多媒体技术课堂提问

1 特征脸(EigenFace)与PCA的关系

特征脸算法是基于主成分分析(PCA)的人脸识别方法。特征子脸技术的基本思想是：从统计的观点，寻找人脸图像分布的基本元素，即人脸图像样本集协方差矩阵的特征向量，以此近似地表征人脸图像。**(总)**

特征脸算法的基本步骤如下：

1. **构建样本集合**：获取包含有M张人脸图像的集合S，每张人脸图片的大小缩放到统一的尺寸，得到图像矩阵集合

$$S = \{\Gamma_1, \Gamma_2, \Gamma_3, ...,\Gamma_M\}$$

2. **求得图像矩阵均值，也就是平均脸**：

$$mean = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$$

3. **图像矩阵零均值化**

$$\Phi_i = \Gamma_i - mean$$

4. **计算协方差矩阵及其特征值、特征向量**

$$\begin{aligned} C &= \frac{1}{M}\sum_{i=1,j=1}^{M}\Phi_i\Phi_j^T \\ &= AA^T \\ A &= \{\Phi_1, \Phi_2, \Phi_3, \cdots\cdots, \Phi_n\} \end{aligned}$$

计算C的特征值$\{\lambda_1, \lambda_2, \ldots, \lambda_{N\times N}\}$与其对应的特征向量$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{\mathbf{N}\times\mathbf{N}}\}$找到前n个特征值所对应的特征向量，组成新的矩阵V，由这些特征向量组成的新图像矩阵，也就是所谓的特征脸。

5. **进行人脸识别**

考虑一张新的人脸，缩放到相同的尺寸，然后进行特征转换，对应的公式为

$$\omega_k = u_k^T(\Gamma - mean)$$

其中$k = 1, 2, \ldots, M$表示对应的特征脸u_k，也就是PCA中的第k个特征映射矢量，通过这M个特征脸，可以将新的人脸转变为在特征脸的坐标系的坐标表示。

$$\Omega^T = [\omega_1, \omega_2, ..., \omega_M]$$

可以通过计算闵式距离，余弦夹角等指标判断新输入的图像和人脸库中图像的相似度。

2 求矩阵的特征值和特征向量

设A是n阶方阵，若存在数λ和非零向量$\boldsymbol{x}$使得

$$Ax = \lambda x$$

则称λ是A的特征值，非零向量x是A的属于特征值λ的特征向量.

$$
\begin{aligned}
\lambda\text{为}A\text{的特征值} &\Leftrightarrow Ax=\lambda x \\
&\Leftrightarrow \lambda x - Ax = 0 \\
&\Leftrightarrow (\lambda E - A)x = 0, x \neq 0 \\
&\Leftrightarrow (\lambda E - A)x = 0 \text{ 有非零解} \\
&\Leftrightarrow |\lambda E - A| = 0.
\end{aligned}
$$

也就是说，求解矩阵的特征值，也就是求解$|\lambda E - A| = 0$的解

求属于λ的特征向量$\Leftrightarrow$求$(\lambda E - A)x = 0$的非零解向量.

例 3 求 $\boldsymbol{A}=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$的全部特征值与特征向量.

【分析】求具体矩阵特征值及其特征向量，一般通过解特征方程求解及方程组求解(具体步骤见上要点)

解 $\boldsymbol{A}$ 的特征方程为

$$|\boldsymbol{A}-\lambda\boldsymbol{E}|=\begin{vmatrix}1-\lambda&2&2\\2&1-\lambda&2\\2&2&1-\lambda\end{vmatrix}=(5-\lambda)\begin{vmatrix}1&2&2\\1&1-\lambda&2\\1&2&1-\lambda\end{vmatrix}$$

$$=(5-\lambda)\begin{vmatrix}1&2&2\\0&-1-\lambda&2\\0&0&-1-\lambda\end{vmatrix}=(5-\lambda)(1+\lambda)^2=0.$$

所以 $\boldsymbol{A}$ 的全部特征值为 $\lambda_1=5,\lambda_2=\lambda_3=-1$.

当 $\lambda_1=5$时，$\boldsymbol{A}-5\boldsymbol{E}=\begin{pmatrix}-4&2&2\\2&-4&2\\2&2&-4\end{pmatrix}\sim\begin{pmatrix}1&0&-1\\0&1&-1\\0&0&0\end{pmatrix}$，所以$(\boldsymbol{A}-5\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$可写成

$$\begin{cases}x_1-x_2=0,\\x_2-x_3=0.\end{cases}$$

得基础解系为$\boldsymbol{\xi}_1=(1,1,1)^{\boldsymbol{T}}$，因此属于$\lambda_1=5$的全部特征向量为$k_1\boldsymbol{\xi}_1(k_1\neq 0\text{的实数})$.

当 $\lambda_2=\lambda_3=-1$时，$\boldsymbol{A}-\boldsymbol{E}=\begin{pmatrix}2&2&2\\2&2&2\\2&2&2\end{pmatrix}\sim\begin{pmatrix}1&1&1\\0&0&0\\0&0&0\end{pmatrix}$，所以$(\boldsymbol{A}-\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$可写成

$x_1+x_2+x_3=0$，得基础解系为$\boldsymbol{\xi}_2=(-1,1,0)^{\boldsymbol{T}},\boldsymbol{\xi}_3=(-1,0,1)^{\boldsymbol{T}}$. 因此属于$\lambda_2=\lambda_3=-1$的全部特征向量为$k_2\boldsymbol{\xi}_2+k_3\boldsymbol{\xi}_3$ $(k_2,k_2\text{为不全零的实数})$.

【评注】 1. $\boldsymbol{A}$ 中的 λ_i 对应线性无关的特征向量即为$(\lambda_i\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=0$基础解系；

2. $\boldsymbol{A}$ 中 λ_i 对应线性无关的特征向量个数即为$(\lambda_i\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=0$基础解系所含向量个数的$\boldsymbol{n}-\boldsymbol{r}(\lambda_i\boldsymbol{E}-\boldsymbol{A})$.

3 如何判断两个特征向量的正交性

设两个特征向量为a_1, a_2

$$a_1 = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, a_2 = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

向量正交即$\boldsymbol{a_1}^{\boldsymbol{T}}\boldsymbol{a_2} = 0$

4 什么是矩阵的秩，请举例说明

矩阵A的秩是任意矩阵A经过行变换化成阶梯矩阵后,行的首元为1的个数。

5 如何构成训练样本的特征子空间（特征脸算法的特征子空间）

计算通过样本矩阵得到的协方差矩阵的特征值、特征向量

$$\begin{aligned} C &= \frac{1}{M}\sum_{i=1,j=1}^{M}\Phi_i\Phi_j^T \\ &= AA^T \\ A &= \{\Phi_1, \Phi_2, \Phi_3, \cdots\cdots, \Phi_n\} \end{aligned}$$

计算C的特征值$\{\lambda_1, \lambda_2, \ldots, \lambda_{N\times N}\}$与其对应的特征向量$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{\mathbf{N}\times\mathbf{N}}\}$找到前n个最大特征值所对应的特征向量，组成新的矩阵V，这个矩阵V即特征子空间。

6 证明图像在经过直方均衡化后，图像的信息发生了丢失(或者说图像的信息熵减少了)

假设图像一共有k个灰度级：$0, 1, 2$，$\cdots, k-1$.直方图均衡过程中，每一个灰度级都被映射到新图像中的一个灰度级，把这个映射记为$f: \{0, 1, 2, \cdots, k-1\} \to \{0, 1, 2, \cdots, k-1\}$. 记原图像中对应频率不为0的灰度级的集合为$S$: $S = \{x \in \{1, 2, \cdots, k-1\} |$原图像中灰度级$x$对应的频率不为$0\}$. 现在将$S$中的元素分为2类:

（1）经过映射，只有这一个灰度级被映射到了对应的的灰度级。即灰度级x满足：$x \in S, f(x) = \hat{x}, \forall y \in S$且$y \neq x, f(y) \neq \hat{x}$.
显然，这个灰度级x在原图中对信息熵的贡献，等于新图像中灰度级$\hat{x}$对信息熵的贡献，均为$-p_x log(p_x)$.

（2）经过映射，这个灰度级和S中若干个另外的灰度级，被映射到了新图像中的同一个灰度级。
假设s个灰度级$x_1, \cdots, x_s$都被映射到新图像中的同一个灰度级y_0,即：$f(x_1) = f(x_2) = \cdots = f(x_s) = y_0$.同时假设灰度级$x_i$对应的频率为$p_i$.
考察这个过程里信息熵的变化:
原图中，对应这s个灰度级的信息熵是$H_1 = \sum_{i=1}^{s} p_i log(\frac{1}{p_i})$,新图中对应灰度级$y_0$的信息熵为$H_2 = (\sum_{i=1}^{s} p_i) log(\frac{1}{\sum_{i=1}^{s} p_i})$.

$$\begin{aligned} H_1 - H_2 &= \sum_{i=1}^{s} p_i log(\frac{1}{p_i}) - (\sum_{i=1}^{s} p_i) log(\frac{1}{\sum_{i=1}^{s} p_i}) \\ &= \sum_{i=1}^{s} p_i [log(\frac{1}{p_i}) - log(\frac{1}{\sum_{i=1}^{s} p_i})] \\ &= \sum_{i=1}^{s} p_i log(\frac{\sum_{i=1}^{s} p_i}{p_i}) > 0 \end{aligned}$$

这表明信息熵在直方图均衡过程中有所损失。

7 证明协方差矩阵的三大性质

$$(1)\ Cov\,(Y,X)\ =\ Cov(X,Y)^T$$

$$\begin{aligned}
&Cov\,(X,Y)\\
&=\frac{1}{m}\begin{bmatrix}(x_1-E\,(x_1))\,(y_1-E\,(y_1)) & \cdots & (x_1-E\,(x_1))\,(y_n-E\,(y_n))\\ (x_2-E\,(x_2))\,(y_1-E\,(y_1)) & \cdots & (x_2-E\,(x_2))\,(x_n-E\,(y_n))\\ . & . & .\\ . & . & .\\ . & . & .\\ (x_n-E\,(X_n))\,(y_1-E\,(y_1)) & \cdots & (x_n-E\,(x_n))\,(y_n-E\,(y_n))\end{bmatrix}\\
&=\frac{1}{m}\begin{bmatrix}(x_1-E\,(x_1))\\ (x_2-E\,(x_2))\\ .\\ .\\ .\\ (x_n-E\,(x_n))\end{bmatrix}[y_1-E\,(y_1),\quad y_2-E\,(y_2),\quad \ldots,\quad y_n-E\,(y_n)]\\
&=E\left((\vec{X}-E(\vec{X}))(\vec{Y}-E(\vec{Y}))^T\right)\\
&\text{故 } Cov\,(Y,X)\ =E\left((\vec{Y}-E(\vec{Y}))(\vec{X}-E(\vec{X}))^T\right)\\
&=Cov(X,Y)^T
\end{aligned}$$

$$(2)\ Cov\,(AX+b,Y)\ =\ ACov\,(X,Y)$$

$$\begin{aligned}
&Cov\,(AX+b,Y)\ =\ E\left((A\vec{X}\ +\ b-E(A\vec{X}\ +\ b))(\vec{Y}-E(\vec{Y}))^T\right)\\
&=\ E\left(\ (A\vec{X}\ +\ b\ -E(A\vec{X})-\ b\ )(\vec{Y}-E(\vec{Y}))^T\right)\\
&=\ E\left(\ (A\vec{X}\ -\ AE(\vec{X}))(\vec{Y}-E(\vec{Y}))^T\right)\\
&=\ AE\left((\vec{X}-E(\vec{X}))(\vec{Y}-E(\vec{Y}))^T\right)\\
&=\ ACov\,(X,Y)
\end{aligned}$$

$$(3)\ Cov\,(X+Y,Z)\ =\ Cov\,(X,Z)\ +\ Cov\,(Y,Z)$$

$$\begin{aligned}
&Cov\,(X+Y,Z)\ =\ E\left((\left(\vec{X}\ +\ \vec{Y}\right)-E\left(\vec{X}\ +\ \vec{Y}\right)\ )(\vec{Z}-E\left(\vec{Z}\right)\ )^T\right)\\
&=\ E\left(\left(\left(\vec{X}\ +\ \vec{Y}\right)\ -\ E\left(\vec{X}\right)\ -\ E\left(\ \vec{Y}\right)\right)(\vec{Z}-E\left(\vec{Z}\right)\ )^T\right)\\
&=\ E\left((\vec{X}-E(\vec{X}))(\vec{Z}-E(\vec{Z}))^T\right)\ +\ E\left((\vec{Y}-E(\vec{Y}))(\vec{Z}-E(\vec{Z}))^T\right)\\
&=\ Cov\,(X,Z)\ +\ Cov\,(Y,Z)
\end{aligned}$$

8 证明大津阈值法的类间方差公式

$$\sigma_B^2=w_0(\mu_0-\mu_T)^2+w_1(\mu_1-\mu_T)^2=w_0w_1(\mu_1-\mu_0)^2$$

$$\begin{aligned}
\sigma_B^2&=w_0(\mu_0-\mu_T)^2+w_1(\mu_1-\mu_T)^2\\
&=w_0(\mu_0-w_0\mu_0-w_1\mu_1)^2+w_1(\mu_1-w_0\mu_0-w_1\mu_1)^2\\
&=w_0(w_1\mu_0-w_1\mu_1)^2+w_1(w_0\mu_1-w_0\mu_0)^2\\
&=w_0w_1(\mu_1-\mu_0)^2
\end{aligned}$$

9 傅里叶描述子基本公式的推导

我们首先假定物体的闭合边界曲线表示为一个总长度为N的离散的坐标序列$\{x(n),y(n)|n=0,1\ldots N-1\}$，则它的复数形式表示为:

$$z(n)=x(n)+jy(n),n=0,1,\ldots,N-1$$

这样这闭合曲线的边界就可以在一维空间上表示。由于闭合曲线的周期数为N(我们取了N个点来表示闭合曲线)，这一边界轮廓曲线的坐标序列也是周期的，周期长度为N。

一维边界轮廓曲线坐标序列的傅里叶变换如下:

$$Z(k)=\sum_{n=0}^{N-1} z(n) \exp \left(-\frac{j 2 \pi k n}{N}\right),(0 \leq n \leq N-1)$$

坐标序列的逆变换为:

$$z(n)=\frac{1}{N} \sum_{k=0}^{N-1} Z(k) \exp \left(\frac{j 2 \pi k n}{N}\right),(0 \leq n \leq N-1)$$

$Z(k)$即为傅里叶描述子

10 证明归一化傅里叶描述子的三种不变性

归一化傅里叶描述子有三种不变性，平移不变性、尺度不变性、旋转不变性。

先考虑旋转对傅里叶描述子的影响，旋转后时域表达式为

$$z(n)^{\prime} = e(j\alpha)z(n) \tag{1}$$

由离散傅里叶变换的线性性质，变换后的$Z(k)$为

$$Z(k)^{\prime} = Z(k)e(j\alpha) \tag{2}$$

接着考虑尺度变换对傅里叶描述子的影响，尺度变换后时域表达式为

$$z(n)^{\prime\prime} = \lambda z(n) \tag{3}$$

由离散傅里叶变换的线性性质，变换后的$Z(k)$为

$$Z(k)^{\prime\prime} = \lambda Z(k) \tag{4}$$

最后考虑平移对傅里叶描述子的影响，平移后时域表达式为

$$z(n)^{\prime\prime\prime} = z(n)+z_0 \tag{5}$$

由离散傅里叶变换的线性性质，变换后的$Z(k)$为

$$Z(K)^{\prime\prime\prime} = \sum_{n=0}^{N-1} z(n) \exp \left(-\frac{j 2 \pi k n}{N}\right)+\sum_{n=0}^{N-1} z_0 \exp \left(-\frac{j 2 \pi k n}{N}\right) \tag{6}$$

当k不为0时，$\sum_{n=0}^{N-1} z_0 \exp \left(-\frac{j 2 \pi k n}{N}\right)$等于

$$\begin{aligned}
\sum_{n=0}^{N-1} z_0 \exp \left(-\frac{j 2 \pi k n}{N}\right) &= z_0 \frac{1-exp(-\frac{j2kN\pi}{N})}{1-exp(-\frac{j2k\pi}{N})} \\
&= z_0 \frac{1-exp(-j2k\pi)}{1-exp(-\frac{j2k\pi}{N})} \\
&= z_0 \frac{1-cos(2k\pi)-jsin(2k\pi)}{1-exp(-\frac{j2k\pi}{N})} \\
&= z_0 \frac{1-1}{1-exp(-\frac{j2k\pi}{N})} \\
&= 0
\end{aligned} \tag{7}$$

所以平移变换后傅里叶描述子只有$Z(0)$发生变换

$$\sum_{n=0}^{N-1} z_0 \exp \left(-\frac{j 2 \pi k n}{N}\right) = \sum_{n=0}^{N-1} z_0 = N z_0 \tag{8}$$

$$Z(0)''' = Z(0) + Nz_0 \tag{9}$$

归一化傅里叶描述子即去掉$Z(0)$项，将$Z(k)$除以$Z(1)$

由于去掉$Z(0)$项，显然具有尺度不变性。

由于旋转变换和尺度变换均满足

$$\frac{Z(k)'}{Z(1)'} = \frac{\alpha Z(k)}{\alpha Z(1)} = \frac{Z(k)}{Z(1)} \tag{10}$$

因此归一化傅里叶描述子满足尺度不变性和旋转不变性

11 若A是实矩阵，且AA^T与A^TA存在非零特征值。那么，AA^T与A^TA的非零特征值相同，且对AA^T的特征向量v，必然存在A^TA的一个特征向量v'，满足关系：

$$v = Av'$$

设α是AA^T的特征值，v是AA^T的特征向量，那么(1)式成立

$$AA^Tv = \alpha v \tag{11}$$

设β是A^TA的特征值，v'是特征A^TA的特征向量，那么(2)式成立

$$A^TAv' = \beta v' \tag{12}$$

对(2)式，左右两边同时乘以A，可得

$$AA^T(Av') = \beta(Av') \tag{13}$$

故Av'为AA^T的一个特征向量，β为AA^T的一个特征值。故得证

12 多媒体技术和人工智能的关系

一是多媒体促使人工智能向着更具可解释性的方向发展;

二是人工智能反过来为多媒体研究注入了新的思维方式，促进了多媒体技术推断能力的发展。

这两个方向形成了一个多媒体智能循环,其中多媒体和AI以交互和迭代的方式相互促进增强。

13 视觉知识的三要素

视觉知识的三要素包括**视觉概念、视觉关系和视觉推理。**