

Visão e Processamento de Imagens - Segunda avaliação

Nome: Lucas Sung Jun Hong
NUSP: 8124329

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

Resolução:

Em computação gráfica recebe-se, como entrada, parâmetros para a sintetização de imagens. Já em Processamento de Imagens, a entrada é uma imagem digital existente que sofrerá uma filtração e/ou manipulação para algum propósito. Finalmente em Visão Computacional desenvolve-se sistemas complexos a partir de dados extraídos de imagens digitais.

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).

Resolução:

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?

Resolução:

O problema existe pois a visão humana sofre com interpretações distorcidas criadas pelo próprio cérebro. O olho humano recebe muitas informações que infelizmente a mente não consegue dar conta, podendo produzir ilusões. O que ocorre muitas vezes no desenvolvimento de teorias é que distorce-se a realidade para que ela se encaixe na análise, tal que ainda vemos grandes teorias aceitas sendo substituídas ou descartadas a medida que novas tecnologias são desenvolvidas para complementar as análises humanas.

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

Resolução:

A refração é a propriedade da luz envolvida na cena, sendo que a luz do fundo passa através da janela até aos olhos do personagem Ethan Hunt, sendo que este usa a projeção das imagens para desenhar os prédios. No entanto, os cálculos de Ethan sofrem grandes problemas sendo que ele infere os dados através de uma imagem bi-dimensional sem parecer ter conhecimento das dimensões reais dos prédios e a real distância entre eles: os prédios podem estar distantes para mais ou para menos do que aparentam ser. Mesmo que os prédios das quais ele faz o pulo e precisa invadir fosse perfeitamente paralelos com o prédio da qual ele está presente na cena, ele precisa levar em conta o ângulo horizontal entre ele e as alturas de interesse de cada um dos dois prédios.

- 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 $g(x)$.

Resolução:

$$\text{Temos } \mathcal{F}[f(x)g(x)] = \int_{-\infty}^{\infty} f(x)g(x)e^{-i\omega x}dx$$

$$= \int_{-\infty}^{\infty} f(x) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} F(\nu)e^{i\nu x}d\nu \right] e^{-i\omega x}dx$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\nu) \left[\int_{-\infty}^{\infty} f(x)e^{-ix(\omega-\nu)}dx \right] d\nu$$

tal que $\int_{-\infty}^{\infty} f(x)e^{-ix(\omega-\nu)}dx = G(\omega - \nu)$. Assim,

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\nu)G(\omega - \nu)d\nu$$

Logo,

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

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

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

Resolução:

Tendo $F(k) = \sum_{n=0}^{N-1} f(n)e^{-j2\pi \frac{k}{N}n}$ e $N = 4$, teremos:

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

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

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

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

Temos assim:

$$\begin{bmatrix} -1 + j & -1 - j \\ 4 + 0j & 2 + 0j \end{bmatrix}$$

- 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.

Resolução:

Programa **8124329.m** escrito em Octave 3.8.2. As imagens devem estar no mesmo diretório do programa. O código abre as imagens numeradas de 1 a 10, com a extensão *.bmp*, e escreve imagens com a extensão: *seg.bmp*. Por exemplo, recebendo uma imagem chamada *7.bmp*, ele retorna uma imagem chamada *7seg.bmp*.

O programa aplica as transformações dependendo do tamanho da imagem. Supondo que a imagem tenha dimensão (m, n) : se $\min\{m, n\} < 800$, o comprimento da vértice da função **strel** terá valor 20. Caso contrário, seu comprimento será 5. O valor do threshold varia também dependendo do tamanho: aplicamos um threshold mais intenso para imagens menores. O programa aplica também **imdilate** para a dilatação para imagens maiores para melhor definição dos segmentos.

O programa foi testado em Debian Jessie. Foram instalados os pacotes de Octave usando os comandos: *apt-get install octave; apt-get install liboctave-dev; apt-get install octave-image*

Q5. Auto-avaliação

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

Das que eu lembro, os exercícios resolvidos foram:

- Resolver o histograma;
- Calcular eigen vectors na aula de Aprendizado Estatístico e Computacional;
- Morfologia Matemática, em que tínhamos que desenhar as figuras;

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

No início, estava com quase 25 horas.

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

No início, usava 6h a 8h para brincar com as ferramentas do python para manipular as imagens.

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?

Não tenho mais motivação. Talvez aulas mais práticas, testar os códigos seriam um pouco mais motivador para mim.

5. A disciplina satisfaz suas expectativas?

Sim.

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

Sem comentário, sendo que assisti poucas aulas.

7. Qual dos tópicos você gostaria que fossem aprofundado?

Manipulação de imagens com Python.

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

Não.

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

Acho que o suporte do paca foi suficiente, segundo alguns colegas.

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?

Gostaria que tivesse sido dado mais exercícios similares ao Mandatory Task, o que tinha sido muito legal.

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.

Consegui acompanhar bem até a Filtragem da Mediana. Mas, sinceramente, uma vez que passou desses tópicos, a partir de Aprendizado Estatístico já não entendi e não consegui acompanhar mais. Sendo uma pessoa que depende muito de conteúdo visual, momentos em que envolvem teorias e estatísticas são os fatores decisivos para o meu mau-performance. Eu daria 4 para a minha auto-avaliação.