer*.
- Gersho, A. and Gray, R. M. (1992). *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, MA.
- Jayant, N. S. and Noll, P. (1984). *Digital Coding of Waveforms*, Prentice-Hall, Englewood Cliffs, NJ.
- Leung, F. H. F., Lam, H. K., Ling, S. H. and Tam, P. K. S. (2003). Tuning of the structure and parameters of a neural network using an improved genetic algorithm, Vol. 14, pp. 79–88.
- Li, Z.-N. and Drew, M. S. (2004). Fundamentals of multimedia, *Upper Saddle River: Prentice Hall*.
- Po, L. and Cheung, C. (2000). Normalized partial distortion search algorithm for block motion estimation, Vol. 10, pp. 417–422.
- Richardson, I. E. (2005). H.264 and mpeg-4 video compression. video coding for next-generation multimedia, *Wiley*.
- Sikora, T. (1997). The mpeg-4 video standard verification model, Vol. 7, pp. 19–31.
- Wiegand, T., Sullivan, G. J., Bjntegaard, G. and Luthra, A. (2003). Overview of the h.264/avc video coding standard, Vol. 13, pp. 560–576.