# Can you really trust the sensor's PRNU? How image content might impact the finger vein sensor identification performance

Dominik Söllinger • Luca Debiasi • Andreas Uhl
University of Salzburg, Department of Computer Sciences
Jakob Haringer Str. 2, 5020 Salzburg, Austria

## Abstract

We study the impact of highly correlated image content on the estimated photo response non-uniformity (PRNU) of a sensor unit and its impact on the sensor identification performance. Based on eight publicly available finger vein datasets, we show formally and experimentally that the nature of finger vein imagery can cause the estimated PRNU to be biased by image content and lead to a fairly bad PRNU estimate. Such bias can cause a false increase in sensor identification performance depending on the dataset composition. Our results indicate that independent of the biometric modality, examining the quality of the estimated PRNU is essential before the sensor identification performance can be claimed to be good.

## PRNU Estimation

For a set of $N$ grayscale images $\{I_1, ..., I_N\} \in \mathbb{R}^{w \times h}$ captured with the same sensor, the residual noise matrix $W_k \in \mathbb{R}^{w \times h}$ for each image $I_k$ where $k \in \{1, ..., N\}$ is extracted. $I_k^{(0)} \in \mathbb{R}^{w \times h}$ denotes the "true scene" image that would be captured in the absence of imperfections/distortions such as dark current, shot noise, etc. Since $I_k^{(0)}$ is unknown in practice it can only be approximated by $F(I_k)$ where $F$ is some denoising function.

$$\begin{aligned} W_k &= I_k - I_k^0 \\ &\approx I_k - F(I_k) \end{aligned} \tag{1}$$

The sensor PRNU $K$ is estimated using MLE (Maximum Likelihood Estimation) as shown in Eqn. 2. Note that all matrix operations are understood element-wise.

$$\hat{K} = \frac{\sum_{k=1}^{N} W_k I_k}{\sum_{k=1}^{N} {I_k}^2} \tag{2}$$

## Characteristics of finger vein imagery

To obtain a good PRNU estimate, images should exhibit high luminance, a smooth image content and they should not be correlated. Looking at the sample images in Fig. 1 unveils that these assumptions only hold in certain image regions. On the contrary, images often exhibit what we call *characteristic structures*. These are regions with edge-like structures, such as the finger positioning apparatus or a prominent illumination pattern, which occur in every image of a dataset.

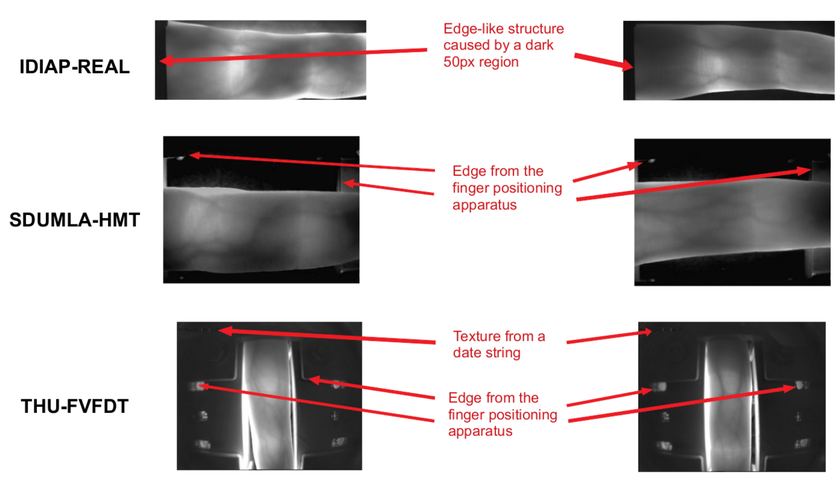


**Figure 1:** Sample images of three different finger vein datasets. The arrows mark sample regions which we expect to exhibit *characteristic structures*.

## Types of image regions and their properties

To study the impact of regions exhibiting different characteristics on the sensor identification performance, we manually defined a set of equally sized image regions ($50 \times 50$ pixels) and categorized them into one out of the following five region types:

- **Fingervein (FV)** - Region mostly contains finger vein tissue
- **Background (BG)** - Mostly contains strongly varying non-finger vein content.
- **Static (ST)** - Contains characteristic structures that occur at the same pixel position
- **Semi Static (SS)** - Contains a characteristic structure but its location varies slightly
- **All Black (AB)** - No image content. Pixels are uniformly black in almost every image.

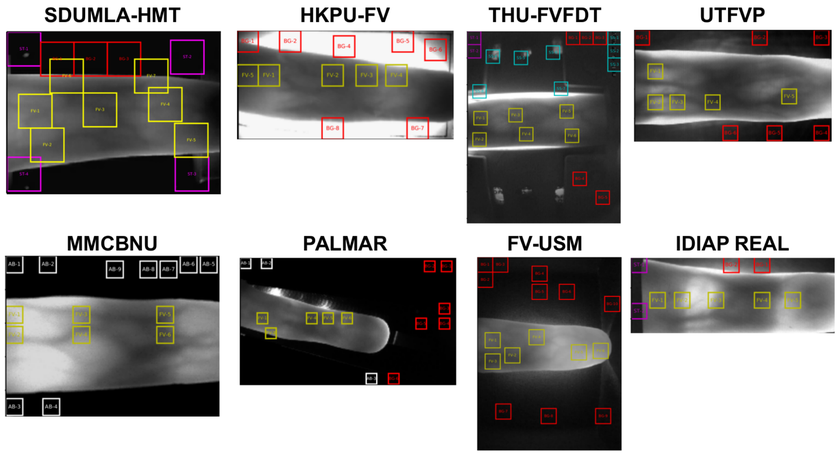


**Figure 2:** Visualization of the different image regions. Color Coding: Yellow - Fingervein (FV). Red - Background (BG). Magenta - Static (ST). Cyan - Semi Static (SS). White - All black (AB)

## The difficulty of PRNU extraction for finger vein imagery: Formal model

To understand the impact of *characteristic structures* on the PRNU estimate, we have to take a look at the underlying model of the residual noise $W_k$ used to derive the PRNU estimator shown in Eqn. 2. The model where $\Xi_k \in \mathbb{R}^{n \times m}$ represents the sum of all noise/error components of the image $I_k$ (such as the dark current, offset or content leftover in $I_k - F(I_k)$) is shown in Eqn. 3.

$$\begin{aligned} W_k &= I_k - F(I_k) \\ &= I_k K + \Xi_k \end{aligned} \tag{3}$$

To derive the ML estimator it is assumed that $\Xi_k$ can be modelled by White Gaussian Noise (WGN) since the content leftover in $\Xi_k$ is small compared to the signal $I_k K$ and $\Xi_k$ is independent of $I_k K$. However this assumption might not hold for areas with characteristic structures in finger vein images. As a result, it is unreasonable to model $\Xi_k$ by WGN. The situation gets even worse if we take into account that our characteristic structures occur in every image at the same position. To highlight the problem we denote the content leftover in $I_k - F(I_k)$ as $\Psi_k$ and exclude it from the term $\Xi_k$. The resulting model shown in Eqn. 4.

$$W_k = I_k K + \Xi_k + \Psi_k \tag{4}$$

Eqn. 4 can then be rewritten to Eqn. 5 by element-wise division through $I_k$.

$$\frac{W_k}{I_k} = \left(K + \frac{\Psi_k}{I_k}\right) + \frac{\Xi_k}{I_k} \tag{5}$$

As can be seen the term $\frac{\Psi_k}{I_k}$ adds to the PRNU $K$. Since we expect $\frac{\Psi_k}{I_k}$ to be large (and almost constant!) in regions with characteristic structures, we can also expect the PRNU to be heavily biased by image content.

## Sensor identification performance of different image regions

To understand the impact of image content on the PRNU fingerprint quality, we compare the sensor identification performance in different image regions, each region exhibiting different properties. We expect PRNUs generated from regions with characteristic structures to (falsely!) outperform regions without such structures in sensor identification. Below we can see the results for PRNUs generated from the THU-FVFDT and the SDUMLA-HMT dataset. Both datasets exhibit regions with characteristic structures.

| Region Pair | Enh. Mode | |
|---|---|---|
| | *LUKAS* | *BM3D* |
| **P-FV / FV** | 0.65 (0.07) | 0.80 (0.08) |
| **P-BG-AB / BG-AB** | 0.88 (0.10) | 0.84 (0.04) |
| **P-ST / BG-AB** | 0.99 (0.01) | 0.85 (0.10) |
| **P-ST / FV** | 1.00 (0.00) | 0.94 (0.07) |
| **P-SS / BG-AB** | 0.93 (0.11) | 0.89 (0.08) |
| **P-SS / FV** | 0.94 (0.11) | 0.96 (0.06) |

**Figure 3:** Sensor identification performance of PRNUs generated from the THU-FVFDT dataset

| Region Pair | Enh. Mode | |
|---|---|---|
| | *LUKAS* | *BM3D* |
| **P-FV / FV** | 0.99 (0.01) | 1.00 (0.01) |
| **P-BG-AB / BG-AB** | 0.88 (0.14) | 0.89 (0.10) |
| **P-ST / BG-AB** | 0.99 (0.01) | 0.99 (0.01) |
| **P-ST / FV** | 1.00 (0.00) | 1.00 (0.00) |
| **P-SS / BG-AB** | n.a. | n.a. |
| **P-SS / FV** | n.a. | n.a. |

**Figure 4:** Sensor identification performance of PRNUs generated from the SDUMLA-HMT dataset

## What if we choose the wrong image region?

Assuming that we obtained a good estimate of the true sensor PRNU, we expect the correlation of the PRNU with noise residuals of the source dataset to be high. Repeating this experiment with BM3D-denoised images, we expect the correlation to decrease strongly since denoising destroys the residual information. However, if the PRNU is contaminated with characteristic structures (non-denoised image), the impact of denoising can be expected to be low as most of them are preserved in the denoised image.

| Dataset | Finger vein region | | | Full (non-cropped) image | | |
|---|---|---|---|---|---|---|
| | *Original* | *Denoised* | *Change* | *Original* | *Denoised* | *Change* |
| **THU-FVFDT** | 0.11 | 0.05 | ↓ 55.7% | 0.13 | 0.14 | ↑ **05.4%** |
| **FV-USM** | 0.26 | 0.10 | ↓ 60.6% | 0.32 | 0.21 | ↓ 35.0% |
| **MMCBNU** | 0.19 | 0.10 | ↓ 46.4% | 0.20 | 0.10 | ↓ 46.7% |
| **PALMAR** | 0.48 | 0.11 | ↓ 77.3% | 0.46 | 0.10 | ↓ 77.6% |
| **UTFVP** | 0.73 | 0.13 | ↓ 82.0% | 0.65 | 0.15 | ↓ 77.4% |
| **HKPU-FV** | 0.15 | 0.08 | ↓ 45.8% | 0.21 | 0.20 | ↓ **05.7%** |
| **IDIAP** | 0.63 | 0.15 | ↓ 75.8% | 0.60 | 0.16 | ↓ 74.0% |
| **SDUMLA-HMT** | 0.26 | 0.10 | ↓ 60.6% | 0.32 | 0.21 | ↓ 35.0% |

**Figure 5:** Average normalized cross correlation score of original and denoised images

As can be seen the relative change in correlation for "Full image" is much lower than for "Finger vein" in case of THU-FVFDT, FV-USM, HKPU-FV and SDUMLA-HMT. This indicates a contamination of the PRNU with image content.