



Boletim de Ciências Geodésicas  
ISSN: 1413-4853  
bcg\_editor@ufpr.br  
Universidade Federal do Paraná  
Brasil

GARCIA TOMMASELLI, ANTONIO MARIA; FERREIRA LOPES, RODRIGO  
REGISTRO SEMIAUTOMÁTICO ENTRE IMAGENS INFRAVERMELHAS E RGB  
COLETADAS POR UM PAR DE CÂMARAS DIGITAIS

Boletim de Ciências Geodésicas, vol. 17, núm. 1, enero-marzo, 2011, pp. 53-74  
Universidade Federal do Paraná  
Curitiba, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=393937719004>

- ▶ Como citar este artigo
- ▶ Número completo
- ▶ Mais artigos
- ▶ Home da revista no Redalyc

# **REGISTRO SEMIAUTOMÁTICO ENTRE IMAGENS INFRAVERMELHAS E RGB COLETADAS POR UM PAR DE CÂMARAS DIGITAIS**

*Semiautomatic register between infrared and RGB bands captured by a digital camera*

**ANTONIO MARIA GARCIA TOMMASELLI  
RODRIGO FERREIRA LOPES**

Programa de Pós-Graduação em Ciências Cartográficas  
Faculdade de Ciências e Tecnologia – FCT  
Universidade Estadual Paulista - UNESP  
[tomaseli@fct.unesp.br](mailto:tomaseli@fct.unesp.br)  
[eng.rodrigo.flopes@gmail.com](mailto:eng.rodrigo.flopes@gmail.com)

## **RESUMO**

A utilização de imagens adquiridas por sensores CCDs de médio formato, em plataformas aéreas, é uma alternativa para a redução de custos de projetos de aerolevantamento. Porém, a utilização de sensores que captam somente a banda do visível (RGB) restringe algumas aplicações, tornando-se relevante integrar um sensor adicional que capte o infravermelho próximo (IR). No mercado há várias soluções para a aquisição simultânea de várias bandas espectrais. Uma alternativa que possibilita uma redução nos custos de integração é coletar as bandas RGB com uma câmara e a infravermelha com uma segunda câmara. É necessário, então, registrar as imagens, o que implica em determinar um conjunto de pontos ou feições correspondentes, calcular uma função de mapeamento polinomial e reamostrar uma das imagens. Um dos problemas mais críticos é a determinação de correspondência entre as imagens, devido às diferenças radiométricas entre as imagens. Neste trabalho são utilizadas técnicas de detecção de pontos de interesse e é proposta uma função de correspondência usando as diferenças nas magnitudes e direções dos gradientes de intensidade entre as imagens RGB e IR. Foram realizados vários experimentos com a técnica proposta, indicando que é possível utilizar esta técnica obtendo-se erros residuais inferiores a 1,5 pixel.

**Palavras Chaves:** Correlação entre Imagens RGB e IR; Registro de Imagens; Resolução Espectral.

### ABSTRACT

Using images acquired by medium format CCD sensors from aerial platforms is an alternative to reduce costs of aerial surveying projects. However, using sensors that capture only the visible band (RGB) it restricts some applications, making the integration of an additional sensor that captures the near infrared (IR) important. There are several commercial solutions for the simultaneous acquisition of multiple spectral bands. A cost effective alternative is the integration of a RGB camera to a second one, to capture the infrared band. It is necessary, then, to register the images, which involve the determination of a set of corresponding points or features, to calculate a polynomial mapping function and to resample one of the images. One of the most critical problems is the determination of corresponding points between the images due to differences in the images radiometry. Some techniques for locating interest points and a correlation function using the differences in magnitudes and directions of the gray levels gradients RGB and IR images, are proposed in this paper. Several experiments using this technique were conducted, indicating that it is possible to use it, achieving residual errors less than 1.5 pixel.

**Keywords:** Correspondence between RGB and IR Images; Image Registration; RGB and IR Registration.

## 1. INTRODUÇÃO

A utilização de câmaras digitais para aquisição de informações espaciais é tendência reconhecida devido a várias vantagens, como a disponibilidade imediata da informação e a redução de custos devidos à eliminação do suporte físico (filme) e das etapas posteriores (revelação e digitalização). Os sistemas com somente um sensor (CCD ou CMOS), têm a desvantagem de captar uma área no terreno com dimensões menores do que aquela que seria registrada por uma câmara de quadro convencional (23x23cm).

Para resolver esta limitação, os fabricantes de câmara recorrem à integração de múltiplos sensores digitais em uma única plataforma. Este tipo de configuração permite um recobrimento maior da área fotografada e quando são utilizados sensores com diferentes características espectrais, há um ganho também na resolução espectral. Podem ser encontrados no mercado diversos sistemas de câmaras aerofotogramétricas digitais que exploram esta potencialidade, como por exemplo, os sistemas de câmaras de sensores CCD lineares como: as câmaras ADS-40 (Leica Geosystems - Hexagon), HRSC-A, HRSC-AX e HRSC-AXW (desenvolvidas pelo centro Alemão de Pesquisas Espaciais - DLR), e os sistemas de sensores com múltiplos CCD matriciais como: o sistema DMC (Z/I Imaging), UltracamX (Microsoft Vexcel), Rolleimetric/Trimble (AICx4) e a DIMAC (Dimac Systems). Por se tratar de sistemas que utilizam câmaras digitais métricas integradas

a sensores de orientação direta de alta precisão, os custos destes sistemas são elevados.

Uma alternativa que vem sendo adotada é a utilização de câmaras digitais convencionais de médio formato. Contudo, estas câmaras não foram originalmente projetadas para propósitos fotogramétricos, e podem apresentar instabilidades em seus parâmetros de orientação interior (POI), embora experimentos realizados por Habib e Morgan (2003) e Machado et al (2003), tenham mostrado que esta pequena instabilidade dos POI não afeta significativamente os resultados. A ausência da banda do infravermelho próximo pode ser suprida usando-se uma segunda câmara, adaptada para captar somente esta porção do espectro eletromagnético. Com a informação adicional da câmara IR (abreviatura de *infrared*, infravermelho, que será usada ao longo do texto), torna-se possível uma melhor identificação de culturas, especificamente realçando o contraste entre cultura/solo e terra/água.

As imagens IR, em conjunto com as imagens RGB, contribuem para o mapeamento de ocorrências específicas, como: monitoramento de cobertura do solo e de culturas agrícolas (pragas, doenças, plantas daninhas), para gerenciamento localizado em sistemas de produção agrícola (agricultura de precisão); estudos ambientais (avaliação de impactos locais); levantamentos florestais de vegetação mais detalhados, estudos de ocupação urbana, localização de depósitos de resíduos sólidos, entre outras. Em síntese, a banda do IR aumenta o potencial para a discriminação de alvos em conjunto com imagens RGB.

Para utilizar este potencial, contudo, é necessário registrar a imagem IR em relação à imagem RGB e, para isto, são necessários vários procedimentos. Uma das etapas críticas é a determinação de pontos correspondentes, o que pode ser feito visualmente por um operador humano, mas isto implica em abdicar da automação do processo. Em projetos de mapeamento, são adquiridos grandes conjuntos de imagem e há necessidade de automatizar estas etapas (ou criar processos semiautomáticos, com pequena intervenção do operador), sob pena de tornar o processo inviável.

Neste sentido, o objetivo deste trabalho é apresentar o desenvolvimento de uma técnica semiautomática para o registro entre imagens IR e RGB, adquiridas por câmaras de médio formato, com um mínimo de interferência do operador. Esta técnica é baseada na determinação preliminar de pontos distinguíveis nas imagens RGB e comparação entre as magnitudes e direções dos gradientes dos tons de cinza em janelas baseadas nestes pontos com janelas de busca na imagem IR.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 Sistema de aquisição de imagens com múltiplas câmaras

A possibilidade de integrar câmaras de médio formato para aquisição de imagens aéreas foi explorada para o desenvolvimento do “Sistema Digital Aerotransportado de Aquisição e Pós-Processamento de Imagens” (SAAPI). Além

de câmaras digitais, o SAAPI incorpora também um sistema de georreferenciamento direto, permitindo a coleta de imagens georreferenciadas RGB e IR (ENGEMAP, 2009). Existem várias configurações possíveis, com duas (uma RGB e outra IR) ou três câmaras (duas câmaras RGB oblíquas e uma IR nadiral). Neste trabalho serão utilizadas imagens obtidas no modo dual-nadiral, ou seja, uma câmara RGB e uma câmara IR, ambas orientadas aproximadamente na vertical. No caso em estudo foram usadas imagens de uma câmara Hasselblad H2D (22 Megapixels) e uma câmara Sony F828, adaptada para captar o infravermelho (ver Tabela 1 para especificações de ambas as câmaras). Atualmente o sistema é composto por duas câmaras RGB do modelo HASSELBLAD H3D de 50 Megapixels, uma das quais adaptada com um filtro IR.

A principal vantagem desse sistema é sua flexibilidade, podendo variar as combinações entre os sensores.

Tabela 1 - Especificações dos modelos de câmaras.

<i>Câmaras</i>	<i>HASSELBLAD H2D</i>	<i>SONY F828</i>
<i>Sensor</i>	CCD – 49,0 x 36,7 mm	CCD – 8,8 x 6,6 mm
<i>Resolução</i>	5440 x 4080 pixels (22 MP)	3264 x 2448 pixels (8 MP)
<i>Dimensão do pixel (mm)</i>	0,009	0,0027
<i>Distância focal nominal (mm)</i>	50	7,35

As imagens IR e RGB, coletadas com um sistema dual, como o especificado na Tabela 1, foram tomadas a partir de posições espaciais e orientações diferentes, embora com diferenças muito pequenas. Além disto, as distorções das lentes e demais parâmetros de orientação interior (POI) também são diferentes, bem como a resolução geométrica das imagens, o que significa que pontos correspondentes nas duas imagens terão diferentes coordenadas, o que impede o uso da imagem IR bruta como um canal adicional da imagem RGB. Para que isto seja possível, é necessário registrar as estas imagens, o que implica em calcular uma função de mapeamento entre as imagens e realizar a reamostragem de uma imagem em relação à outra.

No caso em estudo, a resolução geométrica da imagem IR é inferior à da imagem RGB, e, por esta razão, optou-se por registrar a imagem IR tendo como referência a imagem RGB, ou seja a imagem IR será reamostrada passando a ter a mesma resolução geométrica da imagem RGB, e as mesmas distorções.

## 2.2 Registro de imagens

*“Registro de imagens é o processo de sobreposição (overlaying) de imagens da mesma cena, tomadas em instantes diferentes, de diferentes pontos de vista e/ou por diferentes sensores”* (ZITOVA e FLUSSER, 2003, p. 977, tradução nossa).

Embora alguns autores, como Fonseca e Manjunath (1996), definam o registro como o “casamento” ou correspondência entre as imagens (*matching*), é importante

ressaltar que, nas comunidades fotogramétrica e de visão computacional, a correspondência se refere a uma das etapas do registro (e de outros processos). Neste sentido, o registro de imagens, de acordo com Fedorov (2002), é realizado em três etapas.

- 1) *Obtenção de pontos de controle;*
- 2) *Determinação da função de transformação;*
- 3) *Sobreposição das imagens*” (Fedorov, 2002, p. 21)

Zitova e Flusser (2003) apresentam várias abordagens para a solução do problema de registro, classificando-as de acordo com a natureza da técnica de correspondência e com os procedimentos usados nas quatro etapas do registro: “[...]detecção de feições, correspondência, definição da função de mapeamento e transformação e reamostragem.” (ZITOVA e FLUSSER, 2003). Estas etapas são bastante similares às definidas por Fonseca e Manjunath (1996), embora com denominações diferentes. As duas primeiras etapas estão relacionadas ao problema de correspondência, ou à obtenção dos pontos de controle, como classificam Fedorov (2002).

Existem aplicações de registro de imagens em várias áreas, como Sensoriamento Remoto, Visão de Máquina e diagnóstico médico por imagens.

Diferentes técnicas podem ser empregadas em cada uma das etapas de registro, sendo que as duas primeiras fases são as que apresentam grande número de variações. O problema de correspondência será tratado na seção 2.3, e apresenta maior grau de complexidade quando há diferenças geométricas e radiométricas significativas entre as imagens a serem registradas. Este é o caso em estudo, no qual há diferenças de posição e orientação das câmaras, embora pequenas, e diferenças nas bandas espectrais registradas pelos sensores, o que causa diferenças radiométricas substanciais entre a imagem RGB e IR (ver Figuras 5.a e 5.b). A literatura traz algumas soluções para o registro de imagens coletadas por diferentes sensores. Dare e Dowman (2001) desenvolveram uma técnica para o registro automático de imagens SAR e SPOT que detecta diferentes tipos de feições em ambas as imagens, principalmente grandes segmentos de área e, posteriormente, determina as correspondências entre estes elementos, a partir dos quais são calculados pontos de enlace de alta precisão. Devido às grandes diferenças radiométricas entre estas imagens, faz-se, inicialmente, um alinhamento inicial entre as imagens, usando pontos de enlace selecionados manualmente, para, em seguida, calcular um registro aproximado entre as imagens a partir de segmentos de área (*patches*). Finalmente, o registro preciso é calculado a partir das correspondências entre elementos de bordas.

Hrkać et al (2007) apresentaram uma técnica para o registro de imagens RGB e IR, baseada no mapeamento de nuvens de pontos extraídas por um detector de cantos aplicado à ambas as imagens. Várias transformações foram testadas para o processo de registro, sendo utilizada a distância de *Hausdorff* como medida de

similaridade entre os cantos na imagem RGB e os cantos transformados a partir da imagem IR.

Santos Jr. et al (2007) desenvolveram uma técnica de orientação de imagens RGB, tomadas por câmaras digitais de pequeno formato, utilizando imagens de intensidade de um Sistema Lazer Aerotransportado, imagens estas que estão na banda do infravermelho (IR). A técnica baseia-se na detecção de cantos e bordas e na determinação de correspondência por correlação baseada em áreas.

Como se conclui dos trabalhos revisados a diferença radiométrica entre as imagens deve ser contornada através de uma técnica de correspondência que utilize características, ao invés do sinal original das imagens.

## 2.3 Correspondência de imagens

O processo de correspondência ou “casamento” (*matching*) é um dos problemas fundamentais no processo de automação em Fotogrametria (HEIPKE, 1997).

Heipke (1997) define *matching* como o procedimento para o estabelecimento de correspondências entre vários conjuntos de dados, desde que estes descrevam, pelos menos parcialmente, a mesma cena.

De acordo com Zhang et al. (1995) os algoritmos de correspondência podem ser classificados, de uma forma geral, em duas categorias: correspondência baseada em áreas ou sinal (*area based matching*) e correspondência baseada em feições ou características (*feature based matching*):

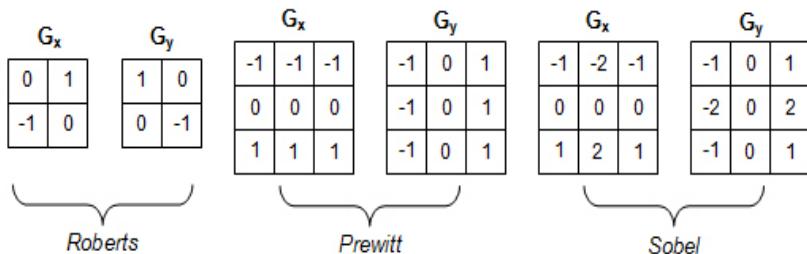
- correspondência baseada em área: é associada à correspondência dos níveis de cinza das imagens. Nesta categoria trabalha-se com janelas onde são realizadas as correspondências, ou seja, são comparados os níveis de cinza de pequenas áreas nas duas imagens (janela de referência e janela de busca ou pesquisa, de maior dimensão) e a similaridade é medida por uma função de correlação estatística;
- correspondência baseada em feições: nos métodos baseados em feições deve-se, inicialmente, extrair primitivas das imagens. Estas primitivas podem ser locais, tais como pontos, bordas, pequenas linhas e regiões, ou globais, como polígonos ou estruturas. Cada feição ou característica reúne um conjunto de atributos, tais como orientação e magnitude das bordas (gradientes), posições (coordenadas), comprimento e curvatura de linhas, entre outros. Além dos atributos, podem ser estabelecidas relações entre as feições, como por exemplo, relações geométricas (como ângulos), radiométricas ou ainda topológicas.

A técnica de correspondência empregada neste trabalho combina algoritmos dos dois tipos, extraíndo-se preliminarmente os gradientes das imagens que são comparados por áreas, usando uma função similar à função erro (HEIPKE, 1997).

## 2.4 Detecção de bordas

Segundo Gonzalez e Woods (1992), borda é o limite entre duas regiões com propriedades relativamente distintas nos níveis de cinza. A detecção de bordas em uma imagem digital baseia-se no cálculo de um operador diferencial local, como os de Roberts, Prewitt e Sobel. Na Figura 1 são mostradas algumas máscaras cuja convolução com a imagem em tons de cinza permite o cálculo dos gradientes em x e y.

Figura 1 - Máscaras de operadores.



A magnitude do gradiente direcional pode ser calculada pela expressão:

$$\nabla f = \sqrt{G_x^2 + G_y^2} \quad (1)$$

E a direção do gradiente por:

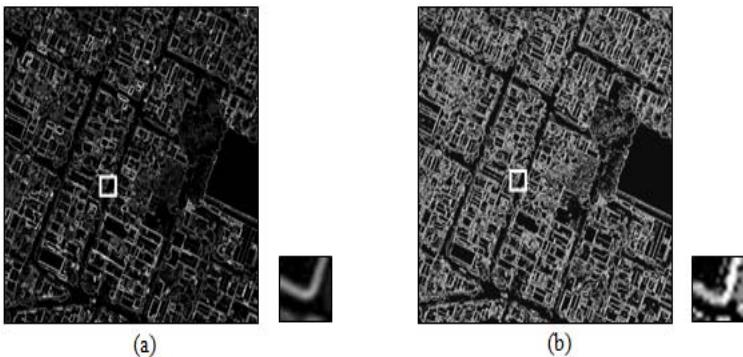
$$\alpha(x, y) = \tan^{-1} \left( \frac{G_y}{G_x} \right) \quad (2)$$

onde  $G_x$  e  $G_y$ , são os gradientes direcionais coluna e linha respectivamente, calculado, por exemplo, com uma das máscaras mostradas na Figura 1.

As magnitudes de direções dos gradientes podem ser quantizadas em valores digitais no intervalo de 0 a 255 (8-bits) e armazenadas como imagens. As Figuras 2(a) e 2(b) ilustram estes tipos de imagens dos gradientes. Nestas imagens foram eliminadas as ocorrências de elementos com magnitude menor que um certo limiar.

Observou-se que as imagens de gradientes e direções, obtidas a partir das imagens (IR e RGB) são similares e podem ser utilizadas como elementos para a comparação entre as imagens. Esta alternativa será apresentada na Seção 3.

Figura 2 - Representação da imagem RGB: (a) Imagem de magnitude do gradiente das bordas (b) Imagem de direção do gradiente das bordas.



Visualmente as imagens são bastante similares, porém a imagem (Figura 2(b)) gerada pelos valores do gradiente de direções das bordas, contém maior quantidade de ruídos e variações indesejadas, o que podem gerar confusões na etapa de comparação entre as imagens (IR e RGB). Isto se deve à incerteza no cálculo das direções a partir de uma janela de apenas 3x3 pixels.

## 2.5 Identificação de pontos de interesse

Esta etapa envolve a localização, seleção e posicionamento de pontos considerados ótimos para correlação entre imagens (pontos distinguíveis, ou pontos de interesse). Estão disponíveis na literatura diversos tipos de operadores, como por exemplo: o operador de Moravec, Förstner, Harris, entre outros (MORAVEC, 1979; FÖRSTNER, 1986; HARRIS e STEPHENS, 1988).

Esses pontos podem ser identificados através da análise da matriz variância covariância de translação da janela que envolve o ponto de interesse (FÖRSTNER, 1986, FÖRSTNER, 1993 e COSTA et al., 2005).

O objetivo desta análise prévia é estimar a qualidade da correlação entre imagens, antes mesmo de realizá-la, a partir dos gradientes dos tons de cinza em uma janela que será usada como referência para a correspondência. Förstner (1993) propôs o seguinte procedimento:

- Cálculo da variância do ruído ( $\sigma_n^2$ ) em função da variância da janela de referência ( $\sigma_f^2$ ) e do máximo coeficiente de correlação ( $\rho_{12}$ );
- Cálculo da matriz de equações normais ( $N$ ) em função dos gradientes nas linhas e colunas da janela de referência;
- Cálculo da precisão dos parâmetros de translação entre as imagens em função da variância do ruído ( $\sigma_n^2$ ) e da matriz de equações normais ( $N$ ).

### 2.5.1 Calculo da Variância do Ruído

De acordo com Förstner (1993) e Costa et al. (2005), ao considerar-se duas imagens não transladas, não rotacionadas e de mesma escala ( $g_1$  e  $g_2$ ), tem-se:

$$g_1(x, y) = f(x, y) + n_1(x, y) \quad (3)$$

$$g_2(x, y) = a[f(x, y) + n_2(x, y)] + b \quad (4)$$

onde  $n_1$  e  $n_2$  são ruídos brancos e independentes, com variância  $\sigma_{n_1}^2 = \sigma_{n_2}^2 = \sigma_n^2$ ,  $f$  é uma função estocástica com variância  $\sigma_f^2$  e independente de  $n_1$  e  $n_2$ ,  $a$  é o fator de escala e  $b$  é a translação radiométrica dos tons de cinza. Ao se realizar uma propagação de covariâncias nas Equações (3) e (4), obtém-se as seguintes variâncias:

$$\sigma_{g_1}^2 = \sigma_f^2 + \sigma_n^2 \quad (5)$$

$$\sigma_{g_2}^2 = a^2(\sigma_f^2 + \sigma_n^2) \quad (6)$$

e a covariância é dada por:

$$\sigma_{g_1 g_2} = a\sigma_f^2 \quad (7)$$

A partir destas grandezas o coeficiente de correlação entre as imagens pode ser estimado pela Equação (8):

$$\rho_{12} = \frac{\sigma_{g_1 g_2}}{\sigma_{g_1} \cdot \sigma_{g_2}} = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_n^2} \quad (8)$$

Conhecendo a variância  $\sigma_f^2$  do sinal observado pode-se obter a variância do ruído:

$$\sigma_n^2 = \sigma_{g_1}^2(1 - \rho_{12}) = \frac{\sigma_f^2(1 - \rho_{12})}{\rho_{12}} \quad (9)$$

O coeficiente de correlação pode ser geometricamente interpretado como sendo o cosseno do ângulo entre os vetores  $g_1$  e  $g_2$  (COSTA et al., 2005).

### 2.5.2 Estimativa dos Parâmetros de Translação

Segundo Costa et al. (2005) ao se considerar apenas uma translação entre duas janelas, o modelo dessa translação pode ser dado por:

$$\begin{aligned} p &= x - \Delta x \\ q &= y - \Delta y \end{aligned} \quad (10)$$

Linearizando esse modelo tem-se que:

$$\Delta g_i = -f_{r_i}(\Delta x - \Delta x_0) - f_{c_i}(\Delta y - \Delta y_0) + n_i \quad (11)$$

sendo:

$f_r$  e  $f_c$  - gradientes dos tons de cinza nas direções das linhas (*row*) e colunas, respectivamente;

$\Delta g_i = g_i(x, y) - g_i(x - \Delta x, y - \Delta y)$ , diferença entre os tons de cinza, para um determinado pixel  $i$ ;

Os parâmetros ( $\Delta x$ ,  $\Delta y$ ) bem como sua matriz variância e covariância podem ser estimados pelo Método dos Mínimos Quadrados (MMQ), utilizando-se a Equação (11) como equação de observação, com  $\Delta g_i$  sendo a observação e as incógnitas ( $\Delta x$ ,  $\Delta y$ ). As equações normais desse modelo são dadas por:

$$N \cdot X = U \quad (12)$$

$$X = \begin{pmatrix} \Delta x - \Delta x_0 \\ \Delta y - \Delta y_0 \end{pmatrix} \quad (13)$$

A matriz  $N$  das equações normais, composta pelos gradientes, é dada por:

$$N = \begin{pmatrix} \sum_{i=1}^m f_{r_i}^2 & \sum_{i=1}^m f_{r_i} f_{c_i} \\ \sum_{i=1}^m f_{r_i} f_{c_i} & \sum_{i=1}^m f_{c_i}^2 \end{pmatrix} \quad (14)$$

Utilizando a mesma notação anterior, Harris e Stephens (1988) propuseram um operador que identifique pontos potencialmente como cantos, com os seguintes elementos:

$$\begin{aligned} C(x, y) &= N / -k(\text{tr}(N))^2 & (15) \\ |N| &= AB - C^2 & (\text{determinante da matriz normal}) \\ \text{tr}(N) &= A + B & (\text{traço da matriz normal}) \\ k &= \text{const.} \end{aligned}$$

onde:

$$N = \begin{pmatrix} A & C \\ C & B \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m f_i^2 & \sum_{i=1}^m f_i f_{c_i} \\ \sum_{i=1}^m f_i f_{c_i} & \sum_{i=1}^m f_{c_i}^2 \end{pmatrix} \quad (16)$$

Este operador também baseia-se na Matriz Variância Covariância (MVC) das translações, e consequentemente, nos desvios-padrão das translações.

## 2.6 Reamostragem de imagens

A reamostragem está relacionada com o processo de interpolação de valores de brilho em uma imagem digital quando estas são submetidas a transformações geométricas (ZITOVA e FLUSSER, 2003). Como exemplos de aplicações deste processo pode-se citar: transformações geométricas básicas (rotações, por exemplo), registro de imagens, retificação de imagens, produção de ortofotos digitais, entre outros. Há dois métodos de reamostragem: o método direto e o método indireto.

O método direto consiste em transferir os valores de brilho dos píxeis da imagem original para uma matriz correspondente à imagem reamostrada, usando a transformação direta.

No método indireto a posição de cada pixel da imagem reamostrada é transformada para o sistema da matriz da imagem original. Como a posição deste pixel na imagem reamostrada será um valor fracionário, pode-se adotar o valor do pixel mais próximo ou, então, realizar uma interpolação a partir dos valores de brilho dos pixels vizinhos.

O método do vizinho mais próximo, não implica em interpolação, pois toma o valor de brilho do pixel mais próximo (GONZALEZ e WOODS, 1992). Esta aproximação do valor do pixel pode ocasionar a disjunção das feições na imagem reamostrada. Existem métodos mais sofisticados para a interpolação do pixel na imagem reamostrada, como por exemplo, a interpolação bilinear e a convolução bicúbica.

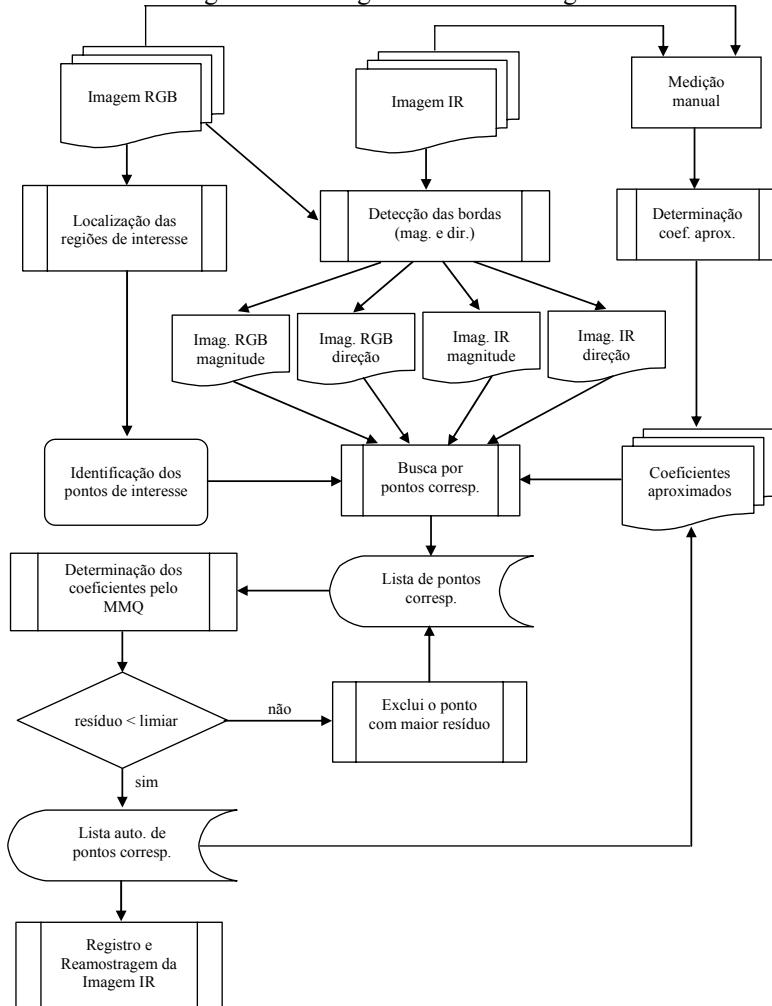
A interpolação bilinear utiliza os valores de brilho dos quatro píxeis mais próximos ao pixel transformado, realizando-se duas interpolações lineares (GONZALEZ e WOODS, 1992). Já na interpolação bicúbica, o valor de brilho do pixel na imagem reamostrada sofre influência dos 16 píxeis vizinhos. Esta ponderação resulta numa ligeira suavização da imagem resultante, maior que a observada na interpolação bilinear.

## 3. MATERIAIS E MÉTODOS

A técnica desenvolvida e avaliada neste trabalho é adequada para o registro de imagens tomadas com o par de câmaras em posições e orientações muito próximas. No caso em estudo, a distância entre as câmaras é menor do que 12cm e, sob estas condições, as variações geométricas causadas pelo deslocamento do relevo podem ser consideradas desprezíveis em uma pequena vizinhança. Uma das hipóteses a

serem testadas neste trabalho, é a de que as diferenças causadas pelas distorções das lentes e pelas pequenas diferenças de posição e orientação podem ser modeladas por uma função polinomial de 2<sup>a</sup> ordem. Esta hipótese é baseada nos modelos de distorção das lentes, que também são funções polinomiais. Para o registro das imagens das imagens, são propostas as etapas descritas nas próximas seções e sintetizadas no fluxograma apresentado na Figura 3.

Figura 3 - Fluxograma da metodologia.



### 3.1 Preparação dos dados

No caso estudado, a imagem IR, adquirida com o sensor da câmara SONY F-828 de 8 Megapixeis, apresenta resolução geométrica inferior à imagem RGB, capturada com uma câmara Hasselblad. Para estabelecer uma correspondência preliminar, a imagem IR foi reamostrada por interpolação bilinear, para uma resolução equivalente de 5440 x 4080 píxeis, que corresponde à resolução do sensor da câmara HASSELBLAD H2D, de 22 Megapixeis.

Em seguida fez-se a medida manual (interativa) de um conjunto de pontos correspondentes em ambas as imagens, cujo objetivo é definir uma função de mapeamento aproximada para reduzir o espaço de busca e evitar falsas correspondências. Esta lista é gerada manualmente apenas para o primeiro par de imagens em um projeto. Considerando-se que a montagem das câmaras é estável, para os demais pares RGB-IR, pode-se usar os parâmetros de mapeamento determinados automaticamente para as demais imagens.

Embora o método não requeira os parâmetros de orientação interior, caso estes dados estejam disponíveis, as imagens podem ser corrigidas de seus erros sistemáticos, o que otimizará o processo de busca por pontos correspondentes, devido à redução de distorções das lentes.

No processo em questão, a imagem captada pelo sensor do espectro do visível (RGB) é utilizada como imagem base para a seleção dos pontos de referência, por apresentar maior riqueza espectral. Essa característica da imagem RGB auxiliará a seleção da melhor janela de referência para a correlação dos pontos.

### 3.2 Cálculo dos gradientes e seleção dos pontos de interesse

A etapa de cálculo dos gradientes foi executada independente do processo de registro entre imagens, sendo posteriormente utilizados nas etapas de seleção de pontos de interesse e comparação das imagens.

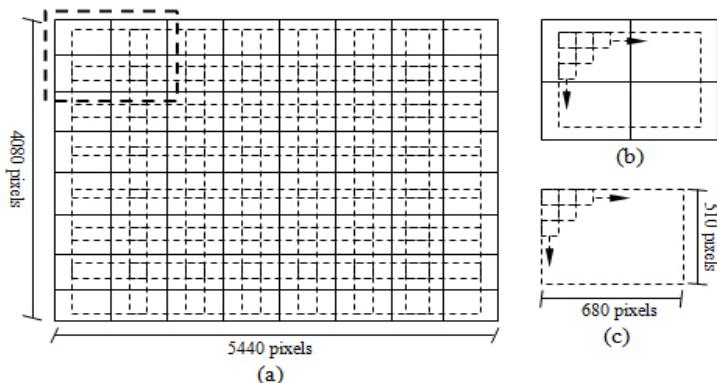
A estratégia de distribuição de pontos é escolhida pelo operador nesta etapa, sendo possível variar a quantidade de pontos que serão distribuídos pela imagem, aumentando-se ou diminuindo o espaçamento entre pontos da grade. Define-se uma grade de sub-imagens e, dentro de cada sub-imagem, a posição dos pontos de interesse é encontrada estratégia de detecção de cantos de Harris e pela análise da matriz variância covariância de translações da janela de análise.

Em cada sub-imagem selecionada é definida uma área útil que será percorrida por uma janela de análise. Em cada posição da janela de análise é realizado o teste de detecção do ponto de interesse e, posteriormente, a análise da matriz variância covariância da translação desta janela. Para exemplificar esta etapa a Figura 4(a) ilustra uma distribuição de 49 pontos (7 linhas por 7 colunas de sub-imagens), em que, para cada sub-imagem, é definida uma área útil de 680 x 510 píxeis (janela de busca), ilustrada na Figura 4(c).

A estratégia de identificação de pontos de interesse aplicada utiliza o operador de Harris. Ela tem a característica de identificar, pelas direções dos gradientes da

imagem, pontos distingíveis. Após a identificação dos possíveis pontos de interesse na sub-imagem é realizada uma análise na matriz variância covariância de todas as janelas de referência identificadas como pontos de interesse, a fim de se selecionar a melhor janela para a correspondência. Por meio da janela de análise, é definida a região ótima para se realizar a correlação entre a imagem RGB e IR. Para isso, são comparados os traços e as variâncias da matriz variância covariância das translações de cada janela. O ponto selecionado será aquele que apresentar menor traço e maior variância.

Figura 4 - Etapas da seleção dos pontos de interesse: (a) Imagem dividida em sub-imagens, (b) Detalhe da sub-imagem, (c) Área útil da busca pelo ponto de referência.



No processo em questão foi estabelecido um limiar igual a 200, isto é, valores calculados pelo algoritmo de Harris ( $C(x,y)$ , Equação 15) menores que 200 não foram considerados. O valor para a constante  $k$  foi adotado como igual 0,05. Esses valores foram definidos empiricamente, após testes com as imagens reais. Valores pequenos para  $C(x,y)$  indicam que os gradientes nas direções das linhas e colunas apresentam pequena magnitude, ou seja, são regiões homogêneas.

A identificação de pontos de interesse por este algoritmo, portanto, exclui regiões homogêneas, como por exemplo, em áreas de pastagens, campos e em grandes extensões de águas, como, lagos e represas, nas quais os valores dos gradientes bidirecionais são muito baixos.

### **3.3 Definição de parâmetros aproximados para a função de mapeamento polinomial**

Partindo do princípio que o par de câmeras sofre pequenas variações durante a aquisição das imagens consecutivas (por exemplo, ao longo de uma faixa ou de um bloco de imagens em um levantamento aéreo), um conjunto de pontos

correspondentes é medido manualmente e, a partir destas correspondências são calculados os coeficientes aproximados do polinômio de mapeamento (Equação 18, na seção 3.5) entre as imagens (IR e RGB), sendo estimados pelo Método dos Mínimos Quadrados (MMQ). Com estes coeficientes aproximados faz-se o mapeamento de cada ponto de interesse selecionado para a imagem IR, na qual será definida a janela de busca. Esse procedimento reduz o tempo de processamento e minimiza falsas correspondências entre as imagens.

A imagem RGB é definida como referência, com uma dimensão da janela de referência (ponto de interesse definido no passo anterior) para cada ponto, de 31 x 31 píxeis. Já a dimensão da janela na imagem de pesquisa (IR) é de 61 x 61 píxeis, o que possibilita 961 comparações entre as janelas de referência e pesquisa. A posição central desta janela é definida pela função de mapeamento aproximada. Não há necessidade de uma janela de pesquisa maior, devido à boa aproximação dos coeficientes do polinômio de mapeamento, determinados com a lista manual de pontos. Os tamanhos das janelas de pesquisa e referência foram determinados empiricamente e podem ser aumentados. Como as diferenças geométricas entre as imagens são pequenas, não há inconveniente em aumentar o tamanho da janela o que certamente reduz o número de falsas correspondências, embora com aumento do custo computacional.

### **3.4 Correspondência com a função erro ponderada entre a Magnitude e a Direção do Gradiente**

Para a identificação da posição de melhor correspondência entre as janelas de referência e pesquisa, são comparados os valores de magnitude e direção dos gradientes nas imagens RGB e IR. O quociente ( $\Delta d$ ) expressa a média ponderada, das discrepâncias entre magnitudes e direções dos gradientes nas imagens RGB e IR, e é descrito pela Equação (17):

$$\Delta d = \frac{w_1 \sum (g_{IR} - g_{RGB}) + w_2 \sum (g_{IR}^d - g_{RGB}^d)}{w_1 + w_2} \quad (17)$$

As variáveis  $g_{IR}$  e  $g_{RGB}$  correspondem às magnitudes dos gradientes (normalizados para o intervalo 0-255) nas imagens IR e RGB, respectivamente e as variáveis  $g_{IR}^d$  e  $g_{RGB}^d$  são as direções das bordas, também normalizadas para o intervalo 0-255;  $w_1$  e  $w_2$  são os pesos atribuídos à discrepâncias entre magnitude e direção respectivamente. A posição selecionada é aquela que apresenta a menor diferença.

### **3.5 Ajuste da função polinomial e análise dos resíduos**

Uma vez estabelecidos os pontos correspondentes entre as imagens, pode-se determinar os parâmetros da função polinomial de mapeamento, que permite

projetar as coordenadas da imagem RGB para a imagem IR. A função de mapeamento utilizada é um polinômio de 2<sup>a</sup> ordem, que permite compensar as diferenças de distorções e pequenas diferenças de orientação entre as câmaras.

$$\begin{aligned}x' &= a_0 + a_1 x'' + a_2 x''^2 + a_3 y'' + a_4 y''^2 + a_5 x'' y'' + a_6 x''^2 y'' \\y' &= b_0 + b_1 x'' + b_2 x''^2 + b_3 y'' + b_4 y''^2 + b_5 x'' y'' + b_6 x''^2 y''\end{aligned}\quad (18)$$

Nas Equações (18), as variáveis ( $x'$ ,  $y'$ ) e ( $x''$ ,  $y''$ ) caracterizam as posições correspondentes nas imagens (RGB e IR, respectivamente), e  $a_0, a_1, \dots, a_6, b_0, b_1, \dots, b_6$  são os 14 coeficientes do polinômio. A escolha do polinômio se deve ao grau de complexidade de transformações necessárias para o melhor registro entre as imagens e foi também realizada empiricamente. Fedorov (2001), utilizou um polinômio de segundo grau semelhante ao apresentado, mas com 12 parâmetros, sem os termos em  $a_6$  e  $b_6$ .

Alguns dos pontos podem ser determinados em posições incorretas pelo processo de correspondência. É o caso de pontos localizados em regiões homogêneas, onde há maior possibilidade de falsas correspondências. A eliminação destes pontos é realizada recursivamente pela análise dos resíduos do ajustamento do polinômio de mapeamento entre as imagens IR e RGB. Os pontos que tiverem um vetor resultante dos resíduos acima de um limiar pré-estabelecido são eliminados sequencialmente. Parte-se, inicialmente, do ponto de maior resíduo, que é eliminado, repetindo-se o cálculo dos parâmetros. Este processo é realizado iterativamente até que a condição estabelecida seja satisfeita, ou seja, que o maior resíduo seja menor que o limite estabelecido. Após esta análise a lista de pontos gerados automaticamente é salva, bem como os valores dos coeficientes ajustados.

Pode-se, então gerar uma imagem reamostrada, que é gerada pixel a pixel, sendo seu valor de brilho interpolado da imagem IR, por interpolação bilinear.

Os parâmetros aproximados da função de mapeamento, determinados a partir de uma lista de pontos medidos manualmente, são utilizados apenas para o registro do primeiro par de imagens (IR-RGB). Considerando um bloco de imagens ou uma faixa de um voo, os parâmetros calculados automaticamente, com a lista de pontos determinados pelo processo de correspondência proposto, podem ser usados como aproximação para o próximo par.

#### 4. EXPERIMENTOS E RESULTADOS

O par de imagens escolhido para o experimento mescla a diversidade de elementos que podem ser encontrados em uma imagem aérea convencional, como por exemplo, a presença de mancha urbana com campo de pastagens e plantações. As imagens exibidas na Figura 5 cobrem parte do município de Tarumã (SP), com um GSD de 30cm e 40cm (respectivamente, RGB e IR) e uma altura de voo de aproximadamente 1620m.

Figura 5 - Imagens aéreas (a) Imagem RGB (b) Imagem IR.



O processo do registro entre imagens IR e RGB desenvolvido, permite diversas variações. Para testar o processo metodológico foram realizados 8 experimentos. Em cada experimento foi realizada uma combinação de técnicas distintas: realizando ou não da retificação das imagens; corrigindo-se ou não a distorção das imagens; alterando-se a quantidade de pontos distribuídos pela imagem e alterando-se o peso dado à magnitude e à direção dos gradientes das imagens, ao se realizar o processo de correspondência. A Tabela 2 sintetiza as características dos experimentos.

Tabela 2 - Características dos experimentos realizados.

Exp.	Retificação	Corr. da Distorção	Nº de pontos	Tolerância do Resíduo (limiar)	Peso Magnitude	Peso Direção
A	N	N	100	1,5	2	1
B	N	N	169	1,5	2	1
C	N	N	400	1,5	2	1
D	N	S	169	1,5	2	1
E	S	S	169	1,5	2	1
F	S	N	169	1,5	2	1
G	S	N	169	1,5	1	1
H	S	N	169	1,5	3	1

Os pesos aplicados às direções foram sempre estabelecidos como a unidade e menores que os pesos das magnitudes, porque a informação de direção é menos precisa em janelas pequenas e também porque janelas distintas com diferentes tons de cinza e diferentes magnitudes dos gradientes podem apresentar mesmas direções (por exemplo, duas rodovias paralelas).

Para cada experimento realizou-se uma estatística dos resultados obtidos, destacando-se: o número de pontos selecionados após a análise dos resíduos e

eliminação de pontos; o sigma a posteriori ( $\hat{\sigma}_0$ ) do ajuste pelo MMQ dos pontos selecionados e a Média da Resultante do Resíduo (MRR) em 9 pontos de verificação medidos nas imagens (após o registro). A Tabela 3 exibe os resultados. Os 9 pontos de verificação foram medidos interativamente, com uma distribuição similar à dos pontos de Grüber.

Tabela 3 - Resultados obtidos.

Exp.	Nº de pontos remanescentes após a eliminação	$\hat{\sigma}_0$ (pixel)	MRR em 9 pontos de verificação (pixel)
A	77	0,71	1,54
B	125	0,65	1,90
C	262	0,59	2,08
D	78	0,68	9,86
E	83	0,70	10,22
F	123	0,65	1,64
G	123	0,66	1,82
H	122	0,65	1,88

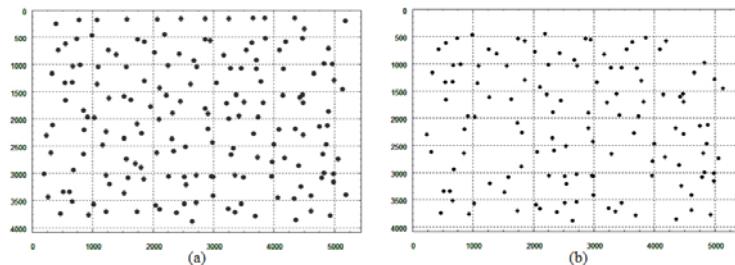
Tomando-se como referência para análise de precisão os valores das médias das resultantes dos resíduos (MRR), verifica-se que os experimentos D e E foram que apresentaram piores resultados. Embora a distribuição de pontos remanescentes tenha sido homogênea pela imagem, os dados utilizados para a correção prévia de distorção das lentes não eram suficientemente acurados, pois não se dispunha de certificados de calibração para o caso específico. Com isto, as correções de distorção, ao invés de tornarem as imagens mais similares, podem ter aumentado as diferenças. Esta alternativa deve ser empregada, portanto, apenas quando os dados de calibração forem confiáveis e recentes.

O experimento A foi o que apresentou o resultado de registro entre as imagens mais preciso, com uma MRR nos pontos de verificação de 1,54 pixel. A diferença da MRR para o experimento F não foi significativa (de 0,10 pixel), o que indica que ambas as estratégias podem ser empregadas, pois foram as que apresentaram melhor registro.

A Figura 6 ilustra a situação para o experimento F, onde a Figura 6(a) exibe a região dos pontos de interesse selecionados e a Figura 6(b) exibe os pontos remanescentes após a eliminação daqueles que apresentaram resíduos acima do limiar, tendo sido mantidos 123 pontos.

Com relação aos melhores pesos as serem aplicados, a análise dos resultados três últimos experimentos (F, G, H) mostra que a magnitude deve ter o dobro do peso das direções, para este caso.

Figura 6 - Malha de pontos (a) Distribuição de pontos (b) Pontos selecionados.



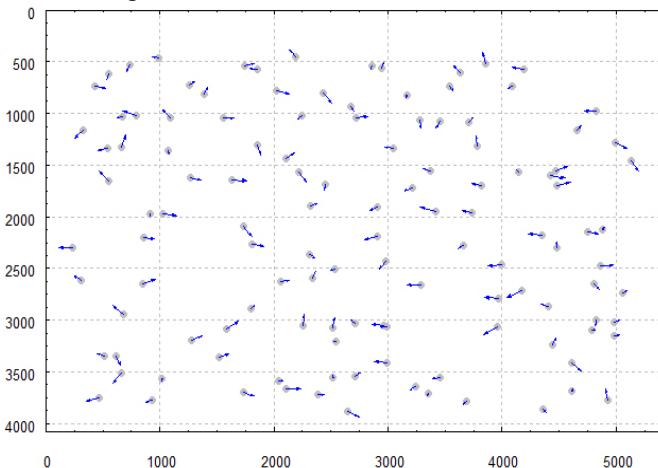
A Tabela 4 exibe os coeficientes do polinômio de mapeamento estimados após o ajuste pelo MMQ dos pontos selecionados.

Tabela 4 - Coeficientes do polinômio de mapeamento.

<b>a<sub>0</sub> (mm)</b>	<b>a<sub>1</sub></b>	<b>a<sub>2</sub> (mm<sup>-1</sup>)</b>	<b>a<sub>3</sub></b>	<b>a<sub>4</sub>(mm<sup>-1</sup>)</b>	<b>a<sub>5</sub> (mm<sup>-1</sup>)</b>	<b>a<sub>6</sub> (mm<sup>-2</sup>)</b>
1,22	1,01	$6,30 \times 10^{-5}$	$3,87 \times 10^{-3}$	$-1,73 \times 10^{-5}$	$-5,16 \times 10^{-5}$	$1,29 \times 10^{-7}$
<b>b<sub>0</sub> (mm)</b>	<b>b<sub>1</sub></b>	<b>b<sub>2</sub> (mm<sup>-1</sup>)</b>	<b>b<sub>3</sub></b>	<b>b<sub>4</sub>(mm<sup>-1</sup>)</b>	<b>b<sub>5</sub> (mm<sup>-1</sup>)</b>	<b>b<sub>6</sub>(mm<sup>-2</sup>)</b>
$1,46 \times 10^{-1}$	$-5,34 \times 10^{-3}$	$6,11 \times 10^{-5}$	1,01	$-1,73 \times 10^{-5}$	$1,88 \times 10^{-4}$	$-2,25 \times 10^{-6}$

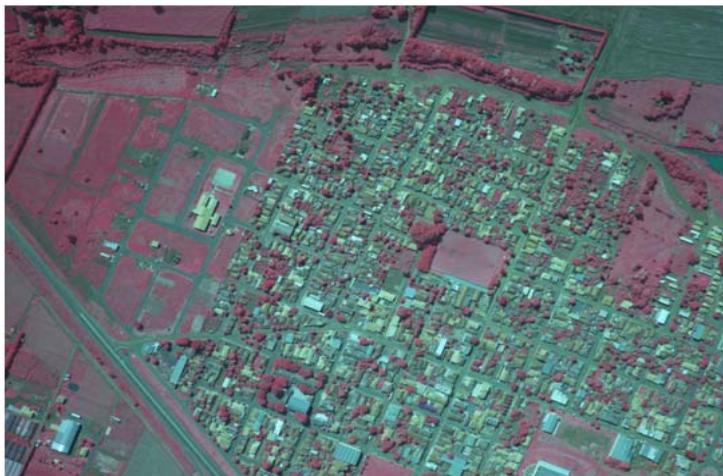
A Figura 7 exibe os vetores dos resíduos resultantes nos 123 pontos selecionados. O vetor com maior magnitude apresenta um resíduo de 1,5 píxeis em sua resultante.

Figura 7 - Vetores dos resíduos resultantes.



A Figura 8 exibe uma composição colorida formada após o registro das imagens, onde a banda R na imagem RGB foi substituída pela banda do IR, a banda G pela R e a banda B pela G.

Figura 8 - Composição de imagem (IR, R, G).



Os resultados obtidos são compatíveis com os encontrados na literatura. Dare e Dowman (2001) obtiveram erros de 4,4 pixels, no melhor caso. Já Fedorov (2002, p. 128) obteve erros nos pontos de verificação de 1.47 pixels, no registro de imagens aéreas. Isto mostra que a técnica proposta neste trabalho é compatível com as demais existentes.

## 5. CONSIDERAÇÕES FINAIS E CONCLUSÕES

Neste trabalho foi proposta e analisada experimentalmente uma técnica semiautomática de registro de imagens RGB e IR adquiridas por um par de câmeras digitais de médio formato.

A qualidade no processo de registro entre a imagem RGB e a IR, é determinada pela distribuição dos pontos correspondentes entre as imagens, para a determinação dos coeficientes do polinômio de mapeamento. Quanto melhor for a distribuição dos pontos, e menor o número de falsas correlações, melhor será o registro entre as imagens. A técnica de detecção de cantos (algoritmo de Harris) auxilia nesta etapa, identificando pontos distinguíveis, que são posteriormente selecionados pela MVC das translações.

Uma primeira aproximação para os coeficientes do polinômio de mapeamento (usando uma lista de pontos medidos manualmente) é importante para redução do

espaço de busca e redução de falsas correspondências. A técnica de determinação de pontos correspondentes foi baseada no cálculo de uma função erro entre os valores de magnitude e direções dos gradientes, com uma ponderação duas vezes maior para as magnitudes dos gradientes.

Verificou-se que o registro com a retificação prévia das imagens não apresentou resultados superiores ao que usou apenas as imagens originais, para as imagens em estudo.

Vários detalhes da técnica podem ser ainda avaliados, o que é recomendado para trabalhos futuros:

- O desenvolvimento e teste de outras funções polinomiais, com análise da significância dos parâmetros estimados;
- O desenvolvimento de técnicas para refinamento da etapa de correspondência, para medição dos pontos correspondentes com precisão subpixel;
- A análise das melhores combinações de tamanhos de janelas de referência e análise para identificação de pontos de interesse;
- A utilização de técnicas com emprego de pirâmides de imagens, que reduzem significativamente os erros de correspondência.

## 6. AGRADECIMENTOS

Os autores deste trabalho agradecem a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), ao Conselho Nacional de Pesquisa (CNPq) pelo apoio ao projeto em forma de bolsa de estudo e à empresa Engemap (Assis - SP) pela cessão das imagens utilizadas no trabalho.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- COSTA, E. R. ; TOMMASELLI, A. M. G ; GALO, M. Técnicas de pré-análise do processo de correlação visando à geração de modelo digital de terreno. *Revista Brasileira de Cartografia*, Rio de Janeiro, v. 57, n. 3, p. 252-262, 2005.
- DARE, P.; DOWMAN, I. An improved model for automatic feature-based registration of SAR and SPOT images. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 56, n.1, p. 13-28, 2001.
- ENGEMAP, Site da Internet. Disponível em <<http://www.engemap.com.br>>. Acesso em: 04 ago. 2009.
- FEDOROV, D. *Sistema semi-automático de registro e mosaico de imagens*. 2002. 150f. Dissertação (Mestrado em Computação Aplicada). INPE, São José dos Campos, 16/08/2002.
- FONSECA, L.; MANJUNATH, B. Registration techniques for multisensor remotely sensed imagery. *Photogrammetric Engineering and Remote Sensing*, v. 62, n. 9, p.1049–1056, 1996.
- FÖRSTNER, W. A feature based correspondence algorithm for image matching. *International Archives of Photogrammetry and Remote Sensing*, Rovaniemi, v.26, part. B3, Com. III, p.150-166, 1986.

- FÖRSTNER, W. Matching Techniques. In: HARALICK, R.; SHAPIRO, L. G. *Computer and Robot Vision*. New York: Addison-Wesley, 1993. v.2. p. 289-378.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento de Imagens Digitais*. São Paulo: Editora Blucher, 1992. 296 p.
- HABIB, A. F.; MORGAN, M. F. Small format digital cameras for mapping applications: Calibration and Stability Analysis. In: Mitishita, E.A. et al.. (Org.). Série em Ciências Geodésicas - Novos Desenvolvimentos em Ciências Geodésicas, 2003, v. 3. Curitiba: Universidade Federal do Paraná, p. 3-25.
- HARRIS, C.; STEPHENS, M. A Combined Corner and Edge Detector. In: Fourth Alvey Vision Conference, 1988, *Proceedings...*, 1988, p. 147-151.
- HEIPKE, C. Automation of interior, relative, and absolute orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 52, n.1, p. 1-19, 1997
- HRKAĆ, T.; KALAFATIĆ, Z.; KRAPAC, J. Infrared-Visual Image Registration Based on Corners and Hausdorff Distance. In: ERSBØLL, B. e PEDERSEN, K. *Lecture Notes in Computer Science: Image Analysis*. Berlin: Springer Berlin / Heidelberg, 2007. p. 383-392.
- MACHADO, A. M. L.; MITISHITA, E. A; SANTOS JR, R. L; BARBOSA, F. P. Verificação da estabilidade geométrica dos parâmetros de calibração uma câmara digital não métrica de pequeno formato Sony Dsc-f717. In: Mitishita, E.A. et al.. (Org.). Série em Ciências Geodésicas - Novos Desenvolvimentos em Ciências Geodésicas, 2003, v. 3. Curitiba: Universidade Federal do Paraná, p. 292-309.
- MORAVEC, H. P. Visual mapping by a robot rover. In: 6th International Joint Conference on Artificial Intelligence, 1979. *Proceedings ...*, 1979, p. 598-600.
- SANTOS JR, R.L.; MITISHITA, E. A. ; Bähr, H.P. ; VÖGLTE, T. Metodologia para Integração Automática de Imagens Aéreas Digitais e Dados SPLA. *Boletim de Ciências Geodésicas*, v. 13, p. 369-394, 2007.
- ZHANG, Z.; DERICHE, R.; FAUGERAS, O.; LUONG, Q.T. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, v. 78, n. 1-2, p. 87-119, 1995.
- ZITOVA, B.; FLUSSER, J. Image registration methods: a survey. *Image and Vision Computing*, v. 21, p. 977-1000, 2003.

(Recebido em julho de 2010. Aceito em março de 2011)