

Uma Nova Técnica de Detecção de Artefatos de Blocagem em Vídeo Comprimido

Wyllian Bezerra da Silva e Alexandre de Almeida Prado Pohl

Resumo—Este trabalho apresenta uma nova técnica de detecção de artefatos de blocagem para vídeos comprimidos que poderá servir como suporte às métricas objetivas de qualidade de imagem ou vídeo. A técnica é baseada na modificação do parâmetro B, que representa a média entre as medidas de blocagem nas direções horizontal e vertical de um bloco.

Palavras-Chave—Compressão de vídeo, avaliação de blocagem.

Abstract—This paper presents a new detection technique for blocking artifacts to images and video compressed that can serve as support for objective metrics of image or video quality. The technique is based on the modification of the parameter B that represents the average between the blocking measurements in the horizontal and vertical direction of a block.

Keywords—Video compression, blocking assessment.

I. INTRODUÇÃO

A Internet atualmente oferta diversos tipos de fluxos de vídeo para aplicações em tempo real ou vídeo armazenado. Devido à limitação de recursos, tais como largura de banda, esses fluxos de vídeo geralmente possuem baixa resolução (LD – *Low Definition*) e estão sujeitos a diversos tipos de degradação. Estas degradações também ocorrem em sistemas de TV Digital (TVD) e impactam diretamente a qualidade dos vídeos em todas as resoluções: baixa, padrão (SD – *Standard Definition*) e alta (HD – *High Definition*). Este impacto ocorre sob a forma de artefatos, dentre os quais cita-se o artefato de blocagem (*blocking*), que é um efeito causado por mudanças abruptas de intensidade em regiões de fronteira entre blocos adjacentes. O Sistema Visual Humano (SVH) interpreta essas descontinuidades entre as bordas de forma extremamente desagradável [1]. Logo, a detecção destes artefatos representa uma parte integrante dos algoritmos de avaliação objetiva de qualidade de vídeo. Dentre esses, destacam-se os algoritmos das métricas de referência reduzida (RR – *Reduced Reference*) e sem referência (NR – *No-Reference*).

Com relação a trabalhos relacionados, este artigo propõe alterar o parâmetro de detecção de blocagem desenvolvido por Wang, Sheikh e Bovik em [2]. Este trabalho desenvolveu uma nova métrica sem referência para avaliação objetiva de qualidade de imagens comprimidas em JPEG (*Joint Photographic Expert Group*). Para tanto, os autores utilizaram algoritmos para detecção de artefatos de blocagem e borramento e compararam os resultados obtidos com a métrica de referência completa PSNR (*Peak Signal-to-Noise Ratio*) e a métrica subjetiva MOS (*Mean Opinion Score*). Os autores em [3] basearam-se no trabalho desenvolvido em [2] para

aprimorar a métrica sem referência para avaliação de imagens em JPEG e estenderam-na para a avaliação objetiva de imagens comprimidas e codificadas em JPEG2000. Para tanto, os autores em [3] introduziram uma constante δ e um fator de correção ϵ , ambos aplicados sobre a técnica de detecção de blocagem desenvolvida por Wang, Sheikh e Bovik. Além disso, os autores apresentaram comparações entre as duas métricas, em termos do coeficiente de correlação de Pearson entre a métrica proposta e a MOS. Os autores em [4] desenvolveram uma métrica sem referência para avaliação objetiva de qualidade de vídeos comprimidos. O algoritmo de detecção do artefato de blocagem é baseado no gradiente local dos coeficientes DCT (*Discrete Cosine Transform*) nas direções horizontal e vertical. Os resultados demonstram que a técnica de detecção de blocagem apresenta forte correlação com a métrica subjetiva MOS.

Neste artigo é apresentada uma nova técnica de detecção de artefato de blocagem, designada por BMs (*Blocking Measure spatial*) que é uma versão alterada da técnica proposta em [2], desenvolvida originalmente para a detecção de blocagem em imagens comprimidas em JPEG, designada por B. Ambas as técnicas exibem comportamentos semelhantes quando aplicadas em imagens comprimidas em JPEG. Contudo, quando empregadas em vídeos codificados e comprimidos em H.264, a técnica BMs apresenta vantagem em relação à proposta em [2]. Isto ocorre porque o *encoder* H.264 ameniza o efeito causado pela blocagem através da aplicação de um filtro passa-baixas ou antiblocagem. Quando este filtro é ativado, o efeito de borramento fica mais evidente e influencia a técnica B. Logo, quanto maior for a compressão, maior será o efeito de borramento causado pelo filtro antiblocagem na compressão H.264. Consequentemente, menor será o índice de detecção da técnica B. Entretanto, a técnica BMs detecta a degradação visual causada pela compressão H.264, mesmo com a aplicação de pós-processamento na saída do *encoder*, através do filtro antiblocagem.

A Seção II descreve a técnica de detecção de artefato de blocagem proposta. Na Seção III apresenta-se uma comparação entre técnicas de detecção de blocagem e a Seção IV contém os resultados com a técnica proposta aplicada em imagem e vídeo de baixa resolução. As conclusões e os trabalhos futuros estão descritos na Seção V.

II. TÉCNICA DE DETECÇÃO DE BLOCAGEM

A técnica de detecção de blocagem proposta neste artigo deriva do trabalho desenvolvido por Wang, Sheikh e Bovik em 2002 [2] que considera o efeito de blocagem em um quadro representado por $x(m,n)$, com a componente de luminância em

Wyllian Bezerra da Silva e Alexandre de Almeida Prado Pohl, Laboratório de Comunicação de Dados, UTFPR – Universidade Tecnológica Federal do Paraná, Brasil, E-mails: wyllianbs@gmail.com, pohl@utfpr.edu.br. Este trabalho foi parcialmente financiado pela CAPES.

blocos de dimensão 8×8 pixels. A variável $m \in [1, M]$ e a variável $n \in [1, N]$, onde M e N representam a resolução do quadro, em termos de colunas e linhas, respectivamente. As diferenças ao longo das linhas (d_h) e das colunas (d_v) são expressas conforme as equações (1) e (2).

$$d_h(m, n) = x(m, n+1) - x(m, n), n \in [1, N-1], \quad (1)$$

$$d_v(m, n) = x(m+1, n) - x(m, n), m \in [1, M-1]. \quad (2)$$

A Figura 1 mostra onde opera a técnica de detecção de blocagem: nas bordas dos blocos adjacentes, cujas diferenças entre as componentes de luminância são calculadas nas direções horizontal e vertical. As Figuras 1(a), 1(b) e 1(c) ilustram as varreduras nas direções horizontal, vertical e a média em ambas as direções, representadas pelas variáveis B_h , B_v e B , respectivamente.

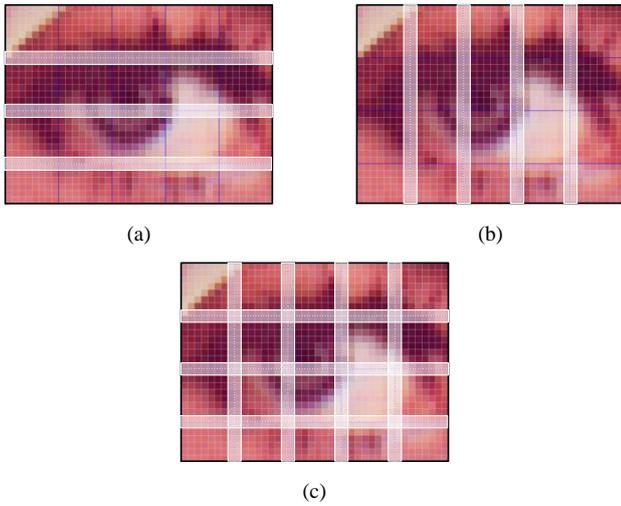


Fig. 1. Varredura entre blocos adjacentes nas direções: (a) Horizontal. (b) Vertical. (c) Horizontal e vertical.

Segundo os autores em [2], o efeito de blocagem pode ser estimado pela variável B , definida pela equação (5) que representa uma média entre as medidas de blocagem nas direções horizontal (B_h) e vertical (B_v), conforme as equações (3) e (4), respectivamente.

$$B_h = \frac{1}{M \left(\frac{N}{8} - 1 \right)} \sum_{i=1}^M \sum_{j=1}^{\left(\frac{N}{8} - 1 \right)} |d_h(i, 8j)|, \quad (3)$$

$$B_v = \frac{1}{N \left(\frac{M}{8} - 1 \right)} \sum_{i=1}^{\left(\frac{M}{8} - 1 \right)} \sum_{j=1}^N |d_v(8i, j)|, \quad (4)$$

$$B = \frac{B_h + B_v}{2}. \quad (5)$$

A Figura 2 ilustra as varreduras que compõem o fator de normalização proposto neste artigo. As Figuras 2(a) e 2(b) representam a variável D_l , composta pela diferença absoluta entre a varredura da primeira à penúltima linha, lado esquerdo do numerador da equação (6) e da segunda à última linha, lado direito do numerador da equação (6), respectivamente. As Figuras 2(c) e 2(d) representam a variável D_c , formada pela

diferença absoluta entre a varredura da primeira à penúltima coluna, lado esquerdo do numerador da equação (7) e da segunda à última coluna, lado direito do numerador da equação (7), respectivamente.

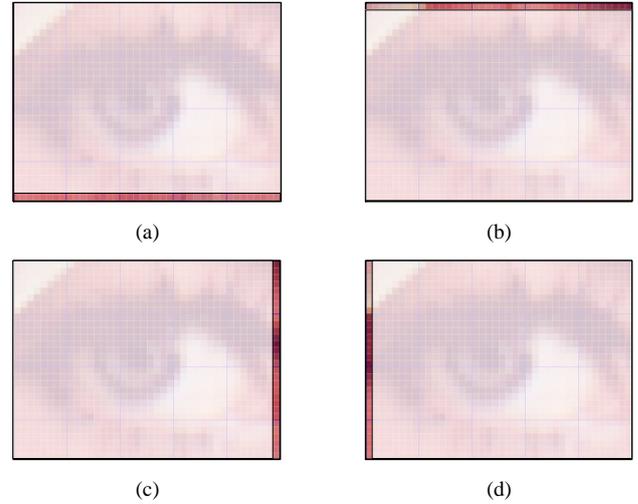


Fig. 2. Varredura do quadro: (a) Da primeira à penúltima linha. (b) Da segunda à última linha. (c) Da primeira coluna à penúltima. (d) Da segunda coluna à última.

Neste trabalho, a medida de intensidade de blocagem não emprega a média entre as equações (3) e (4), pois opera com um fator de normalização formado pelas diferenças absolutas entre as linhas e colunas do quadro/imagem, conforme a equação (8), composta pelas equações (6) e (7).

$$D_l = \frac{\sum_{k=1}^{M(N-1)} \left| \left(\sum_{i=1}^M \sum_{j=1}^{(N-1)} x(i, j) \right) - \left(\sum_{i=1}^M \sum_{j=2}^N x(i, j) \right) \right|}{M(N-1)}, \quad (6)$$

$$D_c = \frac{\sum_{k=1}^{N(M-1)} \left| \left(\sum_{i=1}^{(M-1)} \sum_{j=1}^N x(i, j) \right) - \left(\sum_{i=2}^M \sum_{j=1}^N x(i, j) \right) \right|}{N(M-1)}, \quad (7)$$

$$BMs = \frac{B}{D_l + D_c}. \quad (8)$$

A variável BMs fornece uma medida da intensidade do efeito de blocagem sobre quadros de uma imagem ou de uma sequência de vídeo. Assim, quando $BMs > 1$, implica em uma detecção do artefato de blocagem, ou seja, o numerador da equação (8) apresenta maior valor do que o seu denominador. Entretanto, quando $BMs \leq 1$, indica que o efeito de blocagem apresenta fraca intensidade ou seu efeito é imperceptível.

III. COMPARAÇÃO DAS TÉCNICAS DE DETECÇÃO DE BLOCAGEM

O algoritmo de detecção do artefato de blocagem realiza seu processamento sobre a componente de luminância de imagens comprimidas em JPEG ou de vídeos codificados. Neste trabalho, optou-se pela codificação do vídeo em H.264/AVC (*Advanced Video Coding*), conforme a Norma

ITU-T Rec. H.264 [5]. O *encoder* empregado nesta tarefa foi o x264 [6]. Este *software* livre com sintaxe em linha de comando tem como entrada um vídeo no formato YUV (sem compressão), cuja codificação é realizada com a aplicação de perfis e níveis de codificação, conforme a necessidade da aplicação. Dentre os parâmetros que influenciam na qualidade do vídeo, destaca-se o parâmetro de quantização QP (*Quantization Parameter*) que pode variar entre 0 e 51, onde: QP = 0 corresponde a uma compressão sem perdas e QP > 1 equivale a uma compressão com perdas. Dessa forma, quanto maior o valor do QP, mais evidente se tornarão os artefatos provenientes dessa compressão.

Este trabalho faz uma comparação entre a técnica descrita em [2] e a proposta neste trabalho. A Figura 3 ilustra a imagem da Lena com resolução de 512x512 *pixels* e os vídeos *Controlled Burn* e *Tractor*, disponíveis em HD na base de dados XPIH [7]. A resolução LD nestes vídeos foi obtida através do emprego de sintaxe de linha de comando do *software* FFMPEG [8].

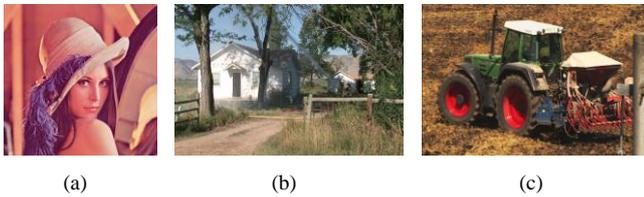


Fig. 3. (a) Imagem da Lena. (b) Vídeo *Controlled Burn* (LD) [7]. (c) Vídeo *Tractor* (LD) [7].

O processo de compressão, representado pela variável Q_{DCT} na equação (9), ocorre efetivamente com o processo de quantização sobre os coeficientes $DCT(u,v)$, cujo fator de qualidade f é aplicado à tabela JPEG padrão (Q_{50}), cujo resultado produzirá a compressão desejada. A variável Q_f é definida pela equação (10). Para $f = 50$, implicará em $Q_f = Q_{50}$ que equivale à tabela JPEG padrão. Logo, o fator de qualidade f pode variar entre 1% e 100%.

$$Q_{DCT} = \text{round} \left(\frac{DCT(u,v)}{Q_f(u,v)} \right), \quad (9)$$

onde:

$$Q_f = \begin{cases} \left(\frac{100 - f}{50} \right) * Q_{50}, & \text{se } Q_f \geq Q_{50} \\ \frac{50}{f} * Q_{50} & , \text{ se } Q_f < Q_{50} \end{cases}. \quad (10)$$

A métrica de referência completa SSIM (*Structural SIMilarity*) [9] indica a qualidade de uma imagem ou de um quadro de vídeo. Dessa forma, quanto mais próximo de um, melhor será o índice de qualidade do quadro ou da imagem. O cálculo dessa métrica leva em consideração a comparação de três informações: luminância, contraste e a estrutura do quadro/imagem. A Figura 4 compara o comportamento das duas técnicas de detecção de artefatos de blocagem e o índice de qualidade SSIM sobre a imagem da Figura 3(a) codificada e comprimida em JPEG, cuja qualidade varia conforme o

parâmetro de compressão definido pelo fator de qualidade f , relacionado ao processo de compressão. Os resultados nesta figura mostram que à medida que a qualidade da imagem aumenta (valor registrado no eixo das abscissas) maior será o índice SSIM e menores serão os valores calculados pelas técnicas B e BMs. A técnica BMs apresenta vantagem em relação à B quanto ao padrão de detecção do artefato de blocagem que é fixo. Dessa forma, quanto melhor a qualidade da imagem/quadro, mais próximo de 1 será o valor de BM, enquanto que a técnica B pode apresentar valores diferentes, devido ao conteúdo do quadro/imagem.

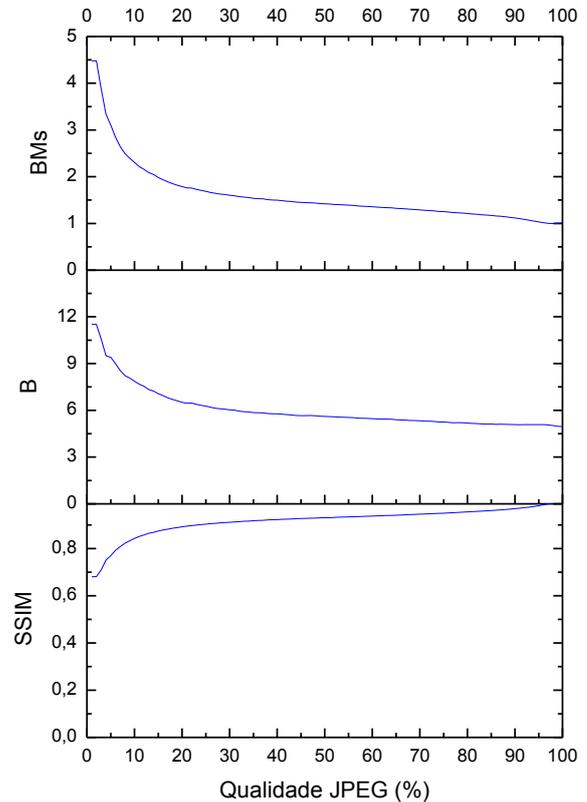


Fig. 4. Comparação entre as técnicas de detecção de artefatos B e BMs e a métrica SSIM.

A técnica BMs apresenta valores próximos a um, quando a qualidade da imagem é maior do que 90%, enquanto que a técnica B apresenta valores maiores do que 4,5. Este comportamento é esperado de ambas as técnicas, pois as mesmas registram o grau de evolução do artefato de blocagem. Todavia, a técnica B quando empregada em vídeo comprimido apresenta um comportamento diferente daquele observado em imagens comprimidas em JPEG.

IV. RESULTADOS

A medida de qualidade de vídeo aplicada em vídeos codificados/comprimidos relaciona o grau de distorção com a compressão com perdas, devido à aplicação do parâmetro de quantização QP. Essa compressão causa efeitos desagradáveis à percepção visual humana. O artefato de blocagem é a principal causa do desconforto na percepção visual de vídeos comprimidos. Esta seção compara os resultados das técnicas de

detecção de artefato de blocagem desenvolvida em [2] e a proposta neste artigo. A Figura 5 ilustra o efeito perceptual visual causado pela compressão em vídeo, através da aplicação do parâmetro QP, cuja qualidade visual decresce com o incremento deste parâmetro. Além disso, esta figura apresenta uma comparação entre as técnicas B e BMs em relação à métrica SSIM. A legenda da Figura 5 mostra os resultados numéricos. Observa-se que nas Figuras 5(c), 5(d) e 5(e) a técnica B expressa um decréscimo no nível de detecção, enquanto que a técnica BMs apresenta comportamento inverso, isto é, detecta o artefato de blocagem à medida que o parâmetro QP aumenta.

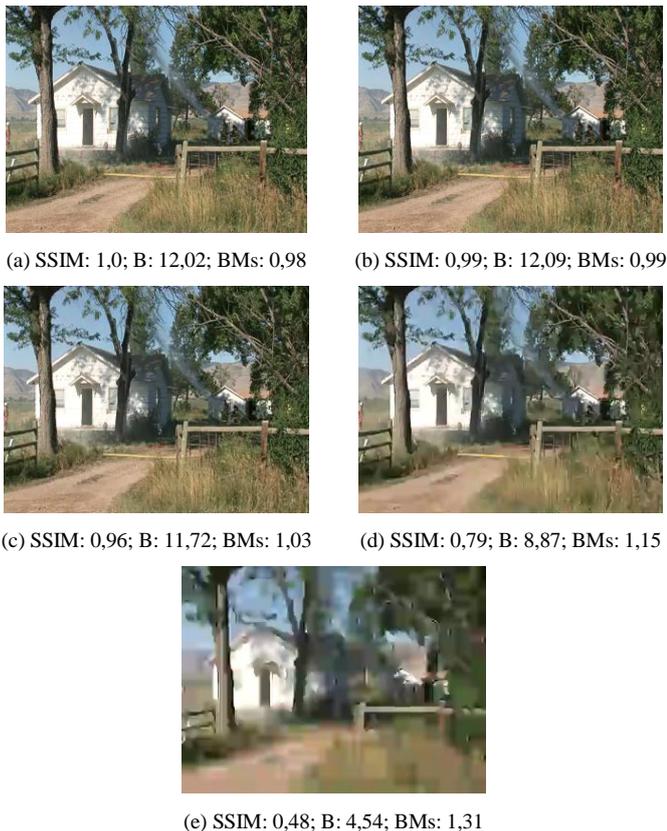


Fig. 5. Vídeo: *Controlled Burn* (LD – quadro 53), QP: (a) 10; (b) 20; (c) 30; (d) 40; (e) 51.

Contudo, observa-se também que o efeito de borramento fica mais evidente, quando a compressão for maior, efeito este causado pelo filtro antiblocagem na saída do codificador H.264. Mesmo assim, a degradação continua a ser “percebida” pelos algoritmos de detecção de blocagem que indicam a percepção do artefato, ou seja, a tendência de acréscimo da detecção de blocagem à medida que o parâmetro QP aumenta. Todavia, a técnica B não acompanha esta tendência, conforme pode ser observado nas Figuras 6 e 8. Nestes vídeos, a técnica B apresenta um decréscimo de seu valor à medida que o parâmetro QP aumenta. Por outro lado, a técnica BMs apresenta vantagens, conforme pode ser observado nas Figuras 7 e 9. Pois, à medida que o valor do parâmetro QP aumenta, a variável BMs acompanha o acréscimo do nível de degradação no vídeo, ou seja, registra a intensidade do artefato de blocagem. Logo, esta técnica apresenta o mesmo

comportamento de detecção do artefato de blocagem tanto para imagens quanto para vídeos codificados em H.264.

V. CONCLUSÕES

Este trabalho apresentou uma modificação à técnica de detecção de artefatos de blocagem desenvolvida por Wang, Sheikh e Bovik [2]. Os resultados mostram que a técnica proposta neste trabalho apresentou vantagens quando empregada em vídeos codificados e comprimidos em H.264, mesmo mediante pós-processamento com a aplicação de um filtro passa-baixas, que ameniza o efeito visual causado pela blocagem, proveniente do processo de compressão. Outras sequências de vídeo também foram submetidas às técnicas descritas neste trabalho e apresentaram resultados semelhantes. A técnica BMs poderá futuramente prover suporte às métricas objetivas de avaliação de qualidade de vídeo, sobretudo, às métricas NR e RR.

AGRADECIMENTOS

A CAPES pelo suporte ao projeto intitulado “Formação de Pessoal Qualificado em Sistemas de Transmissão de TV Digital no Paraná – Processo 23038.23556/2008-16 AUX-PE-RH-TVD 249/2008”.

REFERÊNCIAS

- [1] A. S. Al-Fohoum e A. M. Reza, “Combined edge crispness and statistical differencing for deblocking JPEG compressed images”, *IEEE Trans. Image Processing*, vol. 10, no. 9, pp. 1288-1298, Setembro, 2001.
- [2] Z. Wang; H. R. Sheikh e A. C. Bovik, “No-reference perceptual quality assessment of JPEG compressed images”, *In: Proc. Int. Conf. Image Processing*, vol. 1, pp. 477-480, Dezembro, 2002.
- [3] R. Ferzli e L. J. Karam, “No-reference image quality assessment for JPEG/JPEG2000 coding”, *In. Proc. EUSIPCO*, p. 1301, 2004.
- [4] R. Muijs e I. Kirenko, “A No-reference Blocking Artifact Measure for Adaptive Video Processing”, *In Proc. EUSIPCO*, 2005. Disponível em: <<http://citeseerx.ist.psu.edu>>.
- [5] ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG4-AVC), “Advanced Video Coding for Generic Audiovisual Services”, v5.0, Março, 2010.
- [6] X264, disponível em: <<http://x264.nl/>>.
- [7] XPIH, disponível em: <<http://media.xiph.org/video/derf/>>.
- [8] FFMPEG, disponível em: <<http://www.ffmpeg.org>>.
- [9] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, Abril, 2004.

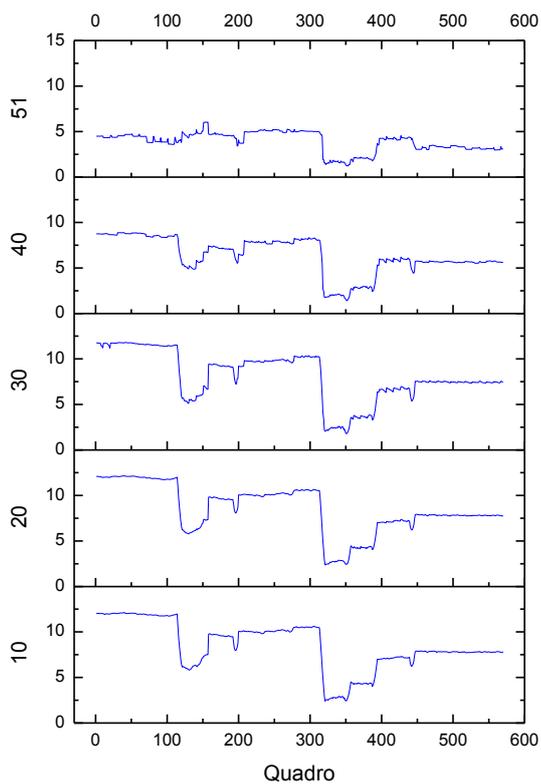


Fig. 6. Técnica B: *Controlled Burn* (LD).

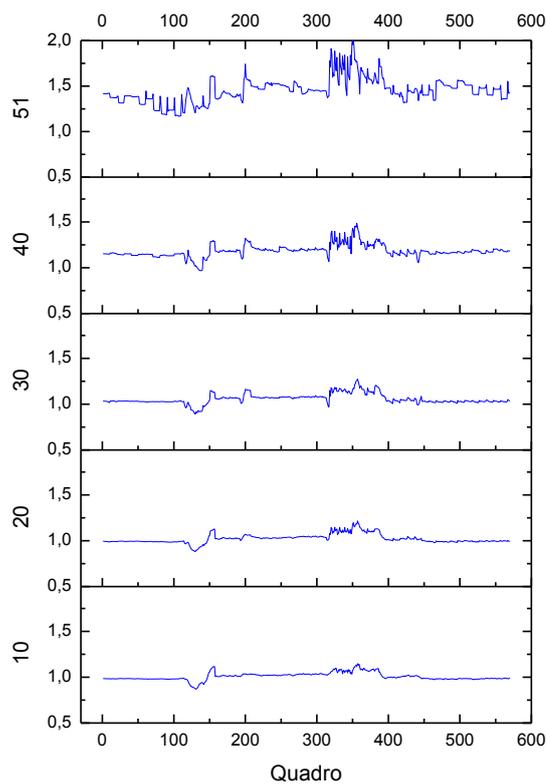


Fig. 7. Técnica BMs: *Controlled Burn* (LD).

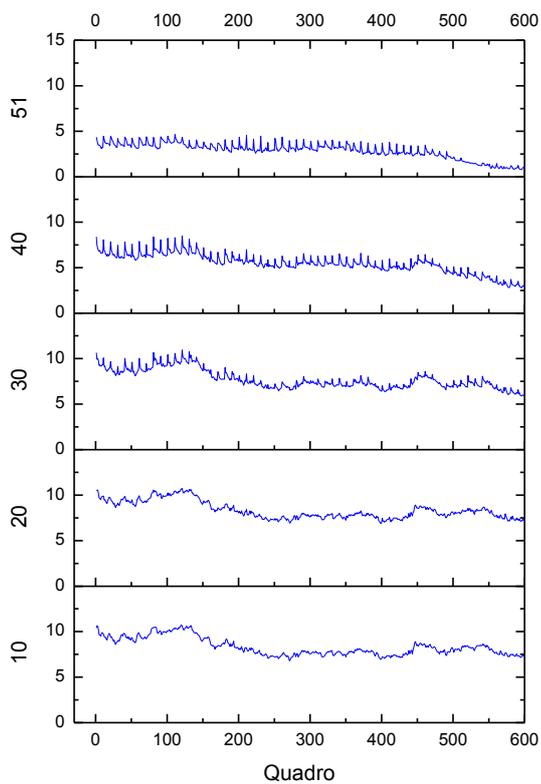


Fig. 8. Técnica B: *Tractor* (LD).

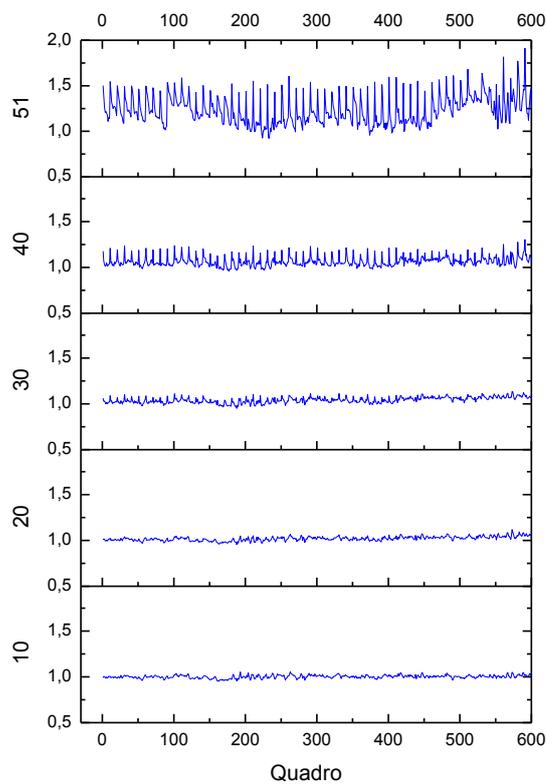


Fig. 9. Técnica BMs: *Tractor* (LD).