

Análise Estatística do Algoritmo LMS Aplicado à Reconstrução de Vídeo com Super-Resolução

Guilherme H. Costa e José C. M. Bermudez

Resumo—A reconstrução com super-resolução de uma seqüência de imagens é fortemente dependente da qualidade da estimação do movimento entre quadros sucessivos. Este trabalho apresenta uma análise estatística do algoritmo *Least Mean Square* (LMS) aplicado à reconstrução com super-resolução. Equações recursivas determinísticas, funções do erro de registro, são derivadas para o comportamento dos erros de reconstrução médio e médio quadrático. O novo modelo descreve o comportamento do algoritmo em situações reais, e apresenta considerável melhora na acurácia em relação ao modelo existente na literatura. Simulações de Monte Carlo demonstram boa concordância entre o comportamento promediado e o modelo proposto.

Palavras-Chave—Super-resolução, LMS, erro de registro, análise estatística, filtragem adaptativa.

Abstract—Super-resolution of image sequences is highly dependent on the quality of the motion estimation between successive frames. This work presents a statistical analysis of the Least Mean Square (LMS) algorithm applied to super-resolution of an image sequence. Deterministic recursions are derived for the mean and mean square behavior of the reconstruction error as a function of the registration errors. The new model describes the behavior of the algorithm in realistic situations, and significantly improves the accuracy of a simple model available in the literature. Monte Carlo simulations show good agreement between actual and predicted algorithm behaviors.

Keywords—Super-resolution, LMS, registration error, statistical analysis, adaptive filtering.

I. INTRODUÇÃO

A reconstrução com super-resolução (SRR — *super-resolution reconstruction*) consiste basicamente em combinar múltiplas imagens de baixa resolução para formar uma imagem de alta (maior) resolução. O processo de SRR pode ser dividido em dois estágios. O primeiro estágio consiste em registrar, ou alinhar, as imagens. Isto é, determinar a posição que cada um dos pixels (objetos) de uma imagem ocupa nas demais. O segundo estágio consiste em unir as múltiplas imagens de baixa resolução (depois de alinhadas) formando uma única imagem de alta resolução.

Em [1], é apresentada uma visão geral sobre os trabalhos publicados em torno desse tema nos últimos anos. Tais trabalhos podem ser divididos em dois grupos: super-resolução de imagens, e super-resolução de seqüências de vídeo. Em aplicações de reconstrução de vídeo (seqüências de imagens), geralmente existe o requisito de processamento em tempo-real. Por esse motivo, diversos algoritmos recursivos têm sido propostos ([2], [3] e referências em [1]). Sabe-se ainda que um registro acurado é crucial para que bons resultados de

SRR sejam alcançados [4], [5], [6]. Quando o movimento entre as imagens é estimado de forma inexata, o algoritmo tende a degradar ainda mais as imagens observadas, ao invés de melhorá-las. Essa degradação é chamada de ruído de erro de registro. Todavia, a qualidade da estimação do movimento não depende apenas do método de estimação utilizado. As características da aplicação e dos sinais de imagem tornam o registro das imagens um problema mal condicionado [7]. Por esse motivo, muitos trabalhos têm procurado por algoritmos mais robustos ao erro de registro. Em [5], Wang e Qi propõem um algoritmo baseado no filtro de Kalman. Nesse trabalho, as incertezas de registro são incluídas (previstas) nas equações do filtro. Em [6], Lee e Kang consideram o ruído de erro de registro para determinar de forma adaptativa o parâmetro de regularização usado na SRR. Em [8], [9], dois algoritmos, baseados na filtragem de Kalman, são propostos. As soluções apresentadas consistem em um estimador *Least Squares* e em um *Least Mean Square*, nos quais a imagem de alta resolução é reconstruída sem a necessidade da inversão de matrizes (ao contrário do algoritmo de Kalman tradicional). Devido à sua reduzida complexidade computacional e conseqüente aplicabilidade em tempo-real, torna-se importante investigar o desempenho de tais algoritmos adaptativos. Esses algoritmos são ditos pertencentes à abordagem de filtragem adaptativa (embora não se tratem de algoritmos de filtragem e sim de estimação) [1], [9]. Entretanto, diferentemente dos tradicionais algoritmos de filtragem adaptativa, existe pouca informação disponível sobre o comportamento desses algoritmos de SRR. A análise existente é limitada ao estudo das condições de convergência e do comportamento em condições bastante restritivas, que pouco nos dizem sobre o desempenho dessas técnicas em situações práticas.

Na maioria dos trabalhos citados acima ([2]-[6]), a proposição e a comparação de soluções são geralmente feitas de uma forma *ad hoc*; pela observação dos resultados ou pela comparação do escopo de aplicação de cada uma das soluções. Embora muitos atentem para o problema do erro de registro, apenas [8] apresenta uma análise, embora muito simples, do desempenho da solução proposta frente a erros de registro.

Este trabalho apresenta uma análise detalhada do algoritmo LMS proposto em [8], [9], evidenciando a influência do erro de registro no resultado da reconstrução das imagens. Na Seção II, o algoritmo de [8], [9] é revisto. Na Seção III o algoritmo é analisado. Um modelo analítico é proposto e sua implementação é discutida. Na Seção IV, são apresentados alguns resultados referentes ao modelo proposto. Finalmente, na Seção V, são apresentadas as conclusões.

Guilherme H. Costa e José C. M. Bermudez, Departamento de Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, Brasil, E-mails: holsbach@eel.ufsc.br, j.bermudez@ieee.org. Este trabalho foi parcialmente financiado pelo CNPq.

II. RECONSTRUÇÃO COM SUPER-RESOLUÇÃO VIA ALGORITMO LMS

O algoritmo *Least Mean Square* (LMS), proposto em [8], [9] será apresentado nesta seção através de uma formulação alternativa que deduz o algoritmo de forma mais direta, a partir do conceito de gradiente estocástico [10].

A. Nomenclatura

De forma a diferenciar o algoritmo LMS de filtragem adaptativa [10] do algoritmo estudado neste trabalho, a solução proposta em [8], [9] será aqui denominada de LMS-SRR.

Em relação às imagens envolvidas em um processo de SRR, define-se como “baixa resolução” a dimensão das imagens observadas, e como “alta resolução” a dimensão das imagens que se deseja obter (reconstruir). O termo “grade de alta/baixa resolução” será tratado como sinônimo de espaço (vetorial) da alta/baixa resolução.

No decorrer deste trabalho pouca referência é feita às características dos sinais de imagem envolvidos no processo de SRR. Assim, o termo “bordas” será utilizado para descrever os limites de uma imagem, e não as regiões com alta frequência espacial.

B. Modelagem dos sinais envolvidos

Os sinais são modelados de forma similar ao proposto em [9]. Dadas uma imagem de baixa resolução (observada) $\mathbf{Y}(t)$, de tamanho $N \times N$, e um imagem de alta resolução $\mathbf{X}(t)$, de tamanho $M \times M$, em que $M > N$, o modelo adotado para a degradação das imagens de alta resolução é

$$\mathbf{y}(t) = \mathbf{D}\mathbf{x}(t) + \mathbf{e}(t), \quad (1)$$

em que $\mathbf{y}(t)$ e $\mathbf{x}(t)$ são, respectivamente, as representações lexicográficas da imagem observada (degradada) e da imagem original (que se deseja reconstruir), no instante de tempo t . A matriz \mathbf{D} modela a degradação (sub-amostragem + distorções ópticas), e possui dimensão $N^2 \times M^2$. O vetor $\mathbf{e}(t)$, de dimensão $N^2 \times 1$, modela o ruído de observação (ruído eletrônico), considerado branco gaussiano e independente de $\mathbf{y}(t)$ e de $\mathbf{x}(t)$.

Sem perda de generalidade, neste trabalho será assumido que a sub-amostragem desempenhada por \mathbf{D} modela a integração espacial que ocorre em sensores CCD com área quadrada e com foto-sensibilidade uniforme, em que a distância interpixel é desconsiderada [11]. As distorções ópticas (causadas pelo sistema de lentes) não serão consideradas mas, desde que conhecidas, podem ser incluídas na matriz \mathbf{D} sem nenhuma consequência para a análise.

A dinâmica do sinal de entrada é modelada por

$$\mathbf{x}(t) = \mathbf{G}(t)\mathbf{x}(t-1) + \mathbf{s}(t), \quad (2)$$

em que $\mathbf{G}(t)$ é a matriz de registro (descreve o movimento entre $\mathbf{x}(t)$ e $\mathbf{x}(t-1)$) e $\mathbf{s}(t)$ modela as inovações em $\mathbf{x}(t)$. O movimento representado pela matriz de registro é considerado global (o mesmo vetor de deslocamento para todos os pixels) e translacional. Assim, $\mathbf{G}(t)$ é função de apenas um vetor de deslocamento $\mathbf{d}(t) = [dc(t), dl(t)]$, em que $dc(t)$ e $dl(t)$ são,

respectivamente, os deslocamentos nas direções das colunas e das linhas, de $\mathbf{x}(t-1)$ para $\mathbf{x}(t)$.

A matriz registro, quando utilizada no algoritmo de reconstrução, obedece às condições de fronteira de Neumann, o que equivale à hipótese de que os pixels externos à cena observada são idênticos aos pixels internos à cena, espelhados em relação às bordas da imagem. Essa condição traz os melhores resultados de reconstrução [12]. Entretanto, no modelo de comportamento aqui proposto, $\mathbf{G}(t)$ é construída considerando-se que a imagem é periódica, i.e., os pixels externos à cena são considerados como repetições da imagem observada. Essa consideração simplifica consideravelmente a implementação do modelo, e o erro causado pela diferença nas matrizes de registro do modelo e do algoritmo de reconstrução não é relevante, como será visto nas próximas seções.

C. Figura de mérito

Tradicionalmente, as soluções para SRR são baseadas na minimização da norma

$$\|\epsilon(t)\| = \|\mathbf{y}(t) - \mathbf{D}\hat{\mathbf{x}}(t)\|, \quad (3)$$

em que $\hat{\mathbf{x}}(t)$ é a imagem estimada (reconstruída). Tal minimização é então feita em um determinado sentido (por exemplo: *mean square*, *least squares*, etc.) e, por vezes, submetida a restrições. Considerem a minimização realizada no sentido médio quadrático e sem nenhum tipo de restrição (regularização). Assim, a função custo, MSE (*mean square error*), será definida por

$$E\{\|\epsilon(t)\|^2\} = E\{\|\mathbf{y}(t) - \mathbf{D}\hat{\mathbf{x}}(t)\|^2\}. \quad (4)$$

É importante ressaltar que o valor esperado em (4) representa o valor médio no sentido do *ensemble* e não a média temporal.

D. Método do Gradiente (Steepest Descent)

De acordo com o método do gradiente, a atualização de $\hat{\mathbf{x}}(t)$, de forma a minimizar a função custo, deve se dar no sentido contrário ao do gradiente. Assim, sendo a superfície de desempenho [10]

$$\mathbf{J}_{\text{MS}}(t) = E\{\|\mathbf{y}(t) - \mathbf{D}\hat{\mathbf{x}}(t)\|^2\}, \quad (5)$$

seu gradiente é dado por

$$\nabla \mathbf{J}_{\text{MS}}(t) = \frac{\partial \mathbf{J}_{\text{MS}}(t)}{\partial \hat{\mathbf{x}}(t)} = -2\mathbf{D}^T \{E[\mathbf{y}(t)] - \mathbf{D}\hat{\mathbf{x}}(t)\}. \quad (6)$$

Portanto, a equação de atualização de $\hat{\mathbf{x}}(t)$ é dada por

$$\hat{\mathbf{x}}_{k+1}(t) = \hat{\mathbf{x}}_k(t) - \mu \frac{1}{2} \nabla \mathbf{J}_{\text{MS}}(t) \quad (7)$$

$$= \hat{\mathbf{x}}_k(t) + \mu \mathbf{D}^T \{E[\mathbf{y}(t)] - \mathbf{D}\hat{\mathbf{x}}_k(t)\}. \quad (8)$$

Note que a recursão de $\hat{\mathbf{x}}(t)$, i.e., o deslocamento sobre a superfície de desempenho $\mathbf{J}_{\text{MS}}(t)$, se dá em k e não em t . Ou seja, uma diferente superfície de desempenho é definida para cada instante de tempo t .

E. O algoritmo LMS-SRR

O algoritmo LMS é uma aproximação estocástica do algoritmo do Gradiente, em que $\nabla \mathbf{J}_{\text{MS}}(t)$ é aproximado por sua

Inicialização:

- Definir núm. de iter. p/cada amostra temporal: K
- $\hat{\mathbf{x}}_0(0) =$ interpolação de $\mathbf{y}(0)$

Algoritmo:

$$\left\{ \begin{array}{l} \text{Laço em } t = 0, 1, 2, \dots \\ \left\{ \begin{array}{l} \text{Laço em } k = 0, 1, \dots, K \\ \hat{\mathbf{x}}_{k+1}(t) = \hat{\mathbf{x}}_k(t) + \mu \mathbf{D}^T [\mathbf{y}(t) - \mathbf{D} \hat{\mathbf{x}}_k(t)] \end{array} \right. \\ \hat{\mathbf{x}}_0(t+1) = \mathbf{G}(t+1) \hat{\mathbf{x}}_{K+1}(t) \end{array} \right.$$

Fig. 1. Algoritmo LMS aplicado à SRR.

estimativa instantânea [10]. Portanto, fazendo-se

$$\begin{aligned} \nabla \mathbf{J}_{\text{MS}}(t) &= -2\mathbf{D}^T \{ \mathbf{E}[\mathbf{y}(t)] - \mathbf{D} \hat{\mathbf{x}}(t) \} \\ &\simeq -2\mathbf{D}^T [\mathbf{y}(t) - \mathbf{D} \hat{\mathbf{x}}(t)] = \hat{\nabla} \mathbf{J}_{\text{MS}}(t), \end{aligned} \quad (9)$$

tem-se diretamente de (7) que

$$\hat{\mathbf{x}}_{k+1}(t) = \hat{\mathbf{x}}_k(t) + \mu \mathbf{D}^T [\mathbf{y}(t) - \mathbf{D} \hat{\mathbf{x}}_k(t)]. \quad (10)$$

A atualização em t , por sua vez, é baseada na equação da dinâmica do sinal (2). Assim, o algoritmo LMS-SRR é definido na Figura 1. Um caso importante desse algoritmo ocorre para $K = 0$ (LMS-SRR-1) [8]. Assim:

$$\hat{\mathbf{x}}_1(t) = \hat{\mathbf{x}}_0(t) + \mu \mathbf{D}^T [\mathbf{y}(t) - \mathbf{D} \hat{\mathbf{x}}_0(t)], \quad (11)$$

em que

$$\hat{\mathbf{x}}_0(t) = \mathbf{G}(t) \hat{\mathbf{x}}_1(t-1). \quad (12)$$

Substituindo-se (12) em (11),

$$\hat{\mathbf{x}}_1(t) = \mathbf{G}(t) \hat{\mathbf{x}}_1(t-1) + \mu \mathbf{D}^T [\mathbf{y}(t) - \mathbf{D} \mathbf{G}(t) \hat{\mathbf{x}}_1(t-1)], \quad (13)$$

ou, eliminando-se o índice 1, agora constante,

$$\hat{\mathbf{x}}(t) = \mathbf{G}(t) \hat{\mathbf{x}}(t-1) + \mu \mathbf{D}^T [\mathbf{y}(t) - \mathbf{D} \mathbf{G}(t) \hat{\mathbf{x}}(t-1)]. \quad (14)$$

III. ANÁLISE DO ALGORITMO

A partir deste ponto, $\mathbf{G}(t)$ será considerada conhecida (determinística) e irá representar a matriz de registro exata (livre de erros de estimação de movimento). A matriz $\hat{\mathbf{G}}(t)$ representará a matriz de registro estimada (com erros). Além disso, a dinâmica do sinal de entrada, descrita em (2), será considerada como

$$\mathbf{x}(t) = \mathbf{G}(t) \mathbf{x}(t-1), \quad (15)$$

em que não são consideradas inovações entre $\mathbf{x}(t)$ e $\mathbf{x}(t-1)$. Esta consideração é razoável para a representação de movimentos globais translacionais, com deslocamentos entre dois quadros consecutivos muito menores do que o tamanho das imagens envolvidas. Aplicações mais gerais também podem ser bem representadas por (15), desde que (15) modele apenas o deslocamento dos pixels do centro (longe das bordas) das imagens. Com base nessa consideração, valores referentes aos pixels próximos às bordas das imagens serão desconsiderados no cômputo dos erros médios modelados, uma vez que não são bem representados pelos modelos propostos.

A. Modelo do erro médio de reconstrução

Considerando agora que apenas $\hat{\mathbf{G}}(t)$ esteja disponível no processo de reconstrução, a equação de recursão do algoritmo LMS-SRR-1 (14) é dada por:

$$\hat{\mathbf{x}}(t) = \hat{\mathbf{G}}(t) \hat{\mathbf{x}}(t-1) + \mu \mathbf{D}^T [\mathbf{y}(t) - \mathbf{D} \hat{\mathbf{G}}(t) \hat{\mathbf{x}}(t-1)], \quad (16)$$

em que $\hat{\mathbf{G}}(t)$ pode ser modelada como

$$\hat{\mathbf{G}}(t) = \mathbf{G}(t) + \Delta \mathbf{G}(t). \quad (17)$$

Segue de (16), de (17) e de (15), que o erro de estimação $\mathbf{v}(t) = \hat{\mathbf{x}}(t) - \mathbf{x}(t)$ é dado por

$$\begin{aligned} \mathbf{v}(t) &= (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \mathbf{G}(t) \mathbf{v}(t-1) \\ &\quad + (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \Delta \mathbf{G}(t) \hat{\mathbf{x}}(t-1) + \mu \mathbf{D}^T \mathbf{e}(t). \end{aligned} \quad (18)$$

Tirando o valor esperado em ambos os lados de (18), chega-se à equação do comportamento do erro médio de reconstrução. Dadas as características de construção de $\mathbf{G}(t)$, e, conseqüentemente, de $\hat{\mathbf{G}}(t)$ e de $\Delta \mathbf{G}(t)$, o vetor $\mathbf{r}(t) = \Delta \mathbf{G}(t) \hat{\mathbf{x}}(t-1)$ possui média zero [6], [8]. Assim, o erro médio de reconstrução fica:

$$\mathbf{E}[\mathbf{v}(t)] = (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \mathbf{G}(t) \mathbf{E}[\mathbf{v}(t-1)]. \quad (19)$$

B. Modelo do erro médio quadrático de reconstrução

Buscando-se agora um modelo para o erro médio quadrático de reconstrução, $\mathbf{E}[\mathbf{v}^T(t) \mathbf{v}(t)]$, será feito uso da propriedade $\mathbf{E}[\mathbf{v}^T(t) \mathbf{v}(t)] = \text{tr}\{\mathbf{E}[\mathbf{v}^T(t) \mathbf{v}(t)]\} = \text{tr}\{\mathbf{E}[\mathbf{v}(t) \mathbf{v}^T(t)]\} = \text{tr}\{\mathbf{K}(t)\}$, em que o operador $\text{tr}\{\cdot\}$ representa o traço de uma matriz. Pós-multiplicando (18) pela sua transposta,

$$\begin{aligned} \mathbf{v}(t) \mathbf{v}^T(t) &= (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \mathbf{G}(t) \mathbf{v}(t-1) \mathbf{v}^T(t-1) \mathbf{G}^T(t) \\ &\quad \times (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) + (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \mathbf{G}(t) \mathbf{v}(t-1) \\ &\quad \times \hat{\mathbf{x}}^T(t-1) \Delta \mathbf{G}^T(t) (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \\ &\quad + \mu (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \mathbf{G}(t) \mathbf{v}(t-1) \mathbf{e}^T(t) \mathbf{D} \\ &\quad + (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \Delta \mathbf{G}(t) \hat{\mathbf{x}}(t-1) \mathbf{v}^T(t-1) \mathbf{G}^T(t) \\ &\quad \times (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) + (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \Delta \mathbf{G}(t) \hat{\mathbf{x}}(t-1) \\ &\quad \times \hat{\mathbf{x}}^T(t-1) \Delta \mathbf{G}^T(t) (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) + \mu (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \\ &\quad \times \Delta \mathbf{G}(t) \hat{\mathbf{x}}(t-1) \mathbf{e}^T(t) \mathbf{D} + \mu \mathbf{D}^T \mathbf{e}(t) \mathbf{v}^T(t-1) \\ &\quad \times \mathbf{G}^T(t) (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) + \mu \mathbf{D}^T \mathbf{e}(t) \hat{\mathbf{x}}^T(t-1) \\ &\quad \times \Delta \mathbf{G}^T(t) (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) + \mu^2 \mathbf{D}^T \mathbf{e}(t) \mathbf{e}^T(t) \mathbf{D}. \end{aligned} \quad (20)$$

Considerando que $\mathbf{e}(t)$ e $\Delta \mathbf{G}(t)$ são descorrelacionados entre si, assim como são descorrelacionados de $\mathbf{v}(t-1)$ e de $\hat{\mathbf{x}}(t-1)$, tirando o valor esperado em ambos os lados de (20) tem-se que:

$$\begin{aligned} \mathbf{E}[\mathbf{v}(t) \mathbf{v}^T(t)] &= (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \mathbf{G}(t) \mathbf{E}[\mathbf{v}(t-1) \mathbf{v}^T(t-1)] \\ &\quad \times \mathbf{G}^T(t) (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) + (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \mathbf{G}(t) \\ &\quad \times \mathbf{E}[\mathbf{v}(t-1) \hat{\mathbf{x}}^T(t-1)] \mathbf{E}[\Delta \mathbf{G}^T(t)] (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \\ &\quad + (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \mathbf{E}[\Delta \mathbf{G}(t)] \mathbf{E}[\hat{\mathbf{x}}(t-1) \mathbf{v}^T(t-1)] \\ &\quad \times \mathbf{G}^T(t) (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) + (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) \\ &\quad \times \mathbf{E}[\Delta \mathbf{G}(t) \hat{\mathbf{x}}(t-1) \hat{\mathbf{x}}^T(t-1) \Delta \mathbf{G}^T(t)] \\ &\quad \times (\mathbf{I} - \mu \mathbf{D}^T \mathbf{D}) + \mu^2 \mathbf{D}^T \mathbf{E}[\mathbf{e}(t) \mathbf{e}^T(t)] \mathbf{D}, \end{aligned} \quad (21)$$

ou, fazendo $E[\mathbf{v}(t)\mathbf{v}^T(t)] = \mathbf{K}(t)$,

$$\begin{aligned} \mathbf{K}(t) &= (\mathbf{I} - \mu\mathbf{D}^T\mathbf{D})\mathbf{G}(t)\mathbf{K}(t-1)\mathbf{G}^T(t)(\mathbf{I} - \mu\mathbf{D}^T\mathbf{D}) \\ &+ (\mathbf{I} - \mu\mathbf{D}^T\mathbf{D})\mathbf{G}(t)E[\mathbf{v}(t-1)\hat{\mathbf{x}}^T(t-1)] \\ &\times \left\{ E[\hat{\mathbf{G}}^T(t)] - \mathbf{G}^T(t) \right\} (\mathbf{I} - \mu\mathbf{D}^T\mathbf{D}) \\ &+ (\mathbf{I} - \mu\mathbf{D}^T\mathbf{D}) \left\{ E[\hat{\mathbf{G}}(t)] - \mathbf{G}(t) \right\} \\ &\times E[\hat{\mathbf{x}}(t-1)\mathbf{v}^T(t-1)]\mathbf{G}^T(t)(\mathbf{I} - \mu\mathbf{D}^T\mathbf{D}) \\ &+ (\mathbf{I} - \mu\mathbf{D}^T\mathbf{D})\mathbf{R}_r(t)(\mathbf{I} - \mu\mathbf{D}^T\mathbf{D}) \\ &+ \mu^2\mathbf{D}^T\mathbf{R}_e\mathbf{D}. \end{aligned} \quad (22)$$

No lado direito de (22) existem três valores esperados a serem resolvidos. Eles são:

- 1) O valor esperado $E[\hat{\mathbf{x}}(t-1)\mathbf{v}^T(t-1)]$, que pode ser escrito como

$$\begin{aligned} E[\hat{\mathbf{x}}(t-1)\mathbf{v}^T(t-1)] &= E\{[\mathbf{v}(t-1) + \mathbf{x}(t-1)]\mathbf{v}^T(t-1)\} \\ &= E[\mathbf{v}(t-1)\mathbf{v}^T(t-1)] + E[\mathbf{x}(t-1)\mathbf{v}^T(t-1)] \\ &= \mathbf{K}(t-1) + E\{\mathbf{x}(t-1)[\hat{\mathbf{x}}(t-1) - \mathbf{x}(t-1)]^T\} \\ &= \mathbf{K}(t-1) + \mathbf{R}_{\mathbf{x}\hat{\mathbf{x}}}(t-1) - \mathbf{R}_{\mathbf{x}}(t-1), \end{aligned} \quad (23)$$

em que $\mathbf{R}_{\mathbf{x}\hat{\mathbf{x}}}(t-1)$ pode ser determinada de forma recursiva, a partir de (15) e de (16):

$$\begin{aligned} \mathbf{R}_{\mathbf{x}\hat{\mathbf{x}}}(t) &= E[\mathbf{x}(t)\hat{\mathbf{x}}^T(t)] \\ &= E\left\{ \mathbf{x}(t) \left\{ \hat{\mathbf{G}}(t)\hat{\mathbf{x}}(t-1) \right. \right. \\ &\quad \left. \left. + \mu\mathbf{D}^T[\mathbf{y}(t) - \mathbf{D}\hat{\mathbf{G}}(t)\hat{\mathbf{x}}(t-1)] \right\}^T \right\} \\ &= E[\mathbf{x}(t)\hat{\mathbf{x}}^T(t-1)\hat{\mathbf{G}}^T(t)] + \mu E[\mathbf{x}(t)\mathbf{y}^T(t)]\mathbf{D} \\ &\quad - \mu E[\mathbf{x}(t)\hat{\mathbf{x}}^T(t-1)\hat{\mathbf{G}}^T(t)]\mathbf{D}^T\mathbf{D} \\ &= \mathbf{G}(t)E[\mathbf{x}(t-1)\hat{\mathbf{x}}^T(t-1)]E[\hat{\mathbf{G}}^T(t)] \\ &\quad + \mu E\{\mathbf{x}(t)[\mathbf{D}\mathbf{x}(t) + \mathbf{e}(t)]^T\}\mathbf{D} \\ &\quad - \mu\mathbf{G}(t)E[\mathbf{x}(t-1)\hat{\mathbf{x}}^T(t-1)]E[\hat{\mathbf{G}}^T(t)] \\ &= \mathbf{G}(t)\mathbf{R}_{\mathbf{x}\hat{\mathbf{x}}}(t-1)E[\hat{\mathbf{G}}^T(t)] + \mu\mathbf{R}_{\mathbf{x}}(t)\mathbf{D}^T\mathbf{D} \\ &\quad - \mu\mathbf{G}(t)\mathbf{R}_{\mathbf{x}\hat{\mathbf{x}}}(t-1)E[\hat{\mathbf{G}}^T(t)] \\ &= (1-\mu)\mathbf{G}(t)\mathbf{R}_{\mathbf{x}\hat{\mathbf{x}}}(t-1)E[\hat{\mathbf{G}}^T(t)] \\ &\quad + \mu\mathbf{R}_{\mathbf{x}}(t)\mathbf{D}^T\mathbf{D}; \end{aligned} \quad (24)$$

- 2) A matriz correlação $\mathbf{R}_r(t)$, que pode ser determinada como segue: assumindo que os elementos de $\mathbf{r}(t)$ são independentes, a matriz de correlação pode ser modelada como $E[\mathbf{r}(t)\mathbf{r}^T(t)] = \mathbf{R}_r(t) = \sigma_r^2(t)\mathbf{I}$. Assim, $\sigma_r^2(t) = \text{tr}[\mathbf{R}_r(t)]/M^2$, em que M^2 é o número de pixels das imagens de alta resolução. Usando a propriedade comutativa do traço de um produto de matrizes, tem-se que

$$\begin{aligned} \text{tr}[\mathbf{R}_r(t)] &= \text{tr}\{E[\Delta\mathbf{G}(t)\hat{\mathbf{x}}(t-1)\hat{\mathbf{x}}^T(t-1)\Delta\mathbf{G}^T(t)]\} \\ &= \text{tr}\{E[\hat{\mathbf{x}}(t-1)\hat{\mathbf{x}}^T(t-1)\Delta\mathbf{G}^T(t)\Delta\mathbf{G}(t)]\} \\ &= \text{tr}\{\mathbf{R}_{\hat{\mathbf{x}}}(t-1)E[\Delta\mathbf{G}^T(t)\Delta\mathbf{G}(t)]\}. \end{aligned} \quad (25)$$

Manipulando-se algebricamente a expressão que define

$\mathbf{R}_{\hat{\mathbf{x}}}(t)$, tem-se que

$$\begin{aligned} \mathbf{R}_{\hat{\mathbf{x}}}(t) &= E[\hat{\mathbf{x}}(t)\hat{\mathbf{x}}^T(t)] \\ &= E\{[\mathbf{v}(t) + \mathbf{x}(t)][\mathbf{v}(t) + \mathbf{x}(t)]^T\} \\ &= \mathbf{K}(t) + E[\mathbf{v}(t)\mathbf{x}(t)^T] + E[\mathbf{x}(t)\mathbf{v}(t)^T] + \mathbf{R}_{\mathbf{x}}(t) \\ &= \mathbf{K}(t) + E\{[\hat{\mathbf{x}}(t) - \mathbf{x}(t)]\mathbf{x}(t)^T\} \\ &\quad + E\{\mathbf{x}(t)[\hat{\mathbf{x}}(t) - \mathbf{x}(t)]^T\} + \mathbf{R}_{\mathbf{x}}(t) \\ &= \mathbf{K}(t) + \mathbf{R}_{\mathbf{x}\hat{\mathbf{x}}}(t) + \mathbf{R}_{\mathbf{x}\hat{\mathbf{x}}}(t) - \mathbf{R}_{\mathbf{x}}(t). \end{aligned} \quad (26)$$

O valor esperado $E[\Delta\mathbf{G}^T(t)\Delta\mathbf{G}(t)]$, em (25), pode ser escrito como

$$\begin{aligned} E[\Delta\mathbf{G}^T(t)\Delta\mathbf{G}(t)] &= E\{[\hat{\mathbf{G}}(t) - \mathbf{G}(t)]^T[\hat{\mathbf{G}}(t) - \mathbf{G}(t)]\} \\ &= E[\hat{\mathbf{G}}^T(t)\hat{\mathbf{G}}(t)] - E[\hat{\mathbf{G}}^T(t)]\mathbf{G}(t) \\ &\quad - \mathbf{G}^T(t)E[\hat{\mathbf{G}}(t)] + \mathbf{G}^T(t)\mathbf{G}(t). \end{aligned} \quad (27)$$

- 3) Finalmente, os valores esperados $E[\hat{\mathbf{G}}(t)]$ e $E[\hat{\mathbf{G}}^T(t)\hat{\mathbf{G}}(t)]$, que aparecem em (22) - (27), podem ser, sob certas condições, facilmente estimados a partir de ruídos brancos gaussianos (ver Seção III-C).

C. Implementação

Como foi visto na subseção anterior, o modelo proposto para o erro médio quadrático é uma função de dois valores esperados não resolvidos, $E[\hat{\mathbf{G}}(t)]$ e $E[\hat{\mathbf{G}}^T(t)\hat{\mathbf{G}}(t)]$. Entretanto, esses valores esperados podem ser estimados, com custo computacional satisfatório, a partir de ruído branco gaussiano, como será visto nesta seção.

Outra maneira de representar a matriz de registro estimada, diferente de (17), é

$$\hat{\mathbf{G}}(t) = \mathbf{G}(t) + \Delta\mathbf{G}(t) = \tilde{\mathbf{G}}(t)\mathbf{G}(t). \quad (28)$$

Note que, uma vez que $\mathbf{G}(t)$ no modelo teórico possui posto completo (mediante as considerações feitas na Seção II-E), existe uma solução $\tilde{\mathbf{G}}(t) = [\mathbf{G}(t) + \Delta\mathbf{G}(t)]\mathbf{G}^{-1}(t)$ que satisfaz (28) para qualquer $\Delta\mathbf{G}(t)$. Fisicamente, isto significa que aplicar o movimento estimado (movimento exato + erro),

$$\hat{\mathbf{d}}(t) = \mathbf{d}(t) + \Delta\mathbf{d}(t), \quad (29)$$

a uma imagem é o mesmo que primeiro aplicar o movimento exato e, depois disso, aplicar o movimento correspondente ao erro de estimação. Além disso, não importa qual movimento (o exato ou o erro) é aplicado primeiro, uma vez que, sob tais condições, as matrizes de registro são bloco-circulantes [13] e podem ser comutadas.

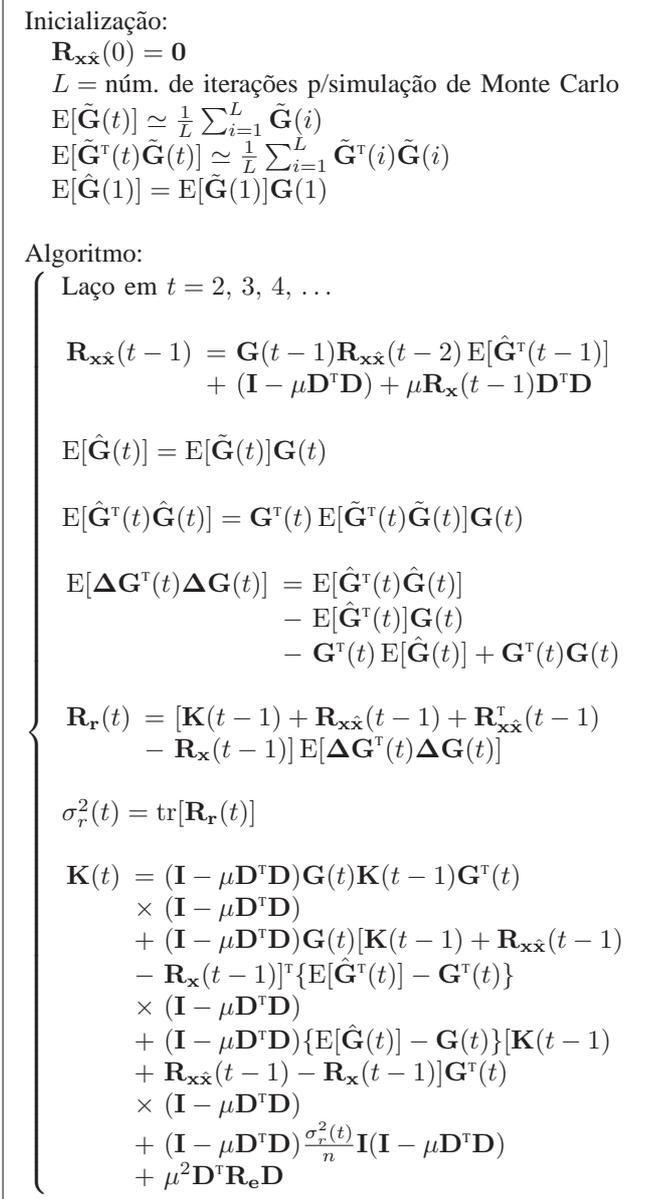
Assim, de forma a resolver (27), segue de (28) que

$$E[\hat{\mathbf{G}}(t)] = E[\tilde{\mathbf{G}}(t)]\mathbf{G}(t) \quad (30)$$

e que

$$E[\hat{\mathbf{G}}^T(t)\hat{\mathbf{G}}(t)] = \mathbf{G}^T(t)E[\tilde{\mathbf{G}}^T(t)\tilde{\mathbf{G}}(t)]\mathbf{G}(t). \quad (31)$$

Considerando que o erro de estimação do vetor de deslocamento é estacionário, os valores esperados em (30) e em (31) são independentes de t e podem ser simulados *a priori* e usados em todas as iterações do modelo recursivo, reduzindo consideravelmente a carga computacional.

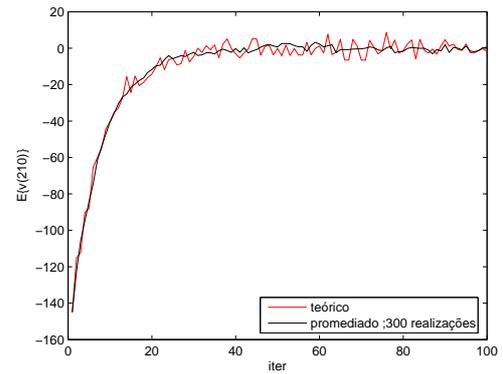
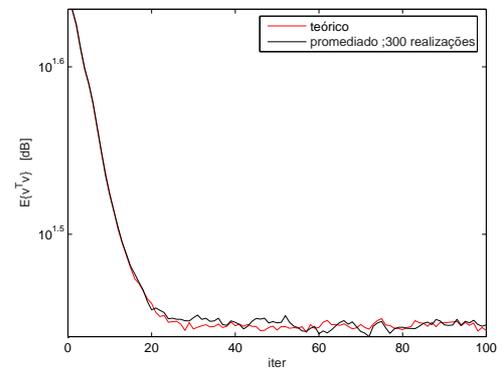

 Fig. 2. Algoritmo para determinar $\mathbf{K}(t)$.

O erro de estimação do vetor de deslocamento, $\Delta\mathbf{d}(t)$, apresenta, sob certas condições, uma função densidade de probabilidade gaussiana [7]. Assim, $\tilde{\mathbf{G}}(t)$ pode ser construído a partir de $\Delta\mathbf{d}(t) = [\Delta d_l(t), \Delta d_c(t)]$, considerando que $\Delta d_l(t)$ e $\Delta d_c(t)$ são variáveis gaussianas independentes (ruído), com média zero e variâncias determinadas pelo método de estimação de movimento utilizado [14].

O algoritmo final (em uma forma didática) para determinar $\mathbf{K}(t)$ é ilustrado na Figura 2.

IV. RESULTADOS

Para todas as simulações apresentadas nesta seção, os vetores de movimento (livres de erros de estimação), $\mathbf{d}(t)$, foram gerados a partir de um processo *random walk* com passos equiprováveis nas direções vertical e horizontal (movimento translacional). Tais vetores foram considerados conhecidos *a*


 Fig. 3. Erro médio de reconstrução: $\mu = 0.5$; $\sigma_e^2 = 10$; $\sigma_{\Delta d}^2 = 0.5$.

 Fig. 4. Erro médio quadrático de reconstrução: $\mu = 0.5$; $\sigma_e^2 = 10$; $\sigma_{\Delta d}^2 = 0.5$.

priori. Note ainda que a hipótese de movimento translacional em passos inteiros é feita em diversos trabalhos [6], [7]. Os vetores de movimento estimado foram montados conforme (29). O erro de estimação de movimento, $\Delta\mathbf{d}(t) = [\Delta d_l(t), \Delta d_c(t)]$, foi considerado como tendo densidade de probabilidade gaussiana, de média nula, com $\Delta d_l(t)$ e $\Delta d_c(t)$ independentes e com variâncias $\sigma_{\Delta d_l}^2 = \sigma_{\Delta d_c}^2 = 0.5$. As imagens de alta resolução utilizadas possuem 20×20 pixels. As imagens de baixa resolução, por sua vez, possuem 10×10 pixels. No cômputo dos erros médios, foram desconsideradas as cinco linhas e colunas mais externas das imagens reconstruídas (de resolução 20×20), visando cancelar a influência das inovações (efeito de bordas) entre elas. Os momentos da matriz de registro estimada, em (27), foram estimados por meio de 500 realizações. A matriz de correlação do sinal de entrada foi estimada a partir das imagens originais, para cada instante de tempo.

As simulações, utilizando os modelos definidos em (19) e na Figura 2, são apresentadas nas Figuras 3 e 4, em que um conjunto de imagens de cenas distintas é utilizado. Na figura 3, é apresentado o erro referente ao pixel central da imagem. O ruído adicionado no modelo de degradação em (1) gerou as seguintes relações médias de sinal-ruído: PSNR \simeq 38dB e SNR \simeq 23dB. Essas relações foram calculadas entre as imagens $\mathbf{y}(t)$ e $\mathbf{D}\mathbf{x}(t)$.

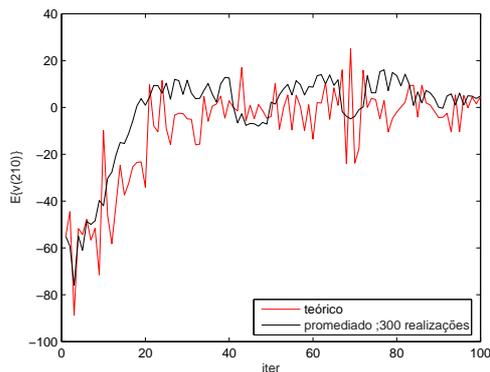


Fig. 5. Erro médio de reconstrução: $\mu = 0.5$; $\sigma_e^2 = 10$; $\sigma_{\Delta d}^2 = 0.5$; imagens de faces.

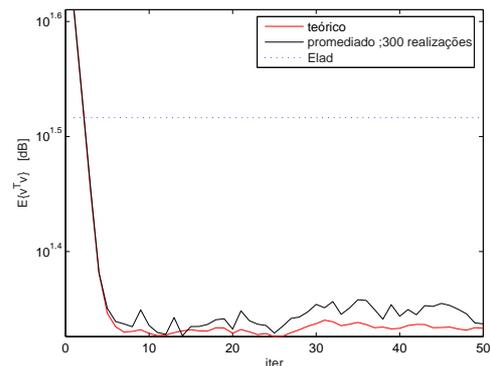


Fig. 7. Erro médio quadrático de reconstrução: $\mathbf{D} = \mathbf{I}$; $\mu = 0.5$; $\sigma_e^2 = 10$; $\sigma_{\Delta d}^2 = 0.5$; imagens de faces.

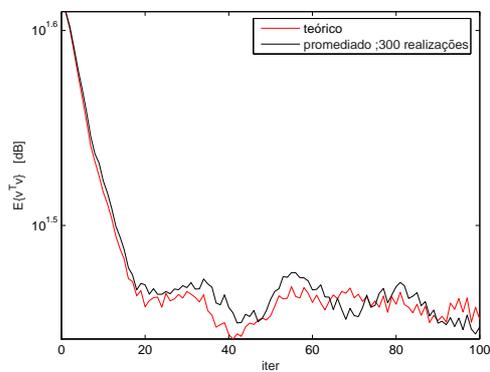


Fig. 6. Erro médio quadrático de reconstrução: $\mu = 0.5$; $\sigma_e^2 = 10$; $\sigma_{\Delta d}^2 = 0.5$; imagens de faces.

Nas Figuras 5 e 6, são apresentadas as simulações utilizando um conjunto de imagens mais homogêneo do que o utilizado nas simulações anteriores. Neste caso, o *ensemble* é constituído por imagens de faces¹.

Considerando $\mathbf{D} = \mathbf{I}$, recai-se no caso modelado em [8]. Para este caso, ambos os modelos para a variância do erro de reconstrução (o proposto neste trabalho e o proposto por Elad, em [8]) podem ser vistos na Figura 7. Verifica-se que a nova análise permite uma previsão bem mais acurada do comportamento do algoritmo.

V. CONCLUSÕES

Neste trabalho foi proposto um modelo analítico para prever o comportamento estocástico do algoritmo LMS-SRR proposto em [9], dando ênfase à influência do erro de registro. Equações recursivas foram derivadas para os erros médio e médio quadrático de reconstrução. O modelo proposto apresenta uma boa conformidade com simulações de Monte Carlo tanto em regime permanente como em regime transitório de adaptação. Finalmente, com relação ao modelo existente [8], o novo modelo contribui com: (i) comportamento do algoritmo considerando super-resolução (subamostragem); (ii)

comportamento médio e variância do erro de reconstrução em regimes transitório e permanente de adaptação do algoritmo.

REFERÊNCIAS

- [1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, May 2003.
- [2] Z. Jiang, T.-T. Wong, and H. Bao, "Practical super-resolution from dynamic video sequences," *IEEE Int'l Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 549–554, June 2003.
- [3] H. He and L. P. Kondi, "Resolution enhancement of video sequences with adaptively weighted low-resolution images and simultaneous estimation of the regularization parameter," *IEEE Int'l Conf. Acoustics, Speech, and Signal Processing*, vol. 3, pp. 213–216, May 2004.
- [4] D. Capel and A. Zisserman, "Computer vision applied to super resolution," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 75–86, May 2003.
- [5] Z. Wang and F. Qi, "Super-resolution video restoration with model uncertainties," *IEEE Int'l Conf. Image Processing*, vol. 2, pp. 853–856, Sept. 2002.
- [6] E. S. Lee and M. G. Kang, "Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration," *IEEE Trans. Image Processing*, vol. 12, no. 7, pp. 826–837, July 2003.
- [7] R. Mester and M. Hotter, "Robust displacement vector estimation including a statistical error analysis," *IEE Int'l Conf. Image Processing and its Applications*, pp. 168–172, July 1995.
- [8] M. Elad, "Super-resolution reconstruction of images," Ph.D. dissertation, Israel Institute of Technology, Dec. 1996.
- [9] M. Elad and A. Feuer, "Superresolution restoration of an image sequence: Adaptive filtering approach," *IEEE Trans. Image Processing*, vol. 8, no. 3, pp. 387–395, Mar. 1999.
- [10] S. Haykin, *Adaptive Filter Theory*, 4th ed. New Jersey: Prentice Hall, 2002.
- [11] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1167–1183, Sept. 2002.
- [12] M. K. Ng and N. K. Bose, "Analysis of displacement errors in high-resolution image reconstruction with multisensors," *IEEE Trans. Circuits and Systems*, vol. 49, no. 6, pp. 806–813, June 2002.
- [13] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Trans. Image Processing*, vol. 10, no. 8, pp. 1187–1193, Aug. 2001.
- [14] G. Dane and T. Q. Nguyen, "The effect of global motion parameter accuracies on the efficiency of video coding," *IEEE Int'l Conf. Image Processing*, pp. 1–4, Oct. 2004.

¹Banco de imagens do Departamento de Ciências da Computação da Universidade de Essex. <http://cswww.essex.ac.uk/mv/allfaces/>