



Introdução ao Processamento Digital de Imagem MC920 / MO443

Prof. Hélio Pedrini

Instituto de Computação

UNICAMP

<http://www.ic.unicamp.br/~helio>

1º Semestre de 2022

- 1 Registro de Imagens
 - Transformadas Geométricas
 - Transformações Espaciais
 - Interpolação de Intensidade

- *Registro de imagens* é o processo de correspondência ou alinhamento entre duas ou mais imagens capturadas da mesma cena, porém, obtidas por diferentes sensores, em diferentes instantes de tempo ou sob diferentes pontos de observação.
- A operação de registro é fundamental em processamento e análise de imagens, auxiliando as etapas de identificação e reconhecimento de objetos nas imagens.

- Diversos domínios de conhecimento utilizam técnicas de registro de imagens.
 - ▶ sensoriamento remoto: aplicações incluem a detecção de mudanças em uma sequência de imagens, a fusão de imagens com características diferentes para permitir que informações sejam adequadamente integradas em termos de resolução espacial ou resolução espectral, a construção de mosaicos a partir de imagens de satélites, a integração de imagens de satélite ou fotografias aéreas com mapas e a localização de regiões de interesse, como edificações e rodovias, em imagens digitais.
 - ▶ área médica: aplicações incluem a combinação de dados obtidos a partir de diferentes técnicas de aquisição, como tomografia computadorizada, tomografia por emissão de pósitrons, ultrassonografia ou ressonância magnética, com o objetivo de obter informações mais completas sobre o paciente.
 - ▶ visão computacional: várias aplicações beneficiam-se das técnicas de registro, tais como a recuperação de informação tridimensional por meio de visão estereoscópica, o rastreamento de objetos em movimento e a reconstrução de objetos.

Registro de Imagens

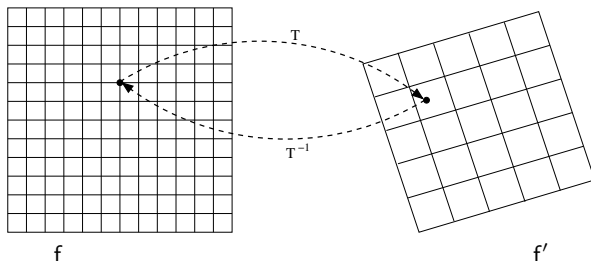
Transformações Geométricas em Imagens

- *Transformações geométricas* são operações cujo principal objetivo é permitir o mapeamento entre posições espaciais dos pixels da imagem de referência e a imagem modificada.
- Uma transformação geométrica consiste em duas operações básicas, uma *transformação espacial* que define a reorganização dos pixels sobre o plano da imagem e uma *interpolação de intensidade* que trata da atribuição dos níveis de cinza ou cores aos pixels da imagem transformada espacialmente.

Registro de Imagens

Transformações Espaciais

- Uma *transformação espacial* é uma função de mapeamento que permite a correspondência entre pontos de duas imagens.
- A figura a seguir ilustra a transformação espacial entre duas imagens bidimensionais f e f' .



- A transformação nem sempre resulta em um mapeamento de pontos com coordenadas inteiras: esse problema é encontrado, por exemplo, quando certas transformações de mudança de escala ou rotação são aplicadas a imagens.

- Seja um ponto P na imagem original f que sofreu uma transformação para uma nova posição P' na imagem resultante f' .
- A transformação espacial entre esses pontos pode ser realizada de duas formas, por *mapeamento direto* ou *mapeamento indireto*.
- No mapeamento direto, determina-se a localização do ponto P na imagem resultante por meio da transformação $T : P \mapsto P'$, a qual pode também ser expressa como

$$P' = T P \quad (1)$$

- Uma dificuldade comum que ocorre no mapeamento direto é que, devido ao problema de discretização da imagem, diferentes pontos da imagem original podem ser mapeados para um mesmo ponto, resultando em falhas na imagem transformada.
- No mapeamento indireto, a transformação $T^{-1} : P' \mapsto P$ pode ser representada como

$$P = T^{-1} P' \quad (2)$$

a qual garante que todos os pontos da imagem transformada terão um pixel associado à imagem original. Contudo, diferentes pixels da imagem resultante poderão ser mapeados em um único pixel da imagem original.

- Para permitir que as transformações espaciais possam ser realizadas por meio de multiplicação de matrizes e que várias transformações possam ser combinadas para produzir um resultado composto, as equações serão expressas na forma de *coordenadas homogêneas*.
- Em coordenadas homogêneas, um novo componente é adicionado ao ponto.
- Assim, um ponto P que tem suas coordenadas expressas por (X, Y, Z) no sistema de coordenadas cartesianas tridimensionais (3D) é descrito em coordenadas homogêneas como (WX, WY, WZ, W) , sendo W um valor diferente de zero.

- Um ponto pode ter várias representações homogêneas diferentes. Por exemplo, $(3, 2, 8, 4)$ e $(6, 4, 16, 8)$ são duas representações diferentes para um mesmo ponto.
- A transformação do sistema de coordenadas homogêneas para o sistema de coordenadas cartesianas é obtida pela divisão dos primeiros componentes das coordenadas pelo último componente, seguida da remoção deste último componente das coordenadas.
- Uma transformação espacial, dada por exemplo pela equação 1, normalmente é expressa em notação matricial, ou seja

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \\ t_{41} & t_{42} & t_{43} & t_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

em que é assumido o valor $W = 1$ nessa representação em coordenadas homogêneas.

Registro de Imagens

Transformações Espaciais

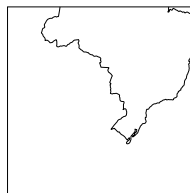
- Exemplos dessas transformações são ilustrados na figura a seguir.



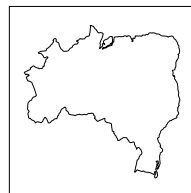
imagem original



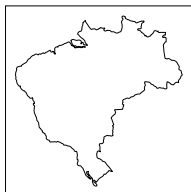
mudança de escala



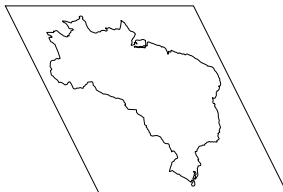
translação



rotação de 27 graus



reflexão em torno do eixo y



cisalhamento em torno do eixo x



perspectiva

- As *transformações afins* constituem uma classe importante de transformações geométricas e podem ser expressas como

$$\begin{aligned}X' &= aX + bY + cZ + j \\Y' &= dX + eY + fZ + k \\Z' &= gX + hY + iZ + l\end{aligned}\tag{3}$$

ou, equivalentemente, na forma matricial utilizando coordenadas homogêneas

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c & j \\ d & e & f & k \\ g & h & i & l \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Registro de Imagens

Transformações Afins

- As transformações afins generalizam as transformações de rotação, translação, escala, reflexão e cisalhamento, descritas a seguir.
- As transformações afins preservam o paralelismo entre retas e planos.
- Em geral, as transformações afins não preservam comprimentos, distâncias, áreas, volumes, ângulos ou perpendicularidade.
- Pode-se mostrar, no entanto, que elas preservam a razão entre os volumes de dois sólidos, a razão entre as áreas de dois objetos planos coplanares ou paralelos, e a razão entre os comprimentos de dois segmentos colineares ou paralelos.
- Em particular, o ponto médio de um segmento é mapeado para o ponto médio da imagem desse segmento.

- A *mudança de escala* pelos fatores S_x , S_y e S_z ao longo dos eixos x , y e z , respectivamente, é definida como

$$\begin{aligned}X' &= S_x X \\Y' &= S_y Y \\Z' &= S_z Z\end{aligned}\tag{4}$$

em que (X', Y', Z') são as coordenadas do novo ponto.

- Na forma matricial, a mudança de escala pode ser representada por

$$S = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- A *translação* de um ponto $P(X, Y, Z)$ para uma nova posição $P(X', Y', Z')$ usando-se deslocamentos t_x , t_y e t_z pode ser expressa como

$$\begin{aligned}X' &= X + t_x \\Y' &= Y + t_y \\Z' &= Z + t_z\end{aligned}\tag{5}$$

- Na forma matricial, a translação pode ser representada por

$$T = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Registro de Imagens

Rotação

- A forma mais simples de realizar uma *rotação* tridimensional é aplicar a rotação de um ponto em torno de cada eixo de coordenadas.
- Os ângulos são medidos no sentido horário, quando a origem é vista pela parte positiva do eixo, conforme figura a seguir.

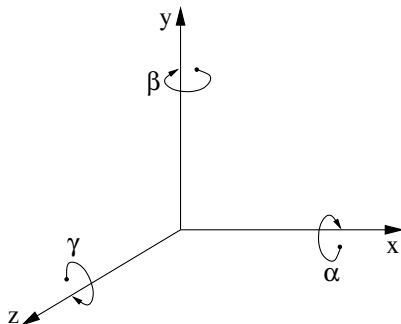


Figura: Ângulos medidos para cada eixo de rotação.

- A rotação de um ponto em torno do eixo x por um ângulo α é realizada pela matriz de transformação

$$R_{\alpha} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha & 0 \\ 0 & \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- A rotação de um ponto em torno do eixo y por um ângulo β é realizada pela matriz de transformação

$$R_{\beta} = \begin{bmatrix} \cos \beta & 0 & \sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- De forma análoga, a rotação de um ponto em torno do eixo de coordenadas z por um ângulo γ é realizada pela matriz de transformação

$$R_{\gamma} = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 & 0 \\ \sin \gamma & \cos \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Em uma *reflexão* com respeito ao plano yz , apenas os valores de coordenadas x dos vetores de posição do objeto sofrem alteração (inversão de sinal). Portanto

$$E_{yz} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Uma reflexão no plano xz é dada por

$$E_{xz} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Finalmente, uma reflexão no plano xy é dada por

$$E_{xy} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- A transformação de *cisalhamento* altera as coordenadas dos pontos de acordo com uma função de direção, de modo a causar uma deformação dos objetos presentes na imagem.
- Cada ponto transformado mantém-se em uma linha paralela à direção do cisalhamento.
- Um cisalhamento no espaço altera os pontos em direção paralela a um plano de coordenadas. Dessa forma, pode-se realizar a transformação em duas direções de eixos de coordenadas e manter a outra fixa.
- Em um cisalhamento com respeito ao plano yz , os fatores c_y e c_z da transformação são aplicados nas direções y e z , respectivamente, mantendo a coordenada x constante.

- Essa transformação pode ser expressa como

$$\begin{aligned}X' &= X \\Y' &= Y + c_y X \\Z' &= Z + c_z X\end{aligned}\tag{6}$$

Na forma matricial, o cisalhamento no plano yz é dado por

$$C_{yz} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ c_y & 1 & 0 & 0 \\ c_z & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Um cisalhamento no plano xz é dado por

$$C_{xz} = \begin{bmatrix} 1 & c_x & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & c_z & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Finalmente, um cisalhamento no plano xy é dado por

$$C_{xy} = \begin{bmatrix} 1 & 0 & c_x & 0 \\ 0 & 1 & c_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Uma *transformação polinomial* no espaço tridimensional é dada por

$$\begin{aligned}x' &= \sum_{i=0}^n \sum_{j=0}^n \sum_{k=0}^n a_{ijk} x^i y^j z^k \\y' &= \sum_{i=0}^n \sum_{j=0}^n \sum_{k=0}^n b_{ijk} x^i y^j z^k \\z' &= \sum_{i=0}^n \sum_{j=0}^n \sum_{k=0}^n c_{ijk} x^i y^j z^k\end{aligned}\tag{7}$$

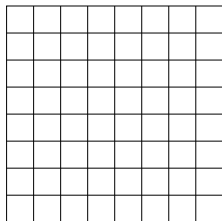
em que n é o grau dos polinômios utilizados na transformação. Os coeficientes a_{ijk} , b_{ijk} e c_{ijk} podem ser determinados a partir de pontos representativos da imagem, também chamados de *pontos de controle*.

- A ordem da transformação é dada por $m = i + j + k$, com $m \geq 1$.
- Para uma transformação polinomial de ordem m , há um total de $(m + 1)(m + 2)(m + 3)/2$ coeficientes.
- As transformações afins, descritas na equação 3, são um caso particular das transformações polinomiais de ordem 1, com 12 coeficientes.
- A transformação polinomial pode ser empregada para compensar alterações geométricas resultantes de mudanças nos ângulos de observação, tais como distorções que ocorrem em equipamentos de aquisição de imagens.

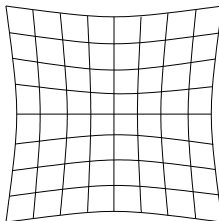
Registro de Imagens

Transformação Polinomial

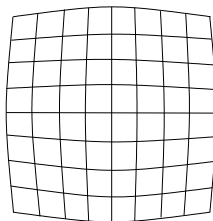
- Duas distorções não lineares que podem ser corrigidas com o modelo polinomial são mostradas a seguir, conhecidas como distorção *almofada* e *barril*, respectivamente.



grade original



distorção *almofada*



distorção *barril*

- As *transformações projetivas* ou *homografias* permitem que cada ponto dos objetos que formam a cena no espaço tridimensional possa ser projetado no plano de imagem.
- As transformações projetivas mais comuns são as projeções ortográfica e perspectiva.

Registro de Imagens

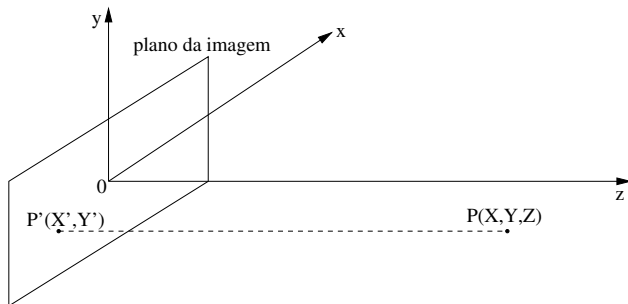
Projeção Ortográfica

- A *projeção ortográfica* é uma transformação que realiza o mapeamento de pontos tridimensionais sobre o plano da imagem, tal que os pontos são projetados ao longo de linhas paralelas na imagem.
- Esse tipo de projeção é bastante utilizado em desenhos de engenharia e arquitetura para representar e visualizar objetos por meio de um conjunto de vistas bidimensionais que mantêm as proporções relativas do objeto.
- Dessa maneira, a forma dos objetos pode ser reconstruída a partir dessas projeções.
- A projeção ortográfica pode ser modelada por uma câmera que captura a cena, cujo centro de projeção esteja localizado no infinito, tornando as linhas de projeção paralelas ao plano da imagem.
- A projeção pode ocorrer em um dos planos $x = 0$, $y = 0$ ou $z = 0$.

Registro de Imagens

Projeção Ortográfica

- Colocando-se a câmera ao longo do eixo z , conforme ilustrado na figura a seguir, todos os pontos são projetados no plano $z = 0$.



- No caso de projeção no plano $z = 0$, o ponto $P(X, Y, Z)$ é projetado como

$$\begin{aligned}X' &= X \\Y' &= Y \\Z' &= 0\end{aligned}\tag{8}$$

Assim, a projeção ortográfica pode ser expressa, em coordenadas homogêneas, como

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ W \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Como uma transformação projetiva mapeia pontos com coordenadas tridimensionais (espaço do objeto) para pontos com coordenadas bidimensionais (plano da imagem), a projeção ortográfica pode ser representada por uma matriz de 3×4 elementos, ou seja

$$\begin{bmatrix} X' \\ Y' \\ W \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Registro de Imagens

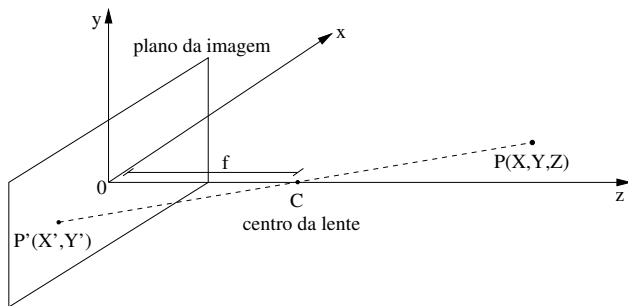
Projeção Perspectiva

- A *projeção perspectiva*, assim como a projeção ortográfica, realiza a transformação de pontos tridimensionais para o plano da imagem.
- Entretanto, o tamanho dos objetos é reduzido com o aumento da distância do centro de projeção.
- Na transformação perspectiva também ocorre uma redução não uniforme de linhas nos objetos como uma função de orientação e distância dos objetos do centro de projeção.
- Esses efeitos estão relacionados à percepção de profundidade do sistema visual humano, embora a forma dos objetos não seja preservada.

Registro de Imagens

Projeção Perspectiva

- A figura a seguir mostra um modelo para o cálculo da projeção perspectiva.
- O sistema de coordenadas (x, y, z) da câmera é posicionado de tal forma que o plano da imagem coincida com o plano xy , enquanto o eixo óptico definido pelo centro da lente situa-se ao longo do eixo z .
- Dessa forma, o centro do plano da imagem localiza-se na origem do sistema, enquanto o centro da lente possui coordenada $(0, 0, f)$, sendo f a distância focal da câmera.



- Seja um ponto P no espaço descrito pelas coordenadas (X, Y, Z) , com $Z > f$, ou seja, localizado na frente da lente.
- A obtenção de (X', Y') a partir da projeção do ponto (X, Y, Z) sobre o plano da imagem é dada pela relação

$$\frac{X'}{f} = -\frac{X}{Z-f} \qquad \frac{Y'}{f} = -\frac{Y}{Z-f} \qquad (9)$$

- O sinal negativo nas equações 9 indica que a imagem está invertida.
- Portanto, as coordenadas do ponto projetado no plano da imagem ($Z' = 0$) são

$$X' = \frac{fX}{f-Z} \qquad Y' = \frac{fY}{f-Z} \qquad (10)$$

- Dessas relações, a projeção perspectiva pode ser expressa, em coordenadas homogêneas, como

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ W \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/f & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Como a projeção é realizada sobre o plano xy , o valor de Z' pode ser desconsiderado nesta transformação.
- Assim, a projeção perspectiva pode ser representada como

$$\begin{bmatrix} X' \\ Y' \\ W \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/f & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ (-Z/f) + 1 \end{bmatrix} = \begin{bmatrix} fX/(f - Z) \\ fY/(f - Z) \\ 1 \end{bmatrix}$$

Registro de Imagens

Projeção Perspectiva

- Nesse modelo de câmera, o plano da imagem está posicionado atrás do centro de projeção a uma distância f e a imagem projetada está invertida. Uma maneira de evitar essa inversão é assumir que o plano da imagem esteja na frente do centro de projeção, como mostrado na figura a seguir.
- O centro do plano da imagem possui coordenada $(0, 0, f)$, na qual f é a distância focal da lente, enquanto o centro da lente localiza-se na origem do sistema.

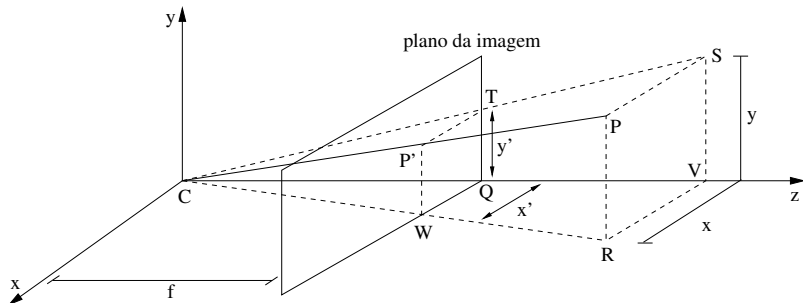


Figura: Formulação alternativa para a projeção perspectiva.

- Considerando um ponto P descrito pelas coordenadas (X, Y, Z) , com $Z > f$, pode-se observar a formação dos triângulos CQW e CVR pela projeção do ponto P no plano xz e dos triângulos CSV e CTQ no plano yz .
- No plano xz , os triângulos CQW e CVR são semelhantes. Da semelhança, extrai-se a seguinte relação

$$\frac{X'}{f} = \frac{X}{Z} \quad (11)$$

que permite calcular a coordenada X'

$$X' = \frac{X}{Z/f} \quad (12)$$

- Analogamente, dos triângulos CSV e CTQ , no plano yz , extrai-se a relação que permite calcular Y'

$$Y' = \frac{Y}{Z/f} \quad (13)$$

- Dessas relações, a projeção perspectiva pode ser expressa, em coordenadas homogêneas, como

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ W \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Como a projeção é realizada sobre o plano xy , o valor de Z' pode ser desconsiderado na transformação.
- Assim, a projeção perspectiva pode ser representada como

$$\begin{bmatrix} X' \\ Y' \\ W \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z/f \end{bmatrix} = \begin{bmatrix} fX/Z \\ fY/Z \\ 1 \end{bmatrix}$$

- Uma outra forma de expressar a projeção perspectiva é a partir da transformação afim, dada pela equação 3, que pode ser reescrita como

$$\begin{bmatrix} W X' \\ W Y' \\ W \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ i & j & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

ou, de forma similar

$$\begin{aligned} X' &= \frac{aX + bY + c}{iX + jY + 1} \\ Y' &= \frac{dX + eY + f}{iX + jY + 1} \end{aligned} \tag{14}$$

- A partir de quatro pontos não colineares em duas imagens correspondentes, conforme ilustrado no exemplo a seguir, os oito coeficientes da equação 14 podem ser determinados.
- Essa transformação projetiva é útil para auxiliar o registro de imagens obtidas sob diferentes vistas.

Exemplo:

Dada a transformação projetiva expressa na equação 14, calcule os parâmetros a , b , c , d , e , f , i e j a partir das quatro coordenadas mostradas nas figuras a seguir.

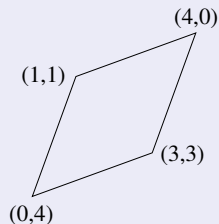
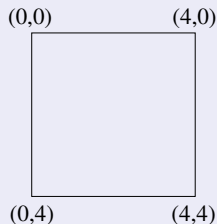


Figura: Mapeamento de coordenadas pela transformação projetiva.

Exemplo: (continuação)

Os parâmetros podem ser calculados tomando-se as coordenadas dos vértices da imagem (a) e seus vértices correspondentes na imagem (b) e substituindo-as na equação da transformação projetiva.

Assim, para o ponto superior esquerdo das imagens, tem-se que

$$1 = \frac{0a + 0b + c}{0i + 0j + 1}$$

$$1 = \frac{0d + 0e + f}{0i + 0j + 1}$$

Portanto, $c = f = 1$. Analogamente, para os outros três vértices, tem-se que os coeficientes da transformação são

$$a = 0.75$$

$$c = 1$$

$$e = 0.75$$

$$i = 0$$

$$b = -0.25$$

$$d = -0.25$$

$$f = 1$$

$$j = 0$$

Registro de Imagens

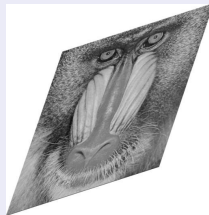
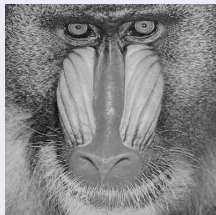
Projeção Perspectiva

Exemplo: (continuação)

Pode-se verificar facilmente que o ponto médio $P(2, 0)$ da aresta superior na imagem a seguir (à esquerda) é mapeado para o ponto médio $P(2.5, 0.5)$ da aresta correspondente na imagem à direita.

Valores fracionários resultantes do cálculo das posições dos pontos em uma imagem devem ser adequadamente aproximados. A intensidade dos pixels deve ser interpolada levando-se em consideração pontos vizinhos.

Uma ilustração da transformação projetiva é mostrada a seguir.



- As transformações espaciais discutidas anteriormente permitem o mapeamento de valores inteiros de coordenadas para valores que podem ser fracionários.
- Portanto, a aplicação satisfatória de transformações geométricas requer estratégias para atribuir intensidades aos pontos que não se encontram na nova grade da imagem.
- Os valores de intensidade de níveis de cinza ou cor para cada pixel na nova grade da imagem transformada podem ser obtidos por um processo de interpolação.
- Quatro métodos comumente utilizados para interpolação de valores de intensidade nos pixels da imagem são:
 - ▶ *vizinho mais próximo.*
 - ▶ *bilinear.*
 - ▶ *bicúbica.*
 - ▶ *polinômios de Lagrange.*

Registro de Imagens

Interpolação pelo Vizinho Mais Próximo

- Na interpolação pelo vizinho mais próximo, o valor da intensidade a ser atribuído ao pixel (x', y') na imagem reamostrada terá o mesmo valor do pixel que estiver mais próximo da posição ocupada pelo pixel (x, y) na imagem original, conforme mostrado na figura a seguir.

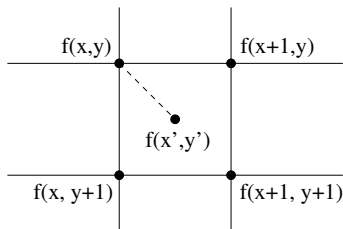


Figura: Interpolação pelo vizinho mais próximo.

- A interpolação pode ser expressa como

$$f(x', y') = \begin{cases} f(x, y), & \text{para } dx < 0.5 \text{ e } dy < 0.5 \\ f(x + 1, y), & \text{para } dx \geq 0.5 \text{ e } dy < 0.5 \\ f(x, y + 1), & \text{para } dx < 0.5 \text{ e } dy \geq 0.5 \\ f(x + 1, y + 1), & \text{para } dx \geq 0.5 \text{ e } dy \geq 0.5 \end{cases} \quad (15)$$

em que dx e dy correspondem às distâncias nas direções x e y , respectivamente, entre os pontos (x', y') e (x, y) , ou seja, apenas as partes fracionárias do ponto (x', y')

$$dx = x' - x = x' - \lfloor x' \rfloor$$

$$dy = y' - y = y' - \lfloor y' \rfloor$$

- A equação 15 pode ser reescrita como

$$f(x', y') = f(\text{round}(x), \text{round}(y))$$

em que *round* é uma função que aproxima um número para seu valor inteiro mais próximo.

- Embora a interpolação pelo vizinho mais próximo seja simples de ser implementada, este método possui a desvantagem de causar distorções em detalhes finos ou criar formas serrilhadas em bordas retas da imagem.

Registro de Imagens

Interpolação Bilinear

- A interpolação bilinear utiliza uma média ponderada de distância dos quatro pixels vizinhos mais próximos para determinar a intensidade de cada pixel (x', y') na imagem transformada, como mostrado na figura a seguir.

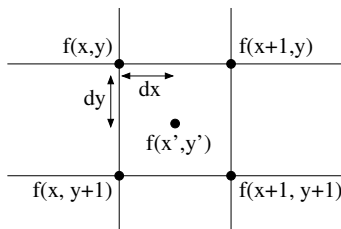


Figura: Interpolação bilinear.

- A interpolação é dada por

$$f(x', y') = (1 - dx)(1 - dy) f(x, y) + dx(1 - dy) f(x + 1, y) + (1 - dx)dy f(x, y + 1) + dx dy f(x + 1, y + 1)$$

- A interpolação bilinear reduz os problemas causados pela interpolação pelo vizinho mais próximo, entretanto, causa borramento devido a sua natureza de suavização.

Registro de Imagens

Interpolação Bicúbica

- A interpolação bicúbica utiliza uma vizinhança de 4×4 pontos ao redor do ponto em questão para calcular seu valor de intensidade, como mostrado na figura a seguir.

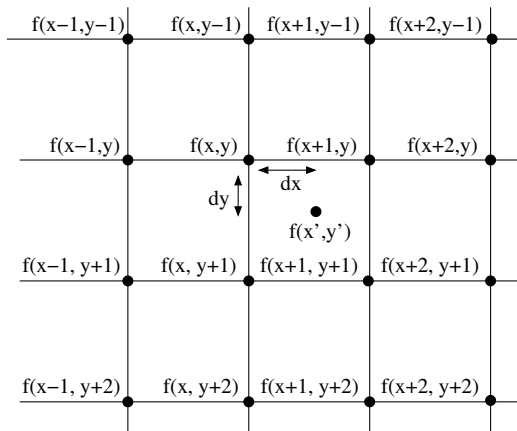


Figura: Interpolação bicúbica.

- Uma função comum para determinar as intensidades do pixel na imagem interpolada é a função B-spline cúbica, definida como

$$f(x', y') = \sum_{m=-1}^2 \sum_{n=-1}^2 f(x + m, y + n) R(m - dx) R(dy - n)$$

sendo

$$R(s) = \frac{1}{6} [P(s+2)^3 - 4P(s+1)^3 + 6P(s)^3 - 4P(s-1)^3]$$

$$P(t) = \begin{cases} t, & t > 0 \\ 0, & t \leq 0 \end{cases}$$

- A interpolação bicúbica não sofre do problema de bordas serrilhadas causado pela interpolação pelo vizinho mais próximo, assim como não causa borramento como ocorre na interpolação bilinear.
- A interpolação bicúbica é frequentemente utilizada para ampliar áreas da imagem com respeito a um ponto arbitrário.
- Essa interpolação também preserva detalhes finos na imagem.

- Outro método que utiliza uma vizinhança de 4×4 pontos para calcular o valor de intensidade de um pixel (x', y') é a interpolação por polinômios de Lagrange, definida como

$$f(x', y') = \frac{-dy(dy-1)(dy-2)L(1)}{6} + \frac{(dy+1)(dy-1)(dy-2)L(2)}{2} + \frac{-dy(dy+1)(dy-2)L(3)}{2} + \frac{dy(dy+1)(dy-1)L(4)}{6}$$

sendo

$$L(n) = \frac{-dx(dx-1)(dx-2)f(x-1, y+n-2)}{6} + \frac{(dx+1)(dx-1)(dx-2)f(x, y+n-2)}{2} + \frac{-dx(dx+1)(dx-2)f(x+1, y+n-2)}{2} + \frac{dx(dx+1)(dx-1)f(x+2, y+n-2)}{6}$$

- Os resultados obtidos a partir da interpolação por polinômios de Lagrange possuem qualidade similar aos obtidos pela interpolação bicúbica, porém, com menor tempo computacional.

Registro de Imagens

Interpolação de Intensidade

- A imagem original é ampliada duas vezes e cada um dos métodos de interpolação de intensidade é aplicado.



original

Registro de Imagens

Interpolação de Intensidade



vizinho mais próximo



bilinear



bicúbica



polinômios de Lagrange

Registro de Imagens

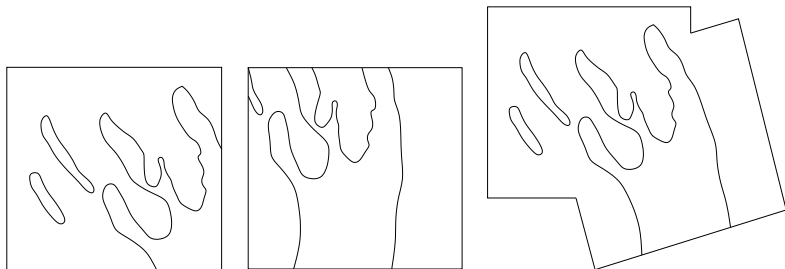
Técnicas para Registro

- Em razão da diversidade de áreas de aplicação e dos vários tipos de distorção a que estão sujeitas as imagens, as técnicas de registro normalmente são específicas para o problema em estudo.
- Cada abordagem deve considerar, de maneira apropriada, a precisão requerida pela operação de registro, as deformações geométricas e radiométricas entre as imagens e as características próprias dos dados da aplicação sob consideração.
- Apesar de não existir um método genérico que apresente bom desempenho em diferentes domínios de aplicação, a maioria das técnicas de registro compartilha alguns passos em comum, descritos a seguir.
- Feições ou alvos representativos extraídos das imagens, como bordas, cantos ou junções, são detectados manual ou automaticamente.
- A correspondência entre essas características é determinada entre o par de imagens por meio de um conjunto de medidas de similaridade.

Registro de Imagens

Técnicas para Registro

- O tipo e os parâmetros da função de mapeamento são estimados a partir da correspondência entre as imagens.
- A transformação para alinhar uma imagem à outra requer a redefinição das coordenadas dos pontos da imagem e a interpolação de seus valores de níveis de cinza ou cores.
- A figura a seguir ilustra o registro de duas imagens.



Registro de Imagens

Técnicas para Registro

- Regiões representativas da imagem, como bordas, pontos isolados, cantos ou junções, devem, idealmente, ser feições distinguíveis e facilmente detectáveis na imagem.
- Exemplos de feições em imagens de sensoriamento remoto são rodovias, rios e topo de edificações.
- Além disso, descritores de objetos, tais como área, assinatura e energia de deformação, podem ser calculados para as regiões.
- Descritores pouco sensíveis a ruído e invariantes com respeito à rotação, translação e escala são particularmente úteis.

- Diversas medidas de similaridade têm sido propostas para avaliar a semelhança (ou discrepância) entre duas imagens.
- Tais medidas podem ser calculadas para uma região específica, em vez de toda a imagem.
- Uma medida muito utilizada na determinação do registro é a *correlação cruzada*, definida como

$$CC(u, v) = \sum_{x \in S} \sum_{y \in S} f(x, y) w(x - u, y - v) \quad (16)$$

tal que a soma é realizada sobre a região S em que a imagem f e a janela w se sobrepõem, cujo centro é dado pelo ponto (u, v) . Deslocando-se a janela por toda a imagem, o valor máximo da correlação cruzada indicará o ponto de melhor registro entre a janela e a imagem de referência.

- A medida de correlação cruzada não é invariante às mudanças na amplitude da imagem causadas, por exemplo, por variações nas condições de iluminação da imagem.

- Uma medida mais apropriada é o *coeficiente de correlação* ou *correlação cruzada normalizada*, definida como

$$CCN(u, v) = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - \bar{f}] [w(x - u, y - v) - \bar{w}]}{\sqrt{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - \bar{f}]^2 [w(x - u, y - v) - \bar{w}]^2}} \quad (17)$$

em que \bar{w} é o valor médio dos pixels na janela w , calculado apenas uma vez, e \bar{f} é a média das intensidades de f na região sob a janela.

- O valor de CCN varia no intervalo de -1 a 1 e é independente da escala de valores de intensidade de f e w .
- O valor máximo de CCN indica a posição de melhor registro entre a janela e a imagem. Valores próximos a zero indicam baixo grau de registro entre as imagens.

- A *informação mútua* é um conceito baseado na entropia que também pode ser utilizado no registro de imagens.
- Sejam X e Y duas variáveis aleatórias representando as intensidades x_i e y_j das imagens f_1 e f_2 , respectivamente. A informação mútua é definida como

$$IM(X, Y) = \sum_{x_i \in X} \sum_{y_j \in Y} P(x_i, y_j) \log \frac{P(x_i, y_j)}{P(x_i)P(y_j)} \quad (18)$$

em que $P(x_i)$ e $P(y_j)$ são as distribuições de probabilidade dos níveis de cinza para as imagens f_1 e f_2 , respectivamente, e $P(x_i, y_j)$ é a distribuição de probabilidade conjunta.

- A informação mútua também pode ser definida em termos de entropia como

$$IM(X, Y) = H(X) + H(Y) - H(X, Y) \quad (19)$$

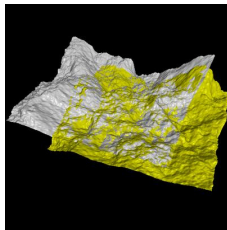
em que $H(X)$ e $H(Y)$ são as entropias para as variáveis aleatórias X e Y , respectivamente, e $H(X, Y)$ é a entropia conjunta.

- O valor máximo da informação mútua $IM(X, Y)$ corresponde ao melhor registro entre as imagens.

Registro de Imagens

Técnicas para Registro

- O algoritmo iterativo de pontos mais próximos tem como objetivo estimar transformações que minimizem distâncias entre um modelo e um dado a ser ajustado.
- O método consiste em encontrar, a cada iteração, uma translação e uma rotação que alinhem o modelo e os dados.
- Conhecendo a correspondência entre pontos dos dois conjuntos, pode-se estimar a transformação relativa entre eles.
- Uma solução para determinar essa correspondência é assumir que o ponto mais próximo é o par correspondente, determinar a melhor transformação e iterativamente encontrar e refinar o alinhamento.



Registro de Imagens

Técnicas para Registro

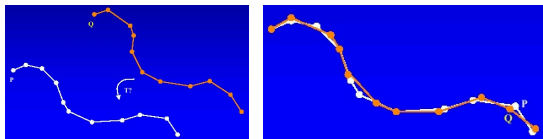
- Algoritmo iterativo de pontos mais próximos:

Entrada: Conjuntos de pontos P e Q , em que P é um subconjunto de Q .

- 1: Busca do ponto mais próximo: para cada ponto p de P , encontrar o ponto q mais próximo em Q .
- 2: Cálculo do registro: calcular a transformação rígida T que minimiza as somas dos quadrados das distâncias entre os pares de pontos próximos (p, q) .
- 3: Transformação: aplicar a transformação rígida T a todos os pontos do conjunto P .
- 4: Iteração: repetir os passos 1 a 3 até o método convergir.

- Minimização da função objetivo

$$E = \sum_{i=1}^{N_p} \|T(q_i) - p_i\|^2$$



- O método da correlação de fase baseia-se na propriedade de translação da transformada de Fourier.
- Dadas duas imagens, f_1 e f_2 , que diferem apenas por uma translação (x_0, y_0) , as transformadas de Fourier correspondentes F_1 e F_2 são relacionadas como

$$F_2(u, v) = e^{i2\pi(x_0u + y_0v)} F_1(u, v)$$

- As duas imagens possuem então a mesma magnitude da transformada, entretanto, diferem em fase cujo grau é diretamente proporcional à translação.

- O espectro de potência cruzado das duas imagens f_1 e f_2 , com transformadas de Fourier F_1 e F_2 respectivamente, fornece a diferença de fase e é definido como

$$\frac{F_1(u, v)F_2^*(u, v)}{|F_1(u, v)F_2^*(u, v)|} = e^{i2\pi(x_0u + y_0v)}$$

em que F_2^* denota o complexo conjugado de F_2 .

- Esta propriedade estabelece que a fase do espectro de potência cruzado é equivalente à diferença de fase entre as duas imagens.
- Tomando a inversa da transformada de Fourier, obtém-se uma imagem com valores próximos a zero exceto nas coordenadas que têm o deslocamento necessário para o registro da imagem (função impulso).

- Rotação também pode ser abordada no domínio de Fourier, pois uma rotação pode ser expressa como uma translação em coordenadas polares.
- O espectro de potência cruzada é computado em função de um ângulo de rotação estimado (ϕ) e simplificado com o uso de coordenadas polares

$$G(r, \theta; \phi) = \frac{F_1(r, \theta)F_2^*(r, \theta - \phi)}{|F_1(r, \theta)F_2^*(r, \theta - \phi)|}$$

- Assim, primeiro determina-se ϕ tal que a inversa da transformada de Fourier fornece a aproximação mais próxima a um impulso. A translação restante é então dada pela localização do impulso.

Registro de Imagens

Técnicas para Registro

- Uma variação desta técnica que apresenta vantagens em termos de custo computacional é baseada no *cepstro* de potência das imagens.
- O cepstro é o logaritmo do espectro de potência de uma imagem.
- Inicialmente, o ângulo de rotação é aproximado como anteriormente.
- Então, o espectro de potência é usado para determinar a translação como anteriormente.
- Como logaritmos são utilizados, imagens são adicionadas ao invés de multiplicadas, o que pode aumentar a velocidade de processamento.

- Identificação de pontos correspondentes nas duas imagens (por exemplo, cantos de objetos) por meio de técnicas como:
 - ▶ *Scale-Invariant Feature Transform* (SIFT)
 - ▶ *Speeded up Robust Features* (SURF)
 - ▶ *Features from Accelerated Segment Test* (FAST)
 - ▶ *Binary Robust Independent Elementary Features* (BRIEF)
 - ▶ *Oriented FAST and Rotated BRIEF* (ORB)
- Cálculo da matriz de homografia a partir de 4 ou mais pontos correspondentes nas imagens.
- A matriz de homografia é calculada pela resolução de um sistema linear de equações.