

基于四元数矩阵奇异值分解的彩色图像识别^{*})

冉瑞生 黄廷祝

(电子科技大学计算机科学与工程学院 成都 610054)

摘要 本文首先以实验证明了彩色图像矩阵的奇异值(SVs)仅含有图像的少量信息,大量信息体现在图像矩阵的奇异值分解(SVDQ)的两个四元数酉矩阵中。然后给出了一种新的图像特征提取方法。该方法将图像投影到 SVDQ 的各个正交基上,得到投影系数向量。将此向量作为图像的代数特征并用于彩色图像识别中。实验表明,与奇异值特征向量用于彩色图像识别方法相比,本文方法显著提高了识别率。

关键词 彩色图像识别,四元数矩阵,奇异值分解,投影系数向量

The Recognition of Color Images Based on the Singular Value Decompositions of Quaternion Matrices

RAN Rui-Sheng HUANG Ting-Zhu

(School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054)

Abstract In this paper, with the experiment in color image databases, it is shown that the SVs contain little information of color image and most important information is contained in the two quaternion unitary matrices of SVDQ. So, a new feature extraction method for color image has been proposed. Firstly, the image is projected on to the orthogonal basis of SVDQ, then the projection coefficient vector is used as algebraic feature of image and applied to recognition. With the experiments for image recognition, it is shown that the recognition rate is increased compared to the method by the singular feature vector.

Keywords Color image recognition, Quaternion matrix, Singular value decomposition, Projection coefficient vector

1 引言

1843年,Hamilton首次提出了四元数的概念^[1]。一个四元数 q 由一个实部和三个虚部构成:

$$q = a + bi + cj + dk,$$

其中, a, b, c, d 均是实数, i, j, k 均是虚数单位且满足:

$$i^2 = j^2 = k^2 = -1, ij = -ji = k, ki = -ik = j, jk = -kj = i.$$

如果四元数 q 没有实部($a=0$),则称 q 为纯四元数。一个元素为四元数的矩阵称为四元数矩阵。文[2]详细介绍了四元数及四元数矩阵。1996年,Pei第一次提出了彩色图像的四元数模型^[3]。我们知道,一幅彩色图像每个像素都是由三种基本颜色,即红(R),绿(G),蓝(B)按一定的比例组合而成的。这样,彩色图像的每个像素都可以视为一个纯四元数:

$$q(x, y) = r(x, y)i + g(x, y)j + b(x, y)k$$

其中 $r(x, y), g(x, y)$ 和 $b(x, y)$ 分别表示图像中在 (x, y) 点处的红,绿,蓝色的灰度值。于是,一副 $N \times M$ 的彩色图像就可以看作是一个 $N \times M$ 的纯四元数矩阵。基于此模型,近年来有大量的研究将四元数及四元数矩阵作为一种重要的彩色图像处理工具^[3~6]。

文[6]中,作者将彩色图像矩阵的奇异值向量作为图像特征并用于图像识别中,其识别率可达80%左右。但是,彩色图像的奇异值特征向量包含了图像的足够信息吗?本文的实验表明,奇异值向量仅仅包含了图像的少量信息,而图像的大量信息体现在图像矩阵奇异值分解(SVDQ)的两个四元数酉矩阵中。因而我们将图像投影到模板图像的SVDQ的各个

正交基上,得到投影系数向量,并将此向量作为图像的代数特征用于彩色图像识别中。实验表明,本文方法有效且较文[6]的识别率有显著提高。

2 彩色图像的奇异值分解及图像信息

记 $H^{N \times M}$ 为所有 N 行 M 列的四元数矩阵集合。

定理 1^[2] (SVDQ)对于任一个秩为 r 的四元数矩阵 $Q \in H^{N \times M}$,必有两个四元数酉矩阵 U, V 使得

$$Q = U \Sigma V^{\dagger} = \sum_{i=1}^r \sigma_i u_i v_i^{\dagger},$$

其中,

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0, \dots, 0),$$

$$U = (u_1, u_2, \dots, u_N) \in H^{N \times N},$$

$$V = (v_1, v_2, \dots, v_M) \in H^{M \times M}.$$

符号“ $\dagger$ ”表示共轭转置; $\sigma_i (1 \leq i \leq r)$ 均为实数,即为 Q 的奇异值; U, V 分别称为四元数矩阵 Q 的左、右奇异值矩阵。

设 $x = (\sigma_1, \sigma_2, \dots, \sigma_r, 0, \dots, 0)^T$,我们称向量 x 为矩阵 Q 的奇异值特征向量,由定理1知 x 是一个实向量。文[6]中,作者证明了奇异值向量具有稳定性,旋转不变性,平移和镜像变换不变性等性质,进而将其作为彩色图像的代数特征并用于彩色图像识别中。文中的实验表明其识别率可达80%左右。但是,彩色图像的奇异值向量并非包含了图像的足够信息,图像的大量信息体现在图像矩阵的SVDQ的两个四元数酉矩阵中。下面的实验说明了这一点。

我们获取了40种不同彩色的字符A作成大小均为 45×45 的图片构成彩色图像库1。图1给出了图像库1中的部分

^{*})受到教育部“新世纪优秀人才支持计划”资助。冉瑞生 博士生,主要研究方向:图像处理、模式识别。黄廷祝 教授,博导。

图像。

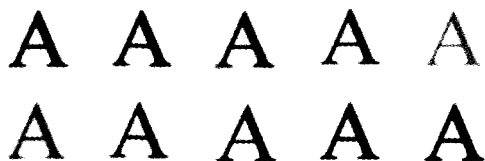


图1 图像库1中的部分图像

设红色字符A图像 Q_1 的奇异值分解为 $Q_1=U_1 \Sigma_1 V_1^d$, 图像库中其余任一图像记为 Q , 设 $Q=U \Sigma V^d$ 。用 Σ_1 替换 Σ , 所得的重建图像记为 $U \Sigma_1 V^d$ 。图2对照图1——列出了部分重建图像。对照图1可见, 所得的重建图像与原图像并无太大的差别, 这表明奇异值的变化对原图像的影响非常小。

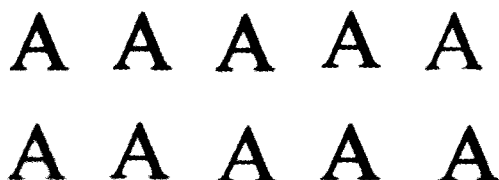


图2 与左上角图像有相同奇异值的部分重建图像

然而, 当保持所有图像的奇异值不变, 而用 Q_1 的两个四元数酉矩阵 U_1, V_1 的部分列向量替换图像库中其余任一图像的SVDQ的两四元数酉矩阵的相应列向量时, 重建图像却发生了很大的变化, 图3对照图1列出了部分重建图像。可见图像的SVDQ的四元数酉矩阵的变化对图像影响较大。



图3 与左上角图像互换四元数酉矩阵的部分列向量后的图像

我们用不同色彩任意20个字符作成大小均为 45×45 的图片构成图像库2, 图4给出了图像库2中的部分图像。在图像库2中做上述相同的实验, 我们也可得到类似的实验结果。



图4 图像库2中的部分图像

上述实验表明彩色图像的奇异值仅含有少量的图像信息, 而大量信息体现在图像矩阵奇异值分解的两个四元数酉矩阵中。而文[6]中将奇异值向量用于图像识别的研究, 其识别率可达80%左右, 其原因在于训练样本库仅含有较少的样本, 或样本库中的图像是在一定的条件下获取的。第5节的实验说明了这一点。

3 图像特征提取

设 $Q \in H^{N \times M} (N < M)$ 为一彩色模板图像, 其奇异值分解设为

$$Q = U \Sigma V^d = \sum_{i=1}^r \sigma_i u_i v_i^d,$$

称 $u_i v_i^d (i=1, 2, \dots, N)$ 为SVDQ的一组正交基。由第2节知图像的大量信息体现在图像矩阵奇异值分解的两个四元数酉矩阵中, 也即体现在向量 u_i, v_i 中。设 $W \in H^{N \times M}$ 是一个彩色图像的四元数矩阵, 我们将 u_i, v_i 作用到 W 上, 也即将 W 投影到模板图像 Q 的正交基 $u_i v_i^d$ 上, 得投影系数

$$\tau_i = u_i^d W v_i.$$

其中 τ_i 是四元数。于是 W 在模板图像 Q 上 N 个基的投影构成一个投影系数向量, 记为

$$\tau = (\tau_1, \tau_2, \dots, \tau_N).$$

我们可将此投影系数向量作为彩色图像 W 关于模板图像 Q 的代数特征并用于图像识别中。事实上, 上述的投影也可以作成

$$\tau_i = v_i^d W^d u_i.$$

而且注意到, 由于 U, V 均是酉矩阵, 于是对任一 $i (i=1, 2, \dots, N)$, 有

$$u_i^d Q v_i = u_i^d \left(\sum_{j=1}^r \sigma_j u_j v_j^d \right) v_i = \sigma_i$$

即模板图像 Q 在其自身的一组正交基 $u_i v_i^d$ 上的投影系数是其相应的奇异值 σ_i , 从而投影系数向量即是 Q 的奇异值特征向量。下面的定理给出了投影系数向量的稳定性。

定理2 设 $W_1, W_2 \in H^{N \times M} (M > N)$, $U = (u_1, u_2, \dots, u_N), V = (v_1, v_2, \dots, v_M)$ 是一模板图像的SVDQ得到的两个四元数酉矩阵。设 W_1, W_2 在正交基 $u_i v_i^d$ 上的投影为 $\tau_i = u_i^d W_1 v_i, \sigma_i = u_i^d W_2 v_i (i=1, 2, \dots, N)$ 。设 $|\cdot|$ 表示两个四元数的模, 则对于四元数矩阵的Frobenius范数 $\|\cdot\|_F$, 有

$$\sqrt{\sum_{i=1}^n |\tau_i - \sigma_i|^2} \leq \|W_1 - W_2\|_F.$$

证明: 由四元数矩阵关于酉不变范数的性质^[7], 有

$$\|W_1 - W_2\|_F^2 = \|U^d (W_1 - W_2) V\|_F^2 \geq \|diag(\tau_1, \tau_2, \dots, \tau_N) - diag(\sigma_1, \sigma_2, \dots, \sigma_N)\|_F^2 = \sum_{i=1}^N |\tau_i - \sigma_i|^2.$$

定理2表明当图像矩阵 W 由于光照变化或噪声干扰而灰度值发生小的变化时, 投影系数对此并不敏感。

4 彩色图像识别

假定训练图像库有 K 类彩色图像, 每类图像有 L 个样本 $Q_j (j=1, 2, \dots, L)$ 。设每一副图像的大小为 $n \times m, p = \min(m, n)$ 。下面给出了将投影系数向量用于彩色图像识别算法的主要步骤:

1. 对第 i 类彩色图像, 计算其平均图像

$$M_i = \frac{1}{L} \sum_{j=1}^L Q_j (i=1, 2, \dots, K),$$

并将其作为第 i 类的模板图像;

2. 对 M_i 作奇异值分解, 得到相应的奇异值向量 S_{M_i} , 四元数酉矩阵 U_{M_i} 和 V_{M_i} 。将 S_{M_i} 作为第 i 类图像的特征向量;

3. 对待识别的彩色图像 X , 将其投影到第 i 类图像的四元数酉矩阵 U_{M_i} 和 V_{M_i} 构成的正交基上, 得相应的投影系数向量 τ , 其中 τ 是一个四元数向量;

4. 用最近邻法分类。计算第 i 类投影系数向量 τ 与奇异

值向量 S_{M_i} 之间的欧氏距离

$$d_i = d(\tau, S_{M_i}) = \|\tau - S_{M_i}\|$$

其中 d_i 是两个四元数向量的距离。若 $d_{i_0} = \min_i(d_i)$, 则彩色图像 X 属于第 i_0 类。

设待识别的图像 X 取自于第 i 类, 由定理 2 知

$$d_i = \|\tau - S_{M_i}\| \leq \|X - M_i\|_F$$

而图像 X 取自于第 i 类, 相对于模板图像 M_i 的变化较小。因而 τ 相对于 S_{M_i} 的变化一般说来也较小, 从而 d_i 较小。而 X 相对于第 j 类 ($j \neq i$) 的模板图像 M_j 的变化较大。因而在 M_j 上的投影系数向量相对于 S_{M_j} 的变化一般说来也较大, 从而 d_j 较大。所以, 将 X 正确地分到第 i 类的可能性较大。而文[6]的方法中, 由于奇异值仅含有图像的少量信息, 因而将其用于识别时的识别率较低, 错分率较高。

5 实验

我们以第 2 节所示的图像库 1 和 2 分别构成训练样本库 1, 2。每一种彩色图像在不同的情况下获取了 10 个不同的样本, 如不同的光照、模糊、一定的倾斜、一定程度的曝光、受到噪声的干扰等。图 5 给出了红色字符 A 作成的图像的部分样本。

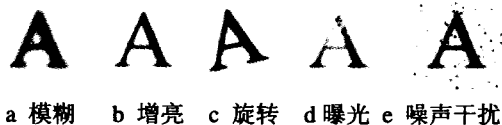


图 5 红色字符 A 作成的图像的部分样本

表 1 在图像库 1 中两种方法识别率的比较

待识别图像类型	文[6]方法识别率	本文方法识别率
无干扰	22.5%	95%
模糊	5%	85%
增亮	5%	80%
旋转	15%	60%
曝光	17.5%	92.5%
噪声干扰	2.5%	87.5%

(上接第 175 页)

- 17 Randell D A, Cohn A G. A Spatial Logic Based on Regions and Connections. In: Nebel B, Richard C, Swartout W. eds. Principles of Knowledge representation and Reasoning. Proceedings of the Third International Conference. Morgan Kaufmann, Los Altos; 1992. 165~76
- 18 Guarino N. Formal Ontology, Conceptual Analysis and Knowledge Representation. International Journal of Human and Computer Studies, 1995, 43(5-6): 625~640,
- 19 Pawlak Z. Rough sets and fuzzy sets. Fuzzy Sets and Systems, 1985, 17: 99~102
- 20 Agrawal R, Imielinski T, Swami A. Mining association rules between sets of items in large databases. Proc of the ACM SIGMOD Conference on Management of Data, Washington, D. C., May 1993
- 21 Chen Bo, Zhou Mingtian. A Pure Mereological Approach to Roughness. In: Wang G, Liu Q, Skowron A, et al. Eds. Proceed-

ings of RSFDGrC 2003, Chongqing, China. LNCS/LNAI 2639, Springer-Verlag, 2003. 425~429

表 2 在图像库 2 中两种方法识别率的比较

待识别图像类型	文[6]方法识别率	本文方法识别率
无干扰	50%	100%
模糊	50%	100%
增亮	25%	80%
旋转	30%	60%
曝光	42.5%	95%
噪声干扰	17.5%	100%

从上表可以看到, 本文方法的识别率有明显提高。可见本文方法是一种有效的图像识别方法。而比较两表可知, 文[6]的方法在样本库 2 中的识别率高于在样本库 1 中的识别率。而文[6]中, 在 9 类图像, 每类 10 个样本构成的样本库里, 识别率可达 80% 左右。可见文[6]的方法适合于样本库较小的情形。在样本库很大时, 文[6]的方法几乎失效。

参考文献

- 1 Hamilton W R. On quaternions. Proceeding of the Royal Irish Academy, 1844
- 2 Zhang F. Quaternion and matrices of quaternion. Lin. Alge. Appl., 1997, 215: 21~57
- 3 Pei S-C, Cheng C-M. A novel block truncation coding of color images by using quaternion moment preserving principle. IEEE International Symposium on Circuits and Systems, Atlanta, 1996
- 4 帅建伟, 陈振湘, 等. 四元神经网络模型: 彩色图像的识别. 计算机学报, 1995, 88(5): 372~379
- 5 李葆青. 基于四元数描述的彩色图像边缘检测器. 中国图形图像学报, 2003, 8(A)(7): 774~778
- 6 Ran R S, Huang T Z. Singular value feature extraction of color image and its application for recognition. In: The 20th congress of the international commission for optics, Changchun, 2005
- 7 刘永辉. 四元数矩阵方程 $AXH^H=B$ 的最小二乘解 (英文). 数学研究, 2003, 36(2): 145~150
- 8 ings of RSFDGrC 2003, Chongqing, China. LNCS/LNAI 2639, Springer-Verlag, 2003. 425~429
- 22 Chen Bo, Zhou Mingtian. Adapting Granular Rough Theory to Multi-agent Context. In: Wang G, Liu Q, Skowron A. Eds. Proceedings of RSFDGrC 2003, Chongqing, China, LNCS/LNAI 2639, Springer-Verlag, 2003. 701~705
- 23 Chen Bo, Zhou Mingtian. Extending Granular Representation Calculus for Internet Media Resources. In: Li J, et al. Eds. Proceedings of the 3rd International Conference on Active Media Technology (ICAMT'04), Chongqing, China. World-Scientific, 2004. 571~576
- 24 Chen Bo, Zhou Mingtian. ODWIS As a Prototype of Knowledge Service Layer in Semantic Grid. In: Liew K M, et al. Eds. Proceedings of the 5th International Conference on Parallel and Distributed Computing, Applications and Technologies, Singapore. LNCS 3320, Springer-Verlag, 2004. 739~744