

4 Registro e Correspondência de Imagens

Duas áreas de interesse em Visão Computacional são o registro e a correspondência de imagens. Estas duas áreas possuem muito em comum por buscarem regiões em comum entre duas imagens. Neste capítulo serão apresentadas algumas teorias e técnicas relativas a estas áreas que serão usadas nos procedimentos experimentais desenvolvidos.

O registro de imagens consiste da sobreposição de imagens de uma mesma cena obtidas em momentos, pontos de vista e até mesmo por sensores diferentes [36]. A correspondência de imagens consiste em classificar imagens mesmo que estas estejam em escalas, rotações, posições e com geometrias diferentes.

A comparação de imagens utilizando-se transformações invariáveis baseadas em Fourier e Mellin está descrita no capítulo 2. Neste capítulo foram apresentados alguns procedimentos de comparação de imagens que podem ser utilizados para se fazer a classificação de duas imagens. Porém, estes procedimentos comparam a imagem de modo global. Os métodos aqui discutidos tratam da correspondência de imagens através de características locais, ou seja, buscam a correspondência de regiões da imagem. Outra grande diferença entre os métodos aqui apresentados e os do capítulo 2 é que através destes é possível se descobrir um modelo que transforma uma das imagens na outra de modo aproximado.

Durante o registro de duas imagens, uma imagem é transformada para poder ser combinada com uma imagem de referência, ou, imagem base. Esta transformação deve remapear as posições dos *pixels* da imagem transformada de modo que a área sobreposta esteja alinhada com a imagem de referência.

Primeiramente, pontos em comum entre as duas imagens devem ser encontrados, estes são chamados de pontos de controle. A partir dos pontos de controle, estimam-se os parâmetros do modelo de transformação que irá gerar a imagem transformada que será sobreposta à imagem base. Por fim é feita a transformação e a sobreposição das imagens.

Uma etapa dos métodos de registro está em se fazer a correspondência de duas imagens. Isto consiste em encontrar pontos em comum entre as duas e a transformação que leva os pontos de uma imagem para a outra de modo que se correspondam. Pode-se entender então a correspondência de imagens como uma sub-área do campo de registro de imagens.

Um outro campo de Visão Computacional ainda não comentado é o campo de Fluxo Ótico [37]. Esta área trabalha com casos que requerem a determinação de um campo vetorial que descreva deslocamentos ocorridos em quadros subsequentes de uma sequência genérica de vídeo. A determinação do Fluxo Ótico é ainda um problema sem solução (como muitos dos problemas descritos no presente trabalho). Um trabalho nacional recente sobre o assunto pode ser visto em [38].

Este capítulo está dividido nas seguintes seções:

- 4.1 - Etapas de um Método de Registro: Apresentação dos passos comuns para métodos de registro;
- 4.2 – Detecção e Casamento de Pontos de Controle: Descreve como são procurados pontos em comum entre diferentes imagens;
- 4.3 – Transformação de Coordenadas de uma Imagem: Explica diferentes maneiras de se mapear uma imagem para que esta corresponda com outra e como se obtém a função que faz o mapeamento de uma imagem em outra. A seção é sub-dividida em:
 - 4.3.1 – Mínimos Quadrados;
 - 4.3.2 – Transformação Procrustes;
 - 4.3.3 – Transformação Afim;
- 4.4 – Refinando o Modelo de Transformação T: São discutidos diferentes métodos de se melhorar a função que mapeia a imagem.

São apresentados:

- 4.4.1 – Transformada de Hough;
- 4.4.2 – RANSAC;
- 4.4.3 – Reajustando a Matriz de Pesos W ;
- 4.4.4 – Determinando-se Dados Inconstantes Através de Limiar;

- 4.5 – Sobreposição de imagens Utilizando Ajuste Radiométrico dos *Pixels*. Esta seção descreve como é aplicado o modelo de transformação de modo a se sobrepor duas imagens;

4.1. Etapas de um Método de Registro

De um modo geral, métodos de registro consistem das seguintes etapas [39]:

- Detecção de pontos ou áreas de interesse: Objetos ou áreas distintos (regiões fechadas, arestas, contornos, intersecção de linhas, cantos) são identificados na imagem a ser transformada. Estes objetos serão representados por suas representações pontuais, denominadas de pontos de controle;
- Casamento dos pontos de controle: Nesta etapa, é encontrada a correspondência dos pontos de controle da imagem a ser transformada na imagem de referência. A idéia do casamento de pontos de controle é ilustrada na Figura 4-1. A detecção de pontos de controle e casamento de pontos entre imagens é apresentada na seção 4.2;
- Estimativa do modelo de transformação: São estimados o tipo e os parâmetros da função de transformação que mapeará a imagem para poder fazer o alinhamento com a imagem de referência. Na seção 4.3 são vistos alguns modelos de transformação possíveis tal como são calculados ;
- Refinando o modelo de transformação: O modelo de transformação visto na etapa anterior pode ser melhorado através da eliminação de pontos de controle inconsistentes como também através de novos cálculos. A seção 4.3 apresenta algumas técnicas que podem ser usadas para se refinar o modelo de transformação obtido;
- Transformação e sobreposição das imagens: A imagem é transformada pela função de transformação obtida. Os valores da imagem em coordenadas não discretas é computado por alguma técnica de interpolação. A transformação das imagens seguida da superposição das mesmas é ilustrada na Figura 4-2 e descrita na seção 4.5;

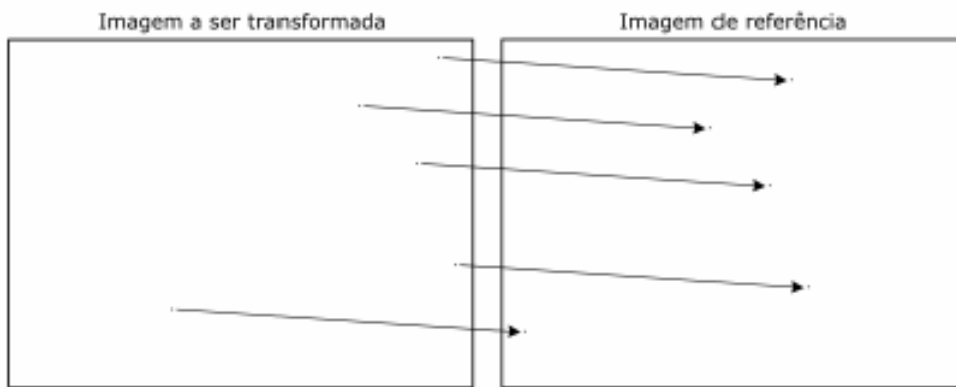


Figura 4-1: Casamento de pontos de controle

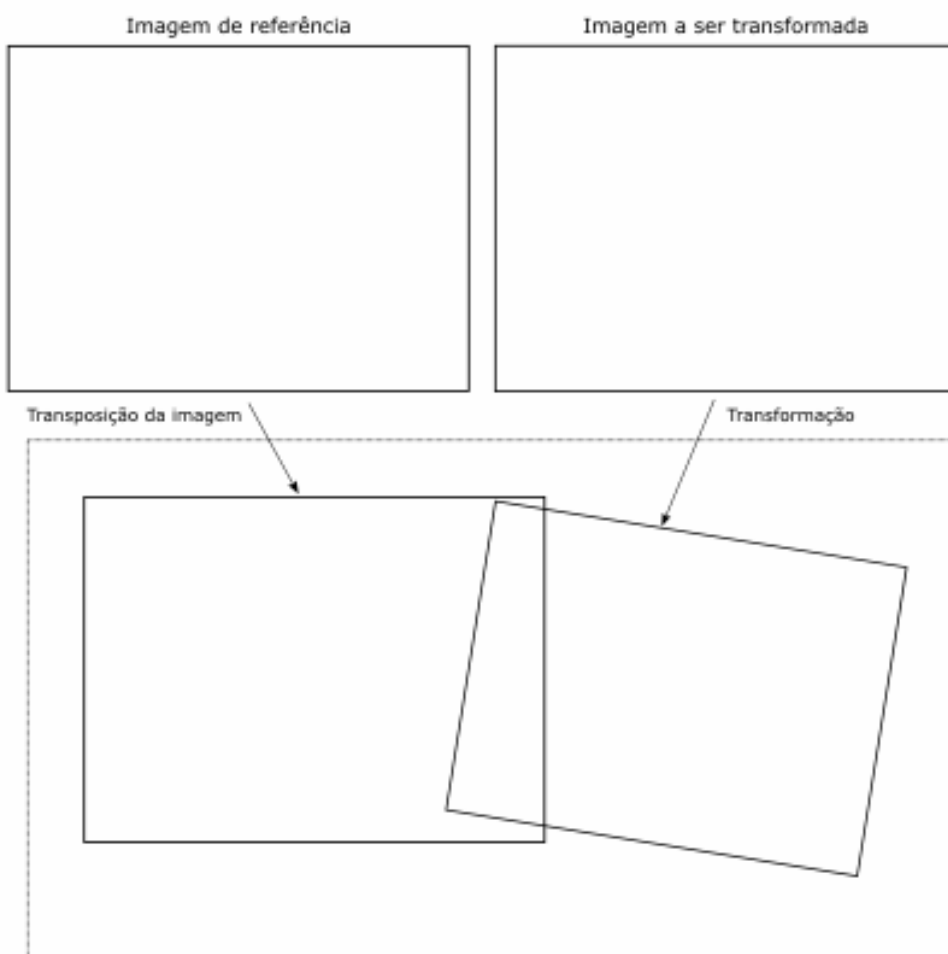


Figura 4-2: Registro de imagens

4.2.

Detecção e Casamento dos Pontos de Controle

Existem diversos métodos para identificar áreas em comum entre as duas imagens. Um método apresentado para se detectar e comparar pontos de controle é apresentado no capítulo 3 através da técnica SIFT. Neste capítulo também são citados alguns outros métodos de obtenção de descritores que podem ser utilizados para se comparar imagens.

Uma das maneiras mais simples em se casar pontos de controle consiste em usar a medida de Correlação Cruzada Normalizada. A Correlação Cruzada Normalizada entre uma imagem I e uma imagem base w é dada por:

$$CC(x',y') = \frac{\sum_x \sum_y [I(x,y) - \bar{I}(x,y)] [w(x-x',y-y') - \bar{w}]}{\left\{ \sum_x \sum_y [I(x,y) - \bar{I}(x,y)]^2 \sum_x \sum_y [w(x-x',y-y') - \bar{w}]^2 \right\}^{1/2}} \quad (129)$$

Obtêm-se a posição x,y de casamento de I na imagem base encontrando-se x' e y' que maximize $CC(x',y')$.

Note que a correlação apresentada aqui é um pouco diferente da vista na seção 2.6.2. Aquela se mostrou mais adequada para a aplicação em questão, mas, os dois meios são válidos como medida de correlação.

Uma maneira de se encontrar pontos de controle de uma imagem em outra, é se definir uma região vizinha ao ponto de controle e procurá-la na imagem base através da correlação cruzada normalizada. Repare que este método possui a desvantagem de não ser invariável a mudanças de rotação e escala, mas somente à translação. Em contrapartida, a correlação cruzada é computada com grande eficiência computacional.

A correlação cruzada é um dos métodos mais simples e comuns de se comparar imagens.

Uma aplicação da correlação cruzada para detecção e casamento de pontos de controle na geração de imagens panorâmicas será proposta na seção 6.2.2.

Após serem encontrados pontos em comum entre duas imagens, estes pontos são usados para se encontrar uma função que mapeie as posições dos pontos de uma imagem para as posições dos pontos correspondentes na outra imagem. O mapeamento de posições é discutido na próxima seção.

4.3. Transformação de Coordenadas de uma Imagem

Uma transformação geométrica mapeia as posições dos *pixels* de uma imagem para novas posições. Em essência, uma transformação geométrica faz o rearranjo da configuração espacial de uma imagem. A Figura 4-3 ilustra este conceito.

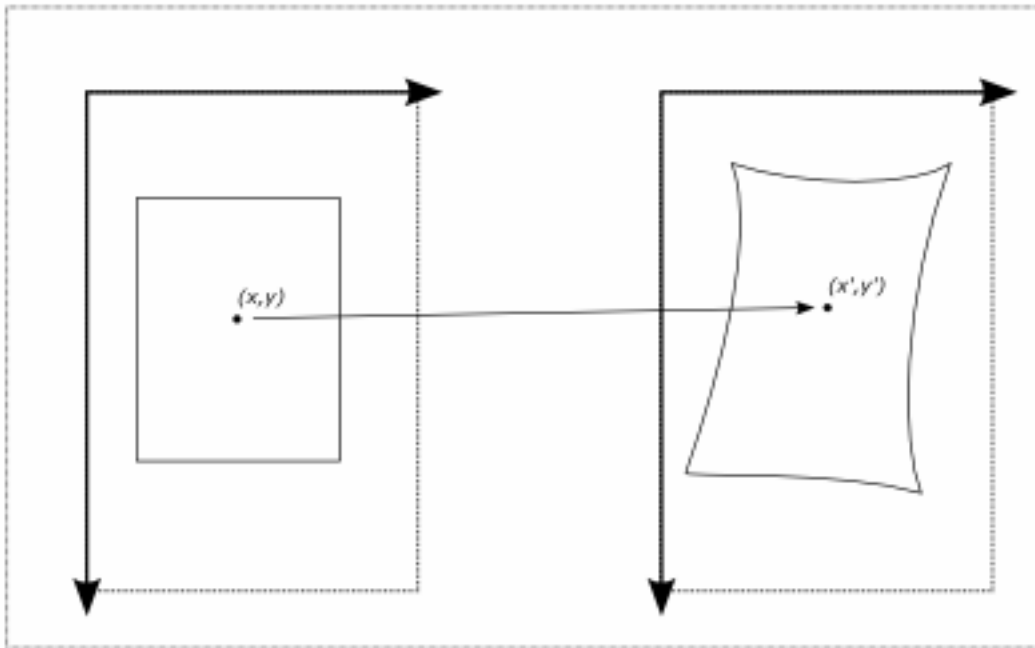


Figura 4-3: Transformação de uma imagem

Para uma transformação de coordenadas que mapeia $p = \begin{pmatrix} x \\ y \end{pmatrix}$ em $p' = \begin{pmatrix} x' \\ y' \end{pmatrix}$

dada por:

$$p' = T(p) \quad (130)$$

Ou,

$$p = T^{-1}(p') \quad (131)$$

A transformação geométrica da imagem $I(p)$ na imagem $I'(p')$ é dada por:

$$\begin{aligned} I'(p') &= I(p) \\ I'(T(p)) &= I(p) \end{aligned} \quad (132)$$

Ou:

$$I'(p') = I(T^{-1}(p')) \quad (133)$$

Perceba que há duas maneiras de se fazer a transformação. A primeira delas é dada por $I'(T(p)) = I(p)$ e é conhecida como método direto. O procedimento consiste em transferir as cores da imagem original para a imagem transformada. Cada *pixel* da imagem original resultará em um *pixel* na imagem transformada. Como o resultado de $T(p)$ pode não ser inteiro, é preciso fazer um rearranjo da posição dos *pixels* resultantes, como por exemplo trabalhar com o *pixel* vizinho mais próximo. Este procedimento tem a desvantagem de poder resultar em pontos em branco em $I'(p')$ pois nem todos os *pixels* de p' serão preenchidos a partir do resultado de $T(p)$.

Outro procedimento possível é trabalhar com a transformação pelo método indireto dado por $I'(p') = I(T^{-1}(p'))$. Através deste método cada *pixel* da imagem transformada será preenchido procurando-se suas cores na imagem original. Como o resultado de $T^{-1}(p')$ pode não ser inteiro, o valor de $I'_{x',y'}$ pode ser calculado pela interpolação bilinear dada por:

$$\begin{aligned} I'_{x',y'} &= ([x] + 1 - x)([y] + 1 - y)I'_{[x],[y]} \\ &+ (x - [x])([y] + 1 - y)I'_{[x+1],[y]} \\ &+ ([x] + 1 - x)(y - [y])I'_{[x],[y+1]} \\ &+ (x - [x])(y - [y])I'_{[x+1],[y+1]} \end{aligned} \quad (134)$$

Várias transformações de coordenadas são apresentadas na

Tabela 4-1 [40, 36].

As transformações Procrustes e Afim serão apresentadas com maiores detalhes nas seções 4.3.2 e 4.3.3 respectivamente. A seção seguinte mostra como calcular os parâmetros das transformações quando se conhece pontos que obedecem a correspondência dada pela transformação através de mínimos quadrados.

Tabela 4-1: Mapeando as coordenadas $p=(x,y)$ de uma imagem em novas coordenadas $p'=(x',y')$

Modelo	Transformação de coordenadas	Parâmetros
Translação	$p' = p + b$	$b \in \mathfrak{R}^2$
Escala e translação	$p' = qp + b$	$q \in \mathfrak{R}, b \in \mathfrak{R}^2$
Rotação e translação	$p' = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} p + b$	$b \in \mathfrak{R}^2$
Transformação Procrustes - Escala, rotação e translação	$p' = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} qp + b$	$q \in \mathfrak{R}, b \in \mathfrak{R}^2$
Transformação Afim	$p' = Ap + b$	$A \in \mathfrak{R}^{2 \times 2}, b \in \mathfrak{R}^2$
Transformação Projetiva	$p' = \frac{Ap + b}{c^T p + 1}$	$A \in \mathfrak{R}^{2 \times 2}, b, c \in \mathfrak{R}^2$
Transformação Bilinear	$\begin{pmatrix} x' \\ y' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + Bxy + b$	$A \in \mathfrak{R}^{2 \times 2}, B, b \in \mathfrak{R}^2$
Transformação Polinomial	$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{bmatrix} \sum a_{ij} x^i y^j \\ \sum b_{ij} x^i y^j \end{bmatrix}$	$a_{ij} \in \mathfrak{R}, b_{ij} \in \mathfrak{R}$

4.3.1. Mínimos Quadrados

Algumas transformações espaciais em imagens podem ser definidas por [40, 36]:

$$X' = TX \quad (135)$$

Para:

$$X = (x \ y \ 1)^T \quad (136)$$

$$X' = (x' \ y' \ 1)^T \quad (137)$$

Para problemas onde se deseja encontrar os parâmetros de uma transformação T , vista na eq. (135), em que se conhecem diversos pontos $p_i = [x_i, y_i]$ e correspondentes $p'_i = [x'_i, y'_i]$, tem-se:

$$P' = \begin{bmatrix} p'_1 \\ \vdots \\ p'_n \end{bmatrix} \quad (138)$$

Se parâmetros da matriz de transformação T são escritos em um vetor Z , e a eq. (135) é reescrita como:

$$P' = JZ \quad (139)$$

Onde J é a Matriz Jacobiana utilizada para descrever P' em termos dos parâmetros da matriz de transformação.

Procura-se a solução que aponta os melhores parâmetros de T que minimiza, para todo i , a soma dos erros entre as localizações mapeadas. Isto pode ser entendido da seguinte maneira, após T ter sido estimado, define-se para cada par de pontos i :

$$X'_i = (x'_i \quad y'_i \quad 1)^T \quad (140)$$

$$X_i = (x_i \quad y_i \quad 1)^T \quad (141)$$

Onde X' é referente à posição do ponto equivalente ao descritor cuja localização é dada por X . Com uma matriz T dada, a localização de X é mapeada para \hat{X} como se segue:

$$\hat{X} = TX \quad (142)$$

Onde:

$$\hat{X}_i = (\hat{x}_i \quad \hat{y}_i \quad 1)^T \quad (143)$$

O erro quadrático residual entre a localização mapeada por T e a localização do descritor dada por X' é calculado como:

$$e_i = \sqrt{(\hat{x} - x')^2 + (\hat{y} - y')^2} \quad (144)$$

Ou:

$$e_i = \sqrt{(X'_i - TX_i)^T (X'_i - TX_i)} \quad (145)$$

Pretende-se estimar os parâmetros de T que minimizem a soma dos erros quadráticos dada por:

$$E = \sum e_i \quad (146)$$

A solução é:

$$Z \equiv J^+ P' \quad (147)$$

Onde J^+ é a matriz pseudoinversa de J :

$$J^+ = \left(J^T J \right)^{-1} J^T \quad (148)$$

Caso sejam atribuídos pesos as correspondências p_i e p'_i , a eq. (139) pode ser reescrita como:

$$W(P') = W(JZ) \quad (149)$$

Onde W é uma matriz diagonal com os pesos escritos em sua diagonal:

$$W = \begin{bmatrix} W_1 & & & & & \\ & \ddots & & & & \\ & & W_i & & & \\ & & & \ddots & & \\ & & & & & W_k \end{bmatrix} \quad (150)$$

Observe que $1 < i < k$.

A solução aqui é dada pela eq. (147) reescrita como:

$$Z \equiv \left((WJ)^T (WJ) \right)^{-1} (WJ)^T W P' \quad (151)$$

Quando a matriz W é a matriz identidade, a eq. (151) se iguala a eq.(147).

A próxima seção apresentará com maiores detalhes a transformação Procrustes e a aplicação da matriz pseudoinversa para cálculo de seus parâmetros.

4.3.2. Transformação Procrustes

A transformação Procrustes [40, 36] é uma transformação de 4 parâmetros que leva em consideração mudanças de escala, rotação e translação entre imagens como ilustrado na Figura 4-4.

A transformação Procrustes pode ser escrita como:

$$\begin{bmatrix} p' \\ 1 \end{bmatrix} = T \begin{bmatrix} p \\ 1 \end{bmatrix}_3 \Rightarrow \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (152)$$

Onde:

$$T = \begin{bmatrix} a \cos(\theta) & a \sin(\theta) & \Delta x \\ -a \sin(\theta) & a \cos(\theta) & \Delta y \\ 0 & 0 & 1 \end{bmatrix} \quad (153)$$

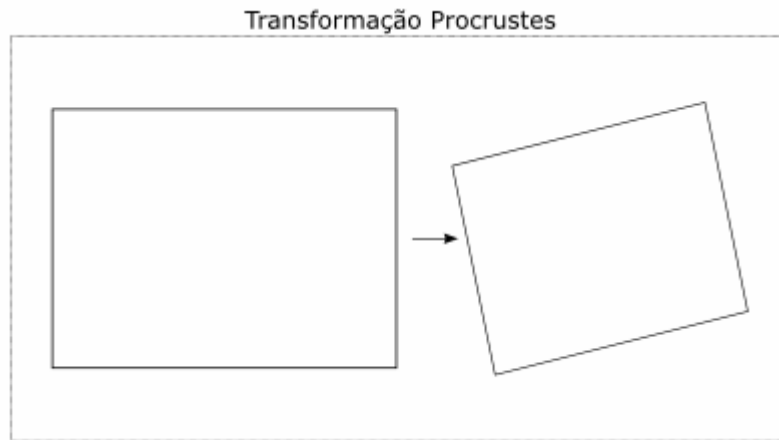


Figura 4-4: Tranformação Procrustes

O parâmetro a realiza mudanças em escala. Quando $a = 1$, não há variação quanto à escala. Para $a > 1$, a imagem é dilatada proporcionalmente. E para $a < 1$, a imagem é reduzida proporcionalmente.

O parâmetro θ cuida da rotação da imagem e está em radianos.

Os parâmetro Δx e Δy realizam as translações horizontal e vertical respectivamente.

Para a transformação Procrustes podemos escrever Z como:

$$Z = \begin{bmatrix} a \sin \theta \\ a \cos \theta \\ \Delta x \\ \Delta y \end{bmatrix} \quad (154)$$

E a matriz Jacobiana J como:

$$J = \begin{bmatrix} y_1 & x_1 & 1 & 0 \\ -x_1 & y_1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ y_n & x_n & 1 & 0 \\ -x_n & y_n & 0 & 1 \end{bmatrix} \quad (155)$$

Outra transformação de interesse é a transformação Afim que será vista na seção a seguir.

4.3.3. Transformação Afim

A transformação Afim [40, 36] é uma generalização de 6 parâmetros da transformação Procrustes que leva em consideração escala, rotação, translação e não ortogonalidade entre eixos, permitindo dilatações e reduções ao longo das linhas e colunas de uma imagem como visto na Figura 4-5.

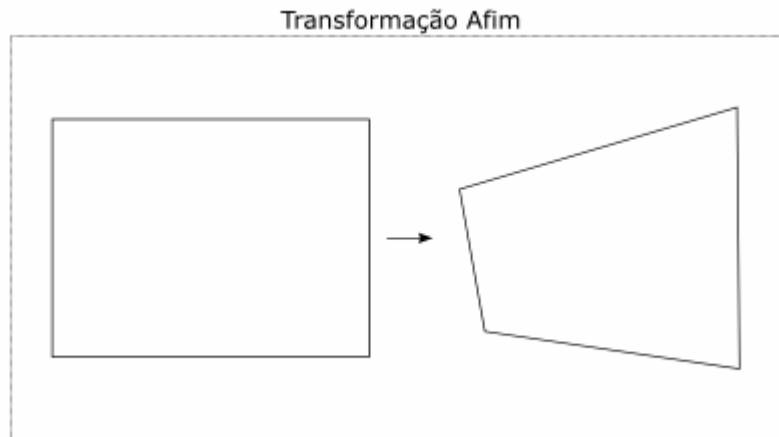


Figura 4-5: Transformação Afim

Esta é a transformação linear mais geral de todas e é definida por:

$$T = \begin{bmatrix} a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \\ 0 & 0 & 1 \end{bmatrix} \quad (156)$$

Para a transformação Procrustes podemos escrever Z como:

$$Z = \begin{bmatrix} a_{10} \\ a_{11} \\ a_{12} \\ a_{20} \\ a_{21} \\ a_{22} \end{bmatrix} \quad (157)$$

E a matriz Jacobiana J como:

$$J = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_n & y_n & 1 \end{bmatrix} \quad (158)$$

Os modelos de transformação T apresentados minimizam os erros quadráticos para os pontos utilizados. Porém, pontos com falsas correspondências induzem ao cálculo de matrizes T que não representam realmente a transformação entre duas imagens. De modo a se refinar o modelo calculado, a seção seguinte apresenta algumas técnicas possíveis.

4.4. Refinando o Modelo de Transformação T

Para procedimentos onde são encontrados pontos em comum entre imagens e se deseja computar a matriz de transformação que mapeia as posições dos pontos de uma imagem para outra imagem, é interessante refinar o modelo T encontrado com auxílio da matriz de pesos W apresentada na eq. (150) ou através de eliminação de pontos mal correspondidos. Para tal, são apresentados aqui alguns procedimentos.

As técnicas aqui apresentadas serão utilizadas para se gerar imagens panorâmicas e também junto aos procedimentos de navegação que trabalham com o método SIFT.

No caso de uso do método SIFT, as técnicas a seguir são utilizadas após o processo de encontrar pontos em comum utilizando a técnica SIFT apresentada na seção 3.2.6.

Serão apresentados métodos relacionados à:

- Transformada de Hough;
- RANSAC (*Random Sample Consensus Algorithm*);
- Reajuste da matriz de pesos W ;
- Eliminação de pontos através de limiar de erro;

4.4.1. Transformada de Hough

A Transformada de Hough [41] é uma técnica de extração de características chaves em uma imagem. A técnica clássica é utilizada para a identificação de linhas e curvas em imagens, porém pode ser estendida para a identificação de posições de formas diversas ou, como no caso aqui descrito, transformações diversas.

A Transformada de Hough pode ser entendida de um modo genérico como uma tabela de parâmetros que compusesse um modelo. A tabela seria preenchida para cada dado de um conjunto de dados apresentado, encontrando todos os modelos possíveis que condissessem com cada ponto (ou sub-conjunto de pontos) e atualizando a tabela incrementando às células referentes aos possíveis parâmetros dos modelos encontrados.

A Transformada de Hough utilizada aqui [25] pode ser entendida como um histograma em 4 dimensões que comporte os parâmetros Δx , Δy , a e θ de uma transformação Afim. Poderia ser pensada também como uma tabela com todas as possíveis combinações dos parâmetros dados, sendo estes discretizados.

Perceba que para cada par de descritores equivalentes encontrado é possível se obter uma transformação Procrustes ou uma transformação Afim. São obtidos então os parâmetros Δx , Δy , a e θ para cada par de descritores encontrado.

A tabela utilizada para se fazer a Transformada de Hough pode ser então criada da seguinte maneira:

- Cria-se uma matriz de 4 dimensões onde cada dimensão representa um parâmetro dentre Δx , Δy , a e θ .
- Para θ , faz-se a discretização de 30° em 30° . Ou seja, cada intervalo na dimensão θ é relativo a 30° . A dimensão definida por θ é dividida em 12 intervalos;
- As escalas a possíveis são dadas por α sendo α inteiro e $-8 \leq \alpha \leq 8$. Veja que são 17 intervalos relativos a a ;
- Δx , Δy tem intervalos de tamanho 0.25 vezes o tamanho da maior imagem;
- A transformada será obtida através dos seguintes passos:
- Para cada ponto em uma imagem, encontre o ponto correspondente para a imagem 2. Encontre então a Transformada Procrustes para o par de coordenadas. Desta maneira são obtidos os parâmetros Δx , Δy , a e θ da transformação deste par de coordenadas;
- Incremente a célula da tabela correspondente aos parâmetros utilizados. A célula é incrementada com a multiplicação de uma série de pesos. Cada peso é calculado pela distância entre cada parâmetro calculado e o valor discreto relativo à célula. Este peso é

normalizado de modo similar à eq. (122), com a diferença de que há um peso para cada parâmetro do modelo;

- Também são incrementadas as células vizinhas com relação a cada parâmetro. Elas são incrementadas da mesma forma que no passo anterior. Para cada parâmetro são adicionados os 4 vizinhos discretos mais próximos do valor obtido. Ao todo são atualizadas $4 \times 4 \times 4 \times 4$ células, ou seja, 16 células para cada par de pontos.

Ao terminar de preencher a tabela, a célula, ou as células com maior peso, são utilizadas para se prosseguir.

Uma alternativa possível é de se trabalhar com um limiar, e aceitar todas as células com valor acima do limiar. Todas as outras são descartadas do processo.

A partir de então, cria-se a matriz de pesos W , onde cada peso, referente a cada par de pontos, é relativo aos pesos encontrados para as células que corresponderam àquele par de pontos. Uma maneira de fazer isso é repetir o mesmo par de pontos, para cada célula referente a este que ficou acima do limiar, e para cada aparição do par de pontos, se usar o peso relativo ao valor da célula.

Esta alternativa utilizada aqui pode ser usada em conjunto com RANSAC, que será apresentado na seção seguinte, e depois pode-se recalcular W recursivamente até se obter T aceitável, como será visto na seção 4.4.3.

Outra alternativa é escolher os pontos relativos à célula, ou as células com maior peso e calcular a transformada afim a partir destes pontos sem usar uma matriz de peso W . Então, pode-se eliminar pontos com erro residual muito grande e repetir o processo por um número de vezes, até se encontrar erros abaixo de um limiar escolhido.

4.4.2. RANSAC (*Random Sample Consensus Algorithm*)

O algoritmo RANSAC é apresentado em [42, 43] como um método para se estimar os parâmetros de um modelo para um conjunto de dados conhecido, porém com presença de diversos dados errôneos.

O algoritmo não é aplicado independentemente porque sua eficiência é limitada quando há um grande número de dados, não levando a uma boa estimativa do modelo. Por isso, é sugerido que se aplique anteriormente Hough,

ou outra técnica, conseguindo-se gerar um conjunto de dados mais robusto que o inicial.

RANSAC é de um algoritmo bem simples definido como se segue.

Dado um modelo com parâmetros \vec{x} , deseja-se estimá-los. Para tal, é assumido:

- Os parâmetros podem ser estimados a partir de um número N de itens em um conjunto de dados conhecido;
- Existe um total de M itens no conjunto de dados;
- A probabilidade de um dado selecionado aleatoriamente fazer parte de um bom modelo é dada por p_g ;
- A probabilidade de que o algoritmo termine sem que se encontre um bom modelo é dada por p_{falha} ;

O algoritmo é então executado através das seguintes etapas:

1. N itens são escolhidos de modo aleatório;
2. A partir dos itens escolhidos, \vec{x} é estimado;
3. Encontra-se o número de itens que se enquadram ao modelo para determinada tolerância especificada. Este número é chamado de K ;
4. Caso K seja grande o suficiente, para um limiar escolhido, o algoritmo termina e foi bem sucedido;
5. O algoritmo é repetido de 1 a 4 um número L de vezes;
6. Caso o algoritmo não tenha terminado após L tentativas, o algoritmo falhou;

L pode ser encontrado através das seguintes maneiras:

- $p_{falha} =$ Probabilidade de L falhas consecutivas;
- $p_{falha} =$ (Probabilidade de que determinado dado seja falho)* L ;
- $p_{falha} = (1 - \text{Probabilidade de sucesso})*L$;
- $p_{falha} = (1 - (\text{Probabilidade de que um item caiba no modelo})*N)*L$;
- $p_{falha} = (1 - (p_g) * N) * L$;

O algoritmo RANSAC é adaptado para se encontrar o modelo da transformação afim de T da seguinte forma:

1. São escolhidos 3 descritores di_1 e dj_2 correspondentes;
2. A partir dos descritores escolhidos, calcula-se a transformação afim T a partir da eq. (151) usando-se pesos calculados para formar W , ou

a matriz identidade caso RANSAC esteja sendo aplicado independentemente;

3. Calcula-se a mediana dos erros residuais em se usar a transformação T dados pela eq. (144);
4. Caso a mediana calculada tenha sido a menor até então, T é guardado ou então substitui a última transformação T guardada;
5. O algoritmo é repetido de 1 a 4 um número L de vezes;
6. O T que apresentou a menor mediana de erros residuais é utilizado;

L é calculado fazendo-se:

$$P_{sucesso} = \left(1 - (P_{item_ruim})^N\right)^L \quad (159)$$

Onde:

- $P_{sucesso}$: Probabilidade de que RANSAC seja bem sucedido. Deseja-se que seja 1;
- N : Número de itens usados para se calcular T a cada passagem de RANSAC. Como visto, N aqui é usado igual a 3;

Portanto:

$$L = \frac{P_{sucesso}}{\left(1 - P_{item_ruim}^3\right)} \quad (160)$$

Obviamente este valor é aproximado. Aqui ele é arredondado para o primeiro inteiro superior.

Normalmente, após se aplicar RANSAC, é feito um procedimento de reajuste da matriz T de modo a se recalculer T . Este procedimento é apresentado a seguir.

4.4.3. Reajustando a Matriz de Pesos W

Este procedimento apresentado consiste em reajustar iterativamente uma matriz de pesos W' de modo a se melhorar a estimativa de T a partir da distribuição gaussiana dos resíduos encontrados na computação do modelo T por mínimos quadrados.

O procedimento aqui apresentado não é altamente eficiente para um número grande de dados inconsistentes, portanto, é interessante a aplicação de Hough e/ou RANSAC anteriormente.

A partir do método proposto em [44] é definido um algoritmo para se encontrar W' que ajude a interpolar um modelo T robusto de modo a se conseguir eliminar dados que levam a uma falsa estimativa do modelo.

De modo a se descrever o qual admissível é um par de pontos i , é estabelecida uma medida de verossimilhança que será usada para se pesar este par de pontos na obtenção de T por mínimos quadrados. Quando se calcula o erro residual dado pela eq. (144), espera-se que aqueles dados que possuem maior erro tenham maior chance de serem inconsistentes. Portanto, é definido uma medida de verossimilhança dada pela distribuição gaussiana do erro residual e_i :

$$p(e_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{e_i^2}{2\sigma^2}} \quad (161)$$

Onde $p(e_i)$ pode ser descrito como a probabilidade do dado i ser consistente.

A estimativa de escala σ é baseada na mediana dos resíduos dados pela eq. (144) como proposto em [44] :

$$\sigma = 1.482 \cdot \text{median}(e_i) \quad (162)$$

Pode-se montar então uma matriz W' dada por:

$$W' = \hat{W}W \quad (163)$$

Onde a matriz de peso W foi computada por procedimentos anteriores, ou então é igual à matriz identidade, e \hat{W} é uma matriz de pesos onde:

$$\hat{W}_i = p(e_i) \quad (164)$$

Para dados em que $e_i/\sigma > 5$, faz-se $\hat{W}_i = 0$.

Para se definir T robustamente, o procedimento a seguir é realizado. \hat{W} é inicialmente usado como matriz identidade.

1. Computa-se T através de mínimos quadrados com matriz de pesos W' ;
2. Calculam-se os erros residuais e_i ;
3. Calcula-se σ , $p(e_i)$ e consecutivamente \hat{W} ;
4. Para dados em que $e_i/\sigma > 5$, faz-se $\hat{W}_i = 0$;

5. Repete-se de 1 a 4 um número n de vezes;

É indicado $n = 10$.

O último procedimento a ser apresentado é o mais simples de todos e pode ser aplicado em conjunto com os métodos já apresentados, trata-se da eliminação de dados inconsistentes através de limiar relativo ao erro residual. Este procedimento é visto a seguir.

4.4.4.

Determinando-se Dados inconsistentes Através de Limiar

O último método aqui apresentado de se descartar pontos inconsistentes para o cálculo da matriz T é através de um limiar definido para o erro residual.

A maneira mais simples de determinar que dados são inconstantes é definir um limiar de erro residual para o qual estes sejam descartados. Após ter se encontrado a matriz T , utilizando ou não algum dos métodos apresentados nas seções 4.4.1, 4.4.2 e 4.4.3, ou outro, calcula-se o erro residual para cada ponto utilizado. Todo ponto que apresentar erro superior ao limiar especificado é eliminado.

Após o descarte, T deve ser recalculada. Este procedimento costuma ser o último aplicado para se melhorar T calculado.

Por fim, a próxima seção descreve como utilizar do modelo T encontrado para se combinar imagens. Este processo é particularmente de interesse na geração automática de imagens panorâmicas que será vista no capítulo 6 , na seção 6.2.

4.5. Sobreposição das Imagens Utilizando Ajuste Radiométrico dos Pixels Superpostos

Ao concatenar duas imagens é preciso decidir como trabalhar a área em que as imagens se superpõem. A

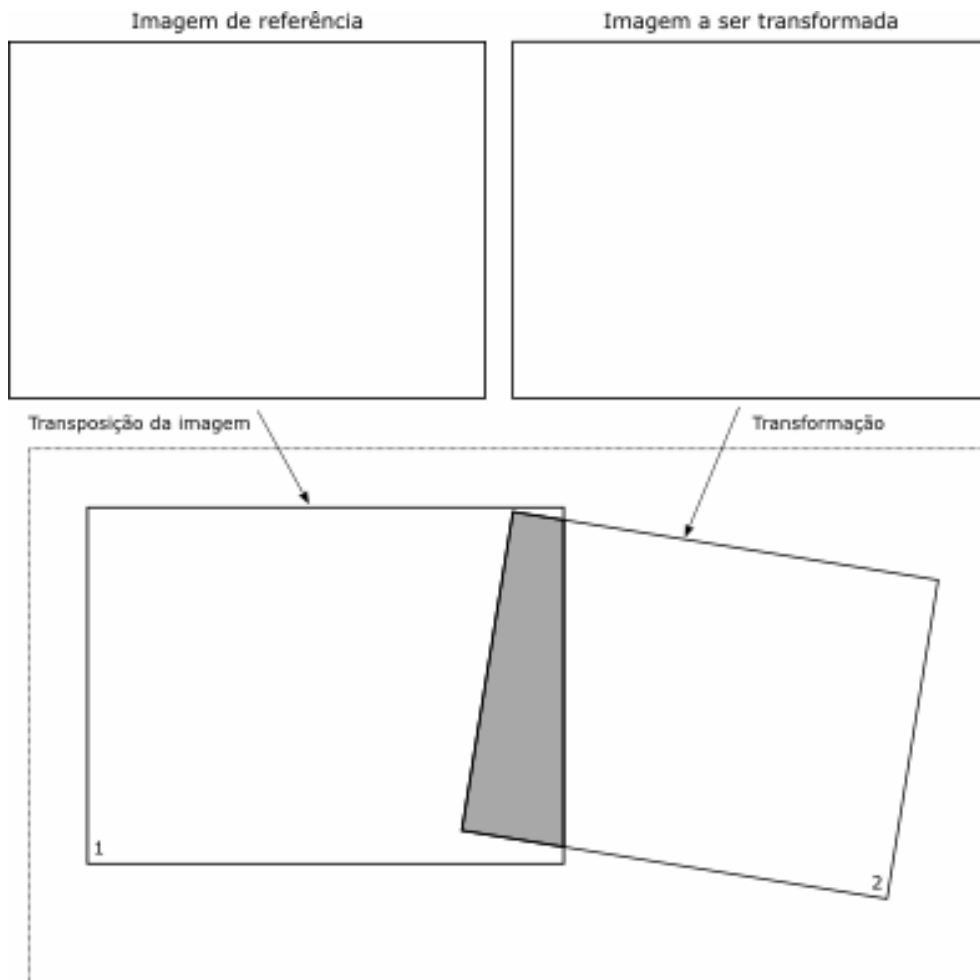


Figura 4-6 ilustra a sobreposição de imagens. Escolher trabalhar com uma imagem sobre a outra leva a possíveis discontinuidades na região de limite entre as imagens. Portanto, uma alternativa é fazer com que as cores de cada *pixel* sejam uma média ponderada entre as duas imagens.

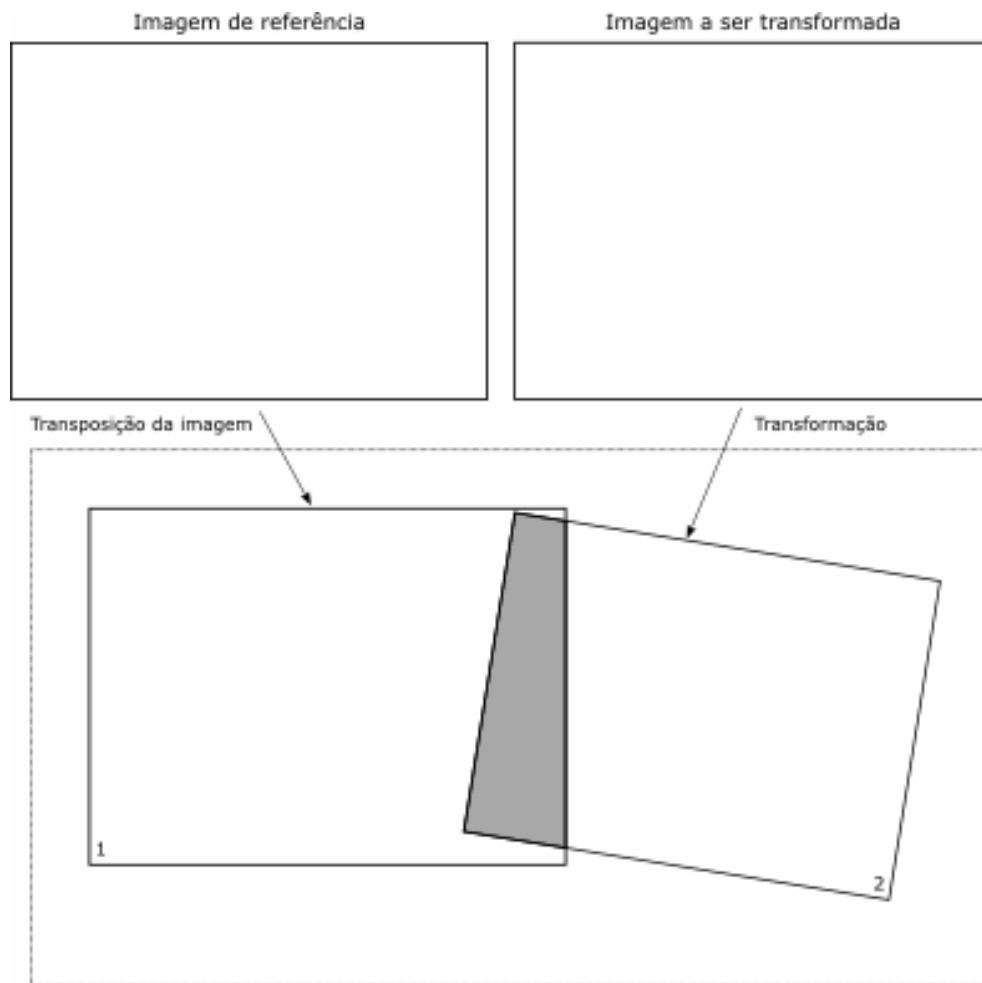


Figura 4-6: Superposição das imagens

Uma solução geral seria ponderar as imagens quanto à distância de cada *pixel* para as bordas de cada imagem. Esta solução ainda leva a algumas discontinuidades visíveis.

Outra possibilidade pode ser utilizada com a transformação Procrustes, criando-se um *dégradée* a partir do começo de uma das imagens até o fim da superposição com inclinação dada por θ . Este ajuste pode ser visto na Figura 4-7 e foi o mais utilizado durante o projeto.

O próximo capítulo apresentará o sistema experimental utilizado e o capítulo 6 passará a utilizar as técnicas discutidas até então para construir o modelo de exploração proposto.

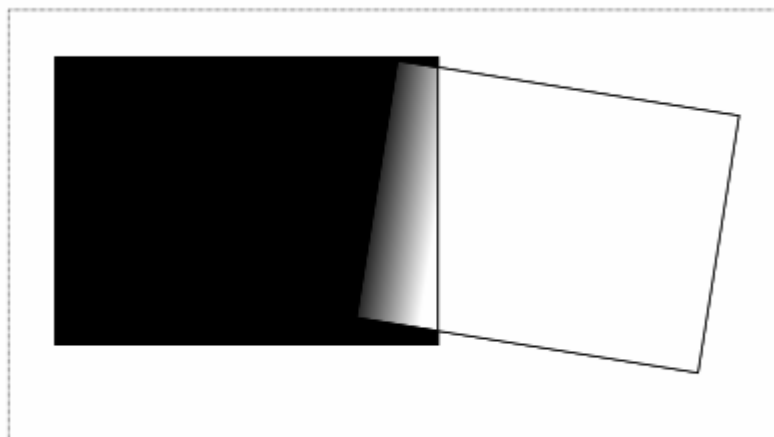


Figura 4-7: Exemplo de ajuste radiométrico