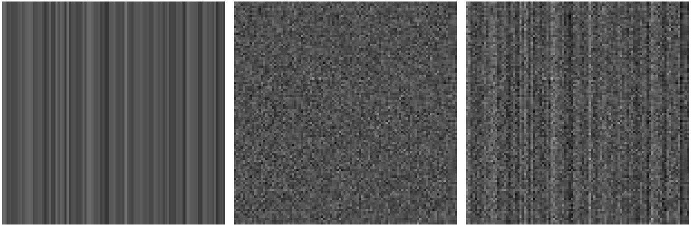
Scene-Based Noise Reduction on a Smart Camera

Faouzi Hamdi, Tomasz Toczek, Barthélémy Heyrman and Dominique Ginhac

LE2I UMR 6306, University of Burgundy, Dijon, France

December – 10 – 2012

- Raw output signal from CMOS image sensors tends to exhibit significant noise called **Fixed Pattern Noise (FPN)**.



(a) columns FPN component

(b) pixels FPN component

(c) global FPN Noise

Figure: Fixed Pattern Noise in CMOS image sensors

- Fixed-Pattern Noise (FPN): output variation under uniform illumination due to mismatched pixels and associating readout circuits.

- Fixed Pattern Noise includes offset FPN and gain FPN.
- Offset FPN:
  - Independent of pixel signal and fixed from frame to frame.
  - Pixel and column noise.
    - pixel noise due to transistors disparities resulting in pixels non-homogeneity.
    - column noise due to sense amplifiers disparities and the analog-to-digital converters which results in vertical stripes.
- Gain FPN
  - Correspond of spatial variation in pixel output values under uniform illumination due to pixel-to-pixel mismatches.
  - Gain FPN is difficult to correct.
- FPN is often corrected by subtracting its value, estimated through calibration, from the sensor's raw signal.

- Calibration-Based Techniques.
  - Calibration step at the factory: "FPN frame" is stored in a ROM and subtracted to each captured frame.
  - Two Point Calibration Technique.

$$X_{ij} = \frac{Y_{ij} - B_{ij}}{A_{ij}} \tag{1}$$

$$A_{ij} = \frac{Y_{ij}^{T2} - Y_{ij}^{T1}}{X_{ij}^{T2} - X_{ij}^{T1}} \quad and \quad B_{ij} = Y_{ij}^{T2} - A_{ij}X_{ij}^{T2} \tag{2}$$

X: Real signal
Y: measured signal
A: Gain FPN
B: Offset FPN
$T1/T2$: Temperature

- Scene-Based Techniques.
  - Algebraic Techniques.
    - Computation based on Global motion between the frames.
    - No statistical assumptions about the FPN.
    - Not suitable for real-time correction because of its weak computation.
  - Statistical Techniques.
    - FPN is modeled as a random spatial noise.
    - Estimates the statistics of the noise to remove it.
    - Lower computational complexity, smaller storage demands and better real-time performance.

– Constant-Statistic (CS) Method (Harris and Chiang) :

- The non-linear behavior of each photodetector is modeled as an affine function :

$$I_k = a_k \tilde{I}_k + b_k \qquad (3)$$

I_k: measured signal for k photodetector

$\tilde{I}_k$: actual luminance received by the k photo-detector

- Same mean and same variance among all the photodetectors *over time* :

$$\tilde{i}_k = \frac{I_k - m_k}{\sigma_k} \quad \textit{and} \quad \tilde{I}_k = A\tilde{i}_k + B \qquad (4)$$

$\tilde{i}_k$: corrected/estimated luminance for k photodetector
m_k: luminance measured average
σ_k: luminance measured variance

A/B: relationship factors between actual and estimated luminance

- Moving camera or otherwise moving objects of the filmed scene.

- CS approach cannot be used to continuously calibrate the sensor.

– Proposed Method (Improved CS Method) :

- Assumptions :
  - FPN has only an offset component (no gain component).
    - FPN is almost only seen on low luminosity and/or long exposure shots.
  - The mean value of the FPN offset is locally zero.
    - manufacturing defects occur randomly.
  - The high spatial frequency component at the corresponding pixel has an average value of zero *over time*.
    - FPN is mostly a high frequency component.

- Principle : Remove the High spatial frequencies from the current frame, apply the CS method on them (with $a_k = 1$) and add them back to obtain the final image.

- Benefits of our method :
  - Range of scenes correctly corrected is much wider.
  - The post-processing dynamic range adjustment step is no longer necessary.
- Correction Equation :

$$I_k^H + I_k^L = a_k \tilde{I}_k + b_k \tag{5}$$

$$\tilde{i}_k^H = I_k^H - m_k^H \qquad (constraint1) \tag{6}$$

– Final correction equation :

$$\tilde{i}_k^H = I_k^H - m_k^H - m'^H_{col(k)} \tag{7}$$

I_k^H / I_k^L: High/Low frequency component

$m'^H_{col(k)}$: average over time of the high spatial frequency component

– Actual luminance equation :

$$\tilde{I}_k = \tilde{i}_k^H + I_k^L \qquad (constraint2) \tag{8}$$

– Exponential window (constraint 3) :

$$\hat{m}_k^H(n) = \alpha I_k^H + (1-\alpha)\hat{m}_k^H(n-1) \tag{9}$$

$$\hat{m'}^H_{col(k)}(n) = \alpha I_k^H + (1-\alpha)\hat{m'}^H_{col(k)}(n-1) \tag{10}$$

- Correction chain : Low pass filter followed by two correction blocks.

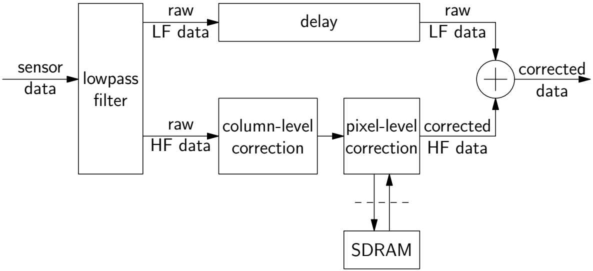


Figure: Proposed FPN correction pipeline

- Pixel correction block needs access to the external memory if the input resolution is too high.
- No access to external memory is required for the column corrector.

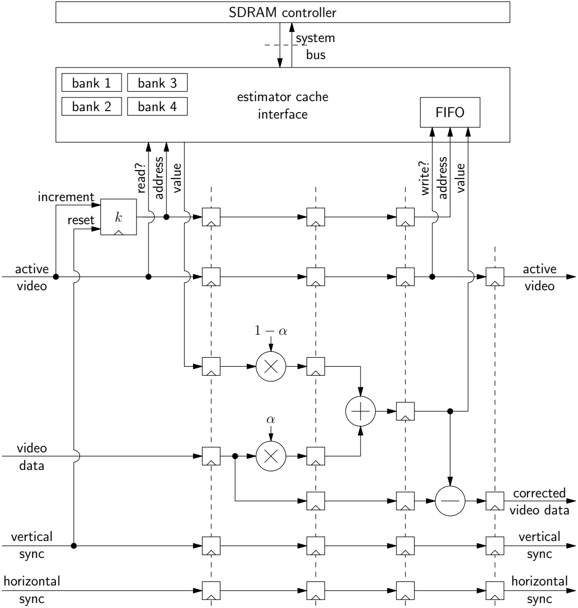


Figure: Pixel correction block principle schematic

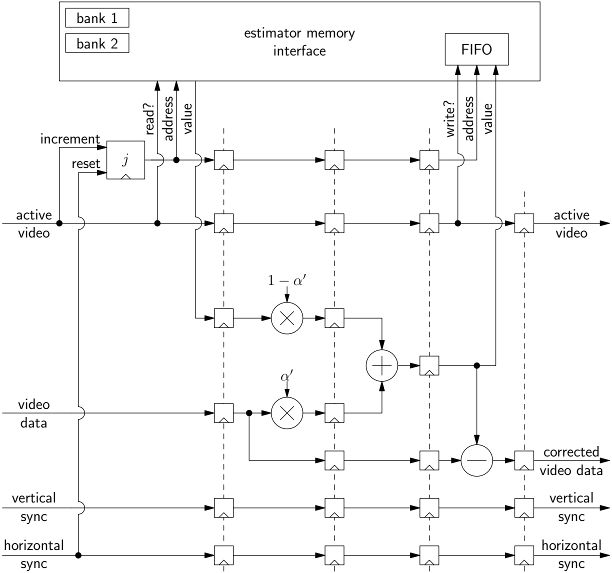


Figure: Column correction block principle schematic

- Pixel Corrector :
  - Main issue : storage of the temporal averages for each photodetector.
  - Solution : using of a small cache capable of prefetching the data from the SDRAM.
- Column corrector :
  - very similar to pixel corrector.
  - the cache is replaced by an on-chip memory containing all the column average values.
  - the estimator address counter is reset on horizontal synchronization instead of vertical synchronization.
- Low pass filter :
  - Sum of all the pixels of the previous line.
  - Intermediate sum of the pixels of the current line read so far.
  - –> "Filtered output" : previous scanline pixel sum, multiplied and shifted.

- Artificial offset noise :
  - uniform Pixel noise : 12.5% of the dynamic range.
  - uniform column noise : 12.5% of the dynamic range.
- Comparison metrics :
  - Universal Quality Index (UQI).
  - Peak Signal to Noise Ratio (PSNR).
- Parameters used for our method :
  - $\alpha = 2^{-5}$.
  - $\alpha' = 2^{-9}$.
  - 64x32 low-pass filter.
  - 16-bits pixel and 20-bits column estimators.
- Basic FPGA-based smart camera :
  - Avnet Spartan-6 LX150T development board (based on the Xilinx XC6SLX150T-3FGG676 FPGA).
  - Omnivision OV9715 image sensor capturing images at resolutions up to 1280x800 at 30fps.

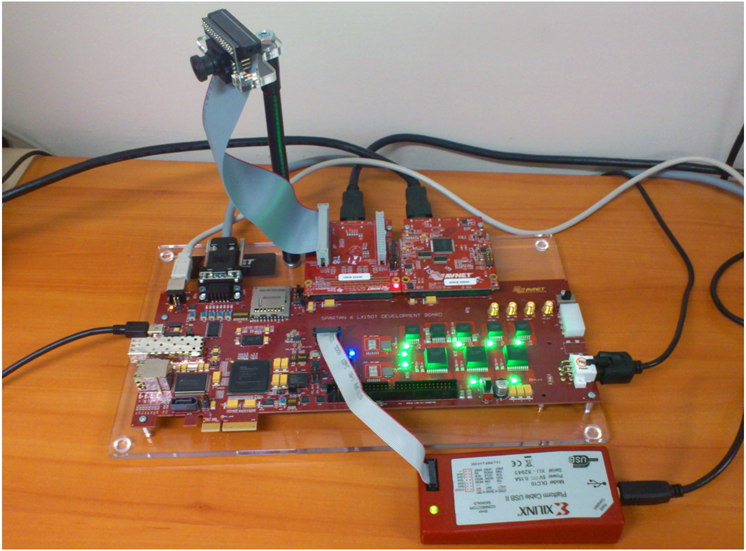


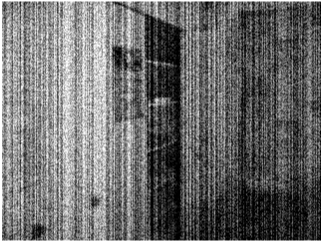
Figure: The Avnet Industrial Video Processing Kit

Faouzi Hamdi, Tomasz Toczek, Barthélémy Heyrman and Dominique Ginhac

# pipeline stages		Frequency	Slice usage	
Lowpass	Corrector		# registers	# LUTs
1	1	80.652 Mhz	580	289
2	2	99.502 Mhz	748	280
4	3	140.076 Mhz	994	762
5	4	149.410 Mhz	1099	745
6	4	169.319 Mhz	1172	596
7	5	129.232 Mhz	1258	786
8	6	158.278 Mhz	1348	854
11	8	148.170 Mhz	1546	671
22	16	134.391 Mhz	1878	1124

Table: FPGA implementation performance and ressource usage, depending on pipeline depths.

(a) Original image



(b) uncorrected image

(c) Corrected image using CS method (Harris and Chiang)

(d) Corrected image using the Proposed method

Figure: 1010^{th} frame of our test scene.

Faouzi Hamdi, Tomasz Toczek, Barthélémy Heyrman and Dominique Ginhac

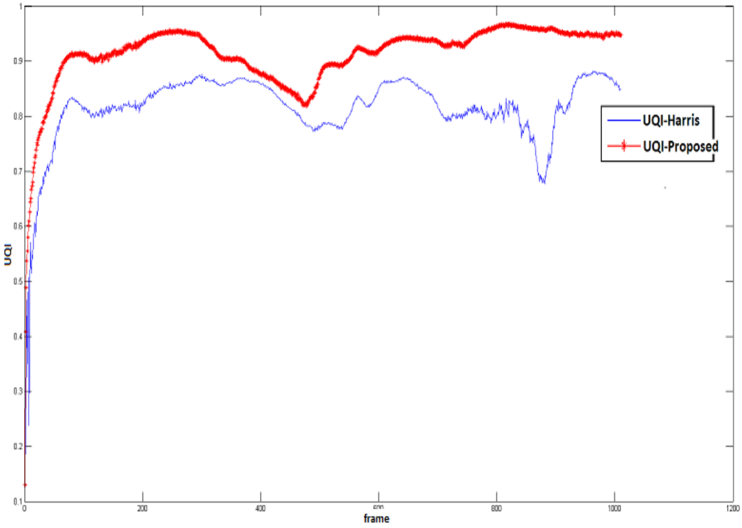


Figure: UQI Quality Metric

- Hardware Complexity : marginally higher than that of a simple reference image.
- Handling a wide variety of scenes correctly.
- Adds little latency in most cases.
- Automatic detection of static scenes.
- Increasing correction robustness: preserve the correctly computed values of two similar consecutive frames.

Faouzi Hamdi, Tomasz Toczek, Barthélémy Heyrman and Dominique Ginhac

Thank You For Your Attention

Any Question ?