

Universidade Estadual de Campinas Instituto de Computação
Introdução ao Processamento Digital de Imagem (MC920 / MO443)
Professor: Hélio Pedrini
Lista II

1. Quais os efeitos causados a representação *quadtree* de uma imagem após sofrer uma mudança de escala, translação ou rotação?

Resposta:

Mudança de escala: a representação continuará a mesma, apesar de estar suscetível a mudanças conforme ocorrer a perda de detalhes da imagem.

Mudança de translação ou rotação: a representação mudará completamente (*quadtree* é muito sensível a esse tipo de mudanças).

2. Suponha a imagem a seguir.

21	21	21	95	169	243	243	243
21	21	21	95	169	243	243	243
21	21	21	95	169	243	243	243
21	21	21	95	169	243	243	243

- a) calcule a entropia da imagem;

Resposta:

Imagem: $4 \times 8 = 32$

Níveis de cinza (i)	0	1	2	3
n_i	12	4	4	12
$p(i)$	0.375	0.125	0.125	0.375

$$H = - \sum_{i=1}^k p_i \log_2 p_i = - (.375 * \log_2(.375) * 2 + .125 * \log_2(.125) * 2) = - (-1.061278 - 0.75) = \mathbf{1.81}$$

- b) calcule o código de Huffman para cada nível de cinza da imagem.

Resposta:

```

      1       0
    0.375   0.625
      1       0
      0.25   0.375
      1       0
    0.125   0.125
  
```

Huffman ~ Binary Search Tree -> Left Node: 1, Right Node: 0

Fontes originais	Redução de fontes
------------------	-------------------

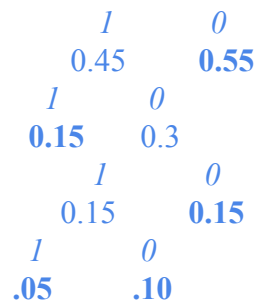
Símbolo	Probabilidade	Código	Etapa 1		Etapa 2	
S_1	.375	1	.375	1	.625	0
S_4	.375	00	.375	00	.375	1
S_2	.125	011	.250	01		
S_3	.125	010				

$$\bar{L} = .375 * 1 + .375 * 2 + .125 * 3 * 2 = \mathbf{1.875 \text{ bits/símbolo}}$$

3. Dada uma fonte de cinco símbolos com probabilidades {0.55,0.15,0.15,0.10,0.05}, obtenha duas codificações de Huffman diferentes. Calcule o comprimento médio L de bits para cada um dos códigos.

Resposta:

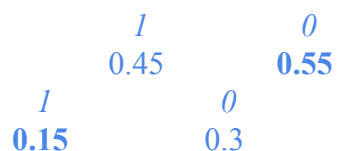
Primeira codificação:



Fontes originais		Redução de fontes						
Símbolo	Probabilidade	Código	Etapa 1		Etapa 2		Etapa 3	
S_1	.55	0	.55	0	.55	0	.55	0
S_2	.15	11	.15	11	.3	10	.45	1
S_3	.15	100	.15	101	.15	11		
S_4	.10	1011	.15	100				
S_5	.05	1010						

$$\bar{L} = .55 * 1 + .15 * 2 + .15 * 3 + .10 * 4 + .05 * 4 = \mathbf{1.9 \text{ bits/símbolo}}$$

Segunda codificação:



$\begin{matrix} 1 & 0 & 1 & 0 \\ 0.05 & 0.10 & 0.15 & 0.15 \end{matrix}$

Fontes originais		Redução de fontes						
Símbolo	Probabilidade	Código	Etapa 1		Etapa 2		Etapa 3	
S_1	.55	0	.55	0	.55	0	.55	0
S_2	.15	100	.15	11	.3	10	.45	1
S_3	.15	101	.15	100	.15	11		
S_4	.10	111	.15	101				
S_5	.05	110						

$$\bar{L} = .55 * 1 + .15 * 3 + .15 * 3 + .10 * 3 + .05 * 3 = \mathbf{1.9 \text{ bits/símbolo}}$$

4. Considere uma fonte com alfabeto $\Omega = \{a,b,c\}$. A partir da sequência 1, 2, 2, 0, 2, 3, 4, 9, 10, 5, 9, construa o dicionário e decodifique essa sequência por meio do algoritmo LZW.

Resposta: bccacccccccccccaccc, d = {0: 'a', 1: 'b', 2: 'c', 3: 'bc', 4: 'cc', 5: 'ca', 6: 'ac', 7: 'cb', 8: 'bcc', 9: 'ccc', 10: 'cccc', 11: 'ccccc', 12: 'cac'}

Input	Dictionary translation	Decode output	New dictionary entry
1	b	b	none
2	b c	bc	bc/3
2	c c	bcc	cc/4
0	c a	bcca	ca/5
2	a c	bccac	ac/6
3	c bc	bccacbc	cb/7
4	bc cc	bccacbcc	bcc/8
9	???	???	???

Input	Dictionary translation	Decode output	New dictionary entry
1	b c	bc	bc/3
2	c c	bcc	cc/4
2	c a	bcca	ca/5
0	a c	bccac	ac/6
2	c bc	bccacbc	cb/7
3	bc cc	bccacbcc	bcc/8
4	cc ccc	bccacbccccc	ccc/9

9	ccc cccc	bccacbccccccccccc	cccc/10
9	ccc cccc	bccacbccccccccccc	cccc/10
10	cccc ca	bccacbccccccccccca	cccc/11
5	ca ccc	bccacbccccccccccccccc	cac/12

5. Descreva os principais tipos de redundância que podem ser explorados pelos métodos de compressão de imagens.

Resposta:

❖ Redundância de codificação:

- Mesma informação (ou símbolo) representada repetidas vezes;
- Em geral, a redundância de codificação está presente quando os códigos atribuídos a um conjunto de eventos (tal como os níveis de cinza da imagem) não foram escolhidos de forma a explorar as probabilidades dos eventos;

❖ Redundância interpixel:

- Explora a característica de que pixels vizinhos em uma imagem possuem alguma relação ou similaridade;
- Contribuição visual de um único pixel é redundante - ela poderia ser prevista com base nos valores dos pixels adjacentes;

❖ Redundância psicovisual:

- Há informações psicovisualmente redundantes - possuem menor importância relativa que outras no processamento visual e podem ser eliminadas sem prejudicar significativamente a percepção da imagem;
- Redundância infere que dados presentes na imagem seriam eliminados de forma **irreversível!**

6. Explique as vantagens e desvantagens em se utilizar blocos de tamanhos diferentes de 8×8 pixels no cálculo da transformada DCT na padronização JPG.

Resposta:

❖ Vantagens:..

- Custo computacional menor em um bloco de tamanho 8×8 para realizar operações como a transformada DCT;
- Matrizes menores que 8×8 não possuem informação suficiente e relevante para a compressão;
- Aplicar a transformada na imagem inteira não é ideal devido ao conteúdo não uniforme da imagem;

❖ Desvantagens:

- Imagem resultante terá tamanho maior do que se utilizados blocos maiores;

7. Qual a vantagem da ordenação ziguezague dos coeficientes DCT na padronização JPG?

Resposta:

A ordenação ziguezague dos coeficientes DCT tem o objetivo de facilitar a etapa de codificação por entropia. Em geral, é mais intuitivo que pixels em um bloco tenham mais em comum do que pixels em uma linha - ou seja, os pixels serão agrupados na ordenação de forma a terem mais características em comum, permitindo que a compressão seja mais bem-sucedida devido a redundância interpixel.

Outra Resposta:

Na matriz DCT, baixas frequências tem os maiores valores de energia. Como as baixas frequências estão nas linhas e colunas de índice baixo, ao percorrer em zigue-zague percorremos na ordem de menor energia esperada, esta ordem deve ter uma longa cauda de zeros o que otimiza a codificação em comprimento de corrida.

8. Descreva dois métodos de compressão de imagens com perdas.

Resposta:

Ambos os métodos partem da redundância psicovisual, a qual justifica as perdas durante a compressão.

Modulação Delta - o preditor supõe que o pixel atual seja igual ao anterior e, para cada pixel, retorna apenas a diferença entre este e o anterior. Na prática, recebe uma constante ζ , o módulo máximo de diferença entre pixels adjacentes. O algoritmo consiste em limitar a diferença entre os pixels adjacentes em $+\zeta$ se a diferença for positiva e $-\zeta$ se a diferença for negativa..

Modulação de Código de Pulso Diferencial - o preditor supõe que o valor de um pixel depende de uma combinação linear de m pixels anteriores e que o erro de quantização seja desprezível. O objetivo é minimizar o erro médio quadrático da predição do codificador.

9. Explique o princípio das técnicas preditivas de compressão de imagens. Descreva a principal diferença entre técnicas preditivas sem perdas e com perdas.

Resposta:

Técnicas preditivas de compressão são técnicas que envolvem uma etapa anterior à codificação dos símbolos, que explora a redundância interpixel na imagem.

As técnicas preditivas sem perdas utilizam conceitos de redundância como redundância de codificação e interpixel - basta representar a informação de uma maneira mais robusta para garantir a sua recuperação. Já as técnicas preditivas com perdas utilizam conceitos de redundância psicovisual, e se baseiam em alterar os valores dos pixels de forma que facilite a compressão da imagem, mantendo a essência da imagem e abstraindo os detalhes (de maneira imperceptível ou tolerada).

Traks:

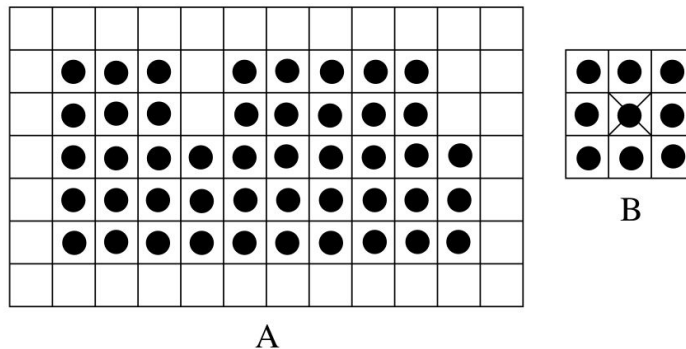
Também acho que faltou explicar o princípio. O princípio das técnicas preditivas é que é mais econômico guardar a diferença entre os pixels do que a intensidade dos pixels. Então numa compressão sem perda, você armazena todas as diferenças e recupera a imagem perfeitamente. Numa compressão com perdas, você estabelece limites que funcionam mais ou menos como: se a diferença entre os pixels for menor que X , não vale a pena armazenar isso e eles serão iguais. por isso tem erro. Ou como na compressão delta, se o erro for maior que X , colocamos um delta fixo e não armazenamos exatamente o erro, mas armazenamos esse delta.

10. A extração da borda F de um objeto A pode ser realizada através do operador morfológico

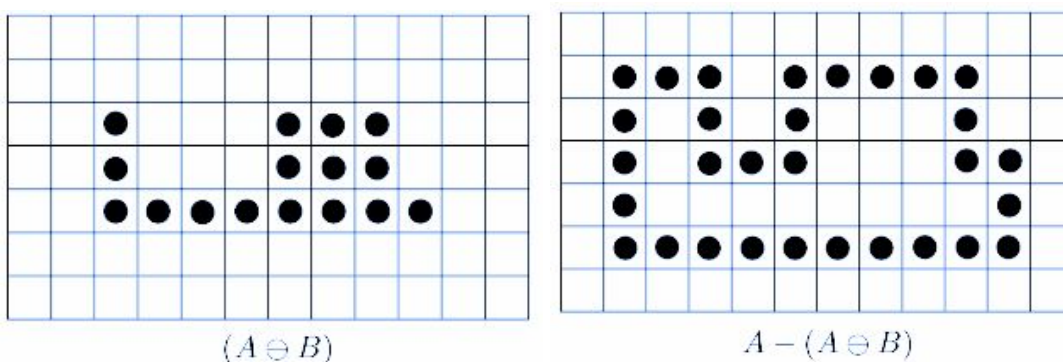
$$F(A) = A - (A \ominus B)$$

em que B é um elemento estruturante adequado.

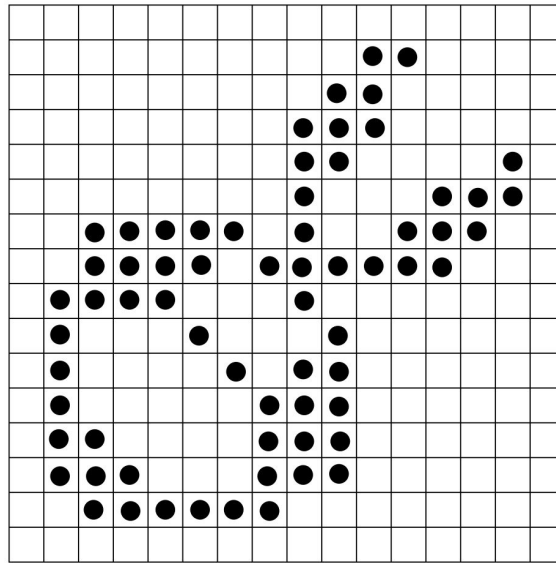
Aplique o operador acima para a imagem A a seguir, utilizando o elemento estruturante B de tamanho 3×3 .



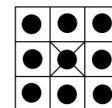
Resposta:



11. Aplique os operadores de abertura e fechamento morfológicos para a imagem a seguir.



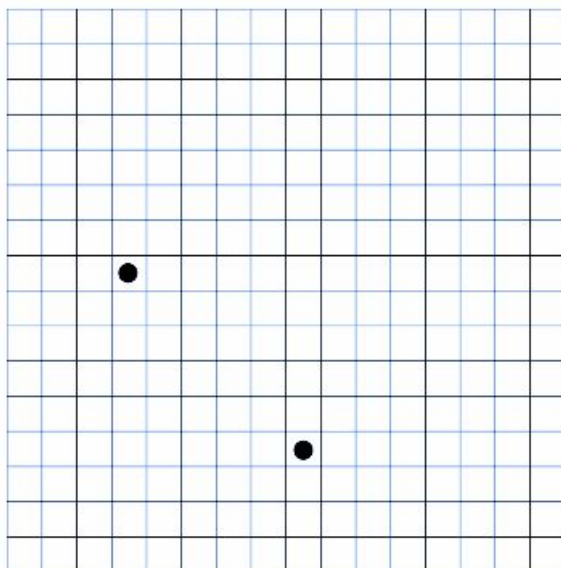
A



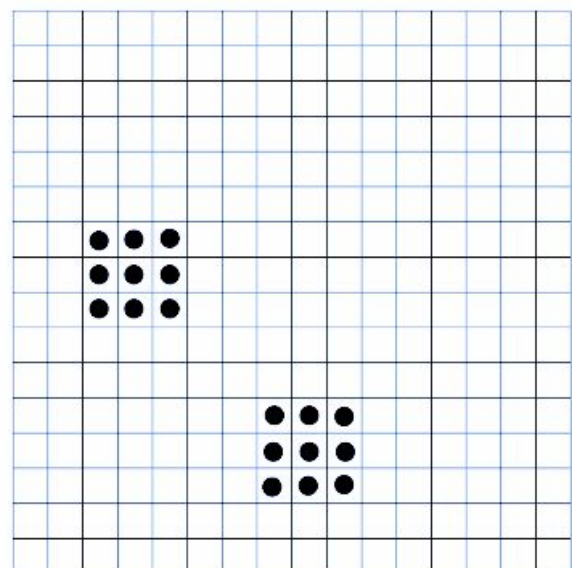
B

Resposta:

Abertura:

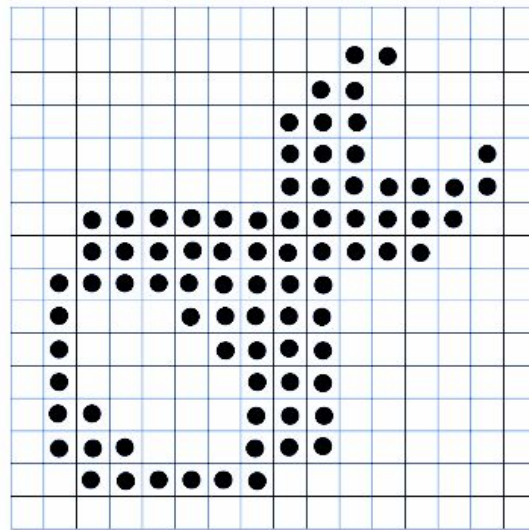


$(A \ominus B)$



$(A \ominus B) \oplus B$

Fechamento:



$$(A \oplus B) \ominus B$$

12. Represente a imagem binária mostrada a seguir pelo código de comprimento de corridas, iniciando-se com o comprimento das corridas de valor 1.

1	1	1	0	0	0	0	1	1	1	1	1	0	0	0	0	1	1	0	1	1	0	1	1	1
0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1
1	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	1	1	0	0	1	0	1	1
0	0	1	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	1	1	0	0	0	0	0
0	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1
1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1

Figura 1: Imagem binária.

Resposta:

3, 4, 5, 4, 2, 1, 2, 1, 3
 0, 4, 4, 3, 6, 1, 2, 1, 4
 3, 7, 3, 4, 2, 2, 1, 1, 2
 0, 2, 1, 2, 6, 8, 2, 4
 5, 6, 4, 3, 2, 5
 0, 1, 3, 2, 3, 1, 10, 3, 2
 3, 10, 5, 2, 5

13. Qual o problema que a utilização da versão binária da unidade de textura, os padrões locais binários (LBP), reduz quando se efetua o cálculo das medidas?

Resposta:

O LBP reduz significativamente o número de entradas do espectro de textura, limitando o valor de cada pixel de 0 a 255.

14. Por que os padrões locais binários (LBP) apresentam-se invariantes a transformações monotônicas aplicadas à imagem? Descreva as vantagens dessa característica.

Resposta:

Porque os valores armazenados são relativos à intensidade entre os pixels e, portanto, o seu valor não sofre modificação desde que as relações entre as cores da imagem permaneça a mesma.

Isso é vantajoso porque permite a extração de características morfológicas da imagem, sem que seja necessário quantificá-la em relação às cores.

15. Utilizando $d = 1$, calcule a matriz de co-ocorrência na orientação 0° para as imagens

1	2	1	2
2	1	2	1
1	2	1	2
2	1	2	1

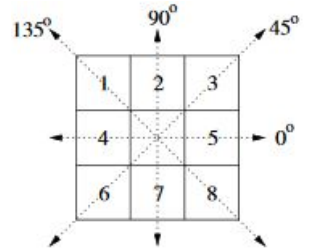
3	2	1	0
2	3	2	1
1	2	3	2
0	1	2	3

Calcule o segundo momento angular para cada uma das matrizes de co-ocorrência. Essa medida pode ser utilizada para discriminar as texturas contidas nas duas imagens?

Resposta:

Segundo momento angular:

$$f_{sma} = \sum_{i=0}^{H_g} \sum_{j=0}^{H_g} p_{i,j}^2$$



(a)

	1	2
1	0	126
2	126	0

$$f_{sma} = 0.5^2 + 0.5^2 = 0.5$$

(b)

1	2	1	2
2	1	2	1
1	2	1	2
2	1	2	1

3	2	1	0
2	3	2	1
1	2	3	2
0	1	2	3

	0	1	2	3
0	0	2	0	0
1	2	0	4	0
2	0	4	0	6
3	0	0	6	0

$$f_{sma} = 2 \cdot (1/12)^2 + 2 \cdot (1/6)^2 + 2 \cdot (1/4)^2 = 7/36 = 0.1944...$$

$$f_{sma} = 0.16^2 + 0.16^2 + 0.33^2 + 0.33^2 + 0.5^2 + 0.5^2 = 0.769$$

16. Mostre que duas rotações 2D sucessivas são aditivas.

Resposta:

$$\begin{aligned}
R_\alpha \cdot R_\beta &= \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos \beta & -\sin \beta & 0 \\ \sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \cos \alpha \cdot \cos \beta - \sin \alpha \cdot \sin \beta + 0 \cdot 0 & \cos \alpha \cdot (-\sin \beta) - \sin \alpha \cdot \cos \beta + 0 \cdot 0 & 0 \\ \sin \alpha \cdot \cos \beta + \cos \alpha \cdot \sin \beta + 0 \cdot 0 & \sin \alpha \cdot (-\sin \beta) + \cos \alpha \cdot \cos \beta + 0 \cdot 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) & 0 \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) & 0 \\ 0 & 0 & 1 \end{bmatrix} = R_{\alpha+\beta}
\end{aligned}$$

17. Mostre que duas escalas 2D sucessivas são multiplicativas.

Resposta:

$$S_\alpha \cdot S_\beta = \begin{bmatrix} \alpha_x & 0 & 0 \\ 0 & \alpha_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_x & 0 & 0 \\ 0 & \beta_y & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x \cdot \beta_x & 0 & 0 \\ 0 & \alpha_y \cdot \beta_y & 0 \\ 0 & 0 & 1 \end{bmatrix} = S_{\alpha \cdot \beta}$$

18. Mostre que a rotação e a escala 2D são comutativas se os fatores de escala $S_x = S_y$.

Resposta:

$$\begin{aligned}
R \cdot S &= \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos \theta \cdot S_x & -\sin \theta \cdot S_y & 0 \\ \sin \theta \cdot S_x & \cos \theta \cdot S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
S \cdot R &= \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} S_x \cdot \cos \theta & -S_x \cdot \sin \theta & 0 \\ S_y \cdot \sin \theta & S_y \cdot \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

19. Descreva o conceito de transformada afim.

Resposta:

A transformada afim traduz transformações projetivas de coordenadas, permitindo a transformação de diversos pontos de uma mesma figura para uma determinada perspectiva através de uma matriz \mathbf{M} .

As transformações afins generalizam as transformações de rotação, translação, escala, reflexão e cisalhamento. Elas preservam o paralelismo entre retas e planos, mas não preservam comprimentos, distâncias e áreas.

20. Descreva as principais diferenças entre projeção ortográfica e projeção perspectiva.

Resposta:

A transformação ortográfica realiza o mapeamento de pontos tridimensionais sobre o plano da imagem, tal que os pontos são projetados ao longo de linhas paralelas na imagem. Já a projeção perspectiva, apesar de também realizar transformação de pontos tridimensionais, altera o tamanho dos objetos de acordo com a distância do centro de projeção. Os efeitos de mudança de tamanho estão relacionados à percepção de profundidade do sistema visual humano - e, portanto, a forma dos objetos não é preservada.

21. Qual a vantagem da utilização de coordenadas homogêneas para a representação de transformações geométricas?

Resposta:

As coordenadas homogêneas possuem a vantagem de simplificar e unificar cálculos em matrizes. Por

exemplo, elas permitem: realizar operações de translação e rotação simultaneamente; representação de pontos no infinito (o que permite eliminar casos especiais e tornar os cálculos mais simples e uniformes).