

## Registro Sub-pixel: Uma Solução Rápida para Imagens Obtidas por um Mesmo Sensor.

Claudio Gelelete<sup>1</sup>  
Mariane Rembold Petraglia<sup>2</sup>  
José Gabriel R. C. Gomes<sup>2</sup>

<sup>1</sup> Instituto Militar de Engenharia - IME  
Praça Gen Tibúrcio, 80 – Rio de Janeiro - RJ, Brasil  
cgelelete@yahoo.com.br

<sup>2</sup> Universidade Federal do Rio de Janeiro – PEE - COPPE - UFRJ  
CEP: 21945-970 CP 68501 Rio de Janeiro - RJ, Brasil  
mariane@pads.ufrj.br  
gabriel@pads.ufrj.br

**Abstract.** Sub-pixel registration has been applied in several algorithms that demand high precision in the registration of the involved images. This paper presents a fast solution for the sub-pixel registration between images where the scale correction is not necessary. Rotation and translation are applied separately, with automatic extraction of control points, and the achieved results are adequate for applications such as super-resolution.

**Palavras-chave:** pré-processamento de imagens, registro sub-pixel.

### 1. Introdução

A crescente utilização de imagens de sensoriamento remoto, em suas diversas aplicações, fez com que muitos métodos de registro fossem desenvolvidos. “Por registrar duas imagens de uma mesma cena entende-se superpô-las de modo que pontos das imagens que correspondem aos mesmos objetos coincidam espacialmente” (Mascarenhas et al, 1988).

Registrar duas ou mais imagens se faz necessário quando se objetiva trabalhar com uma sequência temporal adquirida por sensores diferentes (Mather, 1999), ou com aplicações que incluam detecção de mudanças, fusão, construção de mosaicos (Pedrini & Schwartz, 2008), entre outras.

Alguns desses métodos podem ser utilizados para registrar imagens das mais diversas formas (imagens entre sensores diferentes, imagem x mapa, utilizando pontos GPS, entre outras), porém cada método atinge seu melhor desempenho para um caso específico, ficando esta indicação por conta do tipo de sensor utilizado, tipo de imagem obtida ou dados disponíveis para a realização do processamento das imagens.

Na maioria das vezes este registro é feito com a intervenção de um especialista que identifica visualmente pontos, na(s) imagem(ns) envolvida(s) no processo, que representem um determinado alvo de interesse. Estes pontos são chamados de pontos de controle e são utilizados para o ajuste da imagem a ser modificada e o método é denominado de semi-automático. Métodos de ajuste automático (quando pontos de controle são identificados automaticamente pelo sistema) também foram desenvolvidos (Schowengerdt, 1997). Estes métodos apresentam a vantagem de, idealmente, não precisarem da contínua intervenção de um operador na fase de escolha dos pontos de controle; porém é recomendável, sempre que possível, uma verificação dos resultados obtidos, mesmo visualmente, e consequente refinamento da operação caso seja necessário.

O método desenvolvido e abordado neste trabalho objetiva seu emprego em super-resolução de imagens, processamento que tem por finalidade a obtenção de uma imagem de alta resolução espacial partindo de duas ou mais imagens de uma mesma cena de resolução

espacial inferior. Sua aplicação é extensa, podendo ser utilizada para imagens médicas, sensoriamento remoto, aplicações em vídeo entre outras (Park et al, 2003).

As imagens de baixa resolução a serem empregadas, na maioria das vezes, são obtidas por um mesmo sensor, através de seqüências multitemporais, ou por sensores idênticos que adquirem imagens da mesma área simultaneamente. Percebe-se, portanto, que as imagens que servem de insumo ao processo apresentam características muito semelhantes entre si, como deformações peculiares ao sistema óptico, precisão de posicionamento, atitude da plataforma que sustenta o sensor e geometria de aquisição de dados.

Esta semelhança fez com que, nesta primeira abordagem, não fossem consideradas possíveis distorções entre as imagens, supondo-se suficiente considerar movimentos de rotação e translação para que o registro fosse satisfatório. Caso surjam diferenças devidas a esta simplificação, a redução da área útil da imagem a ser trabalhada se torna uma alternativa para sanar o problema, já que as deformações se acentuam do centro para as extremidades das imagens.

No desenvolvimento de alguns métodos de super-resolução, as imagens de baixa resolução espacial, após serem registradas, são submetidas a um processo de interpolação não uniforme (Clark et al., 1985) e (Kim & Bose, 1990), para que passem a ter a mesma resolução da imagem final desejada. Com esta nova resolução as imagens precisam ser registradas novamente, já que a interpolação modifica os valores radiométricos dos pixels das imagens originais, conforme ilustrado na **Figura 1**. As imagens, sob o novo *grid* de melhor resolução, podem apresentar deslocamentos entre os pixels que representam um mesmo alvo. A partir deste novo registro, chamado sub-pixel, o método de super-resolução é aplicado (porém não será abordado neste artigo por não ser este seu objetivo).

125	90	83
80	100	78
130	95	70

(a)

135	123	98	83	82	84
114	109	98	88	83	82
82	88	97	96	85	78
84	90	100	96	82	73
118	114	103	89	75	69
140	129	106	85	72	68

(b)

Figura 1. (a) *grid* de baixa resolução, (b) *grid* de alta resolução, resultado de interpolação utilizando o método Lanczos 3 - Matlab.

Para aplicação e análise dos resultados foram utilizadas imagens IKONOS, LANDSAT, pares de fotografias tiradas simultaneamente por um dispositivo que contém duas câmeras digitais idênticas desenvolvido no Laboratório de Processamento Analógico e Digital de Sinais da COPPE/UFRJ e fotografias digitais tiradas individualmente de alvos cujos detalhes se mostraram de interesse para a análise dos resultados obtidos.

## 2. Desenvolvimento

Como a suposição neste trabalho é a de que existam apenas diferenças de rotação e translação entre as imagens originais, seu desenvolvimento pode ser dividido em cinco etapas: determinação e correção da rotação entre as imagens de baixa resolução; determinação dos pontos de controle entre as imagens de baixa resolução; determinação e correção da translação entre as imagens de baixa resolução; reamostragem das imagens para a resolução final

desejada; e determinação dos pontos de controle e correção da translação entre as imagens reamostradas. Estas etapas serão apresentadas a seguir.

## 2.1. Rotação

Definindo-se como  $\mathfrak{F}(Im)$  a transformada de Fourier de uma imagem  $Im$  e  $|\mathfrak{F}(Im)|$  a representação bi-dimensional das amplitudes da mesma imagem no domínio da frequência,  $|\mathfrak{F}(Im)|$  pode ser utilizada para a determinação do valor de rotação  $\alpha$  entre duas imagens. Vandewalle et al (2006) mostram que o ângulo de rotação  $\alpha$  entre duas imagens,  $Im_1$  e  $Im_2$ , é o mesmo que entre  $|\mathfrak{F}(Im_1)|$  e  $|\mathfrak{F}(Im_2)|$ . A vantagem de se utilizar as representações de amplitude no domínio da frequência reside no fato de que  $\alpha$  pode ser determinado independentemente dos valores de translação  $\Delta x$  e  $\Delta y$  que possam existir entre  $Im_1$  e  $Im_2$ , pois estes deslocamentos lineares somente influenciam os valores de fase da transformada de Fourier. Sendo assim, determinando-se a rotação sofrida pelas linhas que aparecem em  $|\mathfrak{F}(Im_1)|$  e  $|\mathfrak{F}(Im_2)|$ , na parte inferior da **Figura 2**, o ângulo de rotação entre as imagens no domínio do espaço estará determinado e estas poderão ser corrigidas rotacionando-se uma das imagens no sentido contrário.

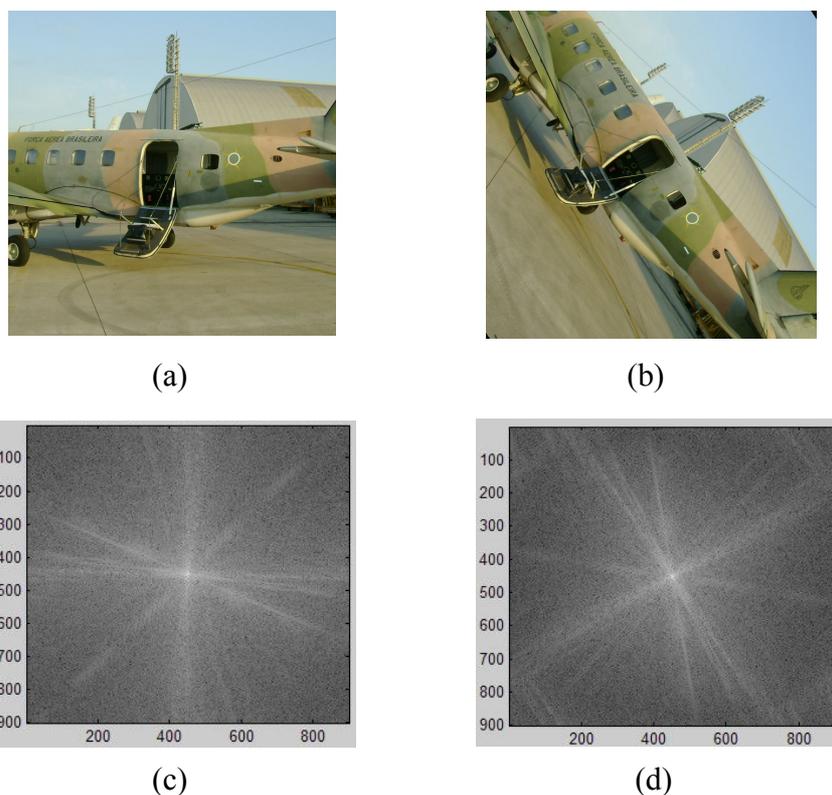


Figura 2. Imagens de um Bandeirante da FAB em sua posição normal (a) e rotacionada em 56,7 graus (b); (c) e (d) representações bi-dimensionais dos valores de amplitude das transformadas de Fourier relativas às figuras (a) e (b) respectivamente. As linhas que partem do centro são formadas pelas altas frequências das imagens originais.

O ângulo de rotação  $\alpha$  pode ser determinado com o auxílio de coordenadas polares, dividindo-se  $|\mathfrak{F}(Im_1)|$  e  $|\mathfrak{F}(Im_2)|$  em 3.600 setores, a partir do centro, para obter-se uma precisão de  $0,1^\circ$ . A máxima correlação entre os setores de  $|\mathfrak{F}(Im_1)|$  e  $|\mathfrak{F}(Im_2)|$  indicará a rotação entre  $Im_1$  e  $Im_2$ . Para se determinar a correlação entre os setores de  $|\mathfrak{F}(Im_1)|$  e  $|\mathfrak{F}(Im_2)|$  somam-se os valores de todos os pixels pertencentes a cada setor, armazenando-os

em um vetor linha relativo a cada imagem  $\aleph(\text{Im}_1)$  e  $\aleph(\text{Im}_2)$ . Mantendo-se  $\aleph(\text{Im}_1)$  fixo e reordenando-se os componentes de  $\aleph(\text{Im}_2)$  de forma que o último elemento, que corresponde ao último setor, de  $\aleph_i(\text{Im}_2)$  passe a ser o primeiro elemento de  $\aleph_{i+1}(\text{Im}_2)$  e todos os demais componentes sejam consequentemente reordenados, calcula-se a norma  $\| \aleph(\text{Im}_1) - \aleph_i(\text{Im}_2) \|$ , onde o menor valor encontrado entre as normas calculadas indicará o valor de  $\alpha$ , conforme ilustrado na **Figura 3**.

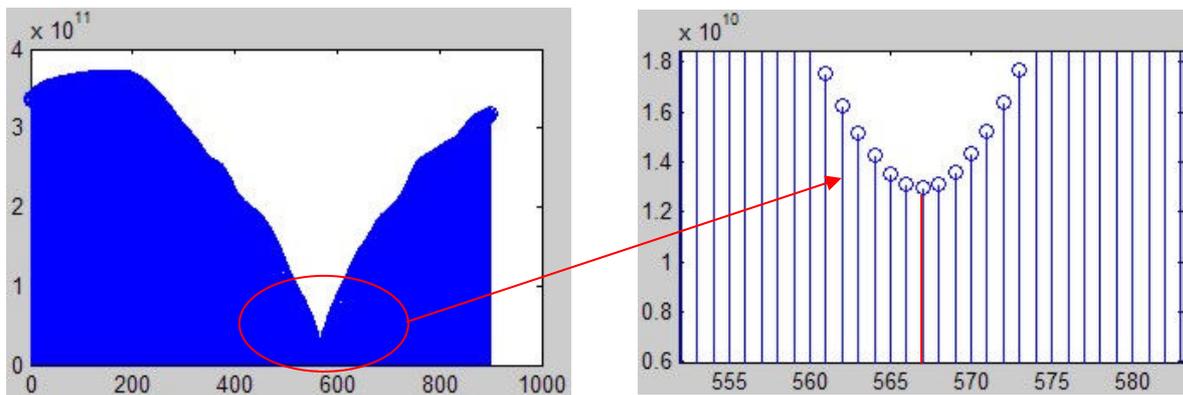


Figura 3. Gráfico mostrando os resultados obtidos com os cálculos de  $\| \aleph(\text{Im}_1) - \aleph_i(\text{Im}_2) \|$ . O menor valor no detalhe do “zoom” indica  $\alpha = 56,7^\circ$  de rotação entre  $\text{Im}_1$  e  $\text{Im}_2$ .

## 2.2. Determinação dos pontos de controle

A determinação de pontos de controle não é imprescindível ao processo de correção da translação entre as imagens, mas sua utilização acelera consideravelmente o tempo computacional necessário, justificando sua utilização.

Os pontos de controle  $P(x,y,c)$  são encontrados comparando-se o valor do nível de cinza “c” de cada pixel a ser avaliado  $c_1=P(x,y)$  com seus oito vizinhos:  $c_2=P(x-1,y-1)$ ;  $c_3=P(x,y-1)$ ;  $c_4=P(x+1,y-1)$ ;  $c_5=P(x-1,y)$ ;  $c_6=P(x+1,y)$ ;  $c_7=P(x-1,y+1)$ ;  $c_8=P(x,y+1)$ ;  $c_9=P(x+1,y+1)$ . A diferença entre o nível de cinza do pixel central e seus vizinhos deve ser igual ou maior que um valor de corte “vc” estipulado ( $P_1 - P_i \geq vc$ , para  $2 \leq i \leq 9$ ). Este valor de corte está ligado com a quantidade de pontos de controle julgada necessária, ou seja, quanto maior a quantidade de pontos de controle desejados, menor a diferença radiométrica aceita entre o pixel candidato a ponto de controle e seus vizinhos, sendo este procedimento realizado apenas na imagem de referência.

Testes realizados mostraram que 500 pontos de controle são suficientes para levarem a uma boa precisão no processo de determinação da translação. Esses testes foram realizados tanto em imagens com grande incidência de altas frequências, como em imagens monótonas (fotografia de uma modelo em um fundo de uma única cor). A utilização de uma quantidade maior de pontos de controle só se justifica em imagens onde sua distribuição se apresenta muito irregular, ou quando as estatísticas de precisão do processo de correção da translação não se mostram favoráveis, conforme explicação no próximo item.

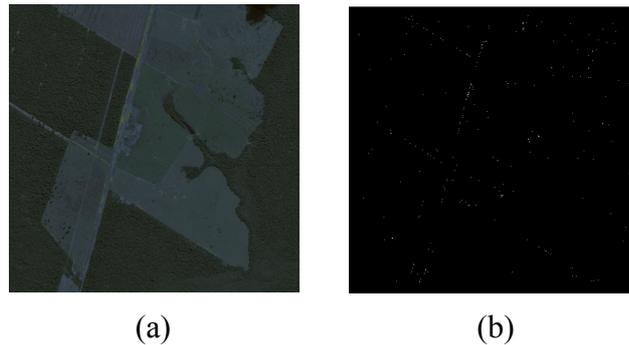


Figura 4. (a) Imagem IKONOS original da região de Sinop-PA, (b) Distribuição dos pontos de controle relativos à imagem original.

### 2.3. Translação

Após a identificação dos pontos de controle, suas coordenadas  $(x,y)$  deverão ser armazenadas em um vetor posição. Uma pequena janela base “J” em torno de cada ponto de controle deve ser determinada e seus valores armazenados, sendo  $J = P_i(x+m,y+n,c_i)$  para  $-m \leq i \leq m$  e  $m = n$  determinando os limites da janela e sendo  $(2m+1)^2$  é a quantidade total de pixels de J.

Testes realizados mostraram que janelas 13x13 ( $m=n=6$ ) atingem resultados satisfatórios, sendo que o aumento deste tamanho não melhora os resultados obtidos. Sua redução até o valor de 9x9 ( $m=n=4$ ) em alguns testes causou pequena piora na precisão do processo. A consequente redução no tempo de computação não foi significativa, devido à quantidade de pontos de controle normalmente utilizada, o que não justifica a utilização destes valores reduzidos.

Baseado na mesma posição  $(x,y)$ , na imagem a ser corrigida, varre-se dentro da área de uma segunda janela “JJ”, tal que  $JJ > J$ , os valores de todas as possíveis janelas de busca, estas do mesmo tamanho da janela base J, subtraindo-se os valores radiométricos  $c_i$  em cada janela de busca, pixel a pixel, respeitando as respectivas posições, dos valores radiométricos da janela base. A diferença encontrada para cada pixel deverá ser elevada ao quadrado e todas as diferenças somadas, obtendo-se um valor diferença final único para cada janela de busca em relação à janela base, sendo este armazenado. Este método é denominado de correlação simples e está exemplificado em (1).

$$\text{Dif Total} = \sum_{i=1}^{m \times n} (P_i(J \text{ base}) - P_i(J \text{ busca}))^2 \quad (1)$$

A menor diferença total encontrada determina a posição do respectivo ponto homólogo na imagem a ser corrigida e a diferença de coordenadas entre o ponto de controle e seu homólogo deve ser armazenada.

$$\text{Pos P Homólogo} = \text{argmin}(\text{diferença total entre todas as janelas de busca}) \quad (2)$$

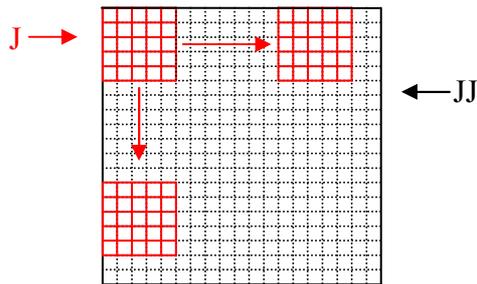


Figura 5. Representação de uma janela de busca 5x5 se deslocando pela janela JJ de dimensões 19x19. O pixel central da janela JJ encontra-se nas coordenadas (x,y) iguais ao do ponto de controle na imagem tomada como referência. Em cada posição ocupada pela janela de busca realizam-se os cálculos dos valores radiométricos e o valor da diferença final é armazenado para ser comparado com os valores das demais. O menor valor indica a posição do ponto homólogo na imagem a ser corrigida.

O artifício do uso da janela JJ tem por objetivo limitar a área de busca do ponto homólogo pela imagem que será transformada, poupando um grande esforço computacional. As dimensões da janela JJ podem ser determinadas de duas formas: através de uma rápida inspeção visual, onde é possível avaliar um valor em “X” e em “Y” onde certamente o ponto homólogo será encontrado, ou fazendo com que os primeiros pontos sejam procurados por toda a imagem e, baseando-se nos valores para a diferença das coordenadas encontradas para os primeiros pontos, o algoritmo automaticamente reduz a área de busca. Uma terceira possibilidade, a mais rápida, é a combinação das duas alternativas anteriores.

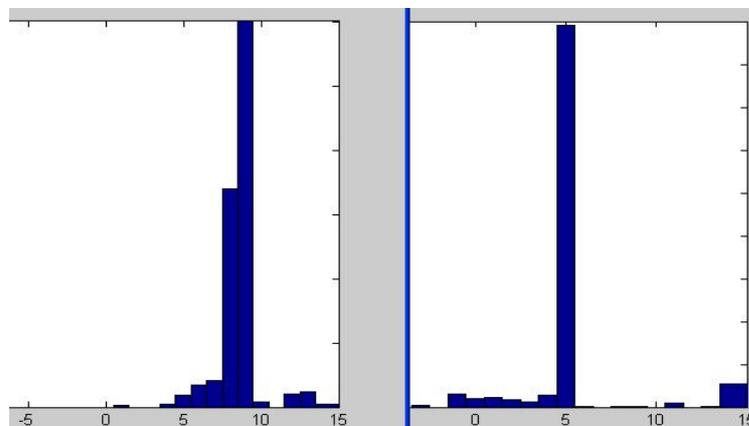


Figura 6. Avaliação do resultado dos valores de translação em “X” e “Y” entre cenas IKONOS de 2001 e 2002 da região de Sinop-PA. Os resultados apontam um deslocamento de 9 pixels para a direita e 5 pixels para cima. O histograma ilustra a incidência do valor modal em relação aos demais valores: 55% para linha e 81% para coluna, indicando o registro satisfatório entre as imagens.

Mesmo em cenas muito comportadas tomadas por sistemas ópticos precisos, as diferenças registradas, relativas aos diversos pontos de controle, não serão idênticas, apresentando diferenças esporádicas e/ou aleatórias que devem ser ignoradas sob pena de imprecisão do processo. Por isso determina-se a moda entre todas as diferenças  $\Delta x$  e  $\Delta y$  encontradas relativas aos pontos de controle. Este procedimento torna o processo mais robusto livrando-o de erros devido à mudança de iluminação da cena (no caso de imagens multitemporais), áreas repletas de altas frequências (como em florestas), dentre outras possíveis variações radiométricas entre as imagens.

Os resultados obtidos foram avaliados quantitativamente através de histogramas exemplificados na **Figura 6**, que registram a incidência para cada valor encontrado como diferença de coordenadas entre os pontos homólogos. Quando o registro está perfeito em toda a imagem o valor encontrado como moda difere bastante dos demais. Uma menor diferença indica que as pequenas distorções existentes entre as imagens farão com que o ajuste não fique perfeito em toda sua extensão, o que poderá impedir que o processo seja utilizado para toda a área da imagem. Neste caso, ou as áreas de registro em cada imagem são reduzidas, ou outro processo que inclua correção de escala deverá ser utilizado.

## 2.4. Interpolação

Com as imagens registradas, o próximo passo é a reamostragem, ou seja, a transformação destas imagens em outras com maior quantidade de pixels. Como não é possível criar informações, não se pode falar em imagens de melhor resolução espacial. Esta melhora nos detalhes das imagens poderá ser obtida através de um processo de super-resolução, que não será abordado neste trabalho. A reamostragem das imagens registradas se traduz em uma redundância de informações com a criação dos novos pixels, necessária para o registro sub-pixel e para outros tipos de processamentos.

Foram testados métodos de interpolação que fazem parte do pacote do Matlab: bilinear, convolução cúbica, lanczos, lanczos 2 e lanczos 3. O teste realizado, simples e objetivo, consistiu em passar um filtro que calculava a média de quatro pixels vizinhos, reduzindo pela metade a quantidade de linhas e colunas de cada imagem. A estas imagens reduzidas foram aplicados os métodos de interpolação citados para se obter imagens interpoladas com a mesma quantidade de linhas e colunas das imagens iniciais. Com as imagens de mesmas dimensões calculou-se a norma entre a imagem original e cada imagem reamostrada, ou seja:

$$\text{Dif Total} = \| \text{Im}_{\text{original}} - \text{Im}_{\text{reamostrada}} \| \quad (3)$$

Os resultados foram muito próximos, com variações em torno de 1% entre os demais métodos. O método lanczos 3 foi o que apresentou menor diferença sendo este o interpolador adotado.

## 2.5. Determinação dos pontos de controle e correção da translação das imagens reamostradas

O mesmo procedimento de registro anterior foi adotado em relação às imagens interpoladas. Para a determinação dos pontos de controle, utilizou-se o critério da diferença radiométrica entre o pixel e seus vizinhos. Ocasionalmente, para esta fase, um número maior de pontos de controle pode ser necessário, mas na maioria dos casos, os 500 pontos foram suficientes.

Assim como no caso da determinação dos pontos de controle, para a determinação e correção dos valores de translação sub-pixel os mesmos procedimentos foram adotados. Os valores estatísticos dos resultados para esta fase normalmente são mais baixos, porém a metodologia também se mostrou adequada para imagens interpoladas. Em um teste didático utilizando a área central de imagens LANDSAT (227-62) do norte do Pará, os valores modais encontrados para o registro das imagens alcançaram 99% para linha e 98% para coluna, enquanto que para o registro sub-pixel estes valores se reduziram para 76% nas linhas e 75% nas colunas.

### 3. Conclusões

A utilização do domínio da frequência para a correção da rotação entre imagens se mostrou como uma alternativa bastante robusta, determinando com a precisão prevista de  $0,1^\circ$  o ângulo de rotação das imagens testadas, apresentando falhas apenas em imagens bastante monótonas, onde a falta de bordas (altas frequências) resultou em valores muito próximos em todos os setores em que  $|\mathfrak{I}(Im)|$  foi dividida.

A busca automática de pontos de controle, valendo-se das diferenças radiométricas entre pixels, obteve resultados rápidos e diminuiu consideravelmente o custo computacional do restante do processo. Os pontos de controle encontram-se, em maior quantidade, em regiões de altas frequências, áreas da imagem onde é mais importante o casamento das diversas feições que compõem a cena. Nas áreas onde há o predomínio das baixas frequências e, conseqüentemente, menor incidência de pontos de controle, o não casamento entre as cenas é dissimulado, ou mesmo desapercibido, devidos às características monótonas destas regiões.

Devido à utilização de valores radiométricos, o método de busca de pontos de controle abordado neste trabalho não se aplica a imagens de radar, devido ao *speckle*, e pode causar imprecisões em cenas cuja área coberta por florestas seja muito extensa.

Conforme já mencionado no início do trabalho, o processo de determinação da translação em X e Y é aplicável: em imagens cujas distorções entre si não causem problemas imperativos de escalas, facilmente identificados em imagens de baixa e média resolução espacial, devido à grande largura da faixa imageada; imagens de visadas oblíquas; dentre outras. A maior vantagem em se trabalhar apenas com rotação e translação de imagens é a manutenção precisa dos valores radiométricos originais das cenas, o que vem a ser muito importante para alguns processos, como super-resolução. Esta vantagem, inevitavelmente, vem acompanhada das restrições geométricas devido à falta de ajuste em escala entre as imagens. Nestes casos, a redução da área de trabalho se torna uma alternativa importante.

### 4. Referências

- Clark, J. J.; Palmer, M. R.; Laurence, P. D. A transformation method for the reconstruction of functions from nonuniformly spaced samples, **IEEE Trans. Acoust., Speech, Signal Processing**, vol. ASSP-33, pp. 1151-1165, 1985.
- Kim, S. P.; Bose, N. K. Reconstruction of 2-D bandlimited discrete signals from nonuniform samples, **Proc. Inst. Elec. Eng.**, vol. 137, part. F, pp. 197-204, June 1990.
- Mascarenhas, N. D. A.; Velasco, F. R. D. **Processamento digital de imagens**. Buenos Aires, Kapelus, 1989. Quarta Escola Brasileiro-Argentina de Informática, Termas de Rio Hondo, Argentina, janeiro 1989.
- Mather, P. M. **Computer Processing of Remotely-Sensed Images: an introduction**. John Wiley & Sons, 1999.
- Park, S. C.; Park, M. K.; Kang, M. G. Super-resolution image reconstruction: a technical overview, **IEEE Signal Processing Magazine**, pp. 21-36, May 2003.
- Pedrini, H.; Schwartz W. R. **Análise de imagens digitais: princípios, algoritmos e aplicações**. São Paulo, Thomson Learning, 2008.
- Schowengerdt, R. A. Remote Sensing, models and methods for image processing. Academic Press, 1997, 522p.
- Vandewalle, P.; Susstrunk, S.; Vetterli, M. A frequency domain approach to registration of aliased images with application to super-resolution, **EURASIP Journal on Applied Signal Processing**, vol. 2006, pp 1-14, 2006.