基于四元数鲁棒主成分分析的彩色视频背景建模

王巍凤[1] 岳亚辉[1] 冯 越[1] 于洪年[2] 廖 亮[1]
WANG Weifeng YUE Yahui FENG Yue YU Hongnian LIAO Liang

摘 要 随着人工智能技术的快速发展，背景建模技术在智能监控、视频分割、视频编辑等领域的应用日益广泛，近年来已成为学术界和工业界共同关注的研究热点。针对现有背景建模方法存在的不足，文章提出了一种基于四元数鲁棒主成分分析 QRPCA 的背景建模方法，将视频序列表示为四元数矩阵，通过将时空信息和颜色信息结合起来，实现了对复杂场景的有效建模。文章通过实验验证了 QRPCA 方法在处理具有挑战性的视频序列中的有效性和鲁棒性。实验结果表明，QRPCA 背景建模方法在克服了传统方法的局限性的同时，能够提高背景建模的准确性和鲁棒性，为视频监控、智能交通等领域的应用提供了有力支持。

关键词 四元数；鲁棒主成分分析；背景建模；彩色视频；奇异值分解

doi：10.3969/j.issn.1672-9528.2025.04.017

0 引言

背景建模是计算机视觉和图像处理中的一项关键技术，其主要目的是通过对视频序列中的像素进行建模，找到视频中的静态背景，从而使后续的对象检测和跟踪更为可靠。由于视频场景的复杂性、光照变化、干扰噪声等因素的影响，静态背景下的建模依然面临着困难。为应对这些挑战，近年来，诸多研究者提出了各种背景建模方法。

深度学习方法在视频背景建模领域得到广泛应用。例如， Braham 等人 [1] 提出了一种基于场景特定的卷积神经网络架构。Gregorio 等人 [2] 提出了无权重神经网络架构。此外，基于稀疏表示也是一种有效的背景建模方法。Candès 等人 [3] 提出了第一个鲁棒主成分分析算法，矩阵图像分解为低秩矩阵与稀疏矩阵；Cao 等人 [4] 提出了全变分正则化鲁棒主成分分析（TVRPCA）方法，将观察到的视频视为由低秩静态背景、稀疏且平滑的前景以及更稀疏的动态背景之和；Lu 等人 [5] 定义了张量谱范数与张量平均秩，并证明了张量核范数的解是张量低秩成分的有效凸近似；Xu 等人 [6] 提出了一种新颖的在线运动感知鲁棒主成分分析算法（OM-RPCAT）；Gowda 等人 [7] 提出了一种结合奇异值分解和鲁棒主成分分析进行奇异谱分析（SSA-SVDRPCA）的分解方法。

对于彩色视频的处理，以上方法均忽略了红、绿、蓝 3 个通道的相互联系，具有一定局限性。针对此问题，本文提出了基于四元数的鲁棒主成分分析算法（QRPCA）。用四元数矩阵表示彩色视频，并将 RPCA 方法推广到四元数系统。QRPCA 在进行彩色视频的分离时，能够以一种整体的方式同时处理彩色图像的每个像素，充分利用彩色视频序列的 R、G 和 B 三个颜色通道之间的耦合信息。通过对比实验验证了 QRPCA 算法相较于传统算法拥有更高准确性，且在多个数据集上表现出较好的稳定性和鲁棒性。

1 四元数

四元数是一种特殊的超复数，由 1 个实数部分和 3 个虚数部分组成，通常表示为：

$$q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k} \tag{1}$$

式中：a、b、c、d 是实数，而 i、j、k 是虚数单位且满足如下关系：

$$\begin{cases} \mathrm{i}^2=\mathrm{j}^2=\mathrm{k}^2=\mathrm{ijk}=-1 \\ \mathrm{ij}=-\mathrm{ji}=\mathrm{k},\mathrm{jk}=-\mathrm{kj}=\mathrm{i},\mathrm{ki}=-\mathrm{ik}=\mathrm{j} \end{cases} \tag{2}$$

若实部 a=0，则 q 称为纯四元数，四元数的共轭和模分别定义为：

$$q^* = a - b\mathrm{i} - c\mathrm{j} - d\mathrm{k} \tag{3}$$

$$\|q\| = \sqrt{a^2+b^2+c^2+d^2} \tag{4}$$

对于任意一幅彩色图像，它的每个像素可用四元数矩阵表示为：

$$\boldsymbol{Q} = \boldsymbol{Q}_1 + \boldsymbol{Q}_2\mathrm{i} + \boldsymbol{Q}_3\mathrm{j} + \boldsymbol{Q}_4\mathrm{k} \in \mathbb{H}^{m\times n} \tag{5}$$

式中：$\boldsymbol{Q}_1$ 为零矩阵；$\boldsymbol{Q}_2$、$\boldsymbol{Q}_3$ 和 $\boldsymbol{Q}_4$ 分别为每个像素的红、绿和蓝三个通道的值。对于任意四元数矩阵$\boldsymbol{Q} = \boldsymbol{Q}_1 + \boldsymbol{Q}_2\mathrm{i} + \boldsymbol{Q}_3\mathrm{j} + \boldsymbol{Q}_4\mathrm{k}$，可以用等价的实矩阵形式表示：

1. 中原工学院信息与通信工程学院 河南郑州 450007
2. 爱丁堡龙比亚大学 英国爱丁堡 EH141DJ

[基金项目] 中原工学院科研团队发展项目（K2022TD001）；河南省杰出外籍科学家工作室（GZS2022012）

$$G:\boldsymbol{Q}\mapsto \boldsymbol{Q}_{(R)} \doteq \begin{pmatrix} \boldsymbol{Q}_1 & \boldsymbol{Q}_2 & \boldsymbol{Q}_3 & \boldsymbol{Q}_4 \\ -\boldsymbol{Q}_2 & \boldsymbol{Q}_1 & -\boldsymbol{Q}_4 & \boldsymbol{Q}_3 \\ -\boldsymbol{Q}_3 & \boldsymbol{Q}_4 & \boldsymbol{Q}_1 & -\boldsymbol{Q}_2 \\ -\boldsymbol{Q}_4 & -\boldsymbol{Q}_3 & \boldsymbol{Q}_2 & \boldsymbol{Q}_1 \end{pmatrix} \tag{6}$$

2 四元数鲁棒主成分分析

四元数鲁棒主成分分析可以看作是传统鲁棒主成分分析问题在四元数数域的推广，即将一个四元数观测矩阵分解为一个四元数低秩矩阵和一个四元数稀疏矩阵 [8]，其数学模型为：

$$\boldsymbol{D}=\boldsymbol{L}+\boldsymbol{S} \tag{7}$$

式中：$\boldsymbol{D}$ 代表输入的四元数观测矩阵；$\boldsymbol{L}$ 代表四元数低秩矩阵；$\boldsymbol{S}$ 代表四元数稀疏矩阵。针对 $\boldsymbol{L}$ 和 $\boldsymbol{S}$ 定义，有两点需要优化：$\boldsymbol{L}$ 的秩和 $\boldsymbol{S}$ 的稀疏性，即找的秩最低的 $\boldsymbol{L}$ 和稀疏的 $\boldsymbol{S}$。因此，优化问题可以表示为：

$$\min_{L,S} \operatorname{rank}(\boldsymbol{L})+\lambda\|\boldsymbol{S}\|_0 \tag{8}$$

式中：$\|\boldsymbol{S}\|_0$ 表示 ℓ_0 范数，即矩阵中非零元素的个数。标量 $\lambda=1/\sqrt{\max(m,n)}$ 是一个非负实数，作为低秩分量和稀疏分量之间的权衡 [9]。

然而，式（8）是非凸非光滑的，一般难以求解。因此，用核范数代替 $\boldsymbol{L}$ 的秩，ℓ_1 范数代替 ℓ_0 范数，将非凸模型转化为凸优化模型：

$$\begin{aligned}&\min_{L,S} \operatorname{rank}\|\boldsymbol{L}\|_*+\lambda\|\boldsymbol{S}\|_1\\ &\text{s.t. } \boldsymbol{D}=\boldsymbol{L}+\boldsymbol{S}\end{aligned} \tag{9}$$

式中：$\|\boldsymbol{L}\|_*$ 表示核范数，即 $\boldsymbol{L}$ 的奇异值之和；$\|\boldsymbol{S}\|_1$ 表示 ℓ_1 范数，即矩阵所有元素绝对值的和。为提高效率与准确率，本文采用交替方向乘子法（ADMM）进行迭代求解 [10]，其优化的代价函数可以写成：

$$\begin{aligned}\mathcal{L}(\boldsymbol{L},\boldsymbol{S},\boldsymbol{Y},\mu)=&\|\boldsymbol{L}\|_*+\lambda\|\boldsymbol{S}\|_1+\langle \boldsymbol{Y},\boldsymbol{D}-\boldsymbol{L}-\boldsymbol{S}\rangle\\ &+\frac{\mu}{2}\|\boldsymbol{D}-\boldsymbol{L}-\boldsymbol{S}\|_F^2\end{aligned} \tag{10}$$

式中：$\boldsymbol{Y}\in\mathbb{H}^{m\times n}$ 表示拉格朗日乘子；<·> 表示矩阵的内积；标量 μ 表示一个非负实数，用来平衡误差项。然而，为使代价函数最小化，标量 μ 需要设置得非常大，以便在 $\boldsymbol{D}$ 和 $\boldsymbol{L}$+$\boldsymbol{S}$ 之间给出较小的误差。借助每次迭代都会更新的拉格朗日乘子 $\boldsymbol{Y}$，无需刻意将其设置为较大值。通过使用 $\boldsymbol{Y}$，优化问题会更快地收敛到最优因为 $\boldsymbol{Y}$ 在每次迭代中都会得到改进。求解过程为：

（1）固定变量 $\boldsymbol{S}$ 和 $\boldsymbol{Y}$，更新变量 $\boldsymbol{L}$：

$$\begin{aligned}\boldsymbol{L}^{k+1}&=\underset{L}{\operatorname{argmin}}\,\mathcal{L}(\boldsymbol{L},\boldsymbol{S}^k,\boldsymbol{Y}^k,\mu_k)\\ &=\underset{L}{\operatorname{argmin}}\|\boldsymbol{L}\|_*+\frac{\mu_k}{2}\left\|\boldsymbol{L}-\left(\boldsymbol{D}-\boldsymbol{S}^k+\frac{\boldsymbol{Y}^k}{\mu_k}\right)\right\|_F^2\end{aligned} \tag{11}$$

这是一个核范数阈值问题，利用奇异阈值收缩算子对变量 $\boldsymbol{L}$ 进行处理，具体表示为：

$$\boldsymbol{L}^{k+1}=\operatorname{QSVT}_{1/\mu_k}\left(\boldsymbol{D}-\boldsymbol{S}^k+\frac{\boldsymbol{Y}^k}{\mu_k}\right) \tag{12}$$

由于四元数不能直接进行 SVT 分解，本文根据四元数的实表示将四元数矩阵转为实矩阵进行运算，其伪代码具体过程如算法 1 所示。

算法 1 QSVT

输入：四元数矩阵 $\boldsymbol{X}\in\mathbb{H}^{m\times n}$，$\tau>0$

输出：软阈值奇异值后的四元数矩阵 $\boldsymbol{X}_{\text{hat}}\in\mathbb{H}^{m\times n}$

1：四元数矩阵实数化：$\boldsymbol{X}\in\mathbb{H}^{4m\times 4n}\leftarrow G(\boldsymbol{X})$

2：$[\boldsymbol{U},\boldsymbol{S},\boldsymbol{V}]=\text{SVD}(\boldsymbol{X})$

3：$\boldsymbol{W}=\boldsymbol{U}\cdot(\boldsymbol{S}-\tau)\cdot\boldsymbol{V}^*$

4：$\boldsymbol{X}_{\text{hat}}\in\mathbb{H}^{m\times n}\leftarrow G^{-1}(\boldsymbol{W})$

（2）固定变量 $\boldsymbol{L}$ 和 $\boldsymbol{Y}$，更新变量 $\boldsymbol{S}$：

$$\begin{aligned}\boldsymbol{S}^{k+1}&=\underset{S}{\operatorname{argmin}}\,\mathcal{L}(\boldsymbol{L}^{k+1},\boldsymbol{S},\boldsymbol{Y}^k,\mu_k)\\ &=\Theta_{\lambda/\mu_k}\left(\boldsymbol{D}-\boldsymbol{L}^{k+1}+\frac{\boldsymbol{Y}^k}{\mu_k}\right)\end{aligned} \tag{13}$$

利用软阈值收缩算子对变量 $\boldsymbol{S}$ 进行处理，其作用是修正四元数稀疏矩阵 $\boldsymbol{S}$ 中的值，具体表示为：

$$\begin{aligned}\Theta_\varepsilon(x)&=\operatorname{sign}(x)\cdot\max(|x|-\varepsilon,0)\\ &=x\cdot\max\left(0,1-\frac{\varepsilon}{|x|}\right)\end{aligned} \tag{14}$$

式中：$\operatorname{sign}(\cdot)$ 为符号函数，可以有效提取数据中的异常或噪声。

（3）更新拉格朗日乘子 $\boldsymbol{Y}$，则：

$$\boldsymbol{Y}^{k+1}=\boldsymbol{Y}^k+\mu_k\left(\boldsymbol{D}-\boldsymbol{L}^{k+1}-\boldsymbol{S}^{k+1}\right) \tag{15}$$

（4）更新惩罚参数 μ，则：

$$\mu_{k+1}=\min(\rho\mu_k,\mu_{\max}) \tag{16}$$

惩罚参数是用来平衡四元数低秩矩阵 $\boldsymbol{L}$ 和四元数稀疏矩阵 $\boldsymbol{S}$ 之间的重要度。较大的 μ 会促使算法生成更稀疏的 $\boldsymbol{S}$，从而达到更好的分离效果。

四元数鲁棒主成分分析伪代码过程如算法 2 所示。

算法 2 QRPCA

输 入： 四 元 数 矩 阵 $\boldsymbol{D}$，$\varepsilon=1e-8$，$N=1000$，$\rho=1e-3$，$\mu=1e-3$，$\mu_{\max}=1e10$，λ

1：初始化：$\boldsymbol{L}$=$\boldsymbol{S}$=$\boldsymbol{Y}$ 为零矩阵

2：通过式（12）更新 $\boldsymbol{L}^{k+1}$

3：通过式（13）更新 $\boldsymbol{S}^{k+1}$

4：通过式（15）更新 $\boldsymbol{Y}^{k+1}$

5：通过式（16）更新 μ_{k+1}

6：判断是否收敛：

（1）$\text{chg}L\leftarrow\|\boldsymbol{L}^{k+1}-\boldsymbol{L}^k\|_\infty$

（2）$\text{chg}S\leftarrow\|\boldsymbol{S}^{k+1}-\boldsymbol{S}^k\|_\infty$

（3）$\text{chg}Z\leftarrow\|\boldsymbol{D}-\boldsymbol{L}^{k+1}-\boldsymbol{S}^{k+1}\|_\infty$

（4）$\text{chg}\leftarrow\|\text{chg}L\ \text{chg}S\ \text{chg}Z\|_\infty$

（5）if $\text{chg}<\varepsilon$ then

break

3 实验

3.1 数据预处理

一个视频由 T 帧彩色图像组成，每一帧的尺寸为 $M\times N\times 3$，其中“3”表示图像的红、绿和蓝三个通道。创建一个全零矩阵作为四元数的实部，同时，将三个通道的像素值分别作为四元数的 3 个虚部，形成四元数张量，其维度为 $T\times M\times N$。

3.2 评估设置

3.2.1 数据集

（scene background initialization, SBI）数据集是一个常用于背景建模与初始化的数据集，由图像序列组成，通常用于测试和评估背景建模算法的性能。SBI 数据集通常包含具有不同复杂度和场景条件的视频序列，这些序列涵盖了各种环境下的背景变化、光照变化和动态物体移动。

3.2.2 评价指标

峰值信噪比（peak signal-to-noise ratio, PSNR）是一种常用的视频评价指标，用于衡量图像或视频质量的好坏。PSNR 的计算基于信号与噪声之间的比率，用公式表示为：

$$\mathrm{PSNR}=20\cdot\log_{10}\left(\frac{\mathrm{MAX}\cdot\sqrt{M\times N\times T}}{\mathrm{MSE}}\right) \tag{17}$$

式中：MAX 为图像或视频数据的最大可能取值（通常为 255，表示 8 位图像的最大像素值）；$M\times N\times T$ 为数据的维度大小；均方误差（mean squared error, MSE）为原始图像和重建图像之间像素差的平方均值。

3.3 结果

图 1 展示了不同算法在 SBI 数据集的第 25 帧的背景建模效果。从上到下分别为原图、背景图、RPCA、TRPCA 和 QRPCA 算法的背景建模图。

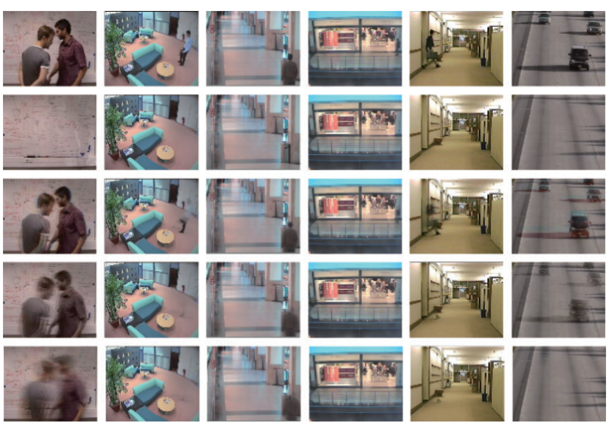

图 1 6 组数据集上不同算法的表现

从图 1 可以看出，QRPCA 相较于 RPCA 与 TRPCA 可以更精确地恢复出视频的背景信息，并更好地抑制噪音的影响。

QRPCA 只考虑视频序列中不随时间变化的部分作为背景模型。例如，在第二行、第四行和第五行的 3 个视频序列中，前景随着时间变化而变化，没有任何静止时刻，因此背景建模效果较好。

表 1 展示了在 6 组数据集上不同算法背景建模的 PSNR 值，可以看到，QRPCA 的效果明显优于其他算法。

表 1 6 组数据集上不同算法的 PSNR

ID	Videos	RPCA	TRPCA	QRPCA
1	Board	13.059 0	13.387 6	13.565 9
2	Candela_m1.10	28.052 1	34.894 2	36.047 6
3	CAVIAR1	25.831 9	27.029 8	28.069 5
4	CAVIAR2	32.777 3	39.234 5	39.984 5
5	HallandMonitor	26.843 0	31.414 2	33.136 8
6	HighwayI	21.789 9	25.951 8	30.548 1

4 结论

本文提出了一种基于四元数的鲁棒主成分分析（QRPCA）方法，利用四元数可以保留彩色视频各通道之间关系的特性，将 RPCA 推广到四元数系统，可以得到更好的背景建模效果。结果表明，与 RPCA、TRPCA 相比，本文方法在彩色视频背景模型重建方面具有优势。

参考文献：

[1] HAM M, DROOGENBROECK M V. Deep background subtraction with scene-specific convolutional neural networks[C/OL]//2016 International Conference on Systems, Signals and Image Processing (IWSSIP). Piscataway: IEEE,2016[2024-06-02].https://ieeexplore.ieee.org/document/7502717. DOI:10.1109/IWSSIP.2016.7502717.

[2] GORIO M D, GIORDANO M. Background estimation by weightless neural networks[J]. Pattern recognition letters, 2017, 96: 55-65.

[3]CANDÈS E J, LI X D, MA Y, et al. Robust principal component analysis?[J]. Journal of the ACM(JACM), 2011, 58(3): 1-37.

[4]CAO X C, YANG L, GUO X J. Total variation regularized RPCA for irregularly moving object detection under dynamic background[J]. IEEE transactions on cybernetics, 2015, 46(4): 1014-1027.

[5]LU C Y, FENG J S, CHEN Y D, et al. Tensor robust principal component analysis with a new tensor nuclear norm[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 42(4): 925-938.

基于 Canny-Devernay 的算法芯片缺陷边缘检测

宋康康[1] 张俊生[2] 仝晓刚[2] 张红宇[1]
SONG Kangkang ZHANG Junsheng TONG Xiaogang ZHANG Hongyu

摘 要 为提高在噪声较多的图像中集成电路封装缺陷图像分析的精度和连续性，文章提出了一种改进的边缘检测方法，即 Canny-Devernay 亚像素边缘检测算法。首先，采用均值滤波对图像进行平滑去噪，并利用直方图均衡化增强图像对比度。其次，计算图像的梯度，确定边缘点，并连接这些边缘点以形成连续的边缘轮廓。最后，通过细化边缘，获得了精确的图像边缘位置信息。与传统 Canny 边缘检测算法相比，Canny-Devernay 亚像素边缘检测算法在准确度、回归率、精确率分别提高了 11.93%、11.06%、15.55%。

关键词 封装缺陷；Canny 边缘检测；Canny-Devernay 亚像素边缘检测；中值滤波；直方图均衡化

doi：10.3969/j.issn.1672-9528.2025.04.018

0 引言

封装芯片的制造过程中，边缘检测是非常关键的一个环节，对芯片的精确放置和焊接质量有直接影响。目前，文献 [1-5] 封装芯片边缘检测技术主要基于图像处理和机器学习算法，但在高精度要求、复杂环境下的检测以及自动化水平的提升等挑战面前，这些技术仍存在一定的局限性。

在边缘检测领域，研究人员提出了多种创新方法，以提高检测的精度和适应性。文献 [6] 采用直方图均衡化与同态滤波的空频结合处理，对图像进行边缘检测，能够准确识别动作，但会出现误判的结果。文献 [7-8] 采用形态滤波和 Otsu 优化、Canny 算子进行边缘检测，梯度差分和 Otsu 等方法缺陷会出现，但是精度不够。文献 [9] 采用利用霍夫变换改进的Canny 算子边缘检测方法，但是会出现运行时间长。文献 [10] 采用小波去噪的算法处理钢表面，检测效果很好，但是精度不够。文献 [11-12] 采用 LOG 算子和 Scharr 算子，检测边缘检测有可能无法有效检测。文献 [13-14] 采用基于多方向各向异性高斯方向导数的 Sobel 算子对 NEQR 量子图像进行边缘提取时，存在边缘区域无法被识别的问题。

本文采用基于 Canny-Devernay 的亚像素边缘检测算法。

1. 太原师范学院计算机科学与技术学院 山西晋中 030619
2. 太原工业学院电子工程系 山西太原 030008
[基金项目] 太原市"双百攻关行动"揭榜挂帅项目（2024TYJB0126）

[6] XU W Y, XIA T, JING C Q. Background modeling from video sequences via online motion-aware RPCA[J]. Computer science and information systems, 2021, 18: 29.

[7]GOWDA V B, GOPALAKRISHNA M T, MEGHA J, et al. Background initialization in video data using singular value decomposition and robust principal component analysis[J]. International journal of computers and applications, 2023, 45(9): 600-609.

[8]TANNER J, VARY S. Compressed sensing of low-rank plus sparse matrices[J]. Applied and computational harmonic analysis, 2023, 64: 254-293.

[9]MOORE B E , GAO C , NADAKUDITI R R .Panoramic robust PCA for foreground-background separation on noisy, free-motion camera Video[J].IEEE transactions on computational imaging,2019,5(2):195-211.

[10]LIN Z C, LI H, FANG C. Alternating direction Method of multipliers for machine learning[J]. Alternating direction method of multipliers for machine learning,2022:1-9.

【作者简介】

王巍凤（1999—），女，河南南阳人，硕士研究生，研究方向：基于四元数 RPCA 的视频背景建模。

（收稿日期：2024-12-11）