

## 基于光响应非均匀性的WhatsApp压缩视频来源识别

陈懿辉 田妮莉 潘晴 苏开清

### WhatsApp compressed video source camera identification based on photo response nonuniformity

CHEN Yihui, TIAN Nili, PAN Qing, SU Kaiqing

引用本文:

陈懿辉, 田妮莉, 潘晴, 等. 基于光响应非均匀性的WhatsApp压缩视频来源识别[J]. 应用光学, 2024, 45(2): 337–345. DOI: 10.5768/JAO202445.0201009

CHEN Yihui, TIAN Nili, PAN Qing, et al. WhatsApp compressed video source camera identification based on photo response nonuniformity[J]. Journal of Applied Optics, 2024, 45(2): 337–345. DOI: 10.5768/JAO202445.0201009

在线阅读 View online: <https://doi.org/10.5768/JAO202445.0201009>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于双树复小波变换的遥感图像去云雾系统设计

Design of cloud and mist removal system from remote sensing images based on dual-tree complex wavelet transform

应用光学. 2018, 39(1): 64–70 <https://doi.org/10.5768/JAO201839.0102004>

#### 基于小波树和哈达玛矩阵的自适应快速三维压缩成像方法

Fast adaptive compressed 3D imaging based on wavelet trees and Hadamard matrix

应用光学. 2018, 39(5): 674–682 <https://doi.org/10.5768/JAO201839.0502003>

#### 红外光谱辐射亮度测量中温度均匀性和源尺寸效应的研究

Research of temperature uniformity and size-of-source effect in infrared spectral radiance measurement

应用光学. 2020, 41(4): 737–742 <https://doi.org/10.5768/JAO202041.0406003>

#### 一种基于恒定统计的红外图像非均匀性校正算法

A non-uniformity correction algorithm of infrared image sequences based on constant-statistics

应用光学. 2017, 38(2): 304–308 <https://doi.org/10.5768/JAO201738.0206002>

#### 基于圆形视场分割的鱼眼相机星图识别方法

Fisheye camera star identification method based on circular FOV segmentation

应用光学. 2018, 39(4): 505–510 <https://doi.org/10.5768/JAO201839.0402004>

#### 天幕立靶探测光幕响应时间一致性测量方法研究

Measurement method of response time consistency for sky screens vertical target

应用光学. 2019, 40(3): 483–488 <https://doi.org/10.5768/JAO201940.0303005>



关注微信公众号, 获得更多资讯信息

文章编号: 1002-2082 (2024) 02-0337-09

# 基于光响应非均匀性的 WhatsApp 压缩视频 来源识别

陈懿辉, 田妮莉, 潘 晴, 苏开清

(广东工业大学 信息工程学院, 广东 广州 510006)

**摘 要:** 光响应非均匀噪声 (photo response nonuniformity, PRNU) 是光学成像传感器成像时引入的一种独特噪声, 可有效识别压缩视频的来源。针对现有算法提取压缩视频的 PRNU 效果并不显著的问题, 论文提出了一种改进 PRNU 提取算法。首先, 去除视频编解码的环路滤波器, 对视频帧使用双密度双树复小波变换进行分解; 然后对高频子带使用基于贝叶斯阈值估计的双变量收缩算法进行估计, 再使用自适应加窗维纳滤波进行二次估计, 得到噪声残差; 最后用基于量化参数值加权的最大似然估计法聚合噪声残差, 再与视频帧估计得到 PRNU。实验结果表明: 该方法在 20 s 时 WhatsApp 视频的识别率为 75%。

**关键词:** 光响应非均匀性; 源相机识别; 压缩视频; 双密度双树复小波变换; 双变量收缩

中图分类号: TP391

文献标志码: A

DOI: 10.5768/JAO202445.0201009

## WhatsApp compressed video source camera identification based on photo response nonuniformity

CHEN Yihui, TIAN Nili, PAN Qing, SU Kaiqing

(School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China)

**Abstract:** Photo response nonuniformity (PRNU) noise is a unique noise introduced to optical imaging sensors during imaging and can be effectively applied to the source camera identification of compressed video. Due to the problem that existing algorithms do not produce significant effect on extracting PRNU of compressed video, an improved algorithm to extract PRNU was proposed. Firstly, the loop filter of video codec was removed, and the video frame was decomposed by double density-dual tree-complex wavelet transform. Then, the high frequency subband was estimated by bivariate shrinkage algorithm based on Bayesian threshold estimation, and the adaptive window Wiener filter was used for secondary estimation. Finally, after the noise residuals were obtained, they were aggregated by the maximum likelihood estimation method based on quantization parameter weighting, and the PRNU was estimated with video frames. Experiments on the VISION dataset show that the accuracy of the proposed PRNU extraction method in WhatsApp compressed video recognition is improved to 75% at 20 s.

**Key words:** photo response nonuniformity; source camera identification; compressed video; double density-dual tree-complex wavelet transform; bivariate shrinkage

## 引言

随着信息技术的发展, 越来越多的人通过网络

分享图像和视频, 一些违法犯罪和造假诽谤的图像或视频不可避免地出现在网络上出现, 这也引发了

收稿日期: 2023-06-05; 修回日期: 2023-07-16

基金项目: 国家自然科学基金 (61901123)

作者简介: 陈懿辉 (1999—), 男, 硕士研究生, 主要从事视频源相机识别研究。E-mail: cheniyihui2026@163.com

通信作者: 田妮莉 (1982—), 女, 博士, 讲师, 主要从事图像视频处理、源相机识别研究。E-mail: tiannili@gdut.edu.cn

一系列网络安全问题。因此,核实网络照片或视频的来源是十分有必要的<sup>[1]</sup>。

众所周知,基于 PRNU 的方法作为一种有效识别手段被广泛研究和应用于相机归属识别。LUKASJ 等人最初开发的数字图像来源识别系统首次使用了 PRNU<sup>[2]</sup>,该系统利用小波去噪滤波器提取图像的噪声残差,然后通过平均同一个相机拍摄的一组图像的噪声残差作为相机的参考 PRNU。文献 [3] 首次引入了 BM3D(block matching and 3D filtering) 滤波器进行相机归属识别,并且取得了很好的效果,但是该算法的搜索和分组过程非常耗时。文献 [4] 提出了一种基于八邻域上下文自适应插值预测器 (predictor based on eight-neighbor context-adaptive interpolation, PCAI8) 提取 PRNU,可以抑制场景和边缘的影响。在文献 [5] 中,作者使用了改进的局部自适应离散余弦变换滤波器和加权平均算法来进一步处理和提取 PRNU。在文献 [6] 中,引入了双树复小波变换 (dual-tree complex wavelet transform, DTCWT) 滤波器提取图像中的 PRNU。文献 [7] 提出一种多阶段渐进式神经网络的 PRNU 提取算法,与传统小波相比,能够提取更多的图像特征。

以往源相机的识别技术主要针对相机拍摄的图像,而对视频识别则没有太多的研究,尤其是压缩视频。随着压缩视频在社交网络平台上的普及,用视频来确定源相机的归属变得更加重要,基于图像 PRNU 的源相机识别方法扩展到视频来源识别是可行的<sup>[8]</sup>。在现有的视频编码标准 (如 H.264、MPEG 系列等标准) 中,一组视频帧的组成包括 I 帧、P 帧和 B 帧,文献 [9] 中表明了 I 帧在视频取证任务中比 P 帧具有更高的可靠性,然而,短视频并不能提供足够数量的 I 帧。此外,一些研究提出了通过减轻 H.264 和 H.265 视频压缩标准的不利影响来提高 PRNU 估计的准确性,如选择性地利用宏块或去除解码后的伪影<sup>[8,10]</sup>。在文献 [11] 中提出了一种同时考虑相机滚动和视频 I 帧的视频取证方法,以提高处理精度和缩短时间。文献 [12] 提出了一种基于方差稳定变换多尺度迭代最小二乘滤波的加权 PRNU 提取模型,以提高多重压缩视频识别的精度。视频稳定功能也会影响视频取证任务,如文献 [13] 提出通过平移和旋转实现稳像。由于 PRNU 匹配需要像素级精度,即使是轻微的几何图像变换也会使相机的噪声残差错位,本文只讨论不具备稳定功能的拍摄设备。

针对压缩程度较轻的视频,上述大多数算法在

提取 PRNU 噪声方面可以取得良好的效果。然而,由于视频被上传到社交网络平台后会被再次压缩,导致提取 PRNU 的难度大幅提高,并且这方面的研究比较匮乏。为了改善这种压缩视频来源识别的问题,本文提出了一种基于贝叶斯阈值估计的双变量收缩结合自适应加窗维纳滤波的多尺度域滤波算法提取噪声残差,并结合去环路滤波操作和基于量化加权的最大似然估计构建了一种 PRNU 提取模型,该模型能有效地从压缩视频提取出更多的 PRNU 噪声。总的来说,本文的贡献主要体现在以下两个方面:

- 1) 引入基于贝叶斯估计的双变量收缩算法到源相机识别领域。通过基于贝叶斯估计的双变量收缩算法与双密度双树复小波的结合,可以很好地去掉视频压缩伪影的干扰,从而提取出质量更好的 PRNU。
- 2) 提出了一个完整的 PRNU 提取模型,可以用于视频来源识别。实验表明,改进的 PRNU 提取模型在面对压缩视频来源识别任务时,效果十分显著。

## 1 改进的 PRNU 提取模型

本文提出的改进的 PRNU 提取模型算法流程如图 1 所示。

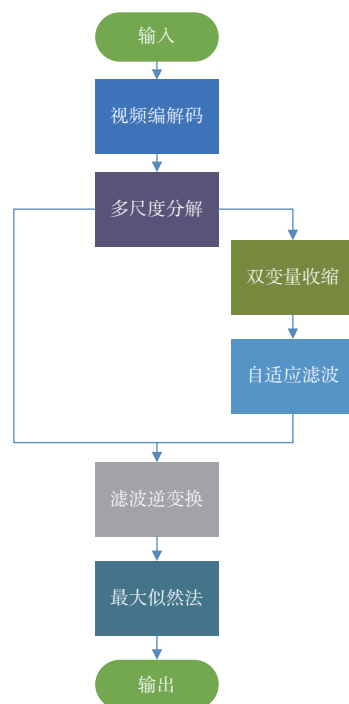


图1 算法流程图

Fig. 1 Flow chart of algorithm

算法总体分为 6 个步骤:

步骤 1) 视频编解码。用去除环路滤波的解码器对视频重新编解码, 得到视频帧<sup>[8]</sup>。

步骤 2) 多尺度分解。用双密度双树复小波变换 (double-density dual-tree complex wavelet transform, DDDTCWT) 对视频帧进行多尺度分解, 把视频帧分解成 8 个高频子带和 1 个低频子带, 并将子带进行归一化处理。

步骤 3) 双变量收缩。使用基于贝叶斯阈值估计的双变量收缩算法对高频子带进行初步估计。

步骤 4) 自适应滤波。用自适应加窗维纳滤波对高频子带做进一步估计。

步骤 5) 滤波逆变换。将归一化的子带进行还原后采用逆双密度双树复小波变换, 得到噪声残差。

步骤 6) 最大似然法。用量化参数 (quantization parameter, QP) 值加权的最大似然估计法估计噪声残差的乘性因子  $K$ , 再使用乘性因子  $K$  与视频帧的均值相乘得到估计的 PRNU 噪声。

在 PRNU 识别阶段, 用带符号的峰值相关能量 (sign peak to correlation energy, SPCE) 对参考 PRNU 和测试 PRNU 进行相关性测量, 得到的相关性值与给定的阈值进行比较, 验证视频是否来自相机<sup>[14]</sup>。

### 1.1 去除环路滤波器的解码器

环路滤波器是视频编解码器中用于减少块效应的一种机制。在视频编码过程中, 逐块量化操作可能导致输出图像出现块效应, 使得相邻宏块的边界出现不连续的现象, 引入环路滤波器来平滑宏块边界, 达到抑制块效应的效果。视频帧的块效应强度主要受压缩级别影响。高质量视频中解码器内的环路滤波对源相机识别影响较小, 但是低质量视频中的解码器环路滤波会导致 PRNU 噪声大部分被去除, 大大增加了识别难度。因此, 使用去除解码器环路滤波对于低质量视频来说可以获得更可靠的 PRNU 估计<sup>[8]</sup>。本文使用去除环路滤波器的解码器, 在解码阶段绕过环路滤波器, 并保持解码器与编码器同步, 同时保留了宏块级别的解码信息, 如大小、位置和量化参数, 以便后续对每帧视频进行可靠的 PRNU 估计。

### 1.2 双密度双树复小波变换与对称边界扩展

PRNU 是一种类似高斯白噪声的噪声, 利用小波分解方法可以有效分离高低频信息, 抑制低频干扰。DDDTTCWT 是一种多尺度、多方向的小波变换方法, 结合了双密度小波变换和 DTCWT 的优

势<sup>[15]</sup>。DDDTTCWT 可以将输入的图像信号在不同尺度上进行分解, 能够提供 16 个尺度的主信号, 并且每个主信号都用实数和虚数 2 个小波表示。在此复转换过程中有 2 个滤波器组, 即主滤波器组和对偶滤波器组, 它们之间的相位差异和滤波器选择, 使得 DDDTCWT 能够提供更好的方向性表示和频率局部化<sup>[16]</sup>。因此, DDDTCWT 能更准确地描述信号的细节和结构, 具有比双密度小波变换和 DTCWT 更好的视频帧分解能力和重构精度能力。

由于一般的图像滤波处理不能很好地处理图像块的边界, 从而导致出现边界效应, 本文使用了一种对称边界扩展的方法<sup>[6]</sup>。为了避免图像边界引入明显的不连续点影响去噪效果, 我们对噪声图像的行列分别进行垂直翻转和水平翻转。翻转的行列数根据小波分解层数  $L$  设置为  $10-2L$ , 这是由于分解层数越深, 小波子带的分辨率越小, 需要扩展的边界越少。考虑到基于贝叶斯阈值估计的双变量收缩算法需要考虑小波系数和其父系数之间的相关性, 因此, 其父矩阵也需要扩展至同样的大小。本文首先对输入的视频帧进行对称边界扩展, 再使用 DDDTCWT 对视频帧进行多尺度分解和滤波处理。

### 1.3 基于贝叶斯阈值估计的双变量收缩和自适应加窗维纳滤波

在提取 PRNU 过程中, 影响最大的是过滤阶段, 本文算法主要改进过滤阶段的处理。文献 [17] 证明了利用小波变换的图像去噪算法可以通过考虑小波系数之间的统计相关性显著改善其性能。双变量收缩算法<sup>[18]</sup>是多尺度变换域滤波的一种有效去噪算法, 该算法可以提取更丰富的 PRNU 噪声。

双变量收缩算法通常设定一个阈值进行去噪, 常用的阈值方法有硬阈值和软阈值, 但是软硬阈值的设定都不能很好地适配双变量收缩算法。贝叶斯估计通过引入后验分布提供了参数的完整概率分布, 能够对不同的小波子带进行参数估计, 还可以将先前的经验与观测数据相结合, 从而提高参数估计的准确性和鲁棒性。DDDTTCWT 通过将时域的含噪图像信号转换到小波域进行分解, 得到多尺度的小波系数, 用下式表示为

$$y = w + n \quad (1)$$

式中:  $y$  表示含噪信号的小波系数;  $w$  表示无噪声图像对应的小波系数;  $n$  表示噪声的小波系数。为了将不同子带的小波系数做相同滤波操作, 需要将不同小波系数进行归一化操作。



随后对小波子带进行滤波估计,即从 $y$ 中估计出 $w$ 。基于贝叶斯阈值估计的双变量收缩算法需要考虑小波系数与其父系数之间的统计相关性。假设用 $w_{1k}$ 表示第 $k$ 层的小波系数, $w_{2k}$ 表示下一层更大尺度中与 $w_{1k}$ 位置相对应的父节点小波系数,则向量矩阵可以表示为

$$y_k = w_k + n_k \quad (k = 1, \dots, m) \quad (2)$$

式中: $y_k=(y_{1k}, y_{2k})$ ;  $w_k=(w_{1k}, w_{2k})$ ;  $n_k=(n_{1k}, n_{2k})$ ;  $m$ 为子小波系数的数目。通过式(2)考虑系数与其父系数之间的统计相关性。未受噪声污染的图像信号的小波系数 $\hat{w}(y_k)$ 可以使用最大后验概率估计得到,可表示为

$$\hat{w}(y_k) = \arg \max_{w_k} p_{w_k|y_k}(w_k|y_k) \quad (3)$$

由贝叶斯概率估计作进一步推导之后,可将式(3)转换为式(4):

$$\hat{w}(y_k) = \arg \max_{w_k} [p_n(y_k - w_k) \cdot p_w(w_k)] \quad (4)$$

$$p_n(n_k) = p_n(y_k - w_k) \quad (5)$$

由式(4)和式(5)可知,只有求得噪声小波系数的概率密度函数 $p_n(n_k)$ 和输出图像信号小波系数的概率密度函数 $p_w(w_k)$ ,才能得到真实信号的系数估计 $\hat{w}(y_k)$ 。贝叶斯规则允许用噪声概率密度和小波系数的先验密度来估计。假设各层噪声信号 $n_{1k}, n_{2k}$ 为独立同分布的高斯白噪声,文献[18]中提出了 $p_n(n_k)$ 和 $p_w(w_k)$ 的模型,用下式表示:

$$p_n(n_k) = \frac{1}{2\pi\sigma_n^2} \exp\left(-\frac{n_{1k}^2 + n_{2k}^2}{2\sigma_n^2}\right) \quad (6)$$

$$p_w(w_k) = \frac{1}{2\pi\sigma_k^2} \exp\left(-\frac{\sqrt{3}}{\sigma_k}\right) \sqrt{w_{1k}^2 + w_{2k}^2} \quad (7)$$

得到 $p_n(n)$ 和 $p_w(w_k)$ 后,可使用双变量阈值收缩函数<sup>[20]</sup>来估计 $w_{1k}$ ,如式(9)所示。其中,基于贝叶斯估计的阈值 $T$ 和一个二元收缩函数 $g$ 可以分别用式(8)和式(10)表示:

$$T = \frac{\sqrt{3}\sigma_n^2}{\sigma_k} \quad (8)$$

$$w_{1k} = \frac{(\sqrt{y_{1k}^2 + y_{2k}^2} - T) +}{\sqrt{y_{1k}^2 + y_{2k}^2}} y_{1k} \quad (9)$$

$$g = (\sqrt{y_{1k}^2 + y_{2k}^2} - T) \quad (10)$$

式中  $g = \begin{cases} 0, & g < 0 \\ g, & g \geq 0 \end{cases}$

由式(7)和式(8)可知,要得到最后的收缩系数 $w_{1k}$ ,还需要得到噪声小波系数的方差 $\sigma_n^2$ 和第 $k$ 层分解系数的方差 $\sigma_k^2$ 。公式(11)可由上述模型观测

得到,所以方差 $\sigma_n^2$ 和 $\sigma_k^2$ 的估计值可由公式(12)~公式(14)求得:

$$\sigma_y^2 = \sigma_n^2 + \sigma_k^2 \quad (11)$$

$$\hat{\sigma}_n^2 = \frac{\text{median}(|y_i|)}{0.6745} \quad (12)$$

$$\hat{\sigma}_y^2 = \frac{1}{M} \sum_{y_{1k} \in N(k)} y_{1k}^2 \quad (13)$$

$$\hat{\sigma}_k^2 = \sqrt{(\hat{\sigma}_y^2 - \hat{\sigma}_n^2)} \quad (14)$$

式中: median()表示中值估计器; $y_i$ 表示 $L_0H_1$ 子带的小波系数; $M$ 表示第 $k$ 层小波系数局部邻域 $N(k)$ 的区域大小; $y_{1k}$ 代表小波变换后的子信号的小波系数。为了从噪声小波系数 $\sigma_n^2$ 估计出噪声方差,使用中值滤波器从最小尺度的小波系数中去估计,如式(12)所示。将得到的方差 $\sigma_n^2$ 与 $\sigma_k^2$ 代入式(7)和式(8)得到估计的 $\hat{w}_{1k}$ 。使用自适应加窗维纳滤波函数进一步估计小波系数,可表示为

$$\hat{w}_{2h}(i, j) = \hat{w}_{1h}(i, j) \frac{\sigma_s^2}{\sigma_s^2 + \sigma_0^2}, \quad (i, j) \in h \quad (15)$$

式中: $\hat{w}_{1h}(i, j)$ 表示估计前的小波系数; $\hat{w}_{2h}(i, j)$ 表示估计后的小波系数。一般设置 $\sigma_0^2$ 为4,  $\sigma_s^2$ 可由下式得到:

$$\hat{\sigma}_s^2(i, j) = \min(\max(0, \frac{1}{w_d^2} \sum_{(i, j) \in h} w_{1h}^2(i, j) - \sigma_0^2)) \quad (16)$$

式中窗口的尺寸 $w_d=3, 5, 7, 9$ ,将估计后的小波系数进行归一化还原,然后进行逆变换得到噪声残差。

#### 1.4 量化参数值加权的最大似然估计

由于噪声残差不仅包含PRNU的乘性因子 $K$ ,还包含了其他过程产生的伪影,因此必须抑制估计的PRNU因子中不必要的伪影。然而压缩视频的每一帧甚至每一个宏块的压缩程度都是不一样的,压缩越严重,宏块包含的伪影越复杂,对PRNU估计的负面影响越大。文献[8]根据PCE值和对应的QP值关系得到了PCE与QP的关系曲线,然后归一化权重矩阵 $U_t$ 。PRNU加权公式如下:

$$K = \frac{\sum_{t=1}^N I_t W_t U_t}{\sum_{t=1}^N (I_t)^2 U_t + \theta} \quad (17)$$

式中: $N$ 表示每个视频中用于估计 $K$ 的视频帧数; $I_t$ 表示视频中第 $t$ 帧图像; $W_t$ 表示视频中第 $t$ 帧图像的噪声残差; $U_t$ 表示每个视频帧的加权系数矩阵; $\theta$ 为一个很小的值,防止出现分母为0的估计错误。

针对压缩程度较小的块, 可以赋予较高的权重系数, 对于压缩程度较大的块, 则赋予较低的权重系数, 不用完全去除。这种方法基本上为每个宏块分配一个对应值的权重, 而不是二进制权重, 因此, 它能够产生更可靠的 PRNU 估计值。

最后相机参考 PRNU 等于  $C$  个视频帧的 PRNU 的平均值, 用公式 (18) 表示为

$$\text{PRNU} = \frac{\sum_{t=1}^C K I_t}{C} \quad (18)$$

### 1.5 相似性检测

为了计算相机的 PRNU 和测试视频 PRNU 的相关性, 本文使用带符号的 PCE(sign peak correlation energy, SPCE) 来统计<sup>[4]</sup>, 并且 SPCE 值越大, 由参考相机拍摄测试图像的可能性就越大。

$$\text{SPCE}(R, Q) = \frac{\text{sign}(C_{RQ}(a, b)) \cdot C_{RQ}(a, b)}{\sqrt{\frac{1}{|MN| - |\beta|} \sum_{(a, b) \in \beta} C_{RQ}^2(a, b)}} \quad (19)$$

式中:  $C_{RQ}(a, b)$  为参考 PRNU 噪声  $R$  和查询 PRNU 噪声  $Q$  之间的二维循环互相关系数;  $\beta$  为  $(a, b)$  周围的一小块区域;  $|\beta|$  为区域尺寸;  $MN$  为  $C_{RQ}$  行列数量的乘积;  $\text{sign}(\cdot)$  为符号函数。

## 2 实验结果与分析

### 2.1 实验设置

为了评估算法识别压缩视频的性能, 这里使用 VISION 公共数据集<sup>[19]</sup>进行测试比较, 数据集中涵盖了本地视频、YouTube 视频和 WhatsApp 视频三种视频类型。其中, YouTube 压缩视频和 WhatsApp 压缩视频是通过网络共享将本地视频重新编码而成的压缩视频。就视频的压缩程度而言, 本地视频具有  $1\,920 \times 1\,080$  pixels 或  $1\,280 \times 720$  pixels 的分辨率, 以及  $9.7 \text{ Mbit/s} \sim 19.7 \text{ Mbit/s}$  的压缩比特率。相比之下, YouTube 视频的压缩程度约为本地视频的 6 倍左右, 压缩比特率降至  $1.7 \text{ Mbit/s} \sim 3.3 \text{ Mbit/s}$ , 同时保持了视频帧的原始分辨率。WhatsApp 视频经过进一步压缩, 压缩比特率为  $1.06 \text{ Mbit/s} \sim 1.29 \text{ Mbit/s}$ , 并将所有视频的分辨率降低至  $848 \times 480$  pixels。我们选择了本地视频分辨率为  $1\,280 \times 720$  pixels 且不具备防抖功能的相机, 共 9 台。实验相机选取结果如表 1 所示, 每台相机选择了 15 个 WhatsApp 视频。其中, 选择 3 个 Flat 视频(静止、移动和旋转运动类型各 1 个)制作参考 PRNU, 其余 12 个 Indoor 和 Outdoor 类型视频用来制作测试 PRNU, 并假设相机识别成功的 PCE 阈

值为 60。通过这些测试, 可以确定用于获取参考 PRNU 噪声的视频类型对识别性能的影响。

表 1 VISION 数据集的相机设备

Table 1 Smartphones models from VISION dataset

编号	相机型号
D01	Galaxy S III Mini GT-I8190N
D07	Lenovo P70-A
D08	Galaxy Tab 3 GT-P5210
D09	iPhone 4
D13	iPad 2
D22	Galaxy Trend Plus GT-S7580
D26	Galaxy S III Mini GT-I8190
D33	Ascend G6-U10
D35	Galaxy Tab A SM-T555

为了验证本文提出的过滤器的性能优于其他过滤器, 将本文方法和 Basic<sup>[2]</sup>、BM3D<sup>[3]</sup>、PCAI8<sup>[4]</sup>、DTCWT<sup>[6]</sup>、WeightMLE<sup>[8]</sup> 方法进行比较, 使用不同的视频长度函数确定了每种 PRNU 估计方法的成功识别率。

### 2.2 评估指标

为了证明本文提出的源相机识别方案的性能, 采用了 3 种评价指标, 即源相机总体识别率<sup>[10]</sup>、总体接收器工作特征(receiver operating characteristic curve, ROC)曲线<sup>[4, 6]</sup>和由混淆矩阵计算出的 Kappa 统计量<sup>[4]</sup>。

为了得到总体的 ROC 曲线, 需要在给定的检测阈值下记录每个相机的真阳性个数和假阳性个数, 然后将这些数字相加求平均。真阳性率(true positive rate, TPR)和假阳性率(false positive rate, FPR)可以简单表示为

$$R_{\text{FP}} = \frac{\sum_{q=1}^H F_q}{(H-1)C} \quad (20)$$

$$R_{\text{TP}} = \frac{\sum_{q=1}^H T_q}{C} \quad (21)$$

式中:  $H$  表示相机的数量;  $C$  表示测试视频的总数;  $T_q$  和  $F_q$  分别为第  $q$  个相机的真阳性和假阳性。

混淆矩阵用  $n$  行  $n$  列的矩阵形式表示, 可以把分类结果的精度显示在一个矩阵中。混淆矩阵的每个元素表示将视频归属于相应相机的数量, 主对角线上的值之和表示正确识别视频的数量。每个混淆矩阵都可以简化为一个表示分类器性能的度量值, 即 Kappa 统计量。Kappa 统计量的值为  $-1 \sim 1$ , 通常 Kappa 统计量的值落在  $0 \sim 1$  之间, 其值越

接近1,性能越好。

### 2.3 性能比较

源相机识别率通常可以直接用来衡量 PRNU 提取算法的识别效果,当参考 PRNU 从不同的视频类型中获取时,相机识别率随剪辑视频长度(1 s~20 s)的变化而变化,如图 2 所示,同时在第 20 s 时源相机识别率如表 2 所示。剪辑视频长度在 2 s 以前,所有算法几乎没有效果。随着视频的增长,所有算法的性能都得到提升,本文方法的识别率相较于其他方法提升更快。从表 2 可以看出,大多数源相机识别准确率集中在 15%~40%,而本文方法基于 Still、Move、Parrot 类型的视频在第 20 s 时的识别率分别达到了 63.9%、75%、75%,识别性能相比于其他 SCI 方法也有较大提升,比识别性能第二的算法分别提升了 20.4%, 20.4%, 27.8%。同时值得注意的是,用 Flat-move 类型视频制作参考 PRNU 时,比 Flat-still, Flat-parrot 类型提取 PRNU 提供了更好的源相机识别率。

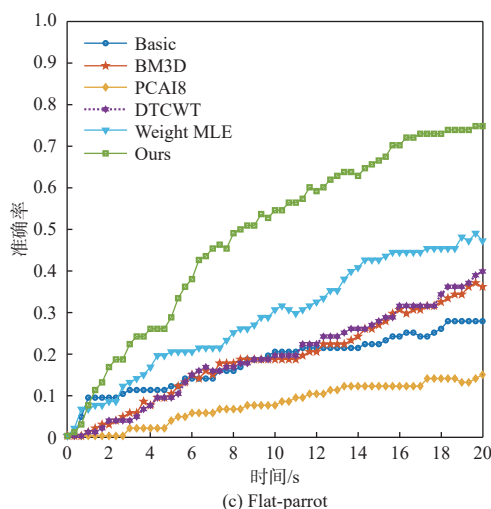
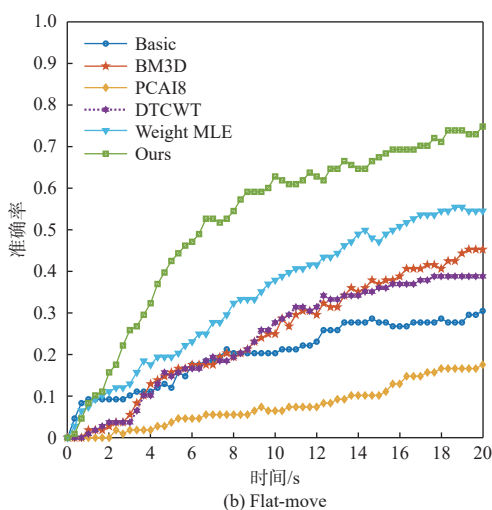
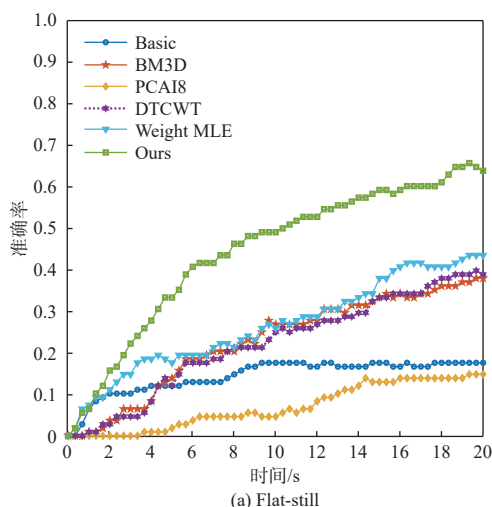


图 2 视频时长为 1 s~20 s 时不同参考 PRNU 下的识别率  
Fig. 2 Recognition rate of different reference PRNU when video duration is 1 s to 20 s

表 2 在 20 s 时 WhatsApp 视频的识别率  
Table 2 Recognition rate of WhatsApp videos at 20 s %

类型	PRNU提取算法					本文
	Basic	BM3D	PCA18	DTCWT	Weight MLE	
静止	17.6	37.9	14.8	38.9	43.5	63.9
移动	30.6	45.4	17.6	38.9	54.6	75.0
旋转	27.8	36.1	14.8	39.8	47.2	75.0

为了评估本文方法在提升识别性能的同时,是否会导致更高的错误识别率,用 Vision 数据集中 WhatsApp 视频进行新的实验,所有 PRNU 都是利用视频的前 20 s 估计。

仅使用视频前 20 s 情况下不同类型视频作为参考 PRNU 时的 ROC 曲线对比图如图 3 所示。图 3(a) 为用 Flat-still 类型视频制作参考 PRNU 时的 ROC 曲线;图 3(b) 为用 Flat-move 类型视频制作参考 PRNU 时的 ROC 曲线;图 3(c) 为用 Flat-parrot 类型视频制作参考 PRNU 时的 ROC 曲线。从图 3 可以看出,本文方法的曲线一直位于其他 ROC 曲线之上,表明本文算法的性能优于其他算法。在源设备识别实际应用中,最小化假阳性概率是十分重要的。表 3 给出了在 WhatsApp 视频识别中当  $R_{FP}=0.01$  时不同算法的  $R_{TP}$  值。从表 3 可以看出,本文方法在 3 种视频的识别中都可以显著地在较低的  $R_{FP}$  时增加  $R_{TP}$ ,这与图 3 中的观察结果相似。

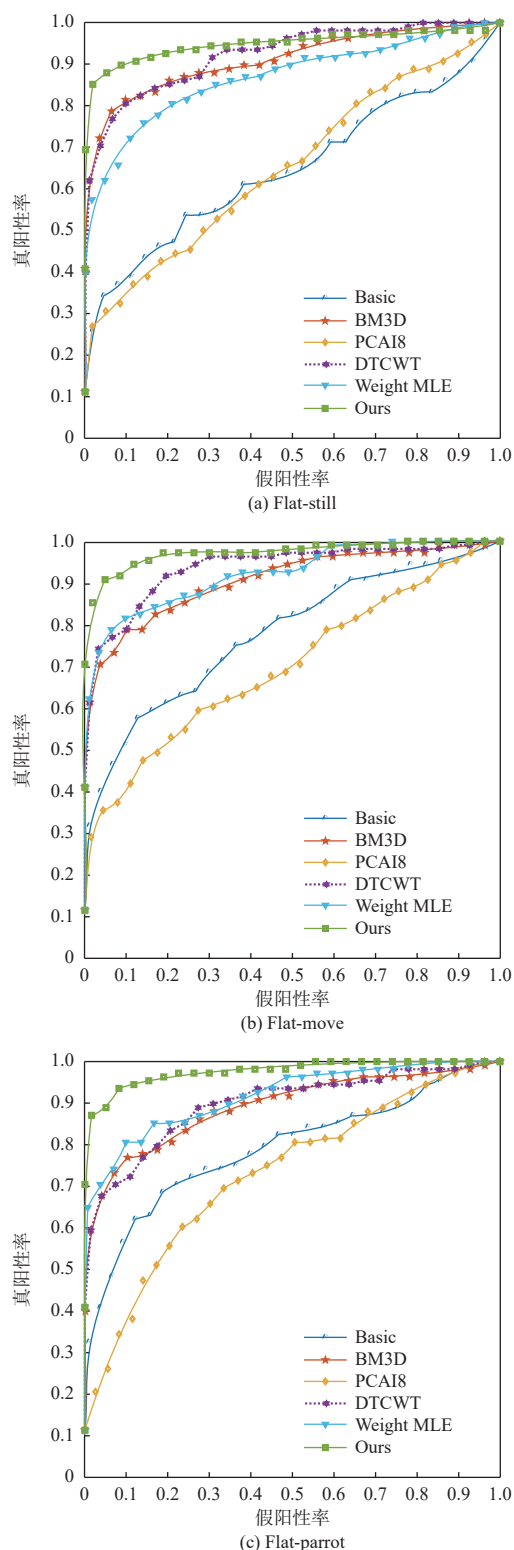


图3 不同参考 PRNU 下的 ROC 曲线

Fig. 3 ROC curves of different reference PRNU

为了进行更加全面比较,表4中列出了不同算法基于 SPCE 值的 Kappa 统计量数据。在 WhatsApp 视频识别中 Kappa 系数分别为 0.712, 0.847, 0.920, 相对于第二性能算法的 Kappa 系数分别提

升了 0.208, 0.170, 0.118。

表3 当  $R_{FP}=0.01$  时各算法的 TPR 值Table 3 TPR value of each algorithm when  $R_{FP}$  is 0.01

类型	PRNU提取算法					本文
	Basic	BM3D	PCAI8	DTCWT	Weight MLE	
静止	0.176	0.491	0.222	0.389	0.361	0.685
移动	0.194	0.435	0.232	0.454	0.509	0.704
旋转	0.250	0.399	0.176	0.500	0.509	0.787

表4 不同算法的 Kappa 统计量

Table 4 Kappa statistic of different algorithms

时间/s	PRNU提取算法					本文
	Basic	BM3D	PCAI8	DTCWT	Weight MLE	
5	0.233	0.469	0.174	0.500	0.504	0.712
10	0.323	0.677	0.243	0.677	0.649	0.847
20	0.431	0.792	0.354	0.802	0.754	0.920

最后,表5列出了不同 PRNU 提取算法在滤波阶段的运行时间。测试图像是一张分辨率为  $1280 \times 720$  pixels 的彩色图像,我们分别用不同算法对测试图像滤波 500 次,并计算了每个方法的平均运行时间。这些模拟是用 Matlab2020a 在一台拥有 3.7 GHz AMD 5900 CPU 和 64G 内存的台式电脑上进行仿真的。相比 Basic, MLE, CAGIF 等速度较快的滤波器,本文方法的计算效率低,其计算效率接近 PCAI8 滤波器,计算成本远低于 BM3D 滤波器。

表5 不同算法滤波阶段的运行时间

Table 5 Running time of filtering phase of different algorithms

PRNU 提取算法	Basic	BM3D	PCAI8	DTCWT	Weight MLE	本文
时间/s	0.35	6.63	2.92	1.13	0.35	3.36

## 2.4 消融实验

为了进一步确定基于贝叶斯估计的双变量收缩算法和自适应加窗维纳滤波在 PRNU 提取模型中的作用,我们在 VISION 数据中进行了消融实验,实验设置见上述 2.1 节。其中,仅使用视频的前 20 s 进行估计。消融实验具体方法是:方法一,在改进的 PRNU 提取模型中移除了基于贝叶斯估计的双变量收缩算法;方法二,在改进的 PRNU 提取模型中移除了自适应加窗维纳滤波。消融实验的测试结果如表6所示。从表6可以看出,在公共数据集上性能的提升验证了本文算法的有效性。



表 6 不同模块组合的消融实验测试结果  
Table 6 Ablation experimental test results of different module combinations %

类型	PRNU提取算法		
	方法一	方法二	本文
静止	15.7	11.1	63.9
移动	25.9	21.3	75.0
旋转	16.7	14.8	75.0

### 3 结论

本文提出了一种基于贝叶斯阈值估计的双变量收缩结合自适应加窗维纳滤波的多尺度域滤波算法提取噪声残差,并结合去环路滤波操作和基于量化加权的最大似然估计构建了一种 PRNU 提取模型。相比于其他源相机识别算法,该算法先去除编解码器的环路滤波器,再用双密度双树复小波变换对视频帧进行多尺度分解并进行归一化;然后先用基于自适应贝叶斯阈值估计的双变量收缩对高频子带进行初步估计;再用自适应加窗维纳滤波对高频子带做进一步估计,将归一化的子带进行还原并逆变换得到噪声残差;最后用量化参数加权的最大似然估计噪声残差中的乘性因子,将乘性因子与多张视频帧的均值相乘得到 PRNU。在 VISION 数据集上的实验结果表明,该算法的效果在识别 WhatsApp 视频方面明显优于现有的算法。

#### 参考文献:

- [1] 张明旺,肖延辉,田华伟,等. 图像中的设备指纹提取技术研究综述[J]. 激光与光电子学进展, 2020, 57(22): 39-48.  
ZHANG Mingwang, XIAO Yanhui, TIAN Huawei, et al. Review of device fingerprint extraction techniques in image[J]. Laser & Optoelectronics Progress, 2020, 57(22): 39-48.
- [2] LUKAS J, FRIDRICH J, GOLJAN M. Digital camera identification from sensor pattern noise[J]. IEEE Transactions on Information Forensics and Security, 2006, 1(2): 205-214.
- [3] CORTIANA A, CONOTTER V, BOATO G, et al. Performance comparison of denoising filters for source camera identification[J]. SPIE:Media Watermarking, Security, and Forensics III, 2011, 7880: 60-65.
- [4] KANG X, CHEN J, LIN K, et al. A context-adaptive SPN predictor for trustworthy source camera identification[J]. EURASIP Journal on Image and Video Processing, 2014, 2014(1): 1-11.
- [5] LAWGALY A, KHELIFI F. Sensor pattern noise estimation based on improved locally adaptive DCT filtering and weighted averaging for source camera identification and verification[J]. IEEE Transactions on Information Forensics and Security, 2016, 12(2): 392-404.
- [6] ZENG H, WAN Y, DENG K, et al. Source camera identification with dual-tree complex wavelet transform[J]. IEEE Access, 2020(8): 18874-18883.
- [7] 高昌锋,肖延辉,田华伟. 基于多阶段渐进式神经网络的图像相机指纹提取算法[J]. 信息网络安全, 2022, 22(10): 15-23.  
GAO Changfeng, XIAO Yanhui, TIAN Huawei. Image camera fingerprint extraction algorithm based on MPR-Net[J]. Netinfo Security, 2022, 22(10): 15-23.
- [8] ALTINISIK E, TASDEMIR K, SENCAR H T. Mitigation of H. 264 and H. 265 video compression for reliable PRNU estimation[J]. IEEE Transactions on Information Forensics and Security, 2019, 15: 1557-1571.
- [9] CHUANG W H, SU H, WU M. Exploring compression effects for improved source camera identification using strongly compressed video[C]//2011 18th IEEE international conference on image processing. NewYork: IEEE, 2011: 1953-1956.
- [10] ALTINISIK E, TASDEMIR K, SENCAR H T. Extracting PRNU noise from H. 264 coded videos[C]//2018 26th European signal processing conference (EUSIPCO). NewYork: IEEE, 2018: 1367-1371.
- [11] YANG W C, JIANG J, CHEN C H. A fast source camera identification and verification method based on PRNU analysis for use in video forensic investigations[J]. Multimedia Tools and Applications, 2021, 80: 6617-6638.
- [12] SU K, TIAN N, PAN Q. Multimedia source identification using an improved weight photo response non-uniformity noise extraction model in short compressed videos[J]. Forensic Science International:Digital Investigation, 2022, 42: 301473.
- [13] 唐武盛,隋京高,张小亚,等. 基于光学相关器的光学稳像器研究[J]. 应用光学, 2014, 35(6): 943-948.  
TANG Wusheng, SUI Jinggao, ZHANG Xiaoya, et al. Research on optical image stabilizer based on optical correlator[J]. Journal of Applied Optics, 2014, 35(6): 943-

- 948.
- [ 14 ] ZENG H, KANG X. Fast source camera identification using content adaptive guided image filter[J]. [Journal of Forensic Sciences](#), 2016, 61(2): 520-526.
- [ 15 ] BARADARANI A J, WU Q M, AHMADI M. An efficient illumination invariant face recognition framework via illumination enhancement and DD-DTCWT filtering[J]. [Pattern Recognition](#), 2013, 46(1): 57-72.
- [ 16 ] SELESNICK I W, BARANIUK R G, KINGSBURY N C. The dual-tree complex wavelet transform[J]. [IEEE Signal Processing Magazine](#), 2005, 22(6): 123-151.
- [ 17 ] FAN G, XIA X G. Image denoising using a local contextual hidden Markov model in the wavelet domain[J]. [IEEE Signal Processing Letters](#), 2001, 8(5): 125-128.
- [ 18 ] SENDUR L, SELESNICK I W. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency[J]. [IEEE Transactions on Signal Processing](#), 2002, 50(11): 2744-2756.
- [ 19 ] SHULLANI D, FONTANI M, IULIANI M, et al. Vision: a video and image dataset for source identification[J]. [EURASIP Journal on Information Security](#), 2017, 2017(1): 1-16.