图像矩阵奇异值旋转不变性的重建

石雅镠[1)] 金声震[2)] 宁书年[1)] 周凤颖[1)]
[1)](中国矿业大学(北京)机电与信息工程学院,北京 100083) [2)](国家天文台空间重点实验室,北京 100012)

摘 要 自从 Hong 于 1991 年把奇异值(SV)代数引入图像识别中以来,奇异值作为良好的模式特征得到了广泛的研究和应用。Hong 阐明了图像矩阵奇异值具有很多优良特性,也得出了图像矩阵奇异值具有旋转不变性这一结论。然而,在分析研究中发现,该结论的论证是建立在错误的前提基础上的。本文首先从空间解析几何与矩阵理论对此进行了详细分析,依据实例计算结果进一步引出了相反的结论。最后,提出了用图像矩阵的极变换来重建奇异值的旋转不变性,并给出了理论和实例证明。该研究结果保证了奇异值可作为通用图像特征用于目标识别。

关键词 图像矩阵 奇异值 旋转不变性 极变换
中图法分类号:TP391.4 **文献标识码**:A **文章编号**:1006-8961(2005)06-0717-04

Restructuring the Rotation Invariance of Singular Values for Image Recognition

SHI Ya-liu[1)], JIN Shen-zhen[2)], NING Shu-nian[1)], ZHOU Feng-ying[1)]
[1)](*College of Mech-Electricity and Information Engineering, China University of Mining Technology, Beijing* 100083)
[2)](*State Space Key Laboratory, National Observatory, Beijing* 100012)

Abstract Since Hong introduces the Singular Value (SV) algebra into Image Recognition in 1991, the SV vector as excellent pattern feature has been widely studied and applied. Hong' article gives many excellent characteristics of the singular values extracted from an image matrix, and reaches the conclusion that the singular values are invariant to rotation. But, based on analysis and study, the authors find out that the argument for this conclusion begins with an incorrect premise. In this article, we, firstly, give detailed discussion on this point from both perspectives of Analytical Geometry and Matrix Theory, and, furthermore, come to inverse conclusion with instance calculation results. Then, a method is proposed to maintain the necessary rotation invariance by transforming image matrix to polar coordinate form followed by the proven of the effectiveness in theory and with numeric calculation results. The results of this study still hold the generality of singular value method for object recognition.

Keywords image matrix, singular values, rotation invariance, polar transform

1 引 言

矩阵奇异值分解是 19 世纪 70 年代分别由 Beltrami 和 Jordan 发现的,起因于微分几何中双实线性的判定问题。一般矩阵的奇异值分解形式是 Eckart 和 Yong 于 1939 年给出的。从 20 世纪 70 年代后期至今,有关奇异值的研究迅速发展[1]。奇异值分解之所以能引起广泛的重视,一方面是它有理论上的重要性,如它可以把一般算子分解成酉算子和半正定算子的积。因为总体上来说,我们对一般算子的性质并不清楚,而对酉算子和半正定算子则有详细的研究。另一方面是随着现代数值计算方法和手段的提高,该方法可以解决各种应用问题,如信号与图像处理。

奇异值在图像中的应用是 Hong 于 1991 年引入

基金项目:国家“863”项目(863-2.5.1.25)
收稿日期:2004-10-20;**改回日期**:2005-03-03
第一作者简介:石雅镠(1974~),男。现为中国矿业大学遥感与图形图像学专业博士研究生。从事计算仿真、信号与图像处理方面的研究。E-mail:yl~shi@sohu.com

的[2]。Hong 把图像识别的方法分成以下 4 种：(1)直观方法，如根据形状和对称性；(2)统计方法，如直方图和矩方法；(3)变换系数方法，如傅里叶描述因子方法；(4)代数方法，主要是奇异值方法。文献[2]证明了图像矩阵奇异值的很多优良特性，这些特性主要包括：稳定性、比例灰度的不变性、转置不变性、镜像不变性、平移不变性和旋转不变性。毫无疑问，该文献对发展(代数)图像识别方面具有重要意义。到目前为止，国内外相关的研究基本上以该文的结论作为依据[3,4]。但是，通过仔细分析会发现，其旋转不变性并不能得到满足，已给出的证明实质是错误的。任何一种图像识别特征，如果这一点不能保证，则基于这一特征的方法就不具备通用性。幸运的是，进一步的研究表明，只要把图像矩阵变换到坐标下(以下简称极变换)，旋转不变性能就得到满足。

2 旋转可变性的证明

奇异值分解有两种等价形式——和形式与积形式。前者是奇异值图像压缩的依据，后者是图像识别的依据。以下仅讨论积形式，它的一般提法是：

定理 1[5,6] 设 $\boldsymbol{A}$ 是 $m\times n$ 的实数矩阵(复矩阵一样)，其秩为 r，即 $rank(\boldsymbol{A})=r$，则存在正交矩阵 $\boldsymbol{U}\in\mathbf{R}^{m\times n}$，$\boldsymbol{V}\in\mathbf{R}^{n\times n}$ 使得

$$\boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{\Sigma}_r & 0\\ 0 & 0\end{pmatrix}\boldsymbol{V}^{\mathrm{T}} \tag{1}$$

其中，$\boldsymbol{\Sigma}_r=\mathrm{diag}(\sigma_1,\cdots,\sigma_r)$，$\sigma_1,\cdots,\sigma_r$ 是 $\boldsymbol{A}$ 的全部非零奇异值，为 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 相应特征值的平方根。

文献[2]第 2 节奇异值的性质 3 断言图像矩阵在旋转作用下奇异值具有不变性，并给出了证明(译文)如下：

设原始图像矩阵为 $\boldsymbol{A}$，对该图像旋转，相当于对矩阵 $\boldsymbol{A}$ 左乘一正交矩阵 $\boldsymbol{P}$，得到的图像为 $\boldsymbol{PA}$。于是有

$$(\boldsymbol{PA})(\boldsymbol{PA})^{\mathrm{T}}=\boldsymbol{P}(\boldsymbol{AA}^{\mathrm{T}})\boldsymbol{P}^{\mathrm{T}} \tag{2}$$

其中，$\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{P}^{-1}$。可见对 $\boldsymbol{A}$ 的正交变换导致了对 $\boldsymbol{AA}^{\mathrm{T}}$ 进行正交相似变换。由于 $\boldsymbol{AA}^{\mathrm{T}}$ 和 $\boldsymbol{P}(\boldsymbol{AA})^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}$ 有相同的特征根。因此 $\boldsymbol{A}$ 和旋转后的图像 $\boldsymbol{PA}$ 有相同的奇异值向量。

上述论证的错误在于记原始图像矩阵为 $\boldsymbol{A}$，对其旋转，相当于对其矩阵 $\boldsymbol{A}$ 左乘一正交矩阵 $\boldsymbol{P}$，得到的图像为 $\boldsymbol{PA}$ 的命题非真。

众所周知，旋转是一个几何操作。其算子也的确可用正交矩阵表示，但作用的对象是坐标向量，而不是图像矩阵本身。如在平面解析几何中有坐标旋转变换公式

$$\begin{bmatrix}x'\\ y'\end{bmatrix}=\begin{bmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

其中的旋转算子 $\boldsymbol{P}=\begin{bmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{bmatrix}$ 作用的是坐标向量。

进一步，不失一般性，假定一幅图像可用二次(曲面)型表示，即

$$f(x,y)=(x,y)\boldsymbol{B}(x,y)^{\mathrm{T}}$$

其中，$\boldsymbol{B}$ 为曲面矩阵。根据二次曲面理论，设旋转后的图像为 $f'(x,y)$，则有

$$f'(x,y)=(x,y)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{BP}(x,y)^{\mathrm{T}}$$

而

$$[f'(i,j)]_{m\times n}=\boldsymbol{P}[f(i,j)]_{m\times n} \tag{3}$$

既不成立也没有意义(因为 $\boldsymbol{P}$ 是 2 维矩阵)，其中 (i,j) 是 (x,y) 的离散坐标。所以表达旋转的算子不能作用于图像矩阵。

退一步，假定 $\boldsymbol{P}$ 不是指 2 维图像旋转算子本身，而仅是某一形式上的正交矩阵，那么旋转图像能表达为 $\boldsymbol{PA}$ 吗？下面从矩阵空间理论来进行一般的分析。

设图像矩阵属于 $\mathbf{R}^{m\times n}$，其中的旋转算子为 $\boldsymbol{Q}$，即

$$\boldsymbol{Q}:\mathbf{R}^{m\times n}\to\mathbf{R}^{m\times n}$$

由于

$$\boldsymbol{R}_1=\begin{bmatrix}1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 0\end{bmatrix}_{m\times n}$$

$$\boldsymbol{R}_2=\begin{bmatrix}0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 0\end{bmatrix}_{m\times n}$$

$$\vdots$$

$$\boldsymbol{R}_{mn}=\begin{bmatrix}0 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 1\end{bmatrix}_{m\times n}$$

是 $\mathbf{R}^{m\times n}$ 的一组正交基，因此对于任意图像矩阵 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$ 有

$$\boldsymbol{A}=[a_{ij}]_{m\times n}=\bar{\boldsymbol{A}}\boldsymbol{R}^{\mathrm{T}}$$

其中

$$\bar{\boldsymbol{A}}=(a_{11},a_{12},\cdots,a_{ij},\cdots,a_{mn})$$

$$\boldsymbol{R}=(\boldsymbol{R}_1,\boldsymbol{R}_2,\cdots,\boldsymbol{R}_{i\times n+j},\cdots,\boldsymbol{R}_{mn})$$

旋转后的图像矩阵($\mathbf{R}^{m\times n}$ 空间的向量)$\boldsymbol{A}'$ 满足坐标

变换公式,所以有

$$A' = [a'_{ij}]_{m\times n} = Q_{mn\times mn}(\bar{A}R^{T}) \tag{4}$$

成立,其中,$Q_{mn\times mn}$为正交矩阵。注意,以上关系式必须先做坐标分量的运算,再做矩阵加运算。式(4)不能退化为

$$A' = Q_{mn\times mn}A$$

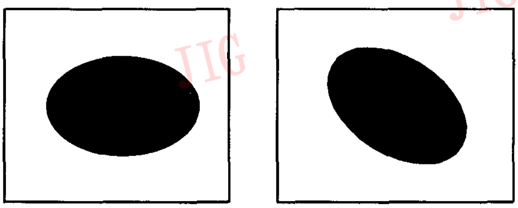
更得不出

$$A' = Q_{m\times m}A \tag{5}$$

所以,旋转图像不能表达为 PA(Q 与 P 只有符号上的区别)。

以上讨论和分析表明,文献[2]证明的前提不成立,相应的结论也就不能保证。

实际上,图像矩阵在旋转作用下奇异值是可变的。如图 1(a)与图 1(b)旋转等价。对它们按 16×16 格式采样进行奇异值分解,所得结果如表 1 所示。可以看出,两图的结果明显不同。一方面,原图像与旋转后图像的非零奇异值个数不一样,前者有 4 个,后者有 7 个,相差接近一倍;另一方面,奇异值向量相差较大:如果用 d 表示两者之差值向量,则距离范数 $\|d\|_2$ 达 3.31(相对差达 50%)。



(a)原图　　(b)旋转 135°图

图 1　计算用图

Fig. 1　Maps for calculation

表 1　图 1 奇异值的计算结果

Tab. 1　Singular values for figure 1

图像	σ_1	σ_2	σ_3	σ_4	σ_5	σ_5	σ_6	$\sigma_7 \sim \sigma_{16}$
原图像	5.44	1.61	0.84	0.33	0.00	0.00	0.00	0.00
旋转后图像	5.29	2.19	1.08	0.89	0.73	0.63	0.42	0.00

3　极变换下的旋转不变性

奇异值的旋转不变性不能得到保证,使其在大多数情况下失去了使用价值。研究发现,通过对图像矩阵进行极变换,可以做到这一点。为此,先证明矩阵奇异值的一条基本性质。把它表述成如下定理:

定理 2　矩阵对换行(或列)后,奇异值保持不变。

证明　设 A 分别对换两行(i 与 j)和两列(k 与 l)所得到的矩阵分别为 A'和 A'',相应的第 1 初等矩阵分别是 $R_{(ij)}(\in Z^{m\times n})$和 $C_{(kl)}(\in Z^{n\times n})$。按定理 1,必有

$$A' = R_{(ij)}A = R_{(ij)}U\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}V^{T}$$

$$= (R_{(ij)}U)\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}V^{T}$$

由初等矩阵的基本性质可知

$$R^{T}_{(ij)} = R^{-1}_{(ij)} = R_{(ij)}$$

故

$$(R_{(ij)}U)^{T}(R_{(ij)}U) = U^{T}(R^{T}_{(ij)}R_{(ij)})U = U^{T}U = I$$

所以存在正交矩阵 $U' = R_{(ij)}U \in R^{m\times m}$,使得

$$A' = U'\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}V^{T} \tag{6}$$

对于列交换,由于

$$A'' = AC_{(kl)} = U\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}V^{T}C_{(kl)}$$

$$= U\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}(V^{T}C_{(kl)})$$

由于

$$C_{(kl)})V^{T} = V^{T}C^{T}_{(kl)} = (C_{(kl)}V)^{T}$$

且同理可证 $C_{(kl)}V$ 是正交矩阵。所以存在正交矩阵

$$V' = C_{(kl)}V \in R^{n\times n}$$

使得

$$A'' = U\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}V'^{T} \tag{7}$$

对比式(1)、式(6)与式(7)可知,交换保持奇异值不变。证毕。

从上面的证明过程可知,当交换矩阵两行(两列)时,只要交换相应奇异值分解的右正交因子(左正交因子)就可得交换后的奇异值分解,并且,对于给定的矩阵,不论交换多少行或多少列,奇异值保持不变。

由于在直角坐标系下,目标的平移可归结为行列交换,因此,图像矩阵在平移作用下,奇异值保持不变。

这就提示我们,为了保证图像识别所要求的旋转不变性,必须改变图像矩阵的形式,使旋转作用表现为图像矩阵交换行(列)。因而根据定理 2,可以保证奇异值的不变性。进一步的分析发现,对于一幅图像,只要在极坐标下采样,以极角为行(列),以极轴为

列(行)构成矩阵,则旋转的效果就是该矩阵的行(列)交换。实际的图像识别针对的是图像矩阵,只要选取一种插值算法进行坐标变换即可。为了反映这种变换的坐标形式和所得矩阵的特点——每一行(列)的第 1 个元素一样,我们称之为极变换。

通过双线性插值下的极变换,对图 1 进行了重新计算,所得结果如表 2 所示。可以看出,原图像与旋转后图像的奇异值之间只存在很小的误差。因此,该方法不仅在理论上是可靠的,而且在实际计算中也具有很高的稳健性。

表 2 极变换后图 1 奇异值的计算结果

Tab. 2 Singular values for figure 1 under polar transform

图像	σ_1	σ_2	σ_3	σ_4	σ_5	σ_5	σ_6	$\sigma_7 \sim \sigma_{16}$
原图像	6.77	2.20	0.83	0.03	0.00	0.00	0.00	0.00
旋转后的图像	6.72	2.19	0.88	0.01	0.00	0.00	0.00	0.00

当然,这一方法显然破坏了奇异值的平移不变性。但这不是原则问题,因为可以在变换前先规格化图像(如使用规则矩方法)。

4 结 论

综上所述,本文得出如下结论:

(1)图像矩阵的奇异值是旋转可变的;为了保证其旋转不变性,必须进行极变换。

(2)当问题是旋转受限时,直接对图像矩阵进行奇异值分解即可用于图像识别;当问题是平移受限时,分解前需要先作极变换;当两者都不受限时,应先进行平移归位,再作极变换。然后再进行奇异值分解获得奇异值向量用于图像识别。

参考文献(References)

1 ZOU H X, WANG D J, DAI Q H, *et al*. Singular Value Decomposition for Extended Matrix[J]. Acta Electronica Sinica, 2001, **29**(3): 289~292. [邹红星,王殿军,戴琼海等. 延拓矩阵的奇异值分解[J]. 电子学报,2001, **29**(3): 289~292.]

2 Hong Z Q, Yang J Y. Algebraic feature extraction of image recognition[J]. Pattern Recognition, 1991, **24**(1): 211~219.

3 ZHOU D L. Face recognition based on singular value decomposition and discriminant KL projection[J]. Journal of Software, 2003, **14**(4): 783~789. [周德龙. 基于奇异值分解和判别式 KL 投影的人脸识别[J]. 软件学报, 2003, **14**(4): 783~789.]

4 Chelleppa R. Human and machine recognition of faces: a survey[J]. Pattern Recognition, 1995, **83**(5): 705~740.

5 ZHANG X D. Linear algebra for signal processing[M]. Beijing: Science Press, 1997: 187. [张贤达. 信号处理中的线性代数[M]. 北京:科学出版社,1997: 187.]

6 LIU H. Matrix Theory and Application[M]. Beijing: Chemical Industry Press, 2003: 166. [刘慧. 矩阵理论与应用[M]. 北京:化学工业出版社,2003: 166.]