

Visão e Processamento de Imagens - Segunda Avaliação

Gustavo Estrela de Matos - 8536051

2 de julho de 2017

Q1. Responda o melhor que puder sobre as questões abaixo de Processamento de Imagens:

1. Explique as diferenças entre Computação Gráfica, Processamento de Imagens e Visão Computacional.

R: A computação gráfica é responsável por criar representações visuais de cenas descritas por uma linguagem. A visão computacional faz o caminho contrário, transformando representações visuais em uma descrição formal. Já o processamento de imagens não envolve descrição formal de uma cena, apenas transforma uma imagem em outra.

2. Explique as semelhanças e as diferenças entre as representações de imagens baseadas em reticulados completos e em espaços vetoriais infinitos. Tome uma imagem binária de 4 pontos (2x2) e represente-a num reticulado completo (para facilitar o seu desenho, sugiro usar a função `Posets.BooleanLattice` do SAGE, um software matemático muito divertido). Represente a mesma imagem no espaço das frequências (sugiro usar a `fft` o `numpy`, como fizemos em classe).

R: Temos como semelhanças:

- É possível aplicar transformações na imagem em ambas representações, como por exemplo a filtragem.
- A menos de erros de precisão e ruído, ambas representações representam unicamente a mesma cena.
- Para ambas representações utilizamos o mesmo número de parâmetros.

Como diferenças, temos:

- Quando representamos uma imagem no reticulado booleano completo, estamos representando a intensidade de cada ponto (píxel) no espaço da cena. Por outro lado, quando representamos a imagem no espaço vetorial, usando a transformada de fourier, estamos representando a mudança de intensidade dos pixels ao longo do espaço da cena.
- As estruturas algébricas são diferentes, portanto existem noções diferentes em cada representação. Por exemplo, não existe um elemento inverso no reticulado booleano e existe um no espaço vetorial. Da mesma forma, não existe complemento no espaço vetorial, mas existe no reticulado.

$$\begin{bmatrix} 2 & -1 - j \\ 0 & -1 + j \end{bmatrix}$$

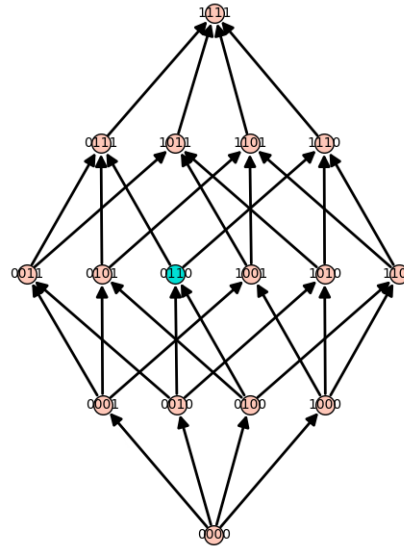


Figura 1: A direita, em ciano, a representação no reticulado booleano completo de uma imagem. A esquerda, a representação da Transformada de Fourier da mesma imagem.

3. Os problemas de visão são mal-postos, isto é, (1) eles podem não ter solução, ou (2), se têm, a solução não é única, ou (3) a solução é muito sensível aos parâmetros iniciais (pouco robusta). Como a visão humana trata isso? Como a visão de máquina poderia tratar isso?

R: A nossa visão trata esses problemas pelas experiências passadas. Se vemos uma cena ambígua, com múltiplas interpretações, então a informação que interpretamos é a mais provável das que conseguimos imaginar, baseados na nossa experiências passadas com objetos parecidos. A mesma ideia pode ser aplicada em visão computacional, com a utilização de um modelo que é capaz de calcular probabilidades das diversas possíveis soluções para cena ou, na prática, que é capaz de estimar essas probabilidades, como se faz em aprendizado de máquina.

- Q2.** Esta questão é inspirada no blog da Tanya Khovanova. No filme Missão Impossível III, Ethan Hunt (Tom Cruise) precisa roubar o “Pé de Coelho” de um prédio muito alto em Shanghai (segundo prédio desenhado na foto abaixo). Do apartamento onde está escondido, ele planeja a estratégia da missão estudando o prédio alvo olhando pela janela do apartamento. Ele pretende entrar pelo topo do prédio e a forma de fazer isso é subir até o topo do prédio que se enxerga no lado esquerdo da foto e usar uma corda, como um pêndulo, para se jogar do prédio mais alto para o prédio alvo (prédio do meio). Para calcular o comprimento da corda, ele usa a projeção dos prédios no vidro da sua janela.

- Que propriedades da luz ele está usando que justificam o seu desenho? Discuta.

R: Ele está usando a propriedade que a luz viaja em linha reta. Considerando um olho humano pontual é possível relacionar dois triângulos que compartilham um ângulo para descobrir o tamanho da corda.

O primeiro triângulo é formado pelo olho e os dois pontos dos prédios; o segundo, pelo olho e os dois pontos dos prédios no desenho no vidro. Como a luz viaja em

linha reta, o ângulo formado no vértice do olho é igual em ambos triângulos.

- Que informações ele ainda precisa para seu plano dar certo? Discuta.

R: Para descobrir o tamanho do pêndulo, Tom deve saber a distância entre a projeção (vidro) e os dois prédios. Outro ponto que Tom deve considerar é saber o ângulo formado entre os dois prédios; veja o exemplo abaixo de duas soluções diferentes para mesma projeção no vidro:

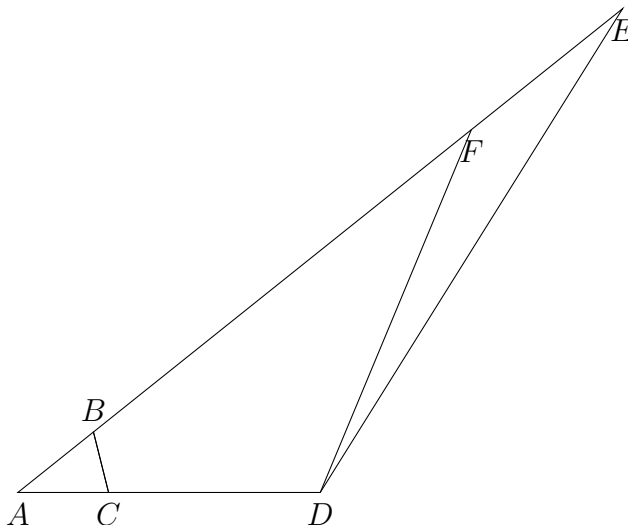


Figura 2: Nessa figura, o ponto A representa o olho de Tom; B e C representam os pontos dos prédios desenhados no vidro; D representa a ponta de um dos prédios; e F e E representam duas possíveis localizações do segundo prédio. Com essa figura, podemos perceber que, para as duas possíveis posições E e F de um dos prédios, a projeção no vidro é igual, mas a distância entre os prédios e o tamanho necessário para a corda são diferentes.

- Q3.** – A convolução das transformadas de Fourier de duas funções é a operação definida por:

$$F(\omega) \otimes G(\omega) = \int_{k=-\infty}^{\infty} F(\nu)G(\omega - \nu)d\nu$$

Mostre que a convolução de $F(\omega)$ e $G(\omega)$ é igual a transformada de Fourier de $f(x)$ e $g(x)$

R: Seja $h(t)$, $f(t)$ e $g(t)$ funções e $H(\omega)$, $F(\omega)$ e $G(\omega)$ suas respectivas transformadas de Fourier. Então, se $H(\omega)$ é a convolução de duas funções $F(\omega)$ e $G(\omega)$, então:

$$\begin{aligned} h(t) &= \frac{1}{2\pi} \int_{\omega=-\infty}^{\infty} H(\omega)e^{i\omega t} d\omega \\ &= \frac{1}{2\pi} \int_{\omega=-\infty}^{\infty} e^{i\omega t} \int_{\nu=-\infty}^{\infty} F(\nu)G(\omega - \nu) d\nu d\omega \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi} \int_{t=-\infty}^{\infty} \int_{\nu=-\infty}^{\infty} e^{i\omega\nu} F(\nu) G(\omega - \nu) d\nu d\omega \\
&= \frac{1}{2\pi} \int_{\nu=-\infty}^{\infty} \int_{\omega=-\infty}^{\infty} e^{i\omega t} F(\nu) G(\omega - \nu) d\omega d\nu \\
&= \frac{1}{2\pi} \int_{\nu=-\infty}^{\infty} e^{i\nu t} F(\nu) \int_{\omega=-\infty}^{\infty} e^{i(\omega-\nu)t} G(\omega - \nu) d\omega d\nu \\
&= \int_{\nu=-\infty}^{\infty} e^{i\nu t} F(\nu) g(t) d\nu \\
&= \int_{\nu=-\infty}^{\infty} e^{i\nu t} F(\nu) d\nu g(t) \\
&= 2\pi f(t)g(t)
\end{aligned}$$

Note que, seguindo os passos anteriores na ordem inversa, a volta também vale. Portanto, provamos que, a menos de uma constante, a transformada de Fourier do produto de duas funções é a convolução das transformadas.

- Calcule, “na mão” (explicitando os cálculos como fizemos em classe) a transformada de Fourier da imagem abaixo:

$$\begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}$$

R: Temos que $F(k) = \sum_{n=0}^3 f(n)e^{-jk2\pi n/4}$ para $k = 0, \dots, 3$.

$$\begin{aligned}
F(0) &= f(0)e^{-j(0)2\pi(0)/4} + f(1)e^{-j(0)2\pi(1)/4} + f(2)e^{-j(0)2\pi(2)/4} + f(3)e^{-j(0)2\pi(3)/4} \\
&= f(0) + f(1) + f(2) + f(3) \\
&= 2
\end{aligned}$$

$$\begin{aligned}
F(1) &= f(0)e^{-j2\pi(0)/4} + f(1)e^{-j2\pi(1)/4} + f(2)e^{-j2\pi(2)/4} + f(3)e^{-j2\pi(3)/4} \\
&= f(0)e^0 + f(1)e^{-j\pi/2} + f(2)e^{-j\pi} + 0 \\
&= f(0) + f(1)(\cos(\frac{\pi}{2}) - j\sin(\frac{\pi}{2})) + f(2)(\cos(\pi) - j\sin(\pi)) \\
&= 1 - 1(0 - j) + 2(-1 - 0j) \\
&= -1 + j
\end{aligned}$$

$$\begin{aligned}
F(2) &= f(0)e^{-j4\pi(0)/4} + f(1)e^{-j4\pi(1)/4} + f(2)e^{-j4\pi(2)/4} + f(3)e^{-j4\pi(3)/4} \\
&= f(0)e^0 + f(1)e^{-j\pi} + f(2)e^{-j2\pi} + 0 \\
&= f(0) + f(1)(\cos(\pi) - j\sin(\pi)) + f(2)(\cos(2\pi) - j\sin(2\pi)) \\
&= 1 - 1(-1 - 0j) + 2(1 - 0j) \\
&= 4
\end{aligned}$$

$$F(3) = f(0)e^{-j6\pi(0)/4} + f(1)e^{-j6\pi(1)/4} + f(2)e^{-j6\pi(2)/4} + f(3)e^{-j6\pi(3)/4}$$

$$\begin{aligned}
&= f(0)e^0 + f(1)e^{-j3\pi/2} + f(2)e^{-j\pi} + 0 \\
&= f(0) + f(1)(\cos(\frac{3\pi}{2}) - j\sin(\frac{3\pi}{2})) + f(2)(\cos(\pi) - j\sin(\pi)) \\
&= 1 - 1(0 - (-1)j) + 2(-1 - 0j) \\
&= -1 - j
\end{aligned}$$

Portanto, a transformada de Fourier da imagem é:

$$\begin{bmatrix} 2 & -1 + j \\ 4 & -1 - j \end{bmatrix}$$

- Q4.** Usando as operações e operadores da Morfologia Matemática, tente segmentar as letras manuscritas (“binarizar”) das imagens de documentos manchadas do arquivo DIP.zip. As melhores soluções são aquelas que funcionam para a maior parte das imagens. **Este é um problema real, não espere resolvê-lo totalmente, mas almeje isto. Por favor, essas imagens são proprietárias, assim, não as distribua..** Sua solução deve usar a toolbox de Morfologia Matemática do prof. Lotufo, dentro de um Jupyter notebook. Envie apenas o notebook, não precisa enviar as imagens.
- Q5.** Questão extra para os alunos de pós (vale nota extra para os alunos de graduação que quiserem fazer).
- Leia o artigo: “Bilateral Filtering for Gray and Color Images” de Tomasi e Manduchi (pegue o artigo aqui: <https://users.cs.duke.edu/~tomasi/papers/tomasi/tomasiIccv98.pdf>). Faça um resumo de uma página do artigo. O resumo deve conter a ideia geral, os principais resultados, as vantagens e desvantagens do método e uma análise dos resultados experimentais. Escolha 10 imagens em níveis de cinza e borre-as com um filtro gaussiano (kernel à sua escolha).
 - Implemente o filtro bilateral, ou use a implementação da OpenCV via python (<http://opencvpython.blogspot.com.br/>, <http://docs.opencv.org/modules/imgproc/doc/filtering.html#bilateralfilter>) e teste o filtro com as imagens borradas. Comente os resultados comparando-os qualitativamente com as imagens originais. Esta parte da questão deve ser feita em notebook do Jupyter e deve ser anexada no zip (tgz) file que será entregue.

Q5. Auto-avaliação

1. Você tentou resolver algum exercício passado em classe? Lembre quais?

R: Sim. Resolvi os seguintes:

- Calcular a equalização de uma imagem monocromática.
- Comparar o método de Pun com método de Otsu.
- Na distância de Mahalanobis, descobrir a forma de Σ quando as variáveis são independentes.
- Exercício sobre dilatação e erosão (no papel quadriculado).
- Provar que \sin e \cos são ortogonais com o produto interno que definimos.

2. Quantas horas por semana você tem para estudar extra-classe?

R: 35 horas em média.

3. Dessas horas, quantas você usa para acompanhar esta disciplina?

R: 4 horas em média.

4. Você tem motivação para assistir às aulas? Se sim, o que aumentaria ainda mais sua motivação? Se não, o que você sugere para que as aulas sejam motivadoras?

R: Sim. Teria mais motivação se tivessem mais exercícios propostos em sala, não necessariamente EPs.

5. A disciplina satisfaz suas expectativas?

R: Sim, minha expectativa para essa matéria era conhecer os principais tópicos de visão computacional e ela foi satisfeita.

6. Além do que foi dado, o que mais você gostaria de aprender?

R:

- Mapeamento tridimensional.
- Biblioteca OpenCV.

7. Qual dos tópicos você gostaria que fossem aprofundado?

R: Provavelmente isso não deveria ser uma obrigação do curso de visão, mas gostaria de ter aprendido com calma séries de Fourier, porque elas aparecem em muitas áreas da computação e, apesar de ter ganhado uma intuição boa sobre ela no curso, não me sinto muito seguro quando estou usando.

8. Você tentou fazer os exercícios do curso do prof. Lotufo? Explícite quais.

R: Resolvi apenas o exercício da lista, que era obrigatório.

9. O suporte do paca foi proveitoso? O que o tornaria mais proveitoso?

R: Foi proveitoso e acho que foi bem utilizado.

10. A quantidade de exercícios foi suficiente? Você sente-se seguro para resolver problemas de processamento de imagens? O que poderia ser melhorado nos exercícios?

R: Não, acho que precisávamos de fazer mais exercícios durante o curso

11. Faça uma breve auto-avaliação de seu desempenho considerando sua facilidade em entender um problema de processamento de imagens, implementar sua solução, seu entendimento dos conceitos, sua assiduidade às aulas, sua participação em aulas e seu desempenho na prova. Se você se sentir confortável para isso, atribua-se uma nota de 0 a 10 de acordo com sua auto-avaliação. Essa nota será considerada no cômputo final da sua nota, caso o professor entenda que você soube se auto-avaliar.

R: Acho que aproveitei bem a matéria. Consegui estudar em casa quase todo o conteúdo que foi passado em sala de aula, relendo os slides e, quando os slides não eram o suficiente, procurando na bibliografia do curso ou na internet. Me interessei bastante por alguns assuntos como análise de aglomerados, aprendizado e morfologia. Entretanto, considero que, apesar de ter lido sobre os assuntos tratados, não me esforcei em fazer, por conta própria, exercícios práticos com os assuntos que tratamos em aula.

Não tive grandes dificuldades de entendimento, e acho que minha maior dificuldade do curso foi com a Transformada de Fourier. Não faltei em nenhuma aula e, acho que fui bem na prova. Tive algumas participações em sala, mas não muitas.