

UM MODELO DE AQUISIÇÃO DE SEQUÊNCIA DE IMAGENS COM MELHORIA DE RESOLUÇÃO ATRAVÉS DE REGULARIZAÇÃO ADAPTATIVA

MARCELO VICTOR WÜST ZIBETTI
JOCELI MAYER

LPDS - Laboratório de Pesquisa em Processamento Digital de Sinais
Departamento de Engenharia Elétrica
Universidade Federal de Santa Catarina
Caixa Postal 476 – 88040-900 – Florianópolis - SC, Brasil
marcelo.zibetti@terra.com.br, mayer@eel.ufsc.br

Abstract: In this work we investigate an acquisition model for image sequences. As one result, we propose an iterative multiframe algorithm for enhancing image resolution. The motion between the low-resolution frames is estimated and the accuracy of motion vectors is computed. The proposed technique is based on the regularized least squares method and use adaptive regularization. In most cases, the motion estimation between the frames is complicated by the presence of outliers. Outlier is a region that belongs only to one image of the sequence. These outlier regions and unreliable motion vectors are detected and eliminated. Smoothing filters are chosen from a set of filters for local adaptive regularization. The choice is based on the degradation system and the contributions of each frame to the process. In this way, the algorithm balances the level of smoothing while preserves the details in the restored image. Results are presented which illustrate a significant perceptual and objective (DSNR) improvement over others methods.

Keywords: resolution enhancement for multiple frames, image interpolation, and adaptive regularization.

1. Introdução:

O aumento de resolução de imagens é um tema que contribui para várias aplicações importantes, como: imagens médicas e científicas, fotos aéreas e de satélites, astronomia e uso militar. No entanto, os sistemas eletrônicos de captura de imagem possuem limitações na resolução de imagem. Melhorar a resolução através do uso de fotosensores mais densos, lentes mais limpas, sistemas de foco mais precisos aumenta substancialmente o custo de um sistema. Uma possível abordagem para o problema é o uso de melhoria digital de resolução usando múltiplas imagens.

A maioria das técnicas de interpolação encontradas na literatura, como, por exemplo, em Bovik (2000) e Jähne, Haubeker e Geibler (1999), não exploram o modelo de degradação do dispositivo de captura e a diferença de informação entre o conjunto de imagens adquiridas. Os métodos de interpolação convencionais como interpolação de ordem zero, bilinear, bicúbica e “*B-spline*” de ordem maior são consideradas funções básicas e apenas expandem a imagem de baixa resolução sem corrigir a degradação. Estes algoritmos têm sido desenvolvidos assumindo que: não há mistura entre os pixels adjacentes no sensor de imagem, não existe distorção por movimento (“*motion blur*”) devido ao tempo de exposição necessário para a captura da imagem, não ocorre distorção isotrópica devido à falta de foco, e não acontece “*aliasing*” no processo de sub amostragem. Desde que as suposições acima mencionadas não são satisfeitas em um sistema de formação de imagem de baixa resolução geral, o uso destes algoritmos é relativamente ineficiente. As técnicas de melhoramento de resolução baseadas em múltiplas imagens são uma alternativa para recuperar os componentes perdidos na sub amostragem, dado que cada imagem da sequência possui uma nova e exclusiva informação sobre a imagem de alta resolução.

Esta diferença de informação é gerada por objetos nas imagens que se movem com incrementos de “*subpixel*” ou por diferenças na degradação entre as imagens, isto também é mencionado por Tom e Katsaggelos (2001) e Schultz, Meng e Stevenson (1998).

Assim o movimento relativo entre as imagens da seqüência é estimado adaptando-se as técnicas tradicionais, encontradas em Stiller e Konrad (1999), para o uso com imagens degradadas. Também, a estimação de movimento é seriamente afetada pela presença de “outliers”. Os “outliers” são:

- Regiões que sofrem movimentos muito complexos que não produzem um movimento aparente que possa ser representados pelo modelo de movimento escolhido.
- Partes de objetos cujo movimento transpassa as bordas da imagem.
- Regiões cobertas ou expostas pelo movimento de um outro objeto.

Devido ao aparecimento dos “outliers” nem toda informação nas imagens pode ser usada na melhoria de resolução. Pode-se eliminar esta informação no procedimento de melhoria de resolução, porém para a obtenção de resultados superiores uma regularização adaptável, dependendo da existência de “outliers”, deve ser usada. A adaptação é realizada de acordo com a contribuição das imagens adicionais para a região. Por isso este trabalho propõe uma regularização que se adapta a contribuição existente entre as imagens.

Na seção 2 o modelo de degradação do dispositivo de aquisição é apresentado. Na seção 3 é feita a apresentação do algoritmo de melhoria de resolução iterativo regularizado. Na seção 4 são mostrados os resultados experimentais. A seção 5 conclui este artigo.

2. Modelo de degradação da aquisição:

Nesta seção será apresentado o modelo de aquisição de um sistema genérico de formação de imagem. No modelo original, também apresentado em Bovik (2000) e Shin, Paik, Price e Abidi (2001), a degradação é apresentada na forma contínua para discreta, como é em sua natureza. Neste trabalho ele será apresentado na forma discreta de alta resolução para discreta de baixa resolução, que é a abordagem com que ele será tratado utilizando processamento digital. A **figura 1** apresenta o diagrama de blocos do modelo de aquisição.

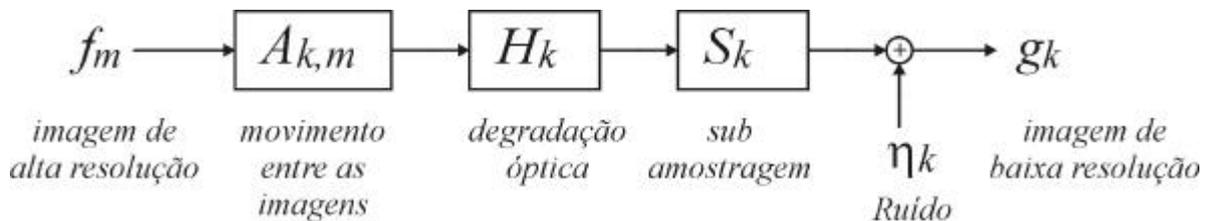


Figura 1: Diagrama de blocos do modelo de aquisição de imagem.

Assume-se que uma imagem contínua de alta resolução, $f_m(x,y)$, é amostrada acima da taxa de Nyquist, então a versão amostrada da imagem contínua de alta resolução não tem “aliasing”. Pela amostragem ideal, acima da taxa de Nyquist, uma imagem discreta de alta resolução, $f_m[m_1,m_2]$, pode ser produzida. Considere o tamanho da imagem discreta de alta resolução como $M_1 \times M_2$. Na **figura 1** a imagem é representada na forma do vetor f_m , de tamanho $M_1 M_2 \times 1$, utilizando ordenação lexicográfica.

Conforme a **figura 1**, a imagem de alta resolução discreta deve possuir uma relação de movimento com as outras imagens da seqüência. Esta relação é expressa na forma de compensação de movimento e é apresentado em (1).

$$f_k = A_{k,m} f_m \quad (1)$$

Então, contando o efeito das distorções ópticas, causadas pelos fotosensores, lentes, atmosfera e demais responsáveis pelo fluxo da luz, podemos modelar a segunda etapa pela função discreta de degradação óptica, como em (2). A distorção é assumida como linear

espaço-invariante. Esta degradação é representada pela matriz bloco-circulante H_k , de tamanho $M_1 M_2 \mathcal{M}_1 \mathcal{M}_2$.

$$\tilde{f}_k = H_k f_k \quad (2)$$

Na terceira etapa, a imagem de alta resolução que sofreu transformação de movimento e distorção óptica é sub amostrada, conforme apresentado em (3).

$$\bar{f}_k = S_k \tilde{f}_k \quad (3)$$

Onde a matriz S_k , de tamanho $N_1 N_2 \mathcal{M}_1 \mathcal{M}_2$, representa a redução de tamanho entre a imagem desejada e a imagem entregue pelo dispositivo. A forma de sub amostragem depende do arranjo do sensor óptico, mas em geral é retangular.

Finalmente, a imagem discreta sub amostrada, é corrompida por ruído em (4).

$$g_k = \bar{f}_k + \mathbf{h}_k \quad (4)$$

Onde \mathbf{h}_k é um vetor de tamanho $N_1 N_2 \mathcal{M}_1 \mathcal{M}_2$, que representa o ruído adicionado no fotosensor.

Como resultado, são observadas L imagens de baixa resolução, $\{g_k\}$ onde $k=1 \dots L$ produzidas pelas L imagens de alta resolução, $\{f_k\}$, onde $k=1 \dots L$ todas elas relacionadas com a imagem f_m , que faz parte do conjunto.

Esta degradação pode ser apresentada, segundo a equação (5).

$$g_k = S_k H_k A_{k,m} f_m + \mathbf{h}_k \quad (5)$$

Onde o vetor g_k , de tamanho $N_1 N_2 \mathcal{M}_1 \mathcal{M}_2$, representa a imagem de baixa resolução $g_k[n_1, n_2]$. Lembrando que $M_1 M_2 @ R^2 N_1 N_2$, onde R é o fator de ampliação ($2\times, 3\times, 4\times$).

O processo pode ser tratado também como (6).

$$g_k = D_{k,m} f_m + \mathbf{h}_k \quad (6)$$

Onde $D_{k,m}$ é uma matriz $N_1 N_2 \mathcal{M}_1 \mathcal{M}_2$, e ainda $D_{k,m} = S_k H_k A_{k,m}$.

3. Melhoramento de resolução utilizando múltiplas imagens:

Nesta seção será apresentada a proposta de solução para a melhoria de resolução utilizando múltiplas imagens. Inicialmente é apresentada a solução iterativa regularizada existente, também encontrada em Katsaggelos (1989) e em Galatsanos, Katsaggelos, Chin e Hillery (1991). Em seguida serão apresentadas as modificações propostas por este trabalho.

3.1. Solução iterativa regularizada:

Para solucionar (6) é necessário achar uma solução inversa viável para a matriz $D_{k,m}$. É conhecido que a matriz $D_{k,m}$ é singular (com deficiência de posto) e mal-condicionada ("ill conditioned"). Desta forma, achar uma simples inversa não traz uma solução adequada. O uso de imagens adicionais aumenta o posto do sistema (6), isto pode ser traduzido como a diminuição do conjunto de soluções possíveis e uma melhor estimativa da imagem de alta resolução. Mesmo assim para solucionar (6) é necessária otimização e regularização.

A otimização é necessária para se buscar a solução ótima segundo algum critério que permita a solução de (6). Como o sistema é mal condicionado a regularização é necessária para a obtenção uma resposta aceitável e evitar a amplificação do ruído. Assim, neste trabalho a técnica de mínimos quadrados regularizados é usada. Os detalhes desta técnica são encontrados em Galatsanos, Katsaggelos, Chin e Hillery (1991). Através deste método encontramos a solução apresentada em (7).

$$\hat{f}_m = \left(\sum_{k=1}^L D_{k,m}^T D_{k,m} + \mathbf{I} C^T C \right)^{-1} \sum_{k=1}^L D_{k,m}^T g_k \quad (7)$$

Um problema ocorre na inversão apresentada em (7). Como o termo a ser invertido é uma matriz de tamanho $M_1 M_2 \times M_1 M_2$, sua inversão na forma tradicional pode ser computacionalmente custosa e inviável para imagens maiores que um tamanho modesto. Desta forma, a busca de (7) é realizada na forma iterativa, apresentada em (8).

$$f_m^{(n+1)} = f_m^{(n)} + \beta \left(\sum_{k=1}^L D_{k,m}^T g_k - \left[\sum_{k=1}^L D_{k,m}^T D_{k,m} + \mathbf{I} C^T C \right] f_m^{(n)} \right) \quad (8)$$

Onde n é a iteração corrente e β é o fator de relaxamento. Ao lidar com algoritmos iterativos, β deve ter uma atenção especial, dado que β controla a convergência, bem como a taxa de convergência. Assim, se β garantir a convergência o mínimo será encontrado.

3.2. Mudança nas equações devido aos “outliers”:

Um dos maiores problemas enfrentados com múltiplas imagens é a impossibilidade do uso das demais imagens devido à presença de “outliers”. Se for feito o uso das regiões de “outliers”, ou das regiões com vetores mal estimados, no melhoramento de resolução resultados muito piores do que os de uma interpolação convencional podem ser produzidos. Sendo assim, estas regiões têm que ser eliminadas.

Matematicamente o movimento entre as imagens m e k era modelado segundo (1). Agora, devido aos “outliers”, ele é modelado segundo (9).

$$f_k^I = I_k A_{k,m} f_m \quad (9)$$

Onde f_k^I é um vetor que representa os pixels que são “inliers” no movimento entre m e k . A matriz I_k é uma matriz diagonal compostas por uns e zeros na diagonal principal. Os uns devem preservar os pixels “inliers”, os zeros devem eliminar os pixels “outliers”. A separação das regiões de “inliers” e “outliers” precisa ser feita em conjunto com a estimação de movimento. As técnicas referentes a este assunto também são encontradas em Schultz, Meng e Stevenson (1998).

Devido à necessidade de mudança no equacionamento do movimento não é mais possível utilizar a equação (1) no modelo de aquisição. Assim a degradação deve ser modificada para incorporar a equação (9). Isto é apresentado em (10).

$$g_k^I = S_k H_k I_k A_{k,m} f_m + \mathbf{H}_k^I \quad (10)$$

O vetor g_k^I representa a informação dos “inliers” na imagem de baixa resolução. Como a imagem de alta resolução, f_k , não é conhecida g_k^I precisa ser gerado de outra forma. A equação (11) apresenta uma solução simples que aproxima com pouco erro os “inliers” na imagem de baixa resolução.

$$g_k^I = I_k^\downarrow g_k \quad (11)$$

Onde a matriz I_k^\downarrow é uma matriz diagonal, cuja diagonal principal é a versão de baixa resolução da diagonal da matriz I_k . Pode-se dizer que se um vetor i_k , que representa o mapeamento de “inliers” em k , compõe a diagonal principal de I_k , então o vetor $i_k^\downarrow = S_k H_k i_k$ irá compor a diagonal principal da matriz I_k^\downarrow . Sendo assim, agora se pode representar o modelo de degradação, considerando apenas os “inliers”, como em (12).

$$g_k^I = D_{k,m}^I f_m + h_k^I \quad (12)$$

Estas mudanças podem ser diretamente incorporadas ao algoritmo iterativo, levando a equação (8) a se tornar igual à equação (13).

$$f_{n+1} = f_n + \mathbf{b} \left(\sum_{k=1}^L (D_{k,m}^I)^T g_k^I - \left[\sum_{k=1}^L (D_{k,m}^I)^T D_{k,m}^I + \mathbf{I} C^T C \right] f_n \right) \quad (13)$$

A equação (13) agora é própria para uso em seqüências de imagens que contenha “inliers” e “outliers”, dado que os “outliers” e vetores mal estimados foram devidamente removidos. Porém, os resultados ótimos provavelmente não serão obtidos. Isto ocorre, principalmente, devido à falta da informação no local onde os pixels foram removidos. Como não há contribuição das imagens adicionais, a regularização precisa compensar a falta de informação. Em geral, há duas soluções para o problema:

- Usar um fator de regularização (λ) adaptativo, ao invés de um constante. Permitindo que a regularização seja localmente selecionada.
- Usar um operador de regularização (C) adaptativo, ao invés de um único para a imagem toda. Permitindo uma filtragem mais seletiva dos componentes indesejados.

A primeira alternativa é mais simples, porém o filtro C é normalmente otimizado para minimizar o ruído e não as altas frequências que surgem devido à sub amostragem. A segunda aproximação pode encontrar um filtro C melhor para o problema, porém, a necessidade de adaptação do filtro C pode tornar o operador inviável para a solução do problema, devido à sua complexidade. Neste trabalho será usado um modo misto. Será adicionado um novo operador de regularização, preservando o operador de regularização do ruído. Este novo operador será composto de um conjunto de filtros base, selecionados para produzir melhores resultados no problema, e terá também o fator de regularização (λ) adaptativo, para atender as necessidades locais de regularização.

3.3. Regularização adaptativa para correção local dos efeitos da sub amostragem:

A maior contribuição que as imagens adicionais podem trazer é a inversão parcial da sub amostragem. Isto acontece quando as imagens possuem movimento de “subpixel”. Devido a este movimento, a imagem adicional possui uma importante informação sobre o espaço entre as amostras da imagem que se deseja melhorar a resolução. Matematicamente esta imagem trará uma pequena contribuição para a inversão da sub amostragem. Entre outros benefícios estão a inversão de um possível “aliasing” e a recuperação de detalhes perdidos na sub amostragem.

Desta forma, se uma imagem com movimento de “subpixel” for utilizado, significa que poderá estar sendo recuperado algum detalhe na imagem, e com isto a regularização não terá que suavizar tanto a região em que o detalhe está sendo recuperado. Porém nas regiões onde não houve contribuição, devido aos “outliers”, devido aos vetores mal estimados ou ainda devido à imagem não possuir nenhum movimento de “subpixel” a regularização deve ser mantida.

Então é adicionada ao algoritmo uma etapa de análise de movimento. Nesta etapa a contribuição local de cada imagem é determinada. Os resultados desta etapa são os valores necessários de regularização que serão usados no conjunto de filtros base de regularização. Assim o novo algoritmo iterativo fica conforme a equação (14).

$$f_{n+1} = f_n + \mathbf{b} \left(\sum_{k=1}^L (D_{k,m}^I)^T g_k^I - \left[\sum_{k=1}^L (D_{k,m}^I)^T D_{k,m}^I + \mathbf{I} C^T C + \mathbf{a} \sum_{i=1}^P W_i G_i^T G_i \right] f_n \right) \quad (14)$$

Onde são usados P filtros bases, mostrados com as matrizes G_i . Normalmente $P=3$, ou seja, a quantidade necessária de filtros para suavizações vertical, horizontal e isotrópico. A matriz W_i determina os pesos em cada pixel de cada filtro base de regularização. Nos casos extremos se tem que:

- Se não houver contribuição de nenhuma imagem adicional a matriz W_j , onde j referencia o filtro isotrópico, terá peso um em todos os pixels. Os demais filtros, onde $i \neq j$, terão peso zeros em todos os pixels.
- Se houver apenas contribuição para a inversão da sub amostragem em uma dimensão em toda imagem, a matriz W_j , onde j referencia o filtro na outra dimensão, terá peso maior, e este será um em todos os pixels se a sub amostragem na dimensão mencionada inicialmente for totalmente invertida.
- Se as imagens adicionais invertem totalmente a sub amostragem, nas duas dimensões e em todos os pixels, então não há necessidade de regularização dos efeitos da sub amostragem, e todos os pesos são zeros.

Seguindo este princípio é possível aplicar a técnica iterativa, com regularização adaptativa, que resultados superiores serão obtidos.

4. Resultados Experimentais:

Para demonstrar o desempenho do algoritmo uma seqüência composta de duas imagens de baixa resolução será usada com o algoritmo proposto. As imagens foram feitas com movimentos de uma câmera aérea. O movimento observado será modelado como movimento do tipo "Affine". Os "outliers" serão excluídos da melhoria de resolução.

Uma simulação do dispositivo de aquisição irá causar uma degradação H_k que aplica uma operação de média móvel de 4×4 pixels, seguido de uma sub amostragem, S_k , retangular de quatro pixels e adição de ruído gaussiano, \mathbf{h}_k , de 5 dB. Esta degradação é uma aproximação discreta da degradação causada em um CCD para o caso de se ampliar a resolução em $4 \times$. A degradação é a mesma as duas imagens. Não estão sendo considerados os efeitos de degradação de movimento e de distorções nas lentes.

A **figura 2** demonstra a imagem de alta resolução original, que é a versão de alta resolução da imagem 01 da seqüência de baixa resolução, e as duas imagens de baixa resolução da seqüência (imagem 01 e imagem 02). Logo abaixo das imagens da seqüência de baixa resolução está o mapa de "inliers" e "outliers" estimado do movimento entre a imagem 01 e a imagem 02. Os "outliers" estão em preto e os "inliers" em branco.

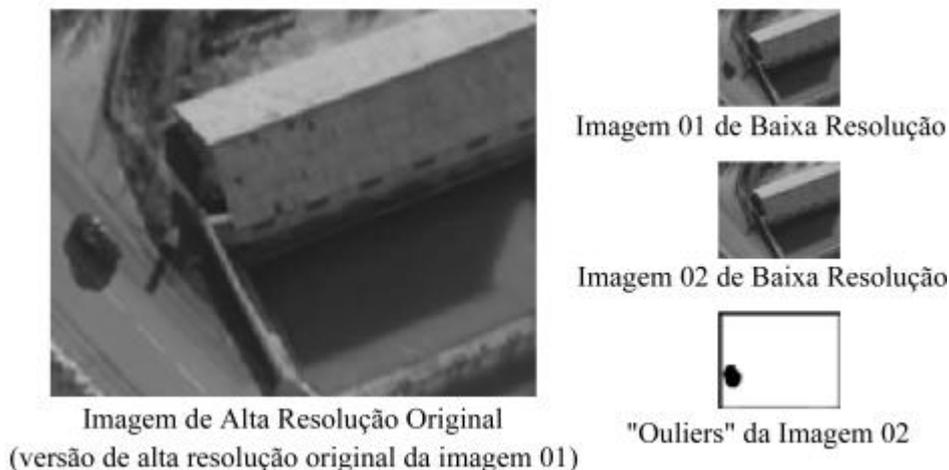


Figura 2: Imagens da seqüência.

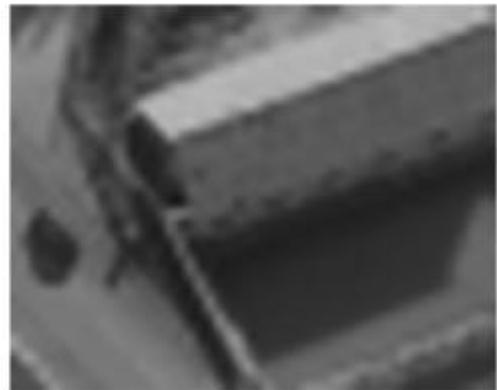
Para comparação, resultados objetivos usando PSNR (“*peak signal-to-noise ratio*”) e DSNR (“*differential signal-to-noise ratio*”) serão apresentados, segundo as equações em (15).

$$PSNR = 10 * \text{Log}_{10} \left(\frac{255^2}{\frac{1}{M_1 M_2} \|f - \hat{f}\|^2} \right) \quad e \quad DSNR = 10 * \text{Log}_{10} \left(\frac{\|f - f_0\|^2}{\|f - \hat{f}\|^2} \right) \quad (15)$$

Onde f é a imagem central original, \hat{f} é a imagem estimada e f_0 é a imagem interpolada segundo a interpolação de ordem zero. A **figura 3** demonstra as imagens obtidas utilizando interpolação convencional de ordem zero e bicúbica.



Interpolação de Ordem Zero.
(PSNR=29,50 dB, DSNR=0dB)



Interpolação Bicúbica.
(PSNR=31,65dB, DSNR=2,51dB)

Figura 3: Imagens obtidas utilizando interpolação convencional.

Em seguida, na **figura 4** as imagens obtidas através do melhoramento de resolução utilizando as duas imagens são apresentadas. Inicialmente a técnica sem regularização adaptativa e em seguida com regularização adaptativa. Os valores de λ e α foram determinados experimentalmente.



Regularização Não-Adaptativa
($\lambda=0.05, \alpha=0.25$)
(PSNR=33,22dB, DSNR=4,08dB)



Regularização Adaptativa
($\lambda=0.05, \alpha=0.25$)
(PSNR=33,84dB, DSNR=4,70dB)

Figura 4: Imagens obtidas utilizando o algoritmo proposto, sem regularização adaptativa e com regularização adaptativa.

5. Conclusão:

Neste trabalho foi demonstrada a necessidade de regularização adaptativa na solução de mínimos quadrados. A regularização deve ajudar na escolha de uma inversa aproximada para uma matriz singular (com deficiência de posto) e mal condicionada. Nos sistemas de aquisição de imagem a degradação óptica e o ruído são os principais responsáveis pelo mal-condicionamento, enquanto a sub amostragem é a principal responsável pela deficiência no posto.

O uso de múltiplas imagens pode solucionar parcialmente a deficiência de posto, mas a inversão total da sub amostragem depende da quantidade certa de imagens e da existência de movimentos de subpixel entre as imagens. Em situações práticas a quantidade correta de imagens necessárias para a inversão da sub amostragem raramente é encontrada. Portanto a solução precisa ser regularizada para compensação dos efeitos da sub amostragem.

O algoritmo proposto busca aplicar a regularização quanto à sub amostragem de forma adaptativa. É feita uma análise sobre a contribuição de cada imagem e determinada a necessidade de regularização. Assim as regiões onde a sub amostragem foi invertida não são regularizadas (quanto aos efeitos da sub amostragem), pois não há necessidade. Nas regiões onde a sub amostragem foi parcialmente invertida ocorre uma regularização parcial. Já nas regiões onde a sub amostragem não foi invertida a regularização é máxima. A técnica utilizando regularização adaptativa mostrou trazer melhores resultados do que as não-adaptativas.

6. Bibliografia:

- Bovik, Alan. *Handbook of Image and Video Processing*, Academic Press, 2000.
- Shin, J.H., Paik, J. K., Price, J. R. e Abidi, M. A. Adaptive regularized image interpolation using data fusion and steerable constraints, *Proc. SPIE Visual Comm, Image Proc.* v. 4310, p.798-808, Jan. 2001.
- Jähne, B., Haubeker, H. e Geibler, P. *Handbook of Computer Vision and Applications: v. 1, Sensors and Imaging*, Academic Press, 1999.
- Katsaggelos, A. K. Iterative image restoration algorithms, *Optical Engineering*, v. 28, n. 7, p. 735-748, Jul. 1989.
- Galatsanos, N. P., Katsaggelos, A. K., Chin, R. T. e Hillery, A. D. Least squares restoration of multichannel images" *IEEE Trans. on Signal Processing*, v. 39, n. 10, p. 2222-2236, Out. 1991.
- Tom, B. C. e Katsaggelos, A. K. Resolution enhancement of monochrome and color video using motion compensation, *IEEE Trans. on Image Processing*, v. 10, n. 2, p. 278-287, Fev. 2001.
- Schultz, R., Meng, R L. e Stevenson, R. L. Subpixel motion estimation for super-resolution image sequence enhancement, *J. of visual communication and image representation*, v. 9, n. 1, p. 38-50, Mar. 1998.
- Stiller, C. e Konrad, J. Estimating Motion in Image Sequences, *IEEE Signal Processing Magazine*, p. 70-91, Jul. 1999.