

Sistema interativo de registro de imagens baseado em feições

Camilo Daleles Rennó

Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12201-970 - São José dos Campos - SP, Brasil
camilo@dpi.inpe.br

Abstract. This work presents an image registration system that uses a feature-based approach during the acquisition of control points. This method minimizes the difficulties caused by the discrepancy and inconsistency among images. The system is capable of handling images with moderated deformation. In addition, an evaluation method of registration, based on PRESS criterion, is described. The system was tested on a real application and the results showed the robustness of the system.

Palavras-chave: Registration, image processing, evaluation method, registro de imagens, processamento de imagens, método de avaliação

1. Introdução

O registro de imagens é uma técnica muito utilizada em aplicações de processamento de imagens que visam a análise conjunta de duas ou mais imagens tomadas em diferentes épocas, em diferentes visadas e/ou por diferentes sensores, como por exemplo, a fusão de imagem, a detecção de mudança, a restauração de imagem, entre outras. Uma revisão completa de métodos de registro de imagens pode ser encontrada em Brown (1992) e Zitova e Flusser (2003).

As técnicas de registro de imagens têm evoluído muito nos últimos anos. Técnicas mais recentes têm permitido o registro automático de imagens, facilitando o processo de georreferenciamento de imagens (Bentoutou et al, 2005), através da seleção e identificação automática ou semi-automática de pontos cujas coordenadas sejam correspondentes entre as imagens a serem registradas. No entanto, este registro automático pode ser dificultado devido a pouca coerência entre as imagens analisadas. Em alguns casos, apenas poucas feições estão presentes em ambas as imagens. Neste caso, mesmo a seleção manual de pontos concordantes nas imagens pode ser difícil.

O objetivo deste trabalho é apresentar um sistema, denominado RegIm (Sistema de Registro de Imagens), que permite fazer o registro interativo de imagens. O grande diferencial neste sistema é que uma das imagens a ser registrada é convertida para o formato vetorial, permitindo que ambas sejam sobrepostas. O ajuste entre as imagens é feita simultaneamente a coleta de pontos, facilitando o processo de registro. A grande motivação deste trabalho foi produzir uma ferramenta de fácil manipulação por usuários de processamento de imagens e sistemas de informação geográfica em geral. O sistema RegIm foi desenvolvido utilizando a linguagem IDL (*Interactive Data Language*) e usa recursos disponibilizados no ENVI (*The Environment for Visualizing Images*).

2. Metodologia para o registro de imagem

O principal objetivo do registro é retificar uma imagem de modo que a imagem resultante possua o mesmo sistema de coordenada da imagem utilizada. Com o registro, todas as imagens passam a ter uma correspondência espacial.

Tradicionalmente, o registro de imagens envolve as seguintes etapas: a) definição da imagem referência usada como base para a transformação de uma outra imagem, denominada imagem distorcida; b) identificação de pontos de correspondência espacial, denominados pontos de controle; c) definição da modelo de transformação espacial e estimação de seus

parâmetros; e d) reamostragem da imagem resultante através de uma função de interpolação qualquer.

A seguir serão discutidos mais profundamente alguns aspectos relativos aos modelos de transformação espacial e à avaliação do resultado do registro. Os procedimentos e as técnicas assim como as funções de interpolação utilizados no processo de reamostragem das imagens não serão abordados neste trabalho.

2.1. Modelos de transformação espacial

A definição da função de transformação espacial deve considerar a natureza das deformações entre as imagens a serem registradas e a precisão desejada do registro. Os casos mais simples de deformações envolvem apenas correções de posição (translação), de escala e rotação. Quando a escala é a mesma para ambos os eixos, tem-se uma transformação denominada corpo rígido, onde tanto distâncias quanto ângulos de um objeto qualquer na imagem são preservados. Esta transformação pode ser expressa por

$$\begin{bmatrix} x^* \\ y^* \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \end{bmatrix} + s \begin{bmatrix} \cos\theta & -\text{sen}\theta \\ \text{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (1)$$

onde x^* e y^* são as coordenadas de um ponto qualquer na imagem referência, x e y as coordenadas do ponto correspondente na imagem a ser registrada, t_x e t_y são os valores de translação nos eixos x e y respectivamente, s é o fator de escala global e θ é o ângulo de rotação. A **Equação 1** pode ser modificada para permitir que as escalas sejam diferentes para cada eixo, resultando numa transformação afim ortogonal que preserva os ângulos de um objeto após a transformação:

$$\begin{bmatrix} x^* \\ y^* \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \end{bmatrix} + \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} \cos\theta & -\text{sen}\theta \\ \text{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

onde s_x e s_y representam os fatores de escala para cada um dos eixos. Esta transformação pode ainda permitir a quebra da ortogonalidade, mas ainda conservando o paralelismo, através da inclusão de fatores de cisalhamento para cada eixo, k_x e k_y :

$$\begin{bmatrix} x^* \\ y^* \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \end{bmatrix} + \begin{bmatrix} 1 & k_x \\ k_y & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} \cos\theta & -\text{sen}\theta \\ \text{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (3)$$

Alternativamente, a **Equação 3** pode ser reescrita através de duas equações lineares bivariadas:

$$\begin{aligned} x^* &= a_x + b_x x + c_x y \\ y^* &= a_y + b_y x + c_y y \end{aligned} \quad (4)$$

onde a_x , b_x , c_x , a_y , b_y , c_y são parâmetros das equações de transformação, sendo esta forma preferida quando não se deseja explicitar as componentes (translação, escala, rotação e cisalhamento) do modelo escolhido para a transformação. A **Figura 1** apresenta os tipos de deformações que podem ser produzidos através da aplicação deste modelo.

Outras transformações geométricas mais complexas podem ser conseguidas com o acréscimo de mais termos à **Equação 4**, como por exemplo a quebra do paralelismo através da inclusão do termo de interação xy , de modo que as equações de transformação tornam-se

$$\begin{aligned} x^* &= a_x + b_x x + c_x y + d_x xy \\ y^* &= a_y + b_y x + c_y y + d_y xy \end{aligned} \quad (5)$$

sendo d_x e d_y os parâmetros de interação. Qualquer outro polinômio pode ser utilizado durante o processo de registro. No entanto, seu uso deve ser feito com cautela, uma vez que, na maioria das aplicações que fazem uso do registro, não se faz necessário fortes deformações

para se fazer o ajuste entre imagens. A **Figura 2** apresenta algumas transformações geométricas onde fortes deformações são impostas com o uso de polinômios de maior ordem.

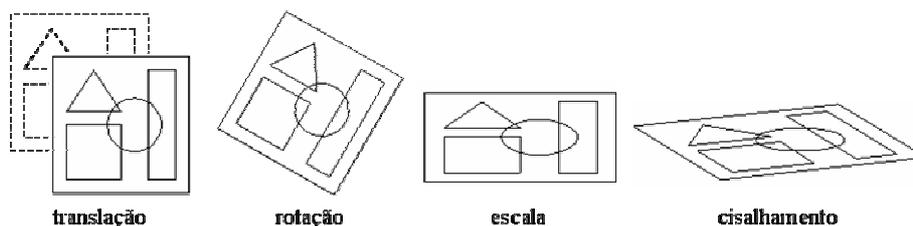


Figura 1 – Exemplos de transformações geométricas.

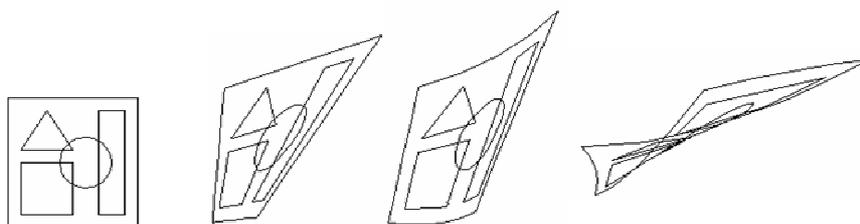


Figura 2 – Exemplos de transformações geométricas mais complexas. Da esquerda para direita: desenho original, deformação usando termo de interação e duas deformações obtidas usando-se um polinômio bivariado de segundo grau.

Uma vez escolhido o modelo de transformação, seus parâmetros devem ser estimados através da identificação e seleção de pontos de correspondência (pontos de controle) entre as imagens que estão sendo registradas. Para cada modelo de transformação, é exigido um número mínimo de pontos de controle definido pelo número de parâmetros a serem estimados. Por exemplo, com apenas um par de pontos, somente é possível fazer duas translações, o que corrigiria problemas de posicionamento. Com dois pares de pontos, é possível se fazer duas translações, uma rotação e uma mudança de escala em ambos os eixos simultaneamente (transformação de corpo rígido). O efeito de cisalhamento é conseguido quando se utiliza três ou mais pares de pontos. Finalmente, a quebra do paralelismo pode ser obtida com o uso de quatro pares de pontos.

2.2. Avaliação do modelo

Quando são coletados mais pontos de controle do que o mínimo exigido pelo modelo de transformação, a estimação dos parâmetros é feita através da minimização dos erros quadráticos. Dessa forma, a transformação não é exata e o erro (ou resíduo) de posicionamento de cada ponto de controle pode ser utilizado para verificar sua adequação ao modelo. Em geral, o erro médio quadrático (*RMS*, *root mean square*) é utilizado como medida da precisão global do registro. Considerando-se, como exemplo, as equações de transformação da **Equação 4**, e supondo-se que n pontos de controle (representando n pares de coordenadas) sejam coletados ($n \geq 3$, nesse caso), a estimação dos parâmetros pode ser obtida por

$$\begin{bmatrix} a_x \\ b_x \\ c_x \end{bmatrix} = \begin{bmatrix} n & \sum x_i & \sum y_i \\ \sum x_i & \sum x_i^2 & \sum x_i y_i \\ \sum y_i & \sum x_i y_i & \sum y_i^2 \end{bmatrix} \begin{bmatrix} \sum x_i^* \\ \sum x_i^* x_i \\ \sum x_i^* y_i \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} a_y \\ b_y \\ c_y \end{bmatrix} = \begin{bmatrix} n & \sum x_i & \sum y_i \\ \sum x_i & \sum x_i^2 & \sum x_i y_i \\ \sum y_i & \sum x_i y_i & \sum y_i^2 \end{bmatrix} \begin{bmatrix} \sum y_i^* \\ \sum y_i^* x_i \\ \sum y_i^* y_i \end{bmatrix} \quad i=1, \dots, n$$

Desse modo, após a aplicação da transformação, as estimativas das coordenadas de cada ponto de controle i e o *RMS* correspondente podem ser obtidos por

$$\begin{aligned}\hat{x}_i^* &= a_x + b_x x_i + c_x y_i \\ \hat{y}_i^* &= a_y + b_y x_i + c_y y_i\end{aligned}\quad (7)$$

e

$$RMS = \sqrt{\frac{\sum [(x_i^* - \hat{x}_i^*)^2 + (y_i^* - \hat{y}_i^*)^2]}{n}}. \quad (8)$$

Idealmente, a avaliação do registro requer um conjunto independente de pares de pontos de controle (comumente chamado de conjunto de teste), cujas coordenadas sejam precisamente conhecidas. Nesse caso, uma avaliação é obtida estimando-se as posições destes novos pontos com base no modelo estimado a partir do conjunto original de pontos de controle (comumente chamado de conjunto de treinamento). O *RMS*, como definido na **Equação 8**, pode ainda ser utilizado como uma medida global de validação. Se o modelo for consistente, espera-se que este novo *RMS* não seja muito diferente daquele previamente encontrado durante a avaliação de ajuste do modelo.

Em casos em que o número de pontos de controle disponível é pequeno ou não se tem informação precisa sobre pontos de controle, esta avaliação torna-se bastante limitada. Nesse caso, o critério PRESS (Neter et al, 1996) pode ser facilmente adaptado a fim de fornecer uma métrica que permita avaliar a capacidade preditiva do modelo selecionado. Este método analisa a importância de cada ponto de controle na definição do modelo.

A métrica proposta neste trabalho, denominada *RMSp* (o *p* faz alusão ao caráter preditivo desta métrica), baseada no critério PRESS, semelhantemente ao *RMS*, calcula uma média das somas quadráticas dos deslocamentos entre os pontos de controle e suas estimativas:

$$RMSp = \sqrt{\frac{\sum [(x_i^* - \hat{x}_{(i)}^*)^2 + (y_i^* - \hat{y}_{(i)}^*)^2]}{n}} \quad (9)$$

onde $\hat{x}_{(i)}^*$ e $\hat{y}_{(i)}^*$ representam as estimativas de predição para o ponto x_i^* e y_i^* . Neste caso, cada estimativa *i* é feita através de um ajuste do modelo sem a inclusão deste ponto no processo de estimação dos parâmetros. Valores baixos de *RMSp* indicam boa distribuição espacial dos pontos de controle e forte robustez do modelo ajustado. A avaliação do erro de predição de cada ponto de controle, $\sqrt{(x_i^* - \hat{x}_{(i)}^*)^2 + (y_i^* - \hat{y}_{(i)}^*)^2}$, também pode indicar o poder deste ponto em forçar o ajuste do modelo. A existência de erros de predição altos pode indicar má distribuição espacial dos pontos de controle ou a adoção inadequada de modelos de transformação de alta complexidade.

3. Sistema RegIm

O sistema RegIm é basicamente composto por um sistema de janelas, duas das quais são usadas para visualização das imagens e uma terceira que gerencia a seleção dos pontos de controle. A **Figura 3** apresenta a interface do sistema durante a aquisição dos pontos dos pontos de controle.

3.1. Entrada de dados

O primeiro passo no processo de registro é a definição da imagem de referência (imagem base) e da imagem a ser corrigida (imagem distorcida). Para permitir a visualização simultânea das imagens numa mesma janela, uma delas (ou ambas) deve ser vetorizada. No processo de vetorização, todas as descontinuidades são convertidas em linhas (**Figura 4**). Por simplificação, as imagens a serem vetorizadas devem representar imagens binárias ou imagens cujos valores representem manchas contínuas no espaço de algum fenômeno que

possa ser identificado na outra imagem. Quando as imagens a serem registradas representarem limites de polígonos e não manchas contínuas, a vetorização deve ser feita de modo que os centros dos pixels marcados como pertencentes a esse limite (cujos valores sejam diferentes de zero) sejam unidos por uma linha (**Figura 4**)

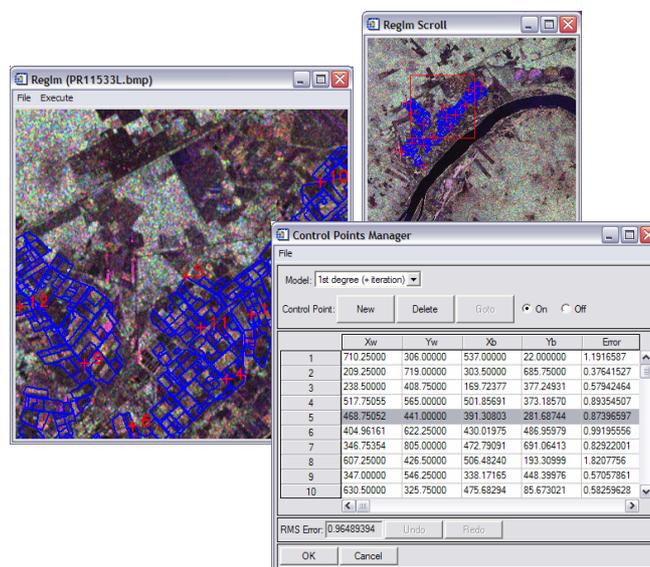


Figura 3 – Interface do sistema RegIm.

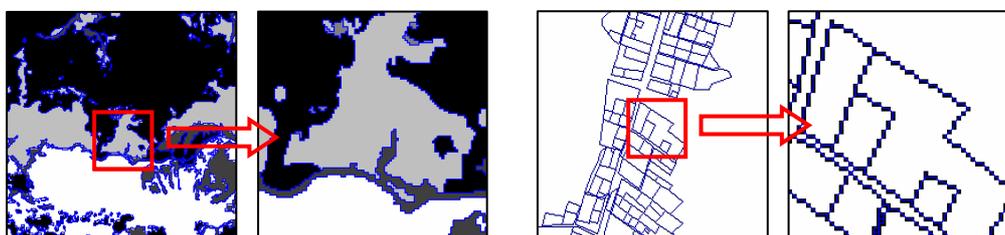


Figura 4 – Vetorização de imagem de manchas e de limites.

Caso as imagens não tenham a mesma projeção cartográfica, a projeção da imagem a ser registrada é automaticamente convertida para a da imagem referência. Caso somente a imagem referência possua projeção cartográfica, a imagem distorcida é vetorizada e sua coordenada modificada de maneira a se ajustar à primeira. Se a imagem referência não possuir projeção cartográfica, então ambas as imagens utilizarão o sistema de coordenadas baseado em linhas e colunas.

3.2. Seleção dos pontos de controle e ajuste dos parâmetros do modelo de transformação espacial

Com as imagens a serem registradas inseridas no sistema, inicia-se o processo de seleção de pontos de controle que permitirão o ajuste dos parâmetros do modelo selecionado. O sistema RegIm possui 4 modelos para o polinômio de transformação: a) primeiro grau com escala global (**Equação 1**); b) primeiro grau com escalas diferentes para cada eixo (**Equação 2**); primeiro grau com termo de interação (**Equação 3**); e d) segundo grau. Além disso, opcionalmente, pode-se ajustar automaticamente o modelo de maior complexidade, dado o número de pontos de controle disponíveis.

A **Figura 5** ilustra as etapas envolvidas no processo de seleção dos pontos de controle. As linhas em azul representam a imagem a ser registrada à imagem de referência (ao fundo). Com a seleção do primeiro par de pontos (imagem do lado superior esquerdo da figura), a imagem distorcida é arrastada (através de duas translações) de modo que os pontos selecionados se coincidam. A partir da seleção do segundo par de pontos (imagem do lado inferior esquerdo), uma etapa intermediária oculta ao usuário é executada: ao selecionar o ponto sobre a imagem distorcida, a coordenada deste ponto é convertida para a projeção original desta imagem (antes de qualquer correção) formando, juntamente com a coordenada do ponto selecionado sobre a imagem referência, o segundo par de pontos, resultando na transformação mostrada no canto superior direito da **Figura 5**.

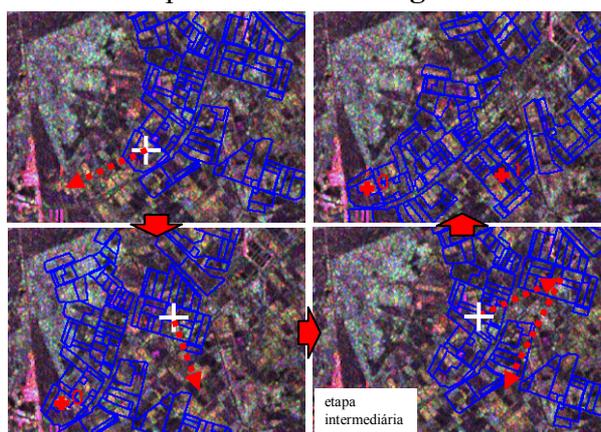


Figura 5 – Etapas da seleção dos pontos de controle.

Como cada ponto de controle pode ser movido livremente, é possível, numa primeira aproximação, selecionar pontos próximos a feições semelhantes e só então fazer um ajuste fino observando a sobreposição das feições no entorno. Este procedimento facilita muito o processo de registro, pois o ponto de controle pode representar centros aproximados de polígonos e não necessariamente cantos e cruzamentos de linhas, como usualmente são os pontos de controle.

3.3. Reamostragem

Uma vez definido o modelo apropriado e estimado seus parâmetros, a imagem distorcida pode ser reamostrada de modo a tornar-se compatível com a imagem de referência. Pelo fato, da imagem distorcida representar uma imagem binária ou manchas contínuas no espaço, a função de interpolação mais apropriada é a do vizinho mais próximo, que preserva os valores de cada elemento da imagem.

A imagem resultante tem as mesmas dimensões, resoluções espaciais e projeção daquela usada como referência, garantindo assim a correspondência espacial necessária para o estudo conjunto das mesmas.

4. Aplicação e discussão dos resultados

Durante o desenvolvimento do sistema, um conjunto de dados obtido de Rennó (1995) foi utilizado. A aplicação consistiu no registro de um croqui, obtido durante um trabalho de campo, a imagens de radar. O croqui representava a divisão em talhões de cultura agrícolas em propriedades dentro de um perímetro irrigado próximo a Petrolina (PE). O conjunto de imagens de radar foi adquirido simultaneamente ao trabalho de campo durante a missão SIR-C, e é composto por três imagens na banda L nas polarizações HH, HV e VV.

O croqui foi elaborado a partir de um mapa em papel com os limites das propriedades, não sendo conhecidas a escala e a projeção deste mapa. Para este trabalho, o croqui foi digitalizado, obtendo-se uma imagem binária com os limites dos talhões destacados (**Figura 6**). A dificuldade principal, nesse caso, foi a identificação de pontos de controle no conjunto de imagens analisadas. Problemas devido à imprecisão na definição exata dos limites dos talhões e à existência de ruídos nas imagens de radar dificultariam a aquisição de pontos de controle e tornariam o processo de registro convencional bastante demorado.

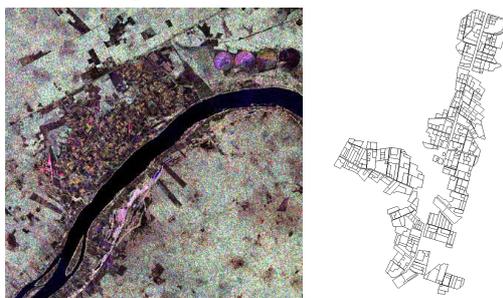


Figura 6 – Imagem SIR-C, banda L, polarizações HH, HV e VV (RGB), e croqui digitalizado mostrando a divisão entre talhões num perímetro irrigado. Ambas as bases não possuem escala nem projeção.

Neste trabalho, foi testada a possibilidade do uso de modelos de transformação mais simples, uma vez que, aparentemente, não havia grandes distorções entre os dois tipos de dados, não justificando a adoção de modelos polinomiais de segunda ordem. Para tanto, foram selecionados 15 pontos de controle (muito superior ao mínimo exigido para o modelo quadrático), o que permitiu a avaliação da adequação do modelo e dos pontos de controle selecionados. A **Figura 7** apresenta o resultado do registro com a utilização do modelo de primeiro grau considerando-se a interação (**Equação 4**). Note, no detalhe da figura, que nem todas as feições do croqui são observadas na imagem de radar, evidenciando o problema de falta de coerência entre os dados.

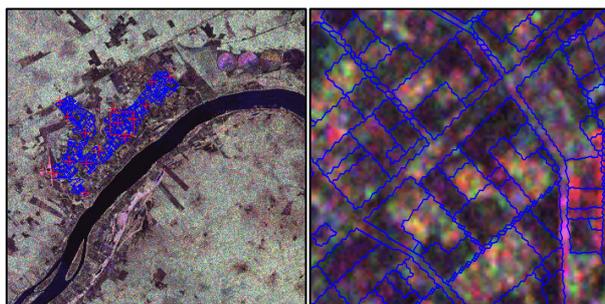


Figura 7 – Resultado do registro usando modelo de primeiro grau com interação.

A fim de avaliar o resultado do registro, foram calculados os valores de deslocamento (distâncias euclidianas) entre cada ponto de controle e seu valor estimado pelo modelo de transformação espacial. A **Figura 8** apresenta os valores de erro de estimação e de predição para cada ponto de controle. De maneira geral, pode-se notar que o modelo mais simples (que considera apenas translação, escala e rotação) não foi suficiente para modelar as distorções existentes entre as imagens analisadas. O modelo de 2^o grau foi o que apresentou os menores erros de estimação. Esse resultado já era esperado, pois este modelo faz uso de mais parâmetros permitindo uma maior flexibilidade de adaptação às distorções. No entanto, observe que este mesmo modelo apresentou valores de erro de predição bastante elevados para alguns pontos de controle, o que indica que o modelo está fortemente condicionado a esses pontos. Modelos de ordem elevada tendem a provocar distorções exageradas em locais com baixa amostragem de pontos. Finalmente, os outros dois modelos apresentaram

resultados bastante semelhantes, ou seja, a inclusão do termo de interação parece não afetar muito o resultado.

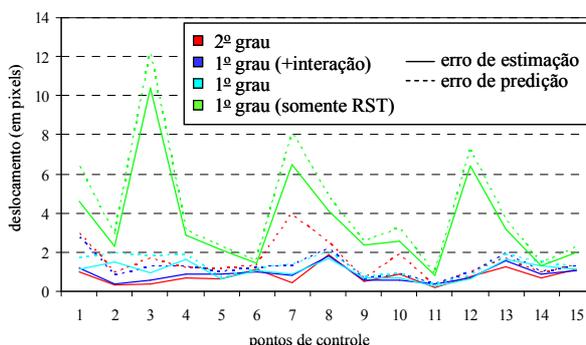


Figura 8 – Avaliação dos erros de estimação e predição dos pontos de controle.

Para complementar a análise de robustez dos modelos, foram calculados o *RMS* e o *RMSp* para cada modelo (**Tabela 1**). Como já discutido anteriormente, o *RMS* diminui a medida que aumenta-se o número de parâmetros do modelo. Assim, o modelo de 2º grau foi o que apresentou o menor valor de *RMS*. No entanto, o valor de *RMSp* para este mesmo modelo foi superior a dos modelos mais simples. Novamente, o modelo mais simples apresentou os maiores valores em ambas as medidas, corroborando a análise anterior.

Tabela 1 – Avaliação global dos modelos.

Modelo	<i>RMS</i>	<i>RMSp</i>
2º grau	0,89	1,80
1º grau (+ interação)	0,96	1,38
1º grau	1,12	1,42
1º grau (somente RST)	4,30	5,12

5. Conclusão

Este trabalho apresentou um sistema interativo de registro de imagens. Este sistema utiliza uma abordagem vetorial que facilita o processo de aquisição de pontos de controle principalmente quando as imagens apresentam baixa concordância. O sistema foi testado num conjunto de imagens mostrando-se robusto. Baseando-se numa avaliação dos erros de estimação e predição foi possível determinar o melhor modelo de transformação espacial para o conjunto testado.

Referências

- Bentoutou, Y.; Taleb, N.; Kpalma, K.; Ronsin, J. An automatic image registration for applications in remote sensing. **IEEE Transactions on Geoscience and Remote Sensing**, v. 43(9), p. 2127-2137, 2005.
- Brown, L.G. A survey of image registration techniques. **ACM Computing Surveys**, v. 24, p. 326-376, 1992.
- Neter, J.; Kutner, M.H.; Nachtsheim, C.J.; Wasserman, W. **Applied Linear Statistical Models**. Boston: McGraw-Hill, 1996. 1408p.
- Rennó, C. D. **Avaliação de medidas texturais na discriminação de classes de uso utilizando imagens SIR-C/X-SAR do perímetro irrigado de Bebedouro, Petrolina, PE**. 1995. 111 p. (INPE-10441-TDI/926). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 1995. Disponível na biblioteca digital URLib: <<http://mtc-m12.sid.inpe.br/rep-sid.inpe.br/jeferson/2004/03.05.11.22>>. Acesso em: 14 nov. 2006.
- Zitová, B.; Flusser, J. Image registration methods: a survey. **Image and Vision Computing**, v. 21, p. 977-1000, 2003.