Universidade Estadual de Campinas
Instituto de Computação

Introdução ao Processamento Digital de Imagem (MC920 / MO443)

Professor: Hélio Pedrini

Trabalho 5

1 Especificação do Problema

A Análise de Componentes Principais (do inglês, *Principal Component Analysis* - PCA) é uma formulação matemática utilizada na redução de dimensionalidade de dados. No contexto de processamento de imagens, a técnica PCA pode ser empregada na compressão de imagens com um certo grau de perda.

Na abordagem PCA, a informação contida no conjunto de dados é representada por meio de uma estrutura com dimensões reduzidas baseada na projeção dos dados em um subespaço gerado por um sistema de eixos ortogonais. O sistema de eixos pode ser obtido por meio da técnica de Decomposição em Valores Singulares (do inglês, *Singular Value Decomposition* - SVD).

O objetivo deste trabalho é aplicar a técnica SVD para comprimir imagens digitais. A SVD de uma matriz real $A_{n \times p}$ corresponde à fatoração

$$A = U\,\Sigma\,V^T$$

em que U é uma matriz unitária real $n \times n$, Σ uma matriz retangular diagonal $n \times p$ com números reais não-negativos na diagonal e V uma matriz unitária real $p \times p$.

Exemplo:

$$\overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \end{bmatrix}}^{A} = \overbrace{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}}^{U} \overbrace{\begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}^{\Sigma} \overbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ 0 & 0 & 0 & 1 & 0 \\ \sqrt{0.8} & 0 & 0 & 0 & -\sqrt{0.2} \end{bmatrix}}^{V^T}$$

Como as matrizes U e V^T são unitárias reais, então $UU^T = I$ e $VV^T = I$, em que I é a matriz identidade. Dessa forma, U e V^T são matrizes ortogonais. As colunas de U são os autovetores da matriz AA^T, as colunas de V são os autovetores da matriz A^TA e os elementos diagonais da matriz Σ são as raízes quadradas dos autovalores de AA^T ou A^TA. Os autovalores estão dispostos na matriz Σ em ordem decrescente, ou seja, se σ_i são os elementos diagonais de Σ, então $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n > 0$. Dessa forma,

$$A = U\,\Sigma\,V^T = \begin{bmatrix} \mathbf{u_1} & \mathbf{u_2} & \dots & \mathbf{u_n} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_n \end{bmatrix} \begin{bmatrix} \mathbf{v_1}^T \\ \mathbf{v_2}^T \\ \vdots \\ \mathbf{v_n}^T \end{bmatrix} = \mathbf{u_1}\sigma_1\mathbf{v_1}^T + \mathbf{u_2}\sigma_2\mathbf{v_2}^T + \ldots + \mathbf{u_n}\sigma_n\mathbf{v_n}^T$$

Como os elementos σ_i estão ordenados decrescentemente, os primeiros termos $\mathbf{u_i}\sigma_i\mathbf{v_i}^T$ contribuem mais significativamente para o somatório. No contexto de processamento de imagens, manter apenas alguns desses termos pode resultar em uma imagem de menor qualidade, entretanto, que requer menor capacidade de armazenamento.

Dada uma imagem colorida, em que cada banda de cor é especificada no intervalo entre 0 e 255, aplique a técnica PCA em cada banda separadamente considerando apenas alguns componentes e, em seguida, gere novamente a imagem colorida. Utilize diferentes quantidades de componentes para a geração das imagens de saída. O Algoritmo 1 apresenta os principais passos do método.

Algoritmo 1: Compressão com Análise de Componentes Principais

input : Imagem f com dimensões $M \times N$ pixels
Número de componentes k
output: Imagem g com dimensões $M \times N$ pixels

\# dividir a imagem RGB em três canais e aplicar a técnica SVD em cada canal
for i = *1 : 3* **do**
    $[U_f(:,:,i), S_f(:,:,i), V_f(:,:,i)]$ = svd(double($f(:,:,i)$))
\# considerar apenas k componentes e combinar novamente os canais
g = zeros(M, N)
for i = *1 : 3* **do**
    $U_g(:,1:k,i) = U_f(:,1:k,i)$
    $S_g(1:k,1:k,i) = S_f(1:k,1:k,i)$
    $V_g(1:k,:,i) = V_f(1:k,:,i)^T$
    $g(:,:,i) = U_g(:,:,i) * S_g(:,:,i) * V_g(:,:,i)$
return g

Alguns resultados da compressão são ilustrados na Figura 1.

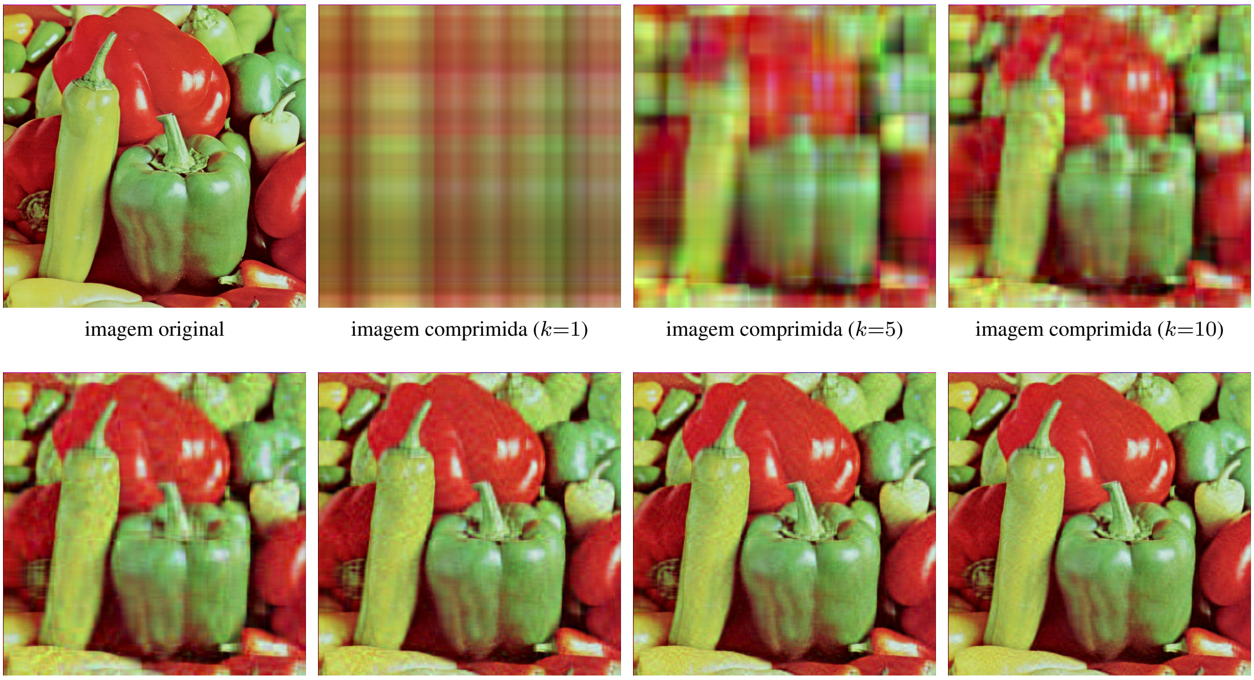

imagem original | imagem comprimida (k=1) | imagem comprimida (k=5) | imagem comprimida (k=10)

imagem comprimida (k=20) | imagem comprimida (k=30) | imagem comprimida (k=40) | imagem comprimida (k=50)

Figura 1: Exemplos de resultados de compressão.

A avaliação da compressão deve ser realizada pela taxa de compressão ρ e pela raiz do erro médio quadrático (RMSE), cujas medidas são definidas como:

$$\rho = \frac{\text{quantidade de memória requerida para representar } g}{\text{quantidade de memória requerida para representar } f}$$

$$\text{RMSE} = \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x,y) - g(x,y)]^2}$$

em que f e g são as imagens original e comprimida, respectivamente.

2 Entrada de Dados

As imagens de entrada estão no formato PNG (*Portable Network Graphics*). Alguns exemplos encontram-se disponíveis no diretório: `http://www.ic.unicamp.br/~helio/imagens_coloridas_png/`

3 Saída de Dados

As imagens de saída devem estar no formato PNG (*Portable Network Graphics*). Resultados intermediários podem ser também exibidos na tela.

4 Especificação da Entrega

- A entrega do trabalho deve conter os seguintes itens:
  - código fonte: o arquivo final deve estar no formato *zip* ou no formato *tgz*, contendo todos os programas necessários para sua execução.
  - relatório impresso: deve conter uma descrição dos algoritmos e das estruturas de dados, considerações adotadas na solução do problema, testes executados, eventuais limitações ou situações especiais não tratadas pelo programa.
- O trabalho deve ser submetido por meio da plataforma *Google Classroom*.
- Data de entrega: 21/11/2019.

5 Observações Gerais

- Os programas serão executados em ambiente Linux. Os formatos de entrada e saída dos dados devem ser rigorosamente respeitados pelo programa, conforme definidos anteriormente. Não serão aceitos trabalhos após a data de entrega.
- Os seguintes aspectos serão considerados na avaliação: funcionamento da implementação, clareza do código, qualidade do relatório técnico.