



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

TEIXEIRA SILVEIRA, MARCELO; QUEIROZ FEITOSA, RAUL; NUNES E SILVA BRITO,  
JORGE LUÍS; JACOBSEN, KARSTEN  
CORRESPONDÊNCIA EFICIENTE DE DESCRIPTORES SIFT PARA CONSTRUÇÃO DE  
MAPAS DENOS DE PONTOS HOMÓLOGOS EM IMAGENS DE SENSORIAMENTO  
REMOTO

Boletim de Ciências Geodésicas, vol. 17, núm. 1, enero-marzo, 2011, pp. 130-160  
Universidade Federal do Paraná  
Curitiba, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=393937719008>

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

# **CORRESPONDÊNCIA EFICIENTE DE DESCRIPTORES SIFT PARA CONSTRUÇÃO DE MAPAS DENSOS DE PONTOS HOMÓLOGOS EM IMAGENS DE SENSORIAMENTO REMOTO**

*Efficient matching steps of the SIFT for constructing a dense map of conjugate points in remote sensing images*

MARCELO TEIXEIRA SILVEIRA<sup>1</sup>

RAUL QUEIROZ FEITOSA<sup>1</sup>

JORGE LUÍS NUNES E SILVA BRITO<sup>2</sup>

KARSTEN JACOBSEN<sup>3</sup>

<sup>1</sup>Pontifícia Universidade Católica do Rio de Janeiro - PUC/RJ

Departamento de Engenharia Elétrica – DEE

Rio de Janeiro - RJ

{marts, raul}@ele.puc-rio.br

<sup>2</sup>Universidade do Estado do Rio de Janeiro - UERJ

Programa de Pós Graduação em Engenharia de Computação – Geomática

Rio de Janeiro, RJ

jsilvabr@gmail.com

<sup>3</sup>Leibniz Universität Hannover

Institut für Photogrammetrie und GeoInformation – IPI

Hannover - Alemanha

jacobsen@ipi.uni-hannover.de

## **RESUMO**

Métodos automáticos de localização de pontos homólogos em imagens digitais baseados em área, combinados com técnicas de crescimento de região, são capazes de produzir uma malha densa e exata de pontos homólogos. Entretanto, o processo de crescimento de região pode ser interrompido em regiões da imagem, cuja paralaxe no eixo horizontal apresenta variação abrupta. Essa situação geralmente é causada por uma descontinuidade na superfície ou espaço-objeto imageado, tal como um prédio numa cena urbana ou um paredão de exploração de uma mina a céu aberto. Nesses casos, novos pares de pontos homólogos (sementes) devem ser

introduzidos, normalmente por um operador humano, a partir dos quais o processo é reiniciado. Dependendo do tipo da imagem utilizada e da estrutura 3D da região mapeada, a intervenção humana pode ser considerável. Uma alternativa totalmente automatizada em que se combinam as técnicas SIFT (*Scale Invariant Feature Transform*), pareamento por mínimos quadrados e crescimento de região foi proposta anteriormente pelos autores. O presente trabalho apresenta uma extensão a essa técnica. Basicamente, propõem-se alterações na etapa de correspondência do SIFT, que exploram características de estereogramas produzidos por sensores aéreos e orbitais. Avaliações experimentais demonstram que as modificações propostas trazem dois tipos de benefícios. Em primeiro lugar, obtém-se um aumento do número de pontos homólogos encontrados, sem aumento correspondente na proporção de falsas correspondências. Em segundo lugar, a carga computacional é reduzida substancialmente.

**Palavras-chave:** Fotogrametria; Correlação por Mínimos Quadrados; Correspondência por Mínimos Quadrados; Crescimento de Região; SIFT; LSM; MDE; MDS.

## ABSTRACT

Area-based automatic image matching combined with a region-growing technique are able to provide a dense and accurate set of corresponding points. However, the region-growing process may stop at image patches where the horizontal  $x$ -parallax has an abrupt change. In such cases new pairs of corresponding points (seeds) must be provided, usually by a human operator. The region growing procedure restarts then from the new seed points. Depending upon the type of image and the 3D-structure of the mapped area, the human intervention may be considerable. A fully automatic alternative for finding conjugate points in stereo pairs was proposed by the authors in a prior work. The method combines the scale invariant feature transform, the Least-Squares matching and the region-growing technique. This work presents an extension of that technique. Basically, improvements in the matching step of the SIFT algorithm are proposed, which explores properties of stereo images produced by aerial and orbital sensors. Experiments conducted on stereo pairs from both airborne and satellite imagery show that the benefit of the proposed changes is twofold. Firstly, the number of true matches increases substantially with no significant increase in the proportion of false matches. Secondly, the computational load is dramatically reduced.

**Keywords:** Photogrammetry; Least Square Correlation; Least Square Matching; Region Growth; SIFT; LSM; DEM; DSM.

## 1. INTRODUÇÃO

A crescente automação dos processos de mapeamento fotogramétrico digital trouxe benefícios práticos para a geração de modelos 3d da superfície terrestre a partir de estereogramas. Esses modelos digitais de superfície (mds) constituem

informação valiosa para aplicações como planejamento de telecomunicações, monitoramento de desastres naturais e planejamento urbano. Mais recentemente, com a disponibilidade de imagens de alta resolução sensoreadas tanto em nível de aeronave quanto em nível orbital, aumentou ainda mais o interesse por métodos que tornem a geração de mds mais automática.

A qualidade dos modelos digitais de elevação (MDE) ou de superfície gerados a partir de estereogramas depende da exatidão com que as entidades homólogas são medidas. Estas entidades podem ser pontos, linhas ou regiões. Uma técnica semi-automática para geração de uma malha densa de pontos homólogos, conhecida como crescimento de região, foi proposta originalmente em (OTTO e CHAU, 1989) e inúmeras variantes foram sugeridas e aplicadas desde então (p.ex. ZEMERLY et al., 1991), (WEI e QUAN 2004) e (ZE-TAO et al., 2009). A técnica deriva de um método baseado em área, conhecido como correspondência por mínimos quadrados (GRUEN, 1985), e é capaz de gerar uma malha densa de pares de pontos homólogos que cubra praticamente toda a imagem. O método requer, contudo, a intervenção humana ao alcançar regiões da imagem com pouca textura ou sujeitas a oclusão ou a distorção acentuada decorrente de variações abruptas de profundidade na área imageada. Nesses casos, o operador humano deve medir manualmente um par de pontos homólogos para que se reinicie o procedimento de crescimento de região. Dependendo da área imageada e das condições de imageamento, a ação do operador humano poderá ser requerida muitas vezes, tornando o processo trabalhoso.

Uma variante do método de crescimento de região, que dispensa a medição manual de sementes, foi proposta inicialmente em (SILVEIRA et al., 2008a) e avaliada para imagens aéreas e orbitais em (SILVEIRA et al., 2008b) (HECKEL, 2009). Essa extensão associa o crescimento de região a um método conhecido como SIFT (do inglês *Scale Invariant Feature Transform*), proposto originalmente em (LOWE, 1999). Nessa nova abordagem, o SIFT substitui a ação humana e provê de modo automático pares de pontos homólogos a serem usados quando o crescimento de região é suspenso por algum motivo.

O presente trabalho apresenta e discute em maiores detalhes o método que combina SIFT e crescimento de região. Este artigo propõe ainda uma nova versão do método, basicamente adaptações ao algoritmo SIFT que exploram propriedades de estereogramas aéreos e orbitais utilizados para a construção de modelos digitais de elevação ou superfície. Relativamente à original, essa nova versão produz uma malha mais densa de pontos homólogos e é mais eficiente computacionalmente.

O texto que se segue está organizado da seguinte forma. A próxima seção traz uma breve revisão da literatura e apresenta o contexto em que se insere o trabalho. A seção seguinte descreve sucintamente as técnicas em que se baseia a presente proposta. A seção subsequente descreve o algoritmo básico e as adaptações sugeridas. Em seguida, apresentam-se os experimentos realizados para avaliação da

proposta e discutem-se os resultados obtidos. O artigo termina com um sumário das principais conclusões e indicações de direções futuras.

## 2. REVISÃO DA LITERATURA

A correspondência de entidades homólogas em estereogramas é um dos temas mais investigados em Visão Computacional e Fotogrametria nas últimas 3 ou 4 décadas, com novos métodos sendo propostos a cada ano. Não sendo, portanto, possível compilar neste texto uma revisão completa dos avanços na área, esta seção se limita a apontar os principais grupos de abordagens e, assim, descrever sucintamente o contexto em que se insere o presente trabalho.

### 2.1 Métodos de Correspondência

Segundo (SCHENK, 1999), as abordagens de correspondência em estereogramas podem ser classificadas em três categorias principais: os métodos baseados em área, os métodos baseados em feições e os métodos baseados em descrições simbólicas das entidades a serem correlacionadas.

#### 2.1.1 Métodos Baseados em Feições

Métodos baseados em feições envolvem pelo menos três passos principais. Primeiramente são detectadas nas imagens pontos de interesse (HARALICK e SHAPIRO, 1992) associados, por exemplo, a arestas (VENKATESWAR e CHELLAPPA, 1995), vértices (HARRIS e STEPHENS, 1988), ou curvas (SCHMID e ZISSERMAN, 1998). Em seguida, calculam-se descritores dessas feições compostos por valores de um conjunto de seus atributos. Com base na similaridade entre descritores, estabelecem-se as correspondências.

Um dos primeiros métodos desta categoria foi proposto por Movarec (MORAVEC, 1977) (MORAVEC, 1979). Seguiram-se outras propostas, entre as quais merecem menção os métodos de Beaudet (BEAUDET, 1978), Förstner (FÖRSTNER e GÜLCH, 1987), Harris (HARRIS e STEPHENS, 1988), Lucas e Kanade (LUCAS e KANADE, 1991), Heitger (HEITGER et al., 1992) e o método Susan (SMITH e BRADY, 1997).

Em (SCHMID et al., 2000), (MIKOLAJCZYK e SCHMID, 2002) e (RODEHORST e KOSCHAN, 2006) são apresentados estudos que compararam vários destes métodos usando como critérios de avaliação a taxa de detecção (razão entre o número de falsas correspondências e o número total de correspondências encontradas), a repetibilidade (capacidade de detectar as mesmas feições sob diferentes condições de aquisição da imagem) e exatidão (desvio em relação à localização real de cada ponto). Esses trabalhos, baseados em imagens tomadas a curta distância, não indicam a superioridade de nenhum método sobre os demais para todos os critérios considerados.

Métodos baseados em feições são ineficientes sobre regiões com pouca textura ou quando as imagens estão afetadas por transformações geométricas

acentuadas, como ocorre freqüentemente nas aplicações de Visão Computacional, o que encorajou, entre outros fatores, o desenvolvimento dos métodos baseados em área.

### 2.1.2 Métodos Baseados em Área

Métodos baseados em área seguem, em sua maioria, essencialmente a mesma sequência de passos dos métodos baseados em feições, com a diferença que, ao invés de feições particulares, a correspondência se faz entre os arranjos das intensidades dos pixels em recortes das imagens.

Nas abordagens mais simples, o recorte é descrito pelo vetor formado pelas próprias intensidades dos pixels no interior de cada recorte e a similaridade é medida pela correlação cruzada normalizada (NCC), pela soma das diferenças quadráticas (SSD), soma das diferenças absolutas (SAD) ou por variantes destas técnicas (SCHARSTEIN e SZELISKI, 2002) aplicadas aos descritores de regiões candidatas a homólogas. Answanden e Guggenbuhl (ASCHWANDEN e GUGGENBUHL, 1993) comparam algumas destas métricas.

Uma dificuldade mais comum nessas técnicas do que nas baseadas em feições são as múltiplas correspondências para um mesmo recorte, que ocorrem em áreas mais homogêneas ou sobre arestas. Para resolver essa ambigüidade, métodos baseados em área impõem restrições aos candidatos a pontos homólogos, como suavidade (a disparidade varia suavemente por sobre a imagem) e restrições relacionadas à geometria epipolar, aplicando técnicas como programação dinâmica (OHTA e KANADE, 1985), *level sets* (FAUGERAS e KERIVEN, 1998), *space carving* (KUTULAKOS e SEITZ, 2000), *graph-cuts* (BOYKOV, 2001) (KOMOLGOROV e SABIH, 2002), equações diferenciais parciais (STRECHA et al., 2003) ou *expectation-maximization* (EM) (STRECHA et al., 2006).

### 2.1.3 Métodos Relacionais

Os métodos do terceiro tipo são também denominados “relacionais”. Em tais métodos cria-se uma descrição simbólica das entidades a serem pareadas. Estas descrições podem ter, por exemplo, a forma de grafos, árvores ou redes semânticas. Contrariamente aos outros métodos que essencialmente compararam propriedades geométricas, os métodos relacionais baseiam-se em relações topológicas entre as entidades candidatas (SCHENK, 1999) e são menos freqüentes na literatura.

## 2.2 Descritores de Região para Mapas Densos de Pontos Homólogos

Ao contrário dos métodos baseados em feições, os métodos baseados em área são potencialmente capazes de produzir pareamento para cada pixel, uma vez que podem ser aplicados a qualquer recorte das imagens. Por esta razão, a maioria das aplicações, que visam à construção de mapas densos de pontos homólogos, utiliza algum método baseado em área (SCHARSTEIN e SZELISKI, 2002) (BROWN et al., 2003).

Em aplicações de Visão Computacional, e mesmo em aplicações mais recentes no campo da Fotogrametria, como sistemas de mapeamento móvel (BARBOSA et al., 2005) (ESPINHOSA et al., 2008) e veículos aéreos não tripulados (MERINO, 2005) (CAI e WALKER, 2010), em que as distorções geométricas entre as imagens do estereograma podem ser acentuadas, os descritores de região e as métricas citadas na seção 2.1.2 apresentam desempenho insatisfatório.

Um método conhecido como correspondência (ou correlação) por mínimos quadrados (GRUEN, 1996) e referido no texto adiante pela sigla LSM (do inglês *Least-Squares Matching*) vem ao encontro dessa dificuldade, ao levar em conta alterações na geometria das regiões homólogas, como se verá com maiores detalhes na seção 3.1.

Uma alternativa a representar cada recorte pelo vetor de intensidades de seus pixels consiste em calcular a partir destas um descritor invariante a uma classe de transformações da imagem. Vários descritores de região propostos para este fim são avaliados em (MIKOLAJCZKY e SCHMID, 2005).

Entre estas abordagens, a técnica conhecida como SIFT (do inglês *Scale Invariant Feature Transform*) é uma das mais utilizadas desde que foi introduzida por David Lowe em 1999 (LOWE, 1999).

O SIFT é invariante quanto à escala e à rotação e parcialmente invariante quanto a pequenas mudanças de perspectiva e iluminação. Por tais virtudes proliferaram as aplicações do SIFT na área de Visão Computacional, tais como localização e mapeamento por robôs (SLAM) (LEE e SONG, 2010), reconhecimento de gestos (MINHAS et al., 2010), reconhecimento facial (TAN et al., 2010), realidade aumentada (SHEN et al., 2010) e análise de imagens médicas (TOEWS et al., 2009).

Aplicações do SIFT em sensoriamento remoto têm também se tornado cada vez mais comuns (FRAUNDORFER et al., 2008) (YANG e NEWSAM, 2008), como, por exemplo, na detecção de objetos (TAO et al., 2011), no registro de imagens aéreas e dados de radar (LI et al., 2008), e no mapeamento em tempo real (FÖRSTNER e STEFFEN, 2008), entre outras aplicações.

Estudos publicados em (MIKOLAJCZKY e SCHMID, 2005) e (SCHMID et al., 2005) apontaram o SIFT como a melhor de um conjunto de técnicas alternativas, tanto em termos de estabilidade quanto do número de pontos detectados. Um estudo similar publicado em (REMONDINO, 2006) confirma as qualidades do SIFT no que tange ao número e à estabilidade dos pontos detectados. O estudo indica, por outro lado, que alguns detectores de feições apresentam melhor exatidão do que o SIFT, e acrescente que a exatidão do SIFT pode ser melhorada utilizando LSM.

Um estudo relativamente recente (LINGUA et al., 2009) compara o SIFT com técnicas comumente utilizadas pela comunidade de Fotogrametria, com o operador de Förstner, LSM e correlação cruzada. O estudo constata a superioridade do SIFT sobre as demais técnicas analisadas, não apenas quanto à exatidão, mas também

quanto ao número de correspondências detectadas. Os autores realçam ainda que o SIFT não exige que se lhe forneça uma solução aproximada, como ocorre com o LSM. Estas constatações fazem do SIFT uma solução atraente tanto para aplicações tradicionais (p.ex., construção de MDS) quanto para aplicações mais recentes na área da Fotogrametria. Os estudos ressaltam, contudo a elevada carga computacional associada ao SIFT.

### 3. TÉCNICAS RELACIONADAS

O método proposto neste artigo combina três técnicas brevemente descritas nas próximas seções. Uma descrição detalhada de cada uma destas técnicas, assim como uma comparação abrangente com técnicas alternativas, foge ao escopo deste trabalho. Esta seção provê tão somente informação suficiente para a compreensão da proposta apresentada mais adiante.

#### 3.1 LSM

Sejam  $g_1$  e  $g_2$  duas imagens de uma mesma cena. Pontos homólogos podem ser definidos como os centros  $(x_{01}, y_{01})$  e  $(x_{02}, y_{02})$  de vizinhanças de  $n \times n$  pixels respectivamente em  $g_1$  e  $g_2$ , no interior das quais a relação (vide, p.ex., SCHENK, 1999)

$$g_1(x, y) = \alpha g_2(x + x_0, y + y_0) + \beta \quad (1)$$

onde  $x_0 = x_{02} - x_{01}$  e  $y_0 = y_{02} - y_{01}$ , é válida em termos de mínimos quadrados para algum valor de  $\alpha$  e  $\beta \in \mathbb{R}$ . A correlação cruzada normalizada admite que pontos homólogos localizados em uma e outra imagem de um estereograma estejam relacionados por uma mera translação. Esse modelo é invariante quanto ao brilho ( $\beta$ ) e ao contraste ( $\alpha$ ), mas se torna ineficiente quando há diferenças geométricas importantes entre as imagens. Tanto a radiometria quanto o brilho de uma cena proveniente de um sensor passivo que atue na faixa óptica do espectro eletromagnético são afetados por diversos fatores, tais como a posição relativa sensor – alvo – fonte de energia, as características da plataforma sensora, as características do sensor, as condições ambientais existentes à época da coleta da imagem e as variações sazonais da intensidade e do ângulo de incidência solares.

O LSM é um refinamento da correlação cruzada normalizada. Essa técnica leva em consideração a distorção geométrica entre as imagens, modelando-a através de uma transformação afim. Formalmente

$$g_1(x, y) = \alpha g_2(ax + by + x_0, cx + dy + y_0) + \beta. \quad (2)$$

Como no caso anterior, o LSM leva em consideração as mudanças de brilho e contraste entre as imagens. Os oito parâmetros  $(x_0, y_0, \alpha, \beta, a, b, c, d)$  são

determinados por um processo que minimiza a diferença média quadrática entre os dois lados da equação 2 para cada um dos pontos da vizinhança analisada por  $N$  iterações sucessivas. O grau de similaridade  $\rho$  entre os pontos  $(x_{01}, y_{01})$  e  $(x_{02}, y_{02})$  é dado pelo coeficiente de correlação entre as vizinhanças neles centradas, após a aplicação em  $g_2$  da transformação encontrada. Para que dois pontos sejam considerados homólogos, o correspondente coeficiente de correlação deve superar um limiar  $\rho_{\min}$  definido pelo usuário. Mais detalhes podem ser encontrados em (SCHENK, 1999), (GRUEN, 1996), (AGOURIS e SCHENK, 1992) e (PERLT, 1984)..

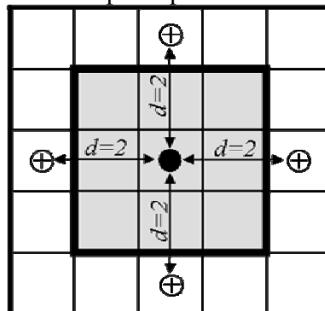
### 3.2 Crescimento de Região

O método conhecido por crescimento de região e referido no texto adiante pela sigla RG (do inglês, *Region Growing*) (OTTO e CHAU, 1989) (PETROU e BOSDOGIANNI, 2004) (PRATT, 2007), foi concebido para gerar um mapa denso de pontos homólogos em estereogramas, cujas imagens possuem pequenas diferenças em termos de escala e rotação. O procedimento parte de um par de pontos homólogos, chamados sementes, que são normalmente identificados por um operador humano. Um método baseado em área é então aplicado para determinar com maior exatidão a posição do ponto homólogo na segunda imagem. Se o valor da correlação for igual ou superior a um limiar pré-definido, o par de pontos encontrado é acrescido a um repositório, caso contrário descartado. Uma vez que a posição exata foi determinada pelo método baseado em área, quatro novos pares de pontos são gerados, localizados a uma distância de  $d$  pixels acima, abaixo, à esquerda e à direita, relativamente ao último ponto localizado, onde  $d$  é um parâmetro, um número natural, que define o espaçamento entre a localização atual e os novos pontos. Esses pares de pontos tornam-se novas sementes, cujas posições são igualmente refinadas por um método baseado em área. O procedimento é repetido recursivamente, espalhando novas sementes em ambas as imagens, fornecendo, assim, um conjunto denso de pontos homólogos.

A figura 1 ilustra parte do processo. As células representam *pixels* da imagem. O símbolo “●” marca um *pixel* de um par de homólogos encontrados pelo LSM. Partindo desse *pixel*, a etapa RG lança mais 4 sementes a uma distância de 2 pixels ( $d=2$ ), indicadas pelo símbolo “⊕”.

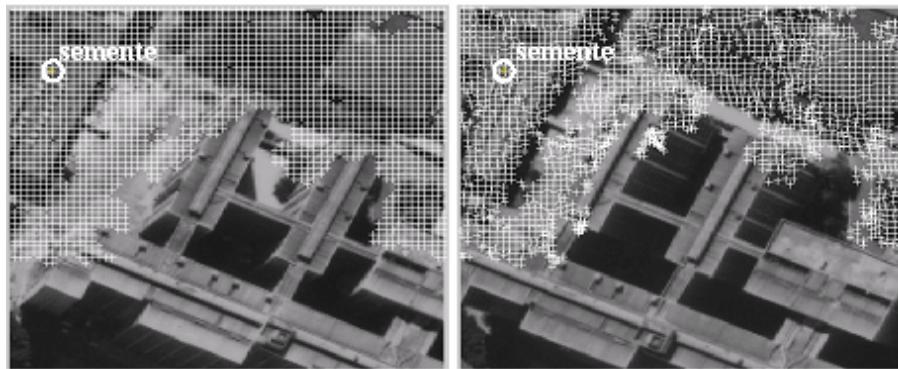
O termo cobertura refere-se neste trabalho à união das vizinhanças de tamanho  $(2d-1) \times (2d-1)$  (corresponde à região sombreada da figura 1, para  $d=2$ ) em torno de cada ponto de uma imagem para o qual se encontrou o homólogo na outra imagem.

Figura 1 - Crescimento de Região: pixel de um par de pontos homólogos ( $\bullet$ ) e sementes lançadas a partir deste, para  $d=2$  ( $\oplus$ ), área de cobertura do ponto  $\bullet$  está representada pelos pixels sombreados.



A figura 2 apresenta o resultado produzido utilizando-se LSM com crescimento de região (LSM+RG), partindo-se de um par de sementes (cruzes brancas) medidas manualmente. Os pontos homólogos encontrados estão indicados por cruzes brancas nas imagens.

Figura 2 - Pontos homólogos (cruzes brancas) determinados pela correlação por mínimos quadrados combinada com crescimento de região. A imagem da esquerda (a) é a imagem de referência, enquanto que da direita (b) é a imagem de busca.



(a)

(b)

Na figura 2, observa-se uma importante limitação dessa abordagem. Nota-se que nenhum ponto foi encontrado no topo do prédio. Com freqüência, o crescimento de região pára sobre regiões onde há grande variação de altura, o que pode provocar oclusão, como ocorre, por exemplo, em fachadas de edifícios altos. Outros fatores que podem bloquear o crescimento de região são sombras causadas pela geometria do imageamento em cenários urbanos, regiões com pouca textura,

áreas de floresta, ou, ainda, pela existência eventual de cobertura ou de sombras de nuvens na cena.

Esse problema pode ser contornado através da medição manual de uma nova semente no interior das regiões não atingidas, a partir da qual o processo de crescimento de região recomeça. Dependendo do tamanho da imagem e do terreno, poderá ser necessário repetir a medição manual de sementes muitas vezes até que se alcance a cobertura desejada na imagem.

### 3.3 SIFT

O método proposto por David Lowe, conhecido pelo acrônimo SIFT (do inglês *Scale Invariant Feature Transform*), foi concebido originalmente para aplicações em visão computacional.

Uma descrição completa deste método é apresentada, por exemplo, em (LOWE, 2004), (LOWE, 2001) e (LOWE, 1999). A seguir, são apresentados de forma sucinta os cinco passos principais que compõem este método.

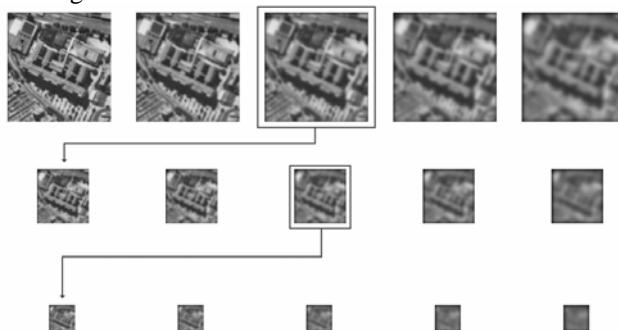
*Passo 1: Detecção dos pontos extremos no espaço-escala*

O primeiro passo consiste na construção da chamada pirâmide Gaussiana, conforme ilustrado na figura 3. Partindo de uma imagem de entrada  $I(x,y)$ , novas imagens  $L(x,y,\sigma)$  são geradas pela aplicação sucessiva de um filtro de suavização Gaussiano  $G(x,y,\sigma)$ , como mostra a equação abaixo,

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y). \quad (3)$$

onde  $\sigma$  denota o desvio padrão do filtro Gaussiano e define a escala ou nível da imagem  $L(x,y,\sigma)$ .

Figura 3 - Pirâmide com 3 oitavas e 5 níveis cada.



Cria-se, assim, uma oitava de imagens com diversos níveis. Uma dessas imagens é decimada gerando uma nova imagem com a metade da resolução da imagem original.

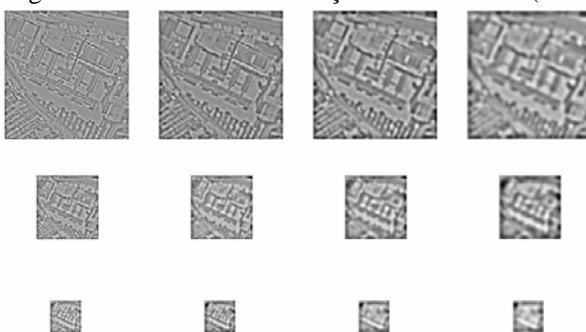
O processo pode ser reiniciado a partir da nova imagem, gerando uma nova oitava, e assim sucessivamente até que se obtenha o número de oitavas desejado.

As imagens da pirâmide Gaussiana assim produzida possuem diferentes escalas ( $\sigma$ ) e, portanto, diferentes níveis de detalhe. O processamento que se segue é aplicado a cada nível da pirâmide, o que confere ao método invariância quanto à escala. Pontos de interesse candidatos a formarem pares homólogos, chamados em (LOWE, 2004), (LOWE, 2001) e (LOWE, 1999) de pontos-chave, são selecionados do seguinte modo. Primeiro, calcula-se a diferença entre cada par de níveis adjacentes da pirâmide Gaussiana, formalmente,

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma). \quad (4)$$

onde  $k\sigma$  e  $\sigma$  representam escalas adjacentes da pirâmide Caussiana. Resulta dessa operação a chamada pirâmide de diferença de Gaussianas (DoG – do inglês *Difference of Gaussians*), conforme ilustra a figura 4. São candidatas a pontos-chave as posições dos máximos e dos mínimos locais da matriz tridimensional formada pela pirâmide DoG.

Figura 4 - Pirâmide de diferenças de Gaussianas (DoG).



Em seguida, os candidatos a pontos-chave localizados em vizinhanças de baixo contraste ou sobre arestas são descartados. As cruzes brancas na figura 5 apresentam o grupo de pontos-chave encontrados num recorte de uma imagem Ikonos da cidade do Rio de Janeiro.

*Passo 2: Refinamento da localização dos pontos-chave*

O próximo passo tem o propósito de melhorar a exatidão da localização de cada um dos pontos-chave selecionados no passo anterior (BROWN e LOWE, 2002). A estratégia consiste basicamente em ajustar uma função quadrática aos pontos da vizinhança  $3 \times 3 \times 3$  da pirâmide DoG em torno do ponto-chave localizado em  $\mathbf{x} = (x, y, \sigma)$ . A localização e a escala em que essa função alcança seu valor extremo nesta vizinhança passam a ser os novos valores de localização e escala do ponto-chave. Em (BROWN e LOWE, 2002) demonstra-se que o deslocamento  $\hat{\mathbf{x}}$  da localização final do ponto chave em relação à posição  $\mathbf{x}$  do extremo correspondente na pirâmide DoG é dado por

$$\hat{\mathbf{x}} = \frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2} \frac{\partial D}{\partial \mathbf{x}}, \quad (5)$$

onde as derivadas são calculadas na posição e escala em que se detectou o extremo na pirâmide DoG.

Figura 5 - Exemplo de pontos-chave (cruzes brancas) associados aos extremos na pirâmide DoG em um recorte de uma imagem Ikonos



*Passo 3: Cálculo da orientação dos pontos-chave*

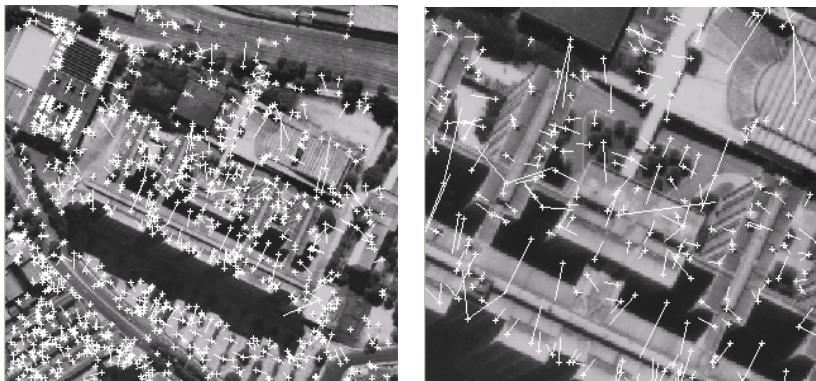
A orientação do ponto-chave ( $\theta$ ) corresponde à direção predominante do gradiente em torno do ponto-chave. É uma informação essencial para conferir invariância à rotação dos descritores SIFT, como se verá mais adiante.

A orientação é calculada no nível da pirâmide Gaussiana mais próxima à escala onde se detectou o ponto-chave. Computa-se o histograma de orientações dos gradientes desta imagem numa vizinhança em torno do ponto-chave.

O histograma cobre  $360^\circ$  em 36 faixas de  $10^\circ$  de largura. A contribuição de cada ponto da vizinhança é ponderada pela magnitude do gradiente e por uma

função que decresce com a distância ao centro da vizinhança. A figura 6 mostra um exemplo de imagem em que estão indicados os pontos-chave e as correspondentes orientações.

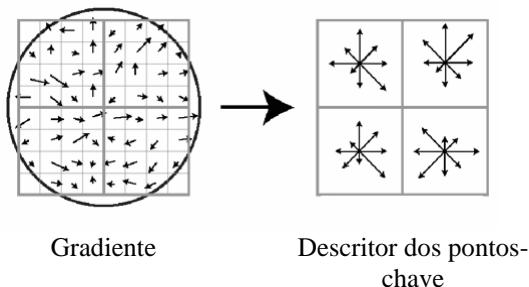
Figura 6 - Orientação de pontos-chave: cruzes e linhas indicam respectivamente a localização e a orientação dos pontos-chave.



#### *Passo 4: Cálculo dos descritores dos pontos-chave*

Neste passo, são calculados os descritores dos pontos-chave determinados no passo anterior. Calculam-se inicialmente os gradientes em uma vizinhança em torno de cada ponto-chave. A vizinhança selecionada é dividida em sub-regiões, conforme mostrado na figura 7.

. Figura 7 - O descriptor do SIFT (fonte: LOWE, 2004).



Em cada sub-região, calcula-se o histograma de direções do gradiente. Na composição desse histograma, os valores acumulados são ponderados pela respectiva magnitude do gradiente. No exemplo mostrado na figura 7, há quatro histogramas compreendendo 8 direções principais. Os valores de contagem dos

histogramas formam um vetor que constitui o descritor daquele ponto-chave particular. É importante notar que as direções dos gradientes são medidas em relação à orientação do ponto-chave. Dessa forma, os descritores SIFT tornam-se invariantes quanto à rotação.

#### *Passo 5: Determinação dos pares de pontos homólogos*

O processo descrito até aqui é aplicado a ambas as imagens do estereograma, produzindo dois conjuntos de pontos-chave, representados por seus descritores. O grau de similaridade entre pontos-chaves de ambos os conjuntos é dado pela distância Euclidiana entre seus descritores. Um par de pontos será considerado homólogo se, simultaneamente,

- a) a distância Euclidiana entre seus descritores for menor do que um certo limiar, e
- b) a menor distância Euclidiana em relação à segunda menor distância for maior do que um segundo limiar.

Esta última restrição diz respeito a situações em que um ponto-chave tem mais de um potencial candidato a homólogo. Nessas situações, a opção do algoritmo é descartar o ponto-chave e seus homólogos potenciais. A seção seguinte apresenta uma proposta que elimina a ambigüidade nesses casos e aumenta o número de pares homólogos encontrados no estereograma.

### **4. O MÉTODO PROPOSTO**

Propõe-se neste trabalho um algoritmo para construção de um mapa denso de pontos homólogos que combina o método de correlação por mínimos quadrados e crescimento de região com um mecanismo automático de geração de sementes. O algoritmo proposto elimina a necessidade de o operador humano fornecer novas sementes em áreas da imagem não alcançadas pelo processo de crescimento de região.

A próxima seção descreve resumidamente a primeira versão do método, conforme proposta em trabalhos anteriores (HECKEL, 2009), (SILVEIRA et al., 2008a) e (SILVEIRA et al., 2008b). Apresenta-se, em seguida, uma nova versão. Basicamente, propõem-se alterações ao SIFT que exploram características da aplicação e implicam numa rede de pontos homólogos mais extensa e numa maior eficiência computacional, como se demonstrará no capítulo 5.

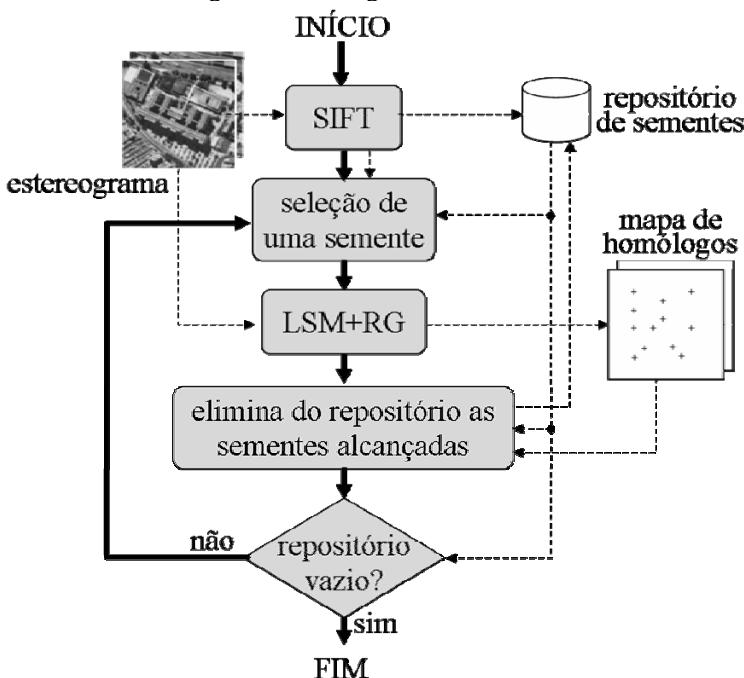
#### **4.1 Algoritmo Básico**

O método proposto para construção de um mapa denso de pontos homólogos está ilustrado na figura 8. Os elementos sombreados da figura correspondem aos passos executados ao longo do algoritmo. Os demais objetos bidimensionais da figura denotam entradas ou saídas de cada passo. As linhas cheias indicam o fluxo de controle, enquanto as linhas tracejadas relacionam cada passo com suas entradas e saídas.

Basicamente, aplica-se o SIFT para gerar um repositório de sementes do qual o algoritmo busca uma nova semente sempre que o processo de crescimento de região encontra uma barreira na imagem, como se descreveu na seção 3.2. O processo completo, referido deste ponto em diante no texto como SIFT+LSM+RG, compõe-se de cinco etapas seqüenciais, a saber:

- (a) Aplica-se inicialmente o SIFT às imagens do estereograma, produzindo, assim, pares de pontos homólogos armazenados num repositório;
- (b) Seleciona-se aleatoriamente um par de pontos do repositório;
- (c) Executa-se o LSM+RG usando como semente o par selecionado no passo anterior; os pares encontrados são armazenados no mapa de homólogos;
- (d) Removem-se do repositório todas as sementes (pares de homólogos) já alcançadas nas etapas anteriores do LSM+RG;
- (e) Repetem-se os passos (b) até (d) até que o repositório fique vazio.

Figura 8 – Fluxograma do método.



O mapa de pontos homólogos constitui o resultado do algoritmo. Na etapa (d), entende-se como “alcançada” as sementes do repositório pertencentes à área de cobertura de algum ponto homólogo identificado em iterações anteriores e que já tenha sido inserido no mapa de homólogos até então. Assim, as sementes cuja

distância Euclidiana ao ponto mais próximo do mapa de homólogos for inferior a  $d$  são consideradas alcançadas e, portanto, eliminadas do repositório de sementes.

#### 4.2 Adaptações ao SIFT

A nova versão do método proposta neste trabalho preserva as mesmas etapas básicas do algoritmo, conforme representados na figura 8. Propõem-se alterações no algoritmo SIFT usado na primeira etapa que exploram características próprias de estereogramas obtidos por sensores aéreos e orbitais, a fim de se obter uma malha de pontos homólogos mais densamente povoada e alcançar maior eficiência computacional. As duas modificações propostas dizem respeito ao passo 5 do algoritmo SIFT. São elas:

- a) Restringir a procura pelo homólogo de cada ponto-chave da imagem de referência ao mesmo nível e aos dois níveis adjacentes (acima e abaixo) da imagem de busca. Em outros termos, se um ponto-chave foi detectado no nível  $k$  da imagem de referência, limitar a busca por seu homólogo aos níveis  $k-1, k, k+1$  da imagem de busca.
- b) Restringir a procura pelo homólogo de cada ponto-chave da imagem de referência aos pontos-chave da imagem de busca, cujas orientações não diferirem em mais do que um valor  $\theta_{\text{máx}}$  a ser fornecido como parâmetro de entrada ao algoritmo. Assim, considerando  $\theta_i$  a orientação de um ponto-chave da imagem de referência, limitar a busca por seu homólogo aos pontos-chave da imagem de busca cuja orientação  $\theta_j$  é tal que  $|\theta_i - \theta_j| < \theta_{\text{máx}}$ .

Essas propostas se justificam em face das seguintes considerações. O algoritmo SIFT foi originalmente concebido para operar em imagens esteroscópicas sensoriadas a curta distância em aplicações de reconhecimento de objetos, tais como robótica e reconhecimento facial. Nessa classe de aplicações, as projeções de um mesmo objeto podem ter grandes diferenças de rotação, escala e também de forma numa e noutra imagem do par estéreo.

Entretanto, em aplicações de mapeamento fotogramétrico digital, independentemente do nível de aquisição das imagens, a geometria do imageamento entre duas exposições, cenas ou linhas de varredura sucessivas é definida de modo a assegurar a necessária diferença de paralaxe esteroscópica ao longo da linha de vôo (paralaxe em "x"). Essa restrição deve ser atendida para que seja possível calcular a interseção espacial de um par de raios projetivos conjugados ou homólogos a partir do conhecimento da posição e dos ângulos de atitude do sensor na tomada ou aquisição de cada uma das imagens que integram o par esteroscópico (BRITO e COELHO FILHO, 2007). Essa condição geométrica da interseção no espaço-objeto de dois raios projetivos fornece como resultado as coordenadas tridimensionais de cada ponto cujas projeções em cada imagem do par esteroscópico tenham sido encontradas pela medição automática ou manual. A

menos da necessidade de computação dos efeitos sistemáticos da correção da refração atmosférica e do efeito da curvatura da Terra, o modelo matemático para o cálculo da interseção espacial pode ser aplicado tanto a estereogramas obtidos por sensores aerotransportados, quanto a estereogramas gerados por sensores orbitais de resolução espacial sub-métrica e cujo sistema de visada do sensor pode ser programado remotamente, de modo a se assegurar a necessária diferença de paralaxe estereoscópica ao longo da direção de deslocamento da plataforma.

Dessa forma, os objetos tendem a apresentar escala e orientação próximas em ambas as imagens do par. Isso implica em que pontos-chave homólogos tendem a ocorrer no mesmo nível da pirâmide DoG e suas orientações a não diferirem em mais do que alguns poucos graus, o que se traduz nas modificações a) e b) propostas no início desta seção.

Essas restrições aplicadas ao passo 5 do SIFT trazem dois benefícios. Em primeiro lugar, reduz-se a complexidade do problema. O passo 5 do algoritmo original envolve a comparação de cada ponto-chave da imagem de referência com todos os pontos-chave da imagem de busca. A primeira das alterações propostas limita as comparações apenas aos pontos-chave do mesmo nível e dos níveis adjacentes, o que reduz substancialmente o tempo de processamento, como se demonstrará adiante.

O segundo benefício diz respeito ao número de pontos homólogos encontrados pelo SIFT. O passo 5 do algoritmo SIFT original considera ambíguos os casos em que a razão entre as distâncias Euclidianas que separam o descritor de um ponto-chave na imagem de referência ao primeiro e ao segundo descritores mais próximos da imagem de busca supera 80%. Nesses casos, o ponto-chave é descartado sem que se lhe associe um homólogo. Dessa forma, elimina-se, segundo estudos empíricos (LOWE, 2004), cerca de 90% dos falsos positivos que surgiriam se essa restrição não fosse aplicada.

Nossos estudos indicaram que, quando a ambigüidade ocorre, freqüentemente o verdadeiro homólogo está entre os pontos-chave correspondentes aos dois descritores mais próximos. A estratégia usada para redução de falsos positivos implica, portanto, num aumento de falsos negativos. As restrições propostas em a) e b) eliminam a referida ambigüidade em muitos desses casos, preservando pares homólogos, que seriam descartados pelo algoritmo SIFT original. Destarte, as alterações a) e b) redundam num maior número de sementes produzidas na etapa SIFT do método descrito na figura 8, contribuindo para que se obtenha uma malha mais extensa de pares de pontos homólogos.

## 5. ANÁLISE EXPERIMENTAL

### 5.1 Materiais

Os experimentos executados para validar o método proposto basearam-se em dois tipos de sensores: aéreo e orbital. Para o sensor aéreo, foi utilizado um par de

imagens digitalizadas a 300 DPI (resolução espacial de aproximadamente 0,6 m), ambas em escala de 1:8000, cobrindo parte da cidade do Rio de Janeiro, com altitudes variando de 0 a 100 m (vide figura 14a). Para o sensor orbital, foi utilizado um par de imagens Ikonos (resolução de aproximadamente 1 m), cobrindo uma área aproximada de 100 Km<sup>2</sup> da cidade do Rio de Janeiro, com altitudes variando de 0 a 500 m (vide figura 15a). A tabela 1 apresenta as características de cada par de imagens utilizadas.

Tabela 1 - Dados das imagens utilizadas nos experimentos.

<b>Imagen</b>	<b>Aérea</b>	<b>Ikonos</b>
Resolução espacial	0,6 m	1,0 m
Dimensões da imagem digital	$2900 \times 2900$ pixels	$2900 \times 2900$ pixels
Área de sobreposição longitudinal	60 %	99 %
Tipo da área imageada	Urbana	Urbana

O algoritmo SIFT foi implementado a partir da biblioteca criada por Rob Hess (HESS, 2008), versão 1.1.1, introduzindo-se as alterações propostas neste trabalho. O SIFT operou para ambos os estereogramas com 5 níveis por oitava e 10 oitavas.

O programa utilizado para a geração de correspondências através do algoritmo LSM+RG foi o DPCOR, que é parte integrante do software fotogramétrico BLUH System, desenvolvido na Universidade Leibniz de Hannover. Os pares de correspondências obtidos através do programa do SIFT foram introduzidos como sementes no programa DPCOR. Os parâmetros (vide seções 2.1 e 2.2) foram ajustados para os seguintes valores: vizinhanças de  $10 \times 10$  pixels ( $n=10$ ), 10 iterações ( $N=10$ ), distância entre sementes de 3 pixels ( $d=3$ ), coeficiente de correlação não inferior a 0,6 ( $\rho_{\min}=0,6$ ), razão entre a menor e segunda menor distância Euclidiana menor ou igual a 0,49 (sem limite superior para a distância Euclidiana entre descritores) e máxima diferença de orientação entre pontos-chave homólogos igual a 5° ( $\theta_{\max}=5^\circ$ ).

## 5.2 Desempenho Computacional

Nesta seção, avalia-se o impacto das alterações propostas neste trabalho em termos da eficiência computacional.

A figura 10 mostra o esforço computacional em cada caso em termos do tempo de processamento da etapa de correspondência apenas (passo 5 do SIFT). Os valores da figura foram medidos em um computador Pentium Quadcore Q6600, 2.4 GHz, memória RAM de 4 Gb e sistema operacional Linux Ubuntu de 64 bits.

A figura 9 mostra o número de comparações realizadas na etapa de correspondência do algoritmo SIFT na versão original e na versão que incorpora cada uma das alterações sugeridas individualmente e combinadas. A redução alcançada foi de

aproximadamente duas ordens de grandeza em ambas as imagens da nossa base de dados.

Figura 9 - Número de comparações de pontos-chave de acordo com a restrição imposta.

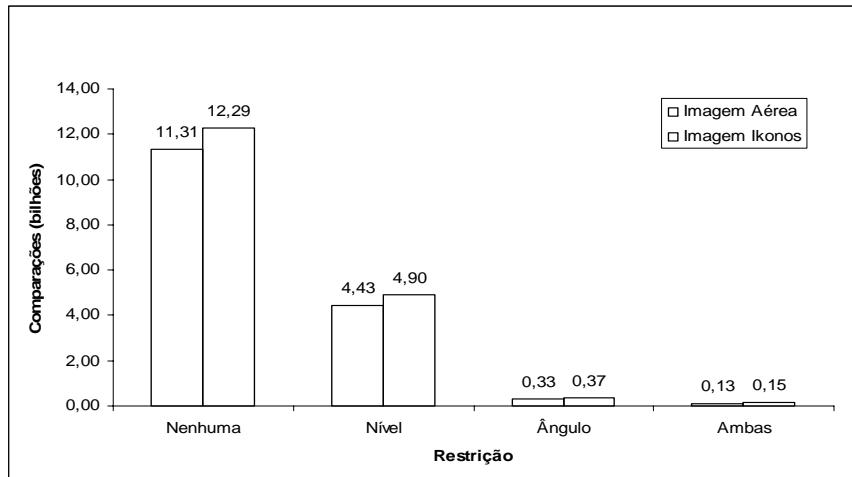
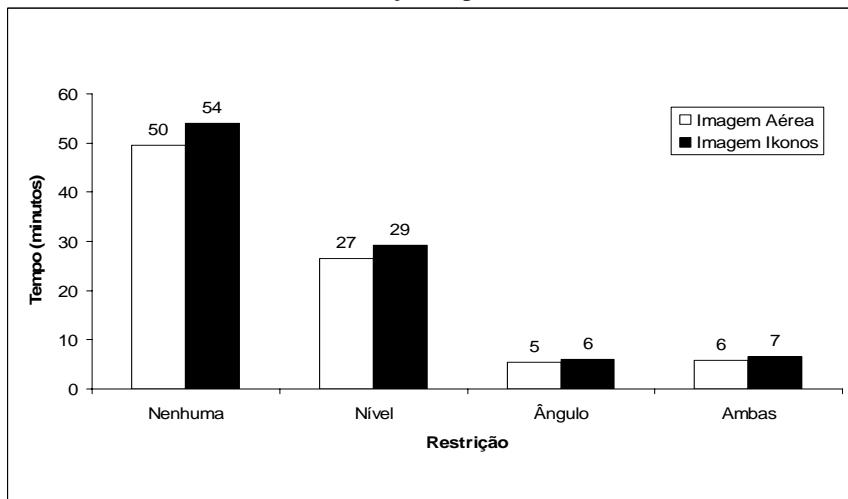


Figura 10 - Esforço computacional na etapa de correspondência, de acordo com a restrição imposta.



O tempo para processar as demais etapas do SIFT foi de aproximadamente 1 minuto em todos os casos. Dessa forma, o tempo gasto na correspondência representou

98% do tempo total da versão original do SIFT. Portanto, as alterações propostas tornaram o processamento do algoritmo aproximadamente 7 vezes mais rápido em ambos os casos avaliados.

### 5.3 Número de Correspondências

A figura 11 apresenta o número de pontos homólogos encontrados pelo SIFT para diferentes versões consideradas na seção anterior. Observa-se que o número de correspondências para a imagem Ikonos é notavelmente menor do que para a imagem aérea.

Buscando a razão para esta diferença, constatou-se que o número de pontos-chave detectados era da mesma ordem de grandeza em ambos os estereogramas. Conclui-se, portanto, que o algoritmo teve comparativamente muito menos sucesso em estabelecer correspondências na imagem Ikonos do que na imagem aérea. Esta diferença deriva das características das áreas imageadas. A imagem Ikonos de nossos experimentos é marcada pela presença de fragmentos de floresta e quadras residenciais arborizadas. Vegetação arbórea em imagens de alta resolução espacial possui uma textura que dá origem a vários pontos-chave, para os quais, em sua maioria, não se encontram os homólogos. Isso explica o número comparativamente reduzido de correspondências encontradas na imagem Ikonos. Este efeito não ocorre na imagem aérea por tratar-se de uma região urbana quase completamente construída, e com raras ocorrências de vegetação. De qualquer forma, é notório o aumento relativo da quantidade de correspondências obtidas pelo SIFT em ambos os estereogramas, quando cada uma das restrições foi imposta individualmente.

Figura 11 - Número de correspondências para diferentes restrições.

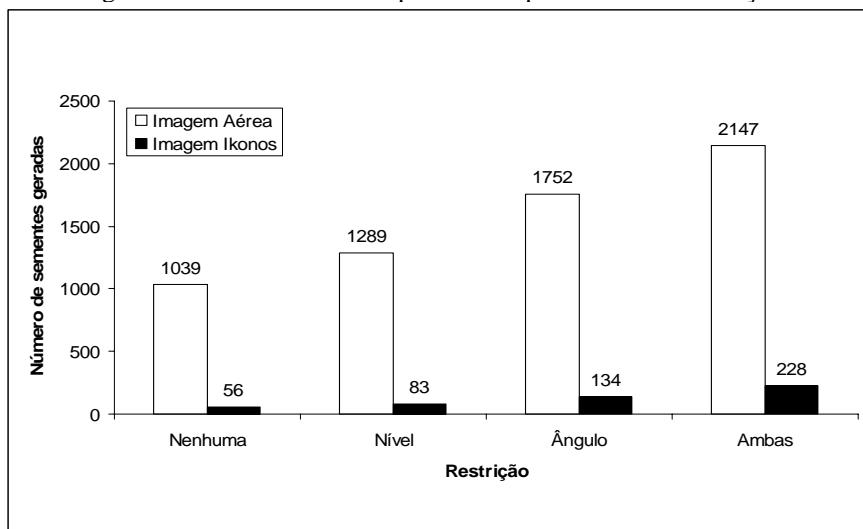
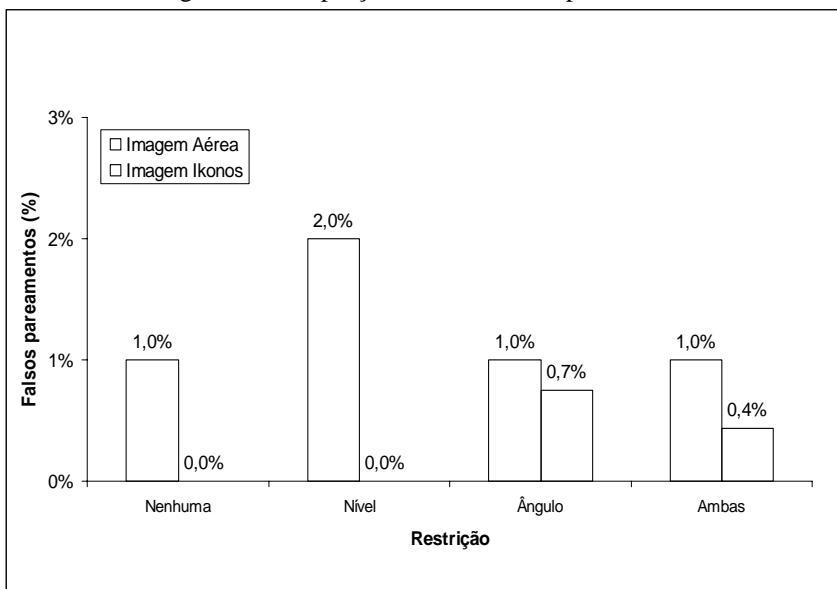


Figura 12 - Proporção de falsas correspondências.



O gráfico da Figura 12 mostra que as adaptações introduzidas no SIFT não resultam num aumento importante da proporção de falsos homólogos. No caso da imagem Ikonos, o gráfico acusa um aumento de 0% para 1% no número de falsas correspondências. Em valores absolutos trata-se aqui de um único falso positivo observado.

Em suma, verificou-se em nossos experimentos, que o aumento superior a 100% do número de correspondências trazido pelas adaptações sugeridas neste trabalho não é acompanhado por um aumento significativo de falsas correspondências.

#### 5.4 Avaliação da Cobertura

Em (SILVEIRA et al., 2008a) e (SILVEIRA et al., 2008b), a eficácia do algoritmo básico descrito na seção 4.1 já foi demonstrada empiricamente. Naqueles trabalhos, utilizou-se como métrica de desempenho a cobertura, conforme definida na seção 3.2.

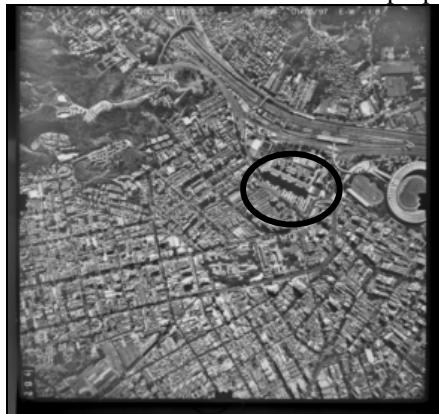
A figura 13a mostra uma das imagens aéreas utilizadas neste estudo. Cabe esclarecer que, nesse caso, somente a região à direita da imagem possui sobreposição com a outra imagem do par estereoscópico. A figura 13b apresenta em cinza a cobertura alcançada a partir de 4 sementes lançadas manualmente nas posições indicadas pelas cruzes negras na imagem. As figuras 13c e 13d mostram os

resultados produzidos pelo algoritmo descrito na figura 8, respectivamente, sem e com as restrições propostas neste trabalho. Não se nota diferença importante entre as figuras 13c e 13d em termos de cobertura. As figuras revelam, contudo, claramente que a nuvem de sementes geradas pelo SIFT com as restrições sugeridas é significativamente mais densa (figura 13c) do que quando se abdica dessas restrições.

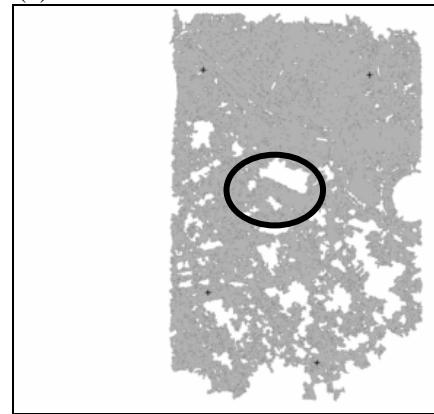
As vantagens das alterações no SIFT propostas neste trabalho se tornam mais evidentes em áreas com muitos edifícios elevados. A figura 14 mostra um recorte das figuras 13b e 13d em torno das fachadas internas e externas de um edifício elevado (indicado pela elipse na figura 13b). A imagem 14a mostra o resultado obtido nesse recorte partindo de 1 semente medida manualmente. Nota-se que o RG não foi capaz de transpor algumas das fachadas e deixou buracos em que não se detectou nenhum ponto com homólogo na outra imagem do estereograma. A figura 14b mostra o resultado obtido pelo algoritmo que incorpora as restrições sugeridas.

Notam-se sementes encontradas pelo SIFT (indicadas pelos anéis negros) sobre a região da imagem não alcançada pelo RG na imagem à esquerda. Essas sementes propiciaram um aumento da malha de pontos homólogos nessa região da imagem. Trata-se de um exemplo do benefício da abordagem SIFT+LSM+RG, quando há variações abruptas de altura na cena, como ocorre com freqüência, por exemplo, em grandes centros urbanos.

Figura 13 - Resultado do LSM+RG para a imagem aérea (a), a partir de 4 sementes manuais (b), sementes SIFT com configuração original (c) e com as modificações propostas (d).



(a)



(b)

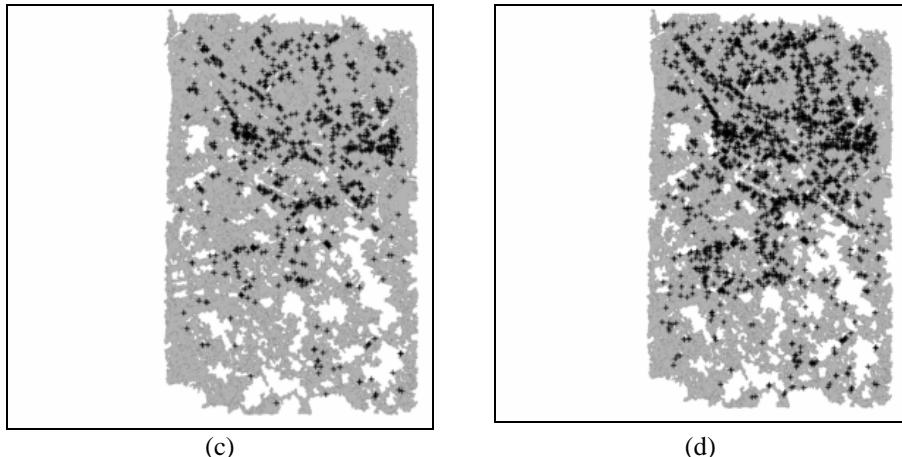
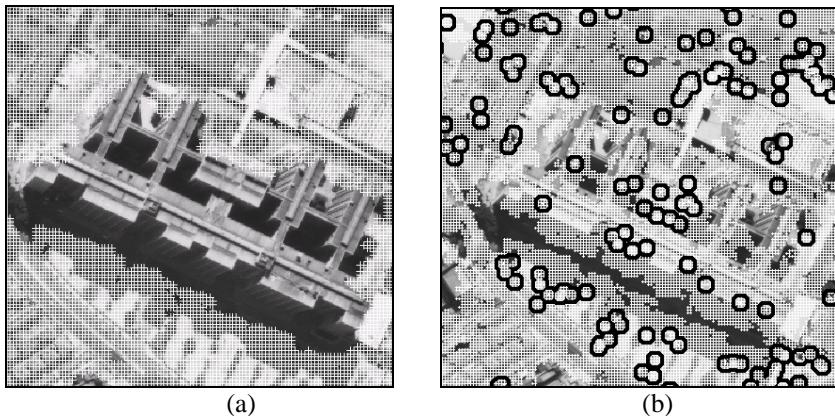


Figura 14 - Detalhe da cobertura do LSM+RG: com 4 sementes manuais (a), não foi possível atingir o topo da edificação, e com as sementes SIFT (nos centros dos anéis negros) com as restrições propostas (b), foi possível atingir o topo da edificação.

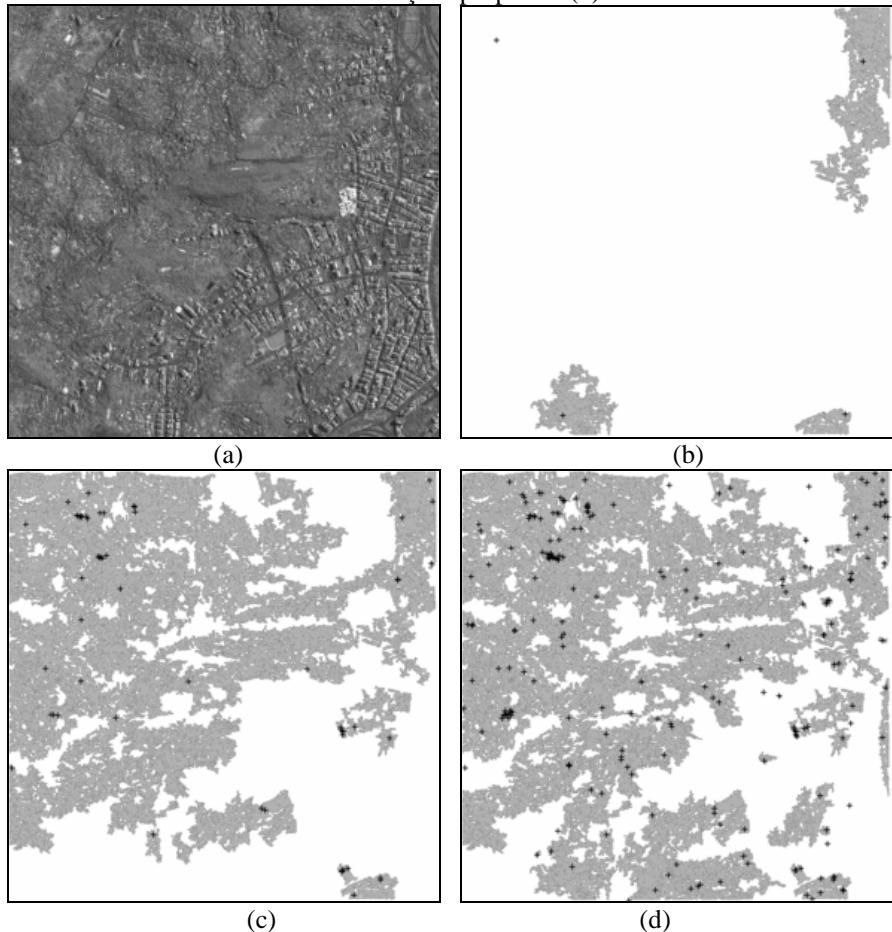


A figura 15 apresenta o resultado de experimento similar realizado sobre a imagem Ikonos mostrada na figura 15a.

Em termos da cobertura, nesse caso é notória a vantagem do método automático (figuras 15c e 15d) sobre o método manual (figura 15b) de seleção de sementes. A figura 15d mostra a contribuição das restrições sugeridas. Nota-se que a cobertura foi superior à produzida pela configuração original (figura 15c). Isso se

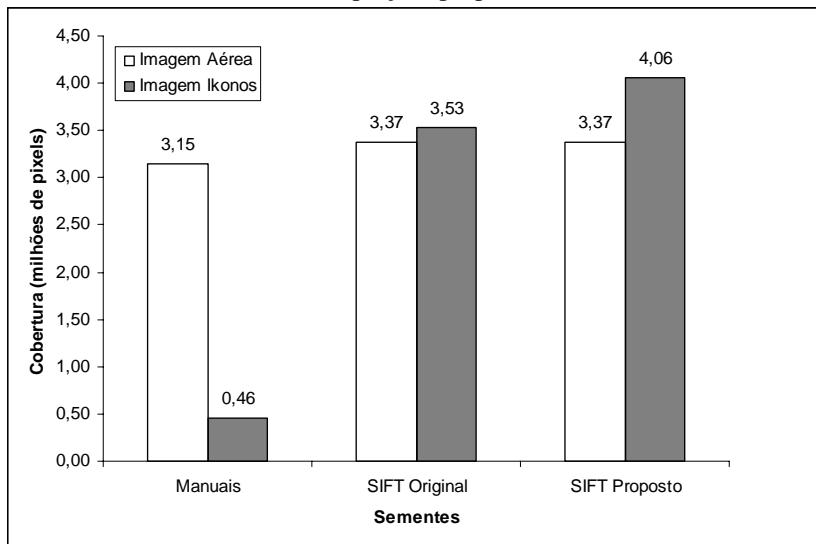
deve ao maior número de sementes geradas, o que permitiu atingir áreas da imagem não alcançadas pelo algoritmo original.

Figura 15 - Resultado do LSM+RG para a imagem Ikonos (a), a partir de 4 sementes manuais (b), sementes SIFT com configuração original (c) e com as modificações propostas (d).



O gráfico da figura 16 apresenta os valores de cobertura medidos em nossos experimentos. Percebe-se que os ganhos em relação à medição manual das sementes foram modestos para a imagem aérea e expressivos para a imagem Ikonos.

Figura 16 – Valores de cobertura expressos em *pixels* para o algoritmo de LSM+RG a partir de sementes manuais, geradas pelo SIFT, versão original e versão contendo as adaptações propostas.



As adaptações sugeridas neste trabalho trouxeram um aumento de cobertura quase nulo e próximo a 15%, respectivamente, para a imagem aérea e Ikonos, em comparação com o algoritmo SIFT+LSM+RG original.

Novamente nesse caso, a diferença observada não deriva do sensor, mas das características das áreas imageadas. Os benefícios do método descrito na Figura 8 dependem da freqüência, forma e distribuição de regiões na imagem, por sobre as quais o RG não se propaga. Nota-se na figura 13b que a partir de poucas sementes medidas manualmente já foi possível cobrir a maior parte da região de superposição das imagens aéreas. No caso da imagem Ikonos, (vide figura 15b) ao contrário, apenas uma pequena parcela da imagem foi alcançada a partir do mesmo número de sementes manuais, devido às ocorrências de vegetação, como se mencionou anteriormente. Como consequência, o ganho em termos de cobertura trazido pelo método para a imagem Ikonos foi maior do que para a imagem aérea.

Fica claro que, em termos de cobertura, os benefícios das adaptações sugeridas dependerão do estereograma de entrada, sendo razoável supor que ganhos ainda superiores aos valores medidos em nossos experimentos poderão ser alcançados, se os fatores que bloqueiam o crescimento de região forem mais freqüentes na cena imageada.

Ao finalizar esta seção, cabe realçar que as propostas deste trabalho não implicam em ganhos de exatidão na localização de pontos individuais, uma vez que

esta continua sendo determinada pelo LSM, como na abordagem mais simples que combina apenas o LSM e RG. Contudo, ao prover uma malha de pontos homólogos mais extensa, o método propicia a construção de modelos digitais mais representativos do terreno. Diminuem-se, assim, as regiões da imagem sem correspondências, onde o valor de altura só pode ser calculado por interpolação.

## 6. CONCLUSÃO E DIREÇÕES FUTURAS

O presente trabalho apresentou uma extensão de um algoritmo automático proposto em publicações anteriores para a geração de malhas densas de pontos homólogos em estereogramas, visando à construção de modelos digitais de superfície. O algoritmo primeiramente aplica o SIFT para produzir um conjunto inicial de sementes utilizadas em um passo subsequente pela correspondência por mínimos quadrados (LSM) combinada com crescimento de região (RG), de forma a gerar uma malha densa de pontos bem distribuída em um estereograma, sem a necessidade de intervenção humana.

A nova proposta basicamente impõe restrições na etapa de correspondência do SIFT que exploram características próprias de estereogramas obtidos por sensores aéreos e orbitais. Experimentos conduzidos sobre um par de imagens aéreas e um par de imagens Ikonos, indicaram que as alterações sugeridas provocaram um aumento de 2 a 4 vezes no número de pares de pontos homólogos encontrados pelo SIFT, sem aumento importante do número de falsas correspondências. Observou-se em nossos experimentos que as modificações sugeridas trouxeram um aumento de cobertura de até 15% .

No que tange à eficiência computacional, as inovações propostas aceleraram em até 50 vezes a etapa de correspondência do SIFT, que é, em geral, responsável pela maior parte do tempo total de processamento do algoritmo. A análise experimental demonstrou, contudo, que os ganhos trazidos pelas alterações sugeridas são fortemente dependentes das características das áreas imageadas.

Muito embora não traga ganhos em termos de exatidão na localização de pares homólogos individuais, o método proposto reduz o número e as áreas de regiões em que não se encontram pontos correspondentes, nas quais a profundidade é calculada por interpolação, contribuindo, assim, para uma melhor exatidão dos modelos digitais de superfície,

Cabe, por fim, mencionar alguns aspectos que motivam uma análise mais detida na continuação deste trabalho. Nota-se que, apesar do aumento na cobertura observado em nossos experimentos, restaram ainda nas imagens de teste várias regiões não alcançadas. Uma razão está em que tanto o SIFT quanto o LSM descartam soluções em recortes com pouca textura. Este comportamento é regulado por parâmetros cujos valores permanecem fixos para toda a imagem. Versões adaptativas do SIFT (LINGUA et al., 2009) e do LSM+RG (Xi et al., 2006) têm sido concebidas para superar esta dificuldade, cujos resultados preliminares

encorajam o estudo de variantes do nosso método em que os algoritmos originais do SIFT e do LSM sejam substituídos por essas novas versões.

Outra desvantagem do SIFT em relação aos demais algoritmos baseados em área é o elevado tempo de processamento. Variantes do SIFT mais eficientes computacionalmente, como o SURF (BAY et al., 2004) ou o DAISY (TOLA et al. 2010), podem em princípio substituir o SIFT no método proposto e absorver as mesmas alterações sugeridas neste trabalho, com possíveis prejuízos em termos de exatidão. Uma investigação destas possibilidades é também considerada no prosseguimento deste trabalho.

## AGRADECIMENTOS

Os autores deste artigo agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Agência Aeroespacial Alemã (DLR) pelo suporte financeiro ao desenvolvimento desta pesquisa.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AGOURIS, P.; SCHENK, T. Multiple image matching. *The International Archives of Photogrammetry and Remote Sensing*. Washington, v. XXIX, Parte B3, p. 802-807, 1992.
- ASCHWANDEN, P.; GUGGENBUHL, W. Experimental results from a comparative study on correlation-type registration algorithms. *Robust Computer Vision*, Wichmann, p. 268-289, 1993.
- BARBOSA, R.L., GALLIS, R.B.A., SILVA, J.F.C., MENEGUETTE JR., M. A computação do fluxo óptico em imagens obtidas por um sistema móvel de mapeamento terrestre. *Revista Brasileira de Cartografia*, v.57, n.2, p. 72-78, 2005.
- BAY, H.; ESS, A.; TUYTELAARS, T.; VAN GOOL, L. Speeded-up robust features (SURF). *Journal of Computer Vision and Image Understanding*, v.110, n.3, p. 346-359, 2008.
- BEAUDET, P. Rotationally invariant image operators. *4 th Int. Joint Conference on Pattern Recognition*, p. 579-583, 1978.
- BOYKOV, Y.; VEKSLER, O.; ZABIH, R. Fast Approximate energy minimization via graph cuts. *IEEE Trans. Pattern Analysis and Machine Intelligence*, v.23, n.11, p. 1222-1239, Nov. 2001.
- BRITO, J.L.N.S; COELHO FILHO, L.C.T. Fotogrametria digital. 2<sup>a</sup> edição. Rio de Janeiro, EdUERJ, 2007, 196 p.
- BROWN, M.; LOWE, D. Invariant features from interest point groups. *British Machine Vision Conference*, Cardiff, Wales, p. 656-665, 2002.
- BROWN, M.Z.; BURSCHKA, D.; HAGER, G.D. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.25, n.8, p. 993 – 1008, 2003.

- CAI, J.; WALKER, R. Height estimation from monocular image sequences using dynamic programming with explicit occlusions. *IET Computer Vision*, v.4, Issue 3, p. 149-161, Set. 2010.
- ESPINHOSA, D. R. S.; SILVA, J. F. C.; GALO, M. Efeitos da injunção da base nas orientações interior e exterior e nas coordenadas do espaço objeto em imagens obtidas com um estéreo vídeo-câmera de um sistema móvel de mapeamento. *Boletim de Ciências Geodésicas*, v. 14, p. 504-522, 2008.
- FAUGERAS , O.D.; KERIVEN, R. Complete dense stereovision using level set methods. *European Conf. Computer Vision*, Jun. 1998.
- FÖRSTNER, W.; STEFFEN R. On visual real time mapping for unmanned aerial vehicles. *XI Congress The International Society for Photogrammetric and Remote Sensing*, Pequim, 2008.
- FÖRSTNER, W. and GÜLCH, E. A fast operator for detection and precise location of distinct points, corners and circular features. *Intercommission Conf. on Fast Processing of Photogrammetric Data*, p. 281-305, 1987.
- FRAUNDORFER, F.; FRAHM, J.M.; SNOEYINK, J.; POLLEFEYS, M.; WU, C. Image localization in satellite imagery with feature-based indexing. *XXI Congress The International Society for Photogrammetric and Remote Sensing*, Pequim, 2008.
- GRUEN, A. Least square matching: a fundamental measurement algorithm. *Close Range Photogrammetry and Machine Vision*. Bristol: Whittle Publishing, Chap. 8, p. 217-255, 1996.
- GRUEN, A. Adaptive least squares correlation: a powerful image matching technique. *South Africa Journal of Photogrammetry Remote Sensing and Cartography*, p. 175-187, 1985.
- HARALICK, R.M. Second directional derivative zero-crossing detector using the cubic facet model. *4 th Scandinavian Conference on Image Analysis*, p.17-30, 1985.
- HARRIS, C.; STEPHENS, M. A combined edge and corner detector. *Alvey Vision Conference*, p. 147-151, 1988.
- HECKEL, Y. *Método híbrido de correspondência para pares estereoscópicos de imagens aéreas e de satélite de alta definição*. 2009. 115 f. Dissertação (Mestrado em Engenharia Elétrica) – Faculdade de Engenharia, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.
- HEITGER, F.; ROSENTHALER, L.; VON DER HEYDT, R.; PETERHANS, E.; Kuebler, O. Simulation of neural contour mechanism: from simple to end-stopped cells. *Vision Research*, v.32, n.5, p. 963-981, 1992.
- HESS, R. SIFT Library. Disponível em <<http://blogs.oregonstate.edu/hess/code/sift/>>. Acesso em 29 de dez. 2010.
- KUTULAKOS, K.N.; SEITZ, S.M. A Theory of Shape by Space Carving. *Computer Vision*, v.38, n.3, p. 197-216, Jul. 2000.

- LEE, Y.J.; SONG, J.B. Autonomous salient feature detection through salient cues in an HSV color space for visual indoor simultaneous localization and mapping. *Advanced Robotics*, v.24, n.11, p. 1595-1613, 2010.
- LI, F.; ZHANG, G.; YAN, J. Coregistration based on SIFT algorithm for synthetic aperture radar interferometry. *XXI Congress The International Society for Photogrammetric and Remote Sensing*, Pequim, 2008.
- LINGUA, A.; MARENCHINO, D.; NEX, F. Performance analysis of the SIFT operator feature extraction and matching in photogrammetric applications. *Sensors*, v.9, p. 3745-3766, 2009.
- LOWE, D.G. Object recognition from local scale-invariant features. *International Conference on Computer Vision*, Corfu, Greece, p. 1150-1157, 1999.
- LOWE, D.G. Local feature view clustering for 3D object recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, Kauai, Hawaii, p. 682-688, 2001.
- LOWE, D.G. Distinctive image features from scale invariant key points. *International Journal of Computer Vision*, p. 91-110, 2004.
- LUCAS, B.D.; KANADE, T. An iterative image registration technique with an application to stereo vision. *Conference on Artificial Intelligence*, p. 674-679, 1981.
- TOEWS, M.; WELLS, W.M.; COLLINS, D.L., ARBEL, T. Feature-based morphometry: discovering group-related anatomical patterns. *NeuroImage*, v.49, n.3, p. 2318-2327, 2010.
- MERINO, L.; WIKLUND, J.; CABALLERO, F.; MOE, A.; DE DIOS, J.R.M.; FORSSEN, P.E.; NORDBERG, K.; OLLERO, A. Vision-based multi-UAV position estimation. *Robotics & Automation Magazine*, v.13, Issue: 3, p. 53-62, 2006.
- MIKOLAJCZYK, K.; SCHMID, C. An affine Invariant Interest Point Detector, Proc. Senth European Conference on Computer Vision, pp. 128-142, 2002. performance
- MIKOLAJCZYK, K.; SCHMID, C. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, v.27, n.10, p. 1615-1630, 2005.
- MINHAS, R.; BARADARANI, A.; SEIFZADEH, S.; JONATHAN Wu, Q.M. Human action recognition using extreme learning machine based on visual vocabularies. *Neurocomputing*, v.73 (10-12), p. 1906-1917, 2010.
- MORAVEC, H.P. Toward automatic visual obstacle avoidance. *5 th Int. Joint Conf. Artificial Intelligence*, Cambridge, p. 584, 1977.
- MORAVEC, H.P. Visual mapping by a robot rover. *6th International Joint Conference on Artificial Intelligence*, p. 598-600, 1979.
- OHTA, Y.; KANADE, T. Stereo by intra- and intra-scanline search using dynamic programming. *IEEE Trans. Pattern Analysis and Machine Intelligence*, v.7, p. 139-154, 1985.

- OTTO, G.P.; CHAU, T.K.W. Region-growing algorithm for matching of terrain images. *Image and Vision Computing*, v. 7, n.2, p. 83-94, 1989.
- PERLT, A. Digital image correlation with the analytical plotter Planicomp C 100. *The International Archives of Photogrammetry and Remote Sensing*. Hamburgo, Comissão III, 25, Parte B3, p. 874-882, 1984.
- PRATT, W.K. Digital image processing. 4<sup>a</sup>. edição. Estados Unidos, John Wiley & Sons, Inc., 2007, 812 p.
- REMONDINO, F. Detectors and descriptors for photogrammetric applications. *Photogrammetric and Computer Vision ISPRS Symposium*, Bonn, 2006.
- RODEHORST, V.; KOSCHAN, A. Comparison and evaluation of feature pointdetectors. *5th International Symposium Turkish-German Joint Geodetic Days*, Technical University of Berlin, Mar. 2006.
- SCHARSTEIN, D.; SZELISKI, R. A taxonomy and evaluation of dense two-frame stereo-correspondence algorithms. *International Journal of Computer Vision*, v.47(1/2/3), p. 7-42, 2002.
- SCHENK, T. Digital photogrammetry. - Volume I. Estados Unidos. Terra Science, 1999, 412 p.
- SCHMID, C; MOHR, R.; BAUCKHAGE, C. Evaluation of interest point detectors. *International Journal of Computer Vision*, v.37, p. 151-172, 2000.
- SCHMID, C.; MOHR, R.; BAUCKHAGE, C. Evaluation of local descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.27, p. 1615-1630, 2005.
- SCHMID, C.; ZISSELMAN, A. The geometry and matching of curves in multiple views. *European Conf. Computer Vision*, p. 104-118, 1998.
- SHEN, J.; LUO, L.; ZHENG, X. Invariant-based augmented reality on mobile phones. *Journal of Multimedia*, v.5, n.6, p. 588-595, 2010.
- SILVEIRA, M.T.; FEITOSA, R.Q.; JACOBSEN, K.; BRITO, J.L.N.S.; HECKEL, Y. Um método híbrido para localização automática de pontos homólogos em pares de imagens estereoscópicas. *II Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação*, Recife, 2008a.
- SILVEIRA, M.T.; FEITOSA, R.Q.; JACOBSEN, K.; BRITO, J.L.N.S.; HECKEL, Y. A hybrid method for stereo image matching. *XXI Congress The International Society for Photogrammetric and Remote Sensing*, Pequim, v.XXXVII , Parte B1 , p. 895-900, 2008b.
- SMITH, S.M.; BRADY, J.M., SUSAN – a new approach to low level image processing. *Journal Computer Vision*, v.23, n.1, p. 45-78, 1997.
- STRECHA, C.; FRANSENS, R.; VAN GOOL, L. Combined depth and outlier estimation in multi-view stereo. *IEEE Conf. Computer Vision and Pattern Recognition*, 2006.
- STRECHA, C.; TUYTELAARS, T.; VAN GOOL, L., Dense matching of multiple wide baseline views. *Conf. Computer Vision*, 2003.

- TAN, C.; WANG, H.; PEI,D. SWF-SIFT approach for infrared face recognition. *Tsinghua Science and Technology*, v.15, n.3, p. 357-362, 2010.
- TAO, C.; TAN, Y.; CAI, H.; TIAN, J. Airport detection from large ikonos images using clustered sift keypoints and region information. *IEEE Geoscience and Remote Sensing Letters*, v. 8, n.1, p. 128-132. 2011.
- TOLA, E.; LEPETIT, V.; FUA, P. DAISY: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.32, n.5, p. 825-830, 2009.
- VENKATESWAR, V.; CHELLAPPA, R. Hierarchical Stereo and Motion Correspondence Using Feature Groupings. *Computer Vision*, v.15, p. 245-269, 1995.
- YANG, Y.; NEWSAM, S. Comparing SIFT descriptors and Gabor texture features for classification of remote sensed imagery. *The 15<sup>th</sup> IEEE Congress on Image Processing*, San Diego, p. 1852-1855, Out. 2008.
- ZEMERLY, M.J.A.; HOLDEN, M.; MULLER, J.-P.; BOFFEY, J. Parallel stereo matching of SPOT satellite images. *IEEE Colloquium on Parallel Architectures for Image Processing Applications*, p. 10/1 - 10/4, 1991.
- ZE-TAO, J.; BI-NA, Z.; MIN, W.; ZHONG-XIANG, C. A 3D Reconstruction Method Based on Images Dense Stereo Matching. *Third International Conference on Genetic and Evolutionary Computing*, p. 319-333, 2009.

(Recebido em julho de 2009. Aceito em março de 2011).