

3

Fundamentos Teóricos

O modelo proposto para esta tese envolve técnicas e conceitos de mais de uma disciplina. Desta forma, este capítulo apresenta os fundamentos teóricos necessários à compreensão do modelo proposto, incluindo noções de Análise de Imagens e Fotogrametria. Os leitores familiarizados com ambas as disciplinas poderão seguir diretamente para o capítulo 4, sem prejuízo para a compreensão do restante desta tese.

3.1

Processamento de Imagens Digitais

Processamento de Imagens Digitais (PDI) é a área do conhecimento que visa à construção de sistemas artificiais para a interpretação de imagens tal como feita pelo Homem, desde a aquisição da imagem até a sua interpretação. O PDI é usualmente dividido em cinco etapas executadas freqüentemente de modo seqüencial: realce, segmentação, pós-processamento, extração de atributos e reconhecimento. A figura 3.1 apresenta um fluxograma dos processos que envolvem o PDI, com base nas etapas de processamento de imagens encontradas em vários livros texto (p. ex. Gonzalez & Woods, 2002; Nixon & Aguado, 2008; Pedrini & Schwartz, 2008; Russ, 2008). Este capítulo, entretanto, somente aborda as etapas de PDI que foram utilizadas na elaboração desta Tese.

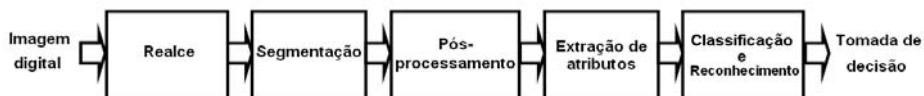


Figura 3.1: Fluxograma de PDI.

3.1.1

Segmentação

Segundo Gonzalez e Woods (2002), "a segmentação de imagem é a divisão

ou separação da imagem em regiões de atributos similares". Os atributos mais comumente utilizados na segmentação são os atributos espectralis, ou seja, os valores associados a cada uma das componentes de uma imagem multiespectral. Podem-se usar ainda atributos de forma e textura para guiar o processo de segmentação.

O método de segmentação conhecido como Crescimento de Regiões (Yakimovsky, 1973) é baseado no agrupamento em regiões ou grupos de *pixels* adjacentes que possuem intensidades similares (de acordo com um limiar estabelecido), através de uma semente introduzida como ponto de partida. O agrupamento pára quando nenhum *pixel* da vizinhança atende a um dado critério de similaridade. A figura 3.2 apresenta um exemplo da aplicação do método de Crescimento de Regiões, onde uma semente (representada pela cruz "+" azul na figura) é introduzida e os *pixels* vizinhos que possuem intensidades similares são agregados.



Figura 3.2: Segmentação por crescimento de regiões.

O método de segmentação conhecido como Divisor de Águas, ou do inglês *Watershed* (Beucher & Meyer, 1993), tem como princípio considerar uma imagem em tons de cinza como uma superfície topográfica, onde os valores de intensidade correspondem à altitude em cada ponto. Tendo em vista esta analogia, uma gota d'água que caísse em qualquer lugar da superfície escorreria em direção ao lugar mais baixo de uma das regiões da imagem. O acúmulo de água nesta região daria origem a uma poça.

Há dois algoritmos que realizam a segmentação por *Watershed*: cascata e inundação.

No método cascata, determinam-se os mínimos de toda a imagem, e cada mínimo local recebe uma identificação única. Mínimos locais adjacentes são combinados em uma única identificação. Em seguida, “gotas d’água” caem sobre cada *pixel* sem identificação e se movem para seu vizinho com menor amplitude, até que atinjam um *pixel* com identificação. Neste instante, a “gota d’água” assume essa identificação (Gonzalez & Woods, 2007).

No método de inundação, são feitos “furos” nos *pixels* correspondentes aos mínimos locais. A partir destes furos emana água que inunda toda a superfície paulatinamente. A água entra pelos furos e começa então a encher as poças. Assim que a poça estiver para transbordar para outra poça vizinha, uma barreira é construída de forma a impedir a passagem da água. Estas barreiras formam o contorno dos segmentos da imagem (Gonzalez & Woods, 2007).

É usual, aplicar-se o método de *Watershed* à magnitude do gradiente da imagem. Em imagens coloridas, os valores de gradiente de cada banda são combinados de algum modo para formar uma única matriz bidimensional em que se aplicará a inundação. A figura 3.18 apresenta um exemplo de resultado de segmentação por *Watershed*.

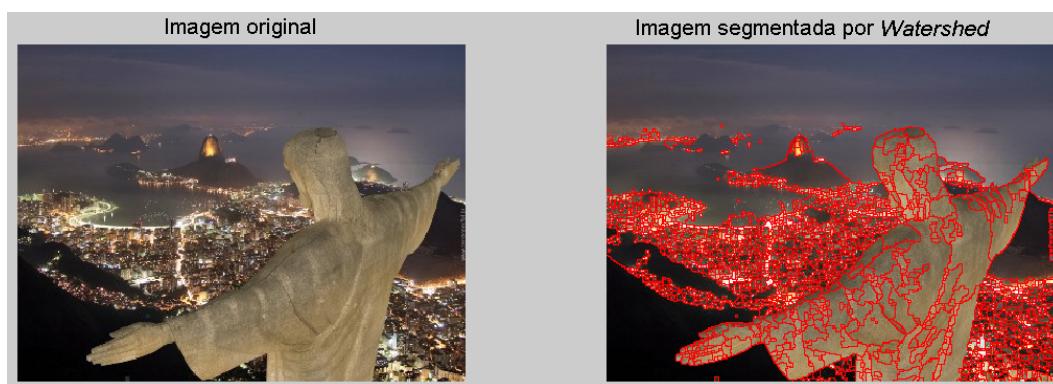


Figura 3.3: Exemplo de segmentação por *Watershed*.

3.1.2 Pós-processamento

O resultado da segmentação de uma imagem geralmente não reproduz com exatidão os contornos dos objetos presentes na imagem ou apresenta ruídos que prejudicam o desempenho das etapas posteriores da análise da imagem. Assim, a

etapa de pós-processamento visa melhorar o resultado da segmentação, bem como minimizar ou eliminar efeitos de ruídos.

Técnicas de morfologia matemática (Serra, 1982; Serra, 1988) que operam sobre imagens binárias resultantes da segmentação são freqüentemente utilizadas para pós-segmentação. Em tais abordagens, opera-se sobre matrizes lógicas com o mesmo número de linhas e colunas das imagens de entrada em que somente os *pixels* de um segmento são representados pelo valor “verdadeiro” ou “1”.

São duas as operações morfológicas básicas: a dilatação e a erosão. Para descrevê-las, convém definir inicialmente os conceitos de “reflexão” e “translação”.

Seja B um conjunto de *pixels* definidos por suas coordenadas na imagem, isto é, $B = \{(x_1, y_1), (x_2, y_2), \dots\}$. A reflexão \hat{B} de B é definida como:

$$\hat{B} = \{(-x_1, -y_1), (-x_2, -y_2), \dots\} \quad (3.1)$$

Por seu turno, a translação B_x de B por $\mathbf{x}=(x,y)$ é definida como:

$$B_x = \{(x_1 + x, y_1 + y), (x_2 + x, y_2 + y), \dots\} \quad (3.2)$$

A partir destes conceitos é possível definir a dilatação $A \oplus B$ de um conjunto A por um conjunto B , chamado “elemento estruturante”, como sendo o conjunto de pixels nas posições $\mathbf{x}=(x,y)$, tais que a translação por \mathbf{x} da reflexão de B tem interseção não vazia com A . Formalmente:

$$A \oplus B = \{\mathbf{x} | (\hat{B})_{\mathbf{x}} \cap A \neq \emptyset\} \quad (3.3)$$

De modo similar, a erosão $A \ominus B$ de um conjunto A por um elemento estruturante B , é dada pelos pixels nas posições $\mathbf{x}=(x,y)$ tais que a translação por \mathbf{x} de B está contida em A . Formalmente:

$$A \ominus B = \{\mathbf{x} | (B)_{\mathbf{x}} \subseteq A\} \quad (3.4)$$

A figura 3.4 ilustra como é feito o processo de dilatação, enquanto que a figura 3.5 ilustra o processo de erosão. O círculo indica a origem do sistema de coordenadas.

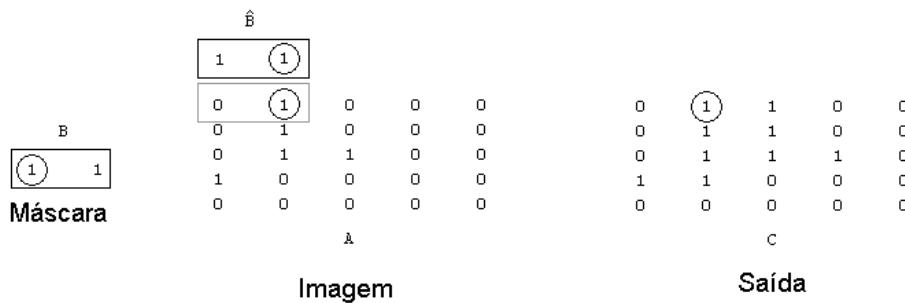


Figura 3.4: Exemplo de operação morfológica de dilatação.



Figura 3.5: Exemplo de operação morfológica de erosão.

Duas operações morfológicas largamente utilizadas em processamento de imagens digitais são a abertura e o fechamento, denotadas respectivamente por $A \circ B$ e $A \bullet B$.

A abertura $A \circ B$ consiste da erosão de A seguida da dilatação utilizando-se em ambas as operações B como elemento estruturante (vide equação 3.5). A abertura age eliminando filamentos estreitos e pequenos segmentos

$$(A \circ B) = (A \ominus B) \oplus B \quad (3.5)$$

De modo similar o fechamento $A \bullet B$ consiste da dilatação de A seguida da erosão, utilizando-se em ambas as operações B como elemento estruturante (vide equação 3.6). O fechamento visa eliminar reentrâncias estreitas ou pequenos buracos nos segmentos.

$$(A \bullet B) = (A \oplus B) \ominus B \quad (3.6)$$

A figura 3.6 mostra um exemplo de operação de abertura e de operação de fechamento.

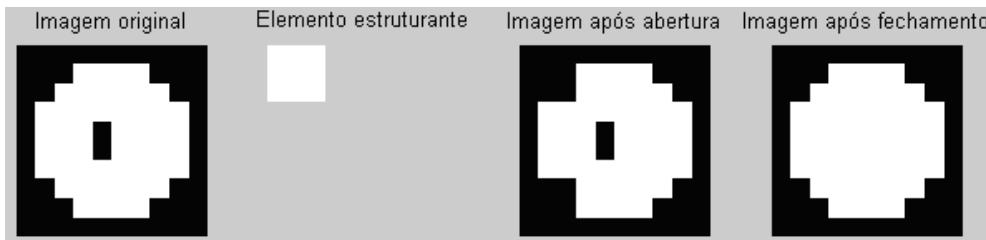


Figura 3.6: Exemplo de operações morfológicas de abertura e de fechamento.

3.1.3 Extração de atributos

O passo seguinte à segmentação consiste em descrever cada um dos segmentos. Tais descrições constituem a informação utilizada na etapa subsequente em que cada segmento será associado a uma das classes de objetos que se esperam encontrar na imagem analisada.

Na abordagem mais usual cada segmento é descrito por um vetor em que cada componente contém o valor de um determinado atributo do segmento. É comum utilizarem-se atributos espectrais, atributos de textura, ou atributos de forma como área, perímetro, excentricidade, entre muitos outros. A escolha do conjunto de atributos que comporão o vetor que descreve cada segmento dependerá das classes de objeto a serem reconhecidas.

A etapa de extração de atributo consiste, portanto, em medir em cada segmento os atributos com base em que serão discriminadas as classes de objetos na aplicação alvo. Russ (2008) e o software de Visão Computacional Halcon (Halcon 2010) apresentam algumas métricas para a extração de atributos geométricos de segmentos.

3.1.4 Classificação e reconhecimento

A etapa de classificação tem como objetivo decidir a que classe pertence cada segmento, baseando-se no correspondente vetor de atributos. Existem duas modalidades de classificação: não-supervisionada e supervisionada (Duda et al., 2001).

Os métodos de classificação não-supervisionados recebem como entrada o número de classes (ou aglomerados) e uma matriz em que cada linha contém o vetor de atributos de um objeto ou segmento e cada coluna corresponde a um atributo (área, comprimento, perímetro, etc). O objetivo é agrupar os objetos em um número predefinido de classes, que reúnem objetos que compartilham certas propriedades. Essas classes não são conhecidas *a priori*, sendo determinadas ao fim do processo.

Nos métodos de classificação supervisionados as classes do problema já são conhecidas *a priori*. Como no caso anterior, o classificador neste caso também recebe uma matriz contendo em cada linha um vetor de atributos de um objeto, ao qual se atribuirá um rótulo que identifica uma das classes do problema.

A diferença essencial entre os métodos não-supervisionados e os supervisionados, está na fase de treinamento. Os métodos supervisionados pressupõem a existência de um conjunto de objetos, chamado “de treinamento”, cujas classes são previamente conhecidas. A partir desta informação, os métodos supervisionados estimam os parâmetros do classificador, que comporão as chamadas “regras de decisão”, que finalmente atribuirão cada objeto não utilizado na etapa de treinamento a uma das classes do problema.

3.2 Fotogrametria

O vocábulo “Fotogrametria” resulta dos radicais gregos *photon*, que significa luz; *graphos*, escrita; e *metron*, medição. Isto significa dizer que Fotogrametria é medição feita em fotografias.

Existem diversas abordagens e discussões em torno do conceito de

Fotogrametria, mas o consenso geral a define como “ciência e tecnologia de se obter informação confiável, através de imagens adquiridas por sensores remotos” (Brito & Coelho Filho, 2007). Ciência, uma vez que os processos de captação da energia eletromagnética, do qual os raios luminosos são formados, são estudados, bem como a análise dos resultados obtidos com os mesmos. Tecnologia, porque a utiliza para tornar os resultados mais rápidos e eficazes para o usuário. Informação confiável, de acordo com as normas e padrões vigentes e dentro de tolerâncias e acurárias desejáveis. Imagens, onde imagens ópticas podem ser classificadas como a reprodução aparente de um objeto, formado por um sistema de lentes ou espelhos, a partir de ondas luminosas refletidas, refratadas ou difratadas”. Sensores Remotos, onde um sensor é “qualquer equipamento capaz de transformar alguma forma de energia em um sinal passível de ser convertido em informação sobre o ambiente, sem contato físico entre este sensor e os alvos de interesse”.

A Fotogrametria traz diversos benefícios aos cartógrafos, tais como a redução do trabalho de campo para levantamento de coordenadas; o menor tempo na obtenção de dados, uma vez que toda informação está na fotografia e o deslocamento no terreno em busca de informação se torna desnecessário; grande área de abrangência; etc. Entretanto, o objetivo principal da Fotogrametria é a reconstrução de um espaço tridimensional, chamado de espaço objeto, a partir de imagens bidimensionais, chamadas de espaço imagem.

A partir da reconstrução do espaço tridimensional do terreno, pode-se reconstruir o espaço físico tridimensional em um computador. Assim, o usuário poderá fazer diversas análises em um terreno qualquer como se estivesse realmente no local, sem custos de deslocamentos e sem enfrentar problemas de acesso ao mesmo.

3.2.1 **Orientação interior**

Uma vez que as imagens em estéreo obtidas se encontram isoladasumas das outras, salvas como arquivos digitais e sem nenhuma informação métrica, faz-se necessário a reconstrução do feixe perspectivo da câmara, ou seja, o

referenciamento desta imagem em relação a esta câmara. Desta forma, poderão ser efetuadas medidas com precisão sobre as imagens fotográficas. Conceitualmente, a orientação interior consiste apenas em colocar as imagens em posição semelhante à que exerciam dentro da câmara, no momento em que foram obtidas. (Brito & Coelho Filho, 2007).

A execução da orientação interior pressupõe a utilização de uma câmara métrica, cujos parâmetros de calibração e distorção das lentes tenha sido previamente calculados. Assim, é possível realizar a transformação do sistema de imagem digital (*pixels*) para o sistema métrico da câmara (milímetros). Dentre os modelos matemáticos utilizados para este fim, utiliza-se comumente a transformação afim plana, com 6 parâmetros, conforme apresentados nas equações 3.7 e 3.8.

$$x = a0 + a1.coluna + a2.linha \quad (3.7)$$

$$y = b0 + b1.coluna + b2.linha \quad (3.8)$$

onde x e y são coordenadas de um ponto no sistema métrico da câmara e *coluna* e *linha* são as coordenadas desse ponto no sistema de imagem digital.

Para se obter as transformações do sistema métrico da câmara para o sistema de imagem digital, utilizam-se as equações 3.7 e 3.8 rearranjadas:

$$coluna = \frac{b2.x - a0.b2 - a2.y + a2.b0}{a1.b2 - a2.b1} \quad (3.9)$$

$$linha = \frac{a1.y - a1.b0 - b1.x + a0.b1}{a1.b2 - a2.b1} \quad (3.10)$$

3.2.2 Orientação exterior

O objetivo primordial da orientação exterior é a obtenção da posição e da atitude do sensor ao coletar cada imagem fotográfica em relação ao referencial do espaço-objeto. Pode-se dizer, então, que uma imagem está orientada exteriormente

se são conhecidos seus seis parâmetros de orientação exterior: as coordenadas no espaço-objeto para o centro de perspectiva (X_0 , Y_0 e Z_0) da câmara ou sensor e os ângulos de rotação ou de atitude do sensor (φ , ω e κ) (Brito & Coelho Filho, 2007).

3.2.3 As equações de colinearidade

As duas equações de colinearidade podem ser consideradas a base da Fotogrametria digital, uma vez que relacionam os parâmetros da orientação exterior, as coordenadas fotográficas de um ponto e as coordenadas tridimensionais do mesmo ponto no sistema referencial do terreno ou do espaço-objeto. A figura 3.7 ilustra o princípio da colinearidade, que estabelece que no momento de tomada da fotografia, o ponto-objeto P , o centro de projeção O e o ponto-imagem p formam uma linha reta. A demonstração dessa formulação pode ser encontrada em (Brito & Coelho Filho, 2007).

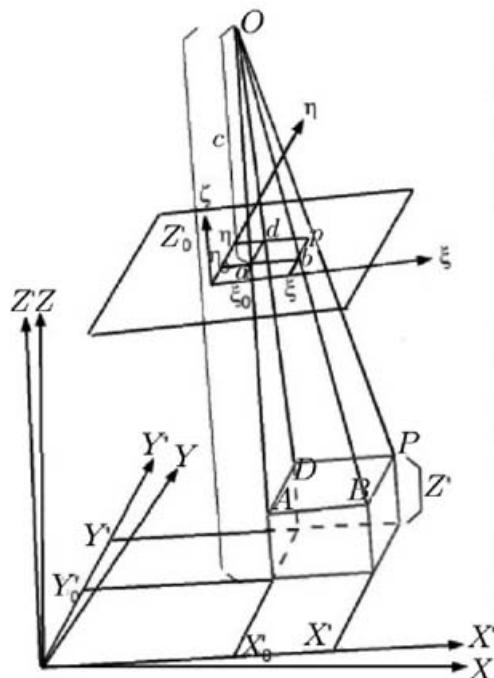


Figura 3.7: Condição de colinearidade para uma imagem. Extraído de (Brito & Coelho Filho, 2007).

Assim, a partir de um ponto 3D, é possível reprojetar este ponto no espaço-objeto a partir das equações de colinearidade:

$$\xi = \xi_0 - c \cdot \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \quad (3.11)$$

$$\eta = \eta_0 - c \cdot \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \quad (3.12)$$

onde:

- ξ é coordenada x de um ponto na imagem;
- ξ_0 é coordenada x do centro de perspectiva da câmara na imagem;
- η é coordenada y de um ponto na imagem;
- η_0 é coordenada y do centro de perspectiva da câmara na imagem; e
- r_{ij} elementos da matriz de rotação R (equação 3.13)

Para obter o valor das coordenadas no sistema de imagem digital, deve-se utilizar as equações 3.9 e 3.10.

A matriz de rotação R do espaço-imagem para o espaço-objeto é definida por:

$$R = \begin{bmatrix} \cos \varphi \cos \kappa & -\cos \varphi \sin \kappa & \sin \varphi \\ \cos \omega \sin \kappa + \sin \omega \sin \varphi \cos \kappa & \cos \omega \cos \kappa - \sin \omega \sin \varphi \sin \kappa & -\sin \omega \cos \kappa \\ \sin \omega \sin \kappa - \cos \omega \sin \varphi \cos \kappa & \sin \omega \cos \kappa + \cos \omega \sin \varphi \sin \kappa & \cos \omega \cos \varphi \end{bmatrix} \quad (3.13)$$

onde os elementos dessa matriz são representados da maneira a seguir:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (3.14)$$

3.2.4 Paralaxe estereoscópica

Paralaxe estereoscópica é definida como o deslocamento aparente da posição de um corpo ou objeto em relação a um ponto ou sistema de referência, causado pela mudança de ponto de observação (ASP, 1960). Em imagens aéreas, os objetos mais próximos em relação ao sensor (objetos mais altos) sofrerão menor deslocamento aparente do que os objetos mais distantes (objetos mais baixos), quando o ponto de vista for mudado. Na figura 3.8 observa-se que o ponto mais baixo da edificação (*A*) sofreu maior deslocamento aparente no espaço-imagem (*a* e *a'*) do que o topo (*b* e *b'*), em relação ao centro da imagem (*o* e *o'*).

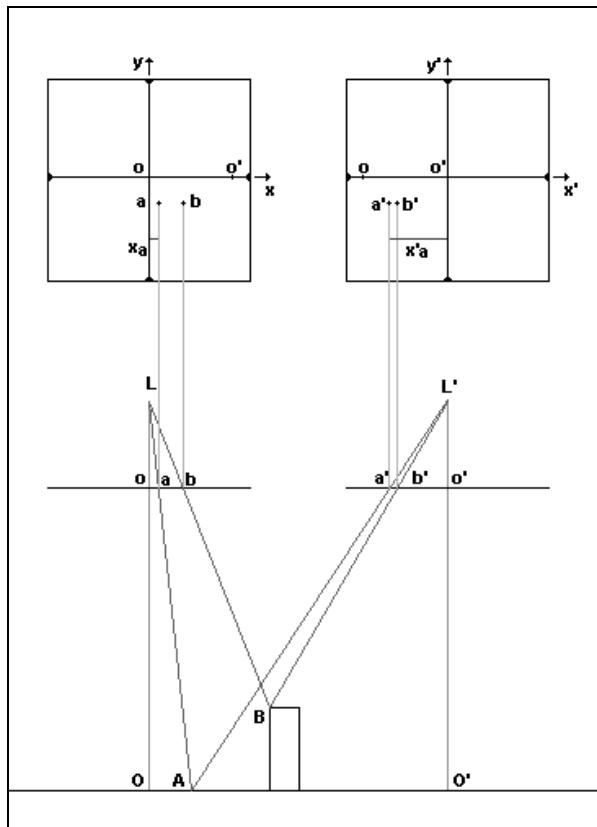


Figura 3.8: Deslocamento aparente da imagem de um objeto devido à paralaxe. Adaptado de: Lillesand & Kiefer (2007).

Em condições ideais, este deslocamento aparente deverá somente ocorrer em relação ao eixo *x* da câmara, conforme mostra a figura 3.8. Entretanto, em condições reais, este efeito também ocorre no eixo *y*. De forma a minimizar este

problema, aplica-se o processo de normalização de imagens (Schenk, 1999), cujo objetivo é reamostrar o par estereoscópico para a geometria epipolar. A geometria epipolar é materializada pela presença de um plano epipolar e de linhas epipolares. O plano epipolar é definido pelos dois centros de perspectiva das imagens e por um ponto no espaço-objeto. As linhas epipolares são as interseções do plano epipolar com os planos das imagens normalizadas (Brito & Coelho Filho, 2007). Entretanto, ressalta-se que o modelo proposto nessa Tese não exige que o par estéreo seja reamostrado para geometria epipolar.

3.2.5 A interseção espacial

Conhecendo os parâmetros da orientação exterior para um par de imagens contíguas que formam uma área em comum, isto é, que formam um modelo ou par estereoscópico, pode-se, a partir das mesmas equações de colinearidade, gerar as coordenadas tridimensionais de um ponto, referidas ao sistema do espaço-objeto. Como o ponto em questão é o mesmo, as incógnitas passam a ser suas coordenadas tridimensionais no espaço-objeto (X , Y e Z). As equações que permitem calcular a coordenada 3D de um ponto a partir de um par de pontos homólogos no espaço-imagem são:

$$X = X_{0_1} + (Z - Z_{0_1}) \cdot \frac{r_{11_1}(\xi_1 - \xi_{0_1}) + r_{12_1}(\eta_1 - \eta_{0_1}) - r_{13_1}c}{r_{31_1}(\xi_1 - \xi_{0_1}) + r_{32_1}(\eta_1 - \eta_{0_1}) - r_{33_1}c} \quad (3.15)$$

$$Y = Y_{0_1} + (Z - Z_{0_1}) \cdot \frac{r_{21_1}(\xi_1 - \xi_{0_1}) + r_{22_1}(\eta_1 - \eta_{0_1}) - r_{23_1}c}{r_{31_1}(\xi_1 - \xi_{0_1}) + r_{32_1}(\eta_1 - \eta_{0_1}) - r_{33_1}c} \quad (3.16)$$

$$X = X_{0_2} + (Z - Z_{0_2}) \cdot \frac{r_{11_2}(\xi_2 - \xi_{0_2}) + r_{12_2}(\eta_2 - \eta_{0_2}) - r_{13_2}c}{r_{31_2}(\xi_2 - \xi_{0_2}) + r_{32_2}(\eta_2 - \eta_{0_2}) - r_{33_2}c} \quad (3.17)$$

$$Y = Y_{0_2} + (Z - Z_{0_2}) \cdot \frac{r_{21_2}(\xi_2 - \xi_{0_2}) + r_{22_2}(\eta_2 - \eta_{0_2}) - r_{23_2}c}{r_{31_2}(\xi_2 - \xi_{0_2}) + r_{32_2}(\eta_2 - \eta_{0_2}) - r_{33_2}c} \quad (3.18)$$

$$Z = \frac{X_{0_2} - Z_{0_2} k_{x_2} + Z_{0_1} k_{x_1} - X_{0_1}}{k_{x_1} - k_{x_2}} \quad (3.19)$$

Onde:

$$k_{x_i} = \frac{r_{11_i}(\xi_i - \xi_{0_i}) + r_{12_i}(\eta_i - \eta_{0_i}) - r_{13_i}c}{r_{31_i}(\xi_i - \xi_{0_i}) + r_{32_i}(\eta_i - \eta_{0_i}) - r_{33_i}c}, \quad i = 1, 2, \dots$$

$$k_{y_i} = \frac{r_{21_i}(\xi_i - \xi_{0_i}) + r_{22_i}(\eta_i - \eta_{0_i}) - r_{23_i}c}{r_{31_i}(\xi_i - \xi_{0_i}) + r_{32_i}(\eta_i - \eta_{0_i}) - r_{33_i}c}, \quad i = 1, 2, \dots$$

X e Y podem ser obtidos a partir de uma das equações ou através das médias das equações 3.15 e 3.17 e às equações 3.16 e 3.18, respectivamente. Uma maneira de verificar se os pontos homólogos estão coerentes, é através da comparação entre os valores obtidos nesses pares de equações. Caso alguns dos pares de valores estejam discrepantes entre si, observa-se então um erro quanto a estes pontos homólogos. A figura 3.9 ilustra a interseção espacial vista no topo do bloco C do prédio da Universidade do Estado do Rio de Janeiro.

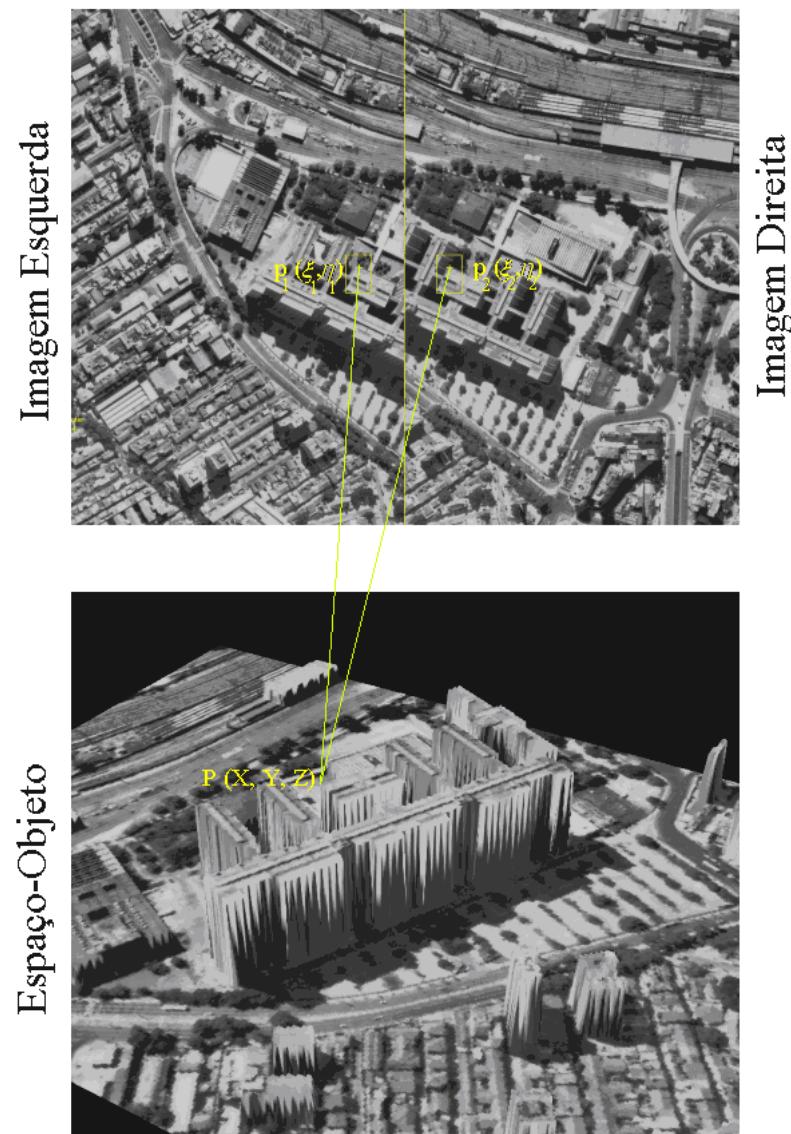


Figura 3.9: Interseção espacial.

3.2.6 Extração do Modelo Digital de Superfícies

Foi visto na seção anterior, o cálculo de uma coordenada 3D a partir de um par estéreo 2D através da interseção espacial. Para a reconstrução 3D completa de uma cena deve-se percorrer as imagens, de forma a encontrar todos os pontos correspondentes na área de sobreposição entre essas imagens.

Para a reconstrução 3D de uma cena, um operador humano deverá localizar os pontos ou contornos homólogos dos objetos presentes e digitalizá-los. Esta é uma tarefa complexa, pois em uma cena geralmente há milhões de pontos ou

centenas de objetos a serem representados em 3D. Em função disso, foram desenvolvidos métodos semi-automáticos que tentam desempenhar o papel humano na reconstrução 3D, sendo, porém, consideravelmente mais rápidos. Entretanto, a exatidão dos resultados e a interpretação da imagem ainda se encontram aquém das qualidades de medição do operador humano.

3.2.6.1

Tipos de representação dos dados 3D do terreno

Linder (2006) define dois tipos de modelagem do terreno: o Modelo Digital de Terreno (MDT), que contém os valores de coordenadas 3D, em especial a coordenada Z, situados diretamente sobre o terreno, e o Modelo Digital de Superfícies (MDS), que contém os valores de Z referidos ao topo dos objetos situados no terreno. A figura 3.10 apresenta as duas modelagens de terreno apresentadas.

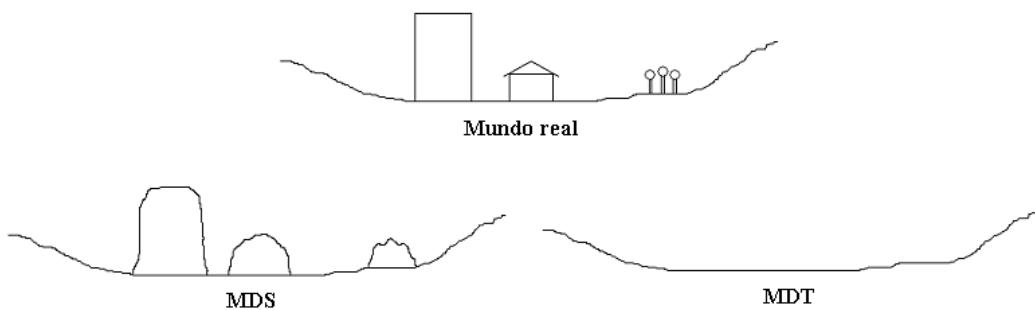


Figura 3.10: Tipos de modelagem de superfície (adaptado de Linder, 2006).

Os resultados da extração do MDS ou MDT é um conjunto de pontos cujas coordenadas 3D no referencial do espaço-objeto (terreno) são determinadas pela intersecção espacial anteriormente descrita. A partir destes, pode-se representar o MDS ou MDT de três formas distintas: nuvem de pontos, aquisição semi-regular e aquisição regular. A nuvem de pontos não segue ordem de aquisição, pois os ternos de coordenadas são adquiridos livremente no espaço a ser representado no MDS ou MDT. A aquisição semi-regular segue algum arranjo específico, sem, no entanto, possuir espaçamento constante entre os pontos. Já a forma regular é aquela em que todos os pontos estão igualmente espaçados em X e Y.

A conversão entre a representação de nuvem de pontos para a grade regular é feita através da interpolação do valor de Z dos pontos desta “nuvem”, de acordo com a coordenada X e Y da grade. Os principais métodos de interpolação de superfícies são:

- Vizinho mais próximo;
- Médias móveis;
- Superfícies de tendência; e
- Superfícies móveis.

O *software* de Sensoriamento Remoto ILWIS (2010) descreve alguns métodos de interpolação de superfícies e também dispõe o código fonte do programa com a implementação desses métodos. Outros métodos de interpolação de superfícies são descritos em Burrough e MacDonnell (1998) e Farin (2001).

3.2.6.2 Extração manual de feições

Na extração manual do modelo de superfícies, o operador humano é o responsável por identificar e digitalizar os objetos presentes em uma cena no par estereoscópico, baseado em três tipos de elementos geométricos: pontos, linhas e polígonos (figura 3.11). O ponto é a unidade de medição, que corresponde a um par de *pixels* homólogos. A linha é um conjunto de pontos conectados a dois outros pontos, com exceção do primeiro e último pontos que estão ligados somente a um ponto. O polígono é uma linha em que o último ponto é ligado ao primeiro.

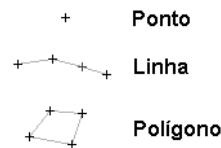


Figura 3.11: Tipos básicos de elementos geométricos utilizados na digitalização de imagens.

A medição de linhas corresponde à medição ponto a ponto de *pixels*

homólogos. Geralmente se utiliza esta forma geométrica para a medição de ruas, linhas férreas, pontes, etc, com base nas linhas de contorno da própria via. Esta medição, juntamente com a medição de pontos isolados, poderá ser utilizada para modelagem do terreno (onde as edificações e vegetações mais altas são ignoradas). A medição de polígonos é utilizada na medição de edificações. Para isso, medem-se também os contornos da edificação, conforme ilustrado na figura 3.12. Os polígonos serão utilizados como linhas de quebra (*break-lines*) na modelagem do terreno, ou seja, o valor da coordenada Z na área no espaço-objeto que estiver dentro do polígono será substituída do valor calculado para o valor de Z correspondente ao do polígono.

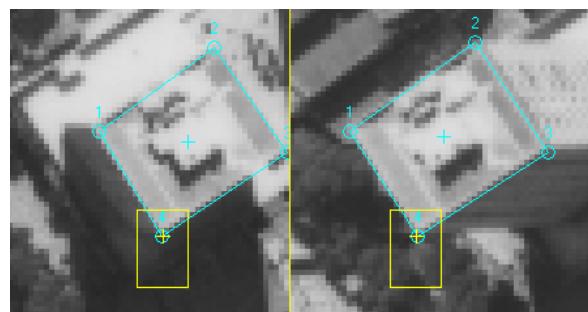


Figura 3.12: Medição manual de polígonos.

A figura 3.13 apresenta um exemplo de extração manual feita utilizando-se o software fotogramétrico livre E-FOTO (2011). Os módulos de Stereoplotter e DEM foram aperfeiçoados como contribuição resultante do desenvolvimento desta Tese.

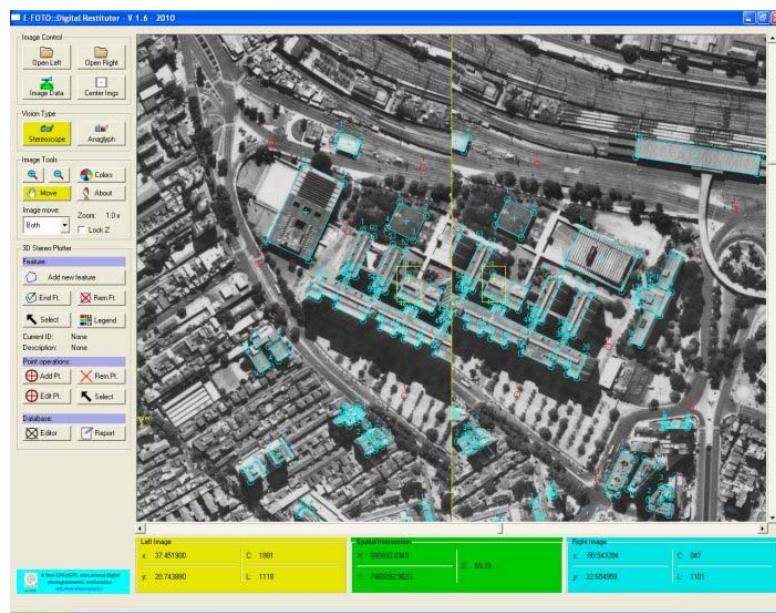


Figura 3.13: Extração manual utilizando o software livre E-FOTO (2011).

Uma vez digitalizados as áreas de interesse, a modelagem do terreno pode ser feita. No módulo de Extração do Modelo Digital de Superfícies do E-FOTO aplicam-se interpolações para os pontos e linhas medidos, enquanto que os polígonos são utilizados como linhas de quebra. A figura 3.14 mostra a extração do MDS através do E-FOTO.

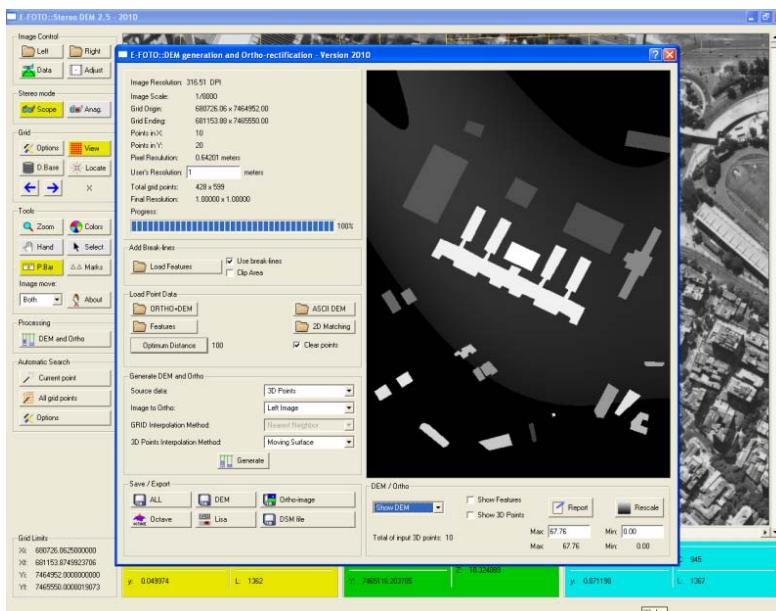


Figura 3.14: Geração do MDS no ambiente E-FOTO (2011).

3.2.6.3

Extração automática de feições

Os métodos de extração automática têm como objetivo substituir o trabalho humano na identificação de pontos e feições homólogas, uma vez que este processo é consideravelmente mais rápido. Entretanto, a extração automática ainda não é capaz de obter resultados tão acurados quanto na extração manual, além de ainda não ser capaz de interpretar e identificar corretamente os objetos presentes em uma cena, isto é, de reconstruir os seus aspectos semânticos, tal como a individualização dos prédios e edificações contidos num cenário urbano. São duas as principais abordagens: os métodos baseados em áreas e os métodos baseados em feições (Schenk, 1999).

Os métodos baseados em área usam como primitivas os níveis de cinza dos *pixels* da imagem. A partir de um ponto em uma imagem de referência, cria-se uma janela em torno deste ponto. Esta janela é chamada de *template*. Como a localização do *pixel* homólogo é desconhecida, cria-se-se uma janela de pesquisa, com dimensões consideravelmente maiores do que as dimensões do *template*, em torno da provável localização do ponto em questão na outra imagem, conforme ilustra a figura 3.15. Assim, um recorte com o mesmo tamanho do *template*, a janela de correspondência é deslizada sobre a janela de busca. Para cada iteração, um valor de correspondência é associado ao *pixel* central do recorte na janela de pesquisa. Ao final, será considerado como candidato a *pixel* homólogo na imagem de pesquisa aquele que tiver o maior grau de correspondência em relação ao *pixel* de referência. Para ser validada a correspondência entre os dois *pixels* o valor da correspondência deverá estar acima de um determinado limiar.

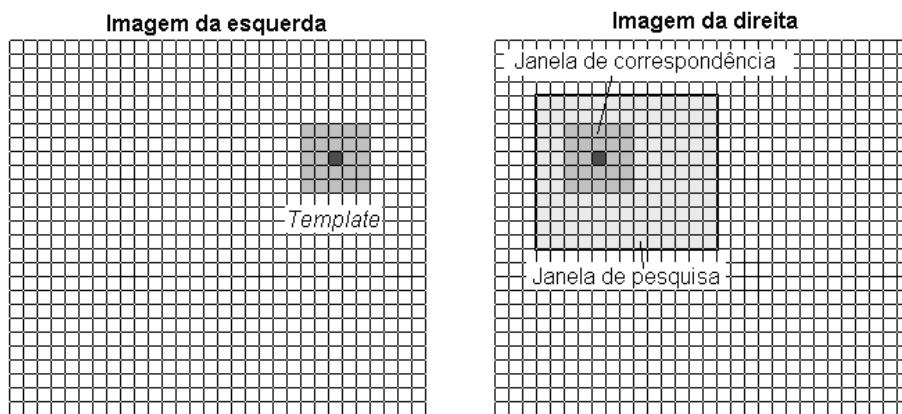


Figura 3.15: Mecanismo de busca por *pixels* homólogos no método baseado

em áreas (adaptado de Schenk, 1999).

A correlação cruzada, também chamada de correlação de Pearson, é baseada na equação 3.20, onde ρ é o coeficiente de Pearson que varia de -1,0 a 1,0. O valor de correlação igual a 1,0 indica que os dados são exatamente iguais, ou seja, que duas amostras são perfeitamente correlacionadas; o valor 0,0 indica que os dados são não correlacionados; e o valor -1,0 indica que um conjunto de dados é o complemento do outro.

$$\rho = \frac{\sum_{i=1}^n (A_i - \mu_a)(B_i - \mu_b)}{\sqrt{\frac{\sum_{i=1}^n (A_i - \mu_a)^2}{n-1}} \cdot \sqrt{\frac{\sum_{i=1}^n (B_i - \mu_b)^2}{n-1}}} \quad (3.20)$$

onde μ_a e μ_b são a média dos níveis de cinza das imagens A e B , respectivamente. Neste caso, considera-se a matriz do *template* e da janela de correspondência como vetores que variam de 1 até n , e i como um elemento do respectivo vetor.

A correlação cruzada não leva em consideração as distorções geométricas existentes de uma imagem do par em relação à outra. Com o objetivo de minimizar este problema, o LSM (Grün, 1985) foi desenvolvido de modo a considerar essas distorções, modelando as imagens através da transformação afim e obtendo uma localização mais exata do ponto homólogo.

Métodos baseados em feições são utilizados especialmente em visão computacional (p. ex. Trucco & Verri, 1998). Esses métodos envolvem pelo menos dois passos principais. Primeiramente são detectadas nas imagens feições de interesse, como pontos, linhas ou cantos (Mikolajczyk & Schmid, 2004). Em seguida, calculam-se descriptores dessas feições compostos por valores de um conjunto de seus atributos. Com base na similaridade entre descriptores, estabelecem-se os pareamentos (Mikolajczyk & Schmid, 2005).

Abordagens baseadas em feições são menos sujeitas a pareamentos múltiplos. A desvantagem em relação aos métodos baseados em área está em que o número de feições de interesse é, em geral, muito inferior ao número de pixels da imagem. Como consequência, a malha de pares de pontos homólogos

produzida é comparativamente mais esparsa. Já 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, razão pela qual a maioria dos algoritmos que produzem mapas densos de pontos homólogos utiliza algum método baseado em área.

Como uma palavra final a respeito da extração automática de feições, cita-se o fato de que esta operação pressupõe a não existência de oclusões no par de imagens estereoscópicas. Essas oclusões podem ter sua origem em ruídos nas imagens provenientes das condições atmosféricas (existência de nuvens) ou podem ter sido ainda causadas pelo deslocamento devido ao relevo. Brito (1997) em sua Tese de Doutorado, por exemplo, aborda o problema de detecção altimétricas de oclusões em imagens digitais.