

Whitepaper über Verbesserung der Bildqualität und Rauschreduzierung durch temporäres Averaging

Dem Rausch entfallen

Das folgende Dokument beschreibt, wie Sie das Entrauschen-Feature (Averaging oder Mittelung) der mvBlueFOX3 und mvBlueCOUGAR-X Familien verwenden können. Hierzu benötigen Sie bei der mvBlueFOX3-1 mindestens die Firmwareversion 2.8.0, bei der mvBlueFOX3-2 mindestens die Version 2.14.2. Im Gegensatz zu Bildverarbeitungsalgorithmen wie Tiefpass- oder Mittelwertfilter verbessert temporäres Entrauschen (Averaging) die Bildqualität, ohne dass Informationen im Bild verloren gehen.

Was macht das Feature?

Nehmen wir an, dass jemand den Inhalt zweier Bilder addiert. Während die Bildhelligkeit verdoppelt wird, wird das Rauschen nicht verdoppelt, da es (meist) statistisch verteilt wird. Statistisch gesehen wird der Rauschpegel nur um den Faktor $\sqrt{2}$ steigen, sobald zwei Bilder addiert werden. Wird das Bild um den Faktor 2 geteilt, um die gleiche Helligkeit wieder zu erreichen, kann gesagt werden, dass der Rauschpegel gesenkt bzw. den Signal-Rausch-Pegel um 3dB erhöht wurde.

Was passiert, wenn mehrere Bilder addiert werden?

Die EMVA1288 definiert eine Reihe von Funktionen und Messregeln für

- den Dynamikbereich,
- das Signal-Rausch-Verhältnis und
- dem Dunkelstrom.

Die Dynamik beispielsweise ergibt sich aus dem Verhältnis zwischen der maximalen Füllmenge mit Ladungsträgern (Full Well Capacity) N_{well} (oder nach EMVA1288 Aufnahmekapazität [Saturation Capacity]) und dem niedrigsten Signal, dem Dunkelstrom n_{dark} (N_{dark} [Dark Current]) gemäß folgender Formel:

Formel 1: Max. DNR eines Sensors

$$DNR_{sensor} \sim \frac{N_{well}}{\sqrt{N_{dark}}} \sim \frac{N_{well}}{n_{dark}}$$

Das Signal-Rausch-Verhältnis ergibt sich aus dem Verhältnis zwischen maximaler Füllmenge mit Ladungsträgern (Full Well Capacity) und dem Rauschen, was eine geometrische Summe aus Dunkelstrom und Signalrauschen ist. Die Physik besagt, dass Signale in der Natur durch sogenanntes

Schrotrauschen (auch Poissonsches Schrotrauschen oder Schottky-Rauschen) beeinflusst werden. Schrotrauschen ist proportional zum Mittelwert des eigenen Signals. In hellen Szenen überwiegt das Schrotrauschen bei Weitem, sodass SNR einfach als $\sqrt{\text{Aufnahmekapazität}}$ bestimmt werden kann.

Formel 2: Vereinfachte max. SNR eines Sensors

$$SNR \sim \frac{N_{well}}{\sqrt{N_{well} + N_{dark}}} \sim \frac{N_{well}}{\sqrt{N_{well}}}$$

Bilder zu addieren ist wie das Addieren der Aufnahmekapazität. Das ergibt für DNR und SNR:

Formel 3: Durchschnitts-DNR eines Sensors

$$DNR_{avg} \sim \frac{avg * N_{well}}{\sqrt{avg * N_{dark}}} = \frac{avg * N_{well}}{\sqrt{avg * n_{dark}^2}} = \sqrt{avg} * DNR$$

Formel 4: SNR-Erhöhung durch Entrauschen (Averaging)

$$SNR_{avg} \sim \frac{avg * N_{well}}{\sqrt{avg * N_{well}}} = \sqrt{avg * N_{well}} = \sqrt{avg} * SNR$$

Sowohl die Dynamik als auch das SNR werden durch \sqrt{avg} erhöht.

Dies ist $20 * (\log(\sqrt{avg}))$ dB

Beispiel: $avg=16$: DNR und SNR verbessern sich um den Faktor 4.

Das sind 12.04dB oder entsprechend eine 2 Bit höhere Dynamik.

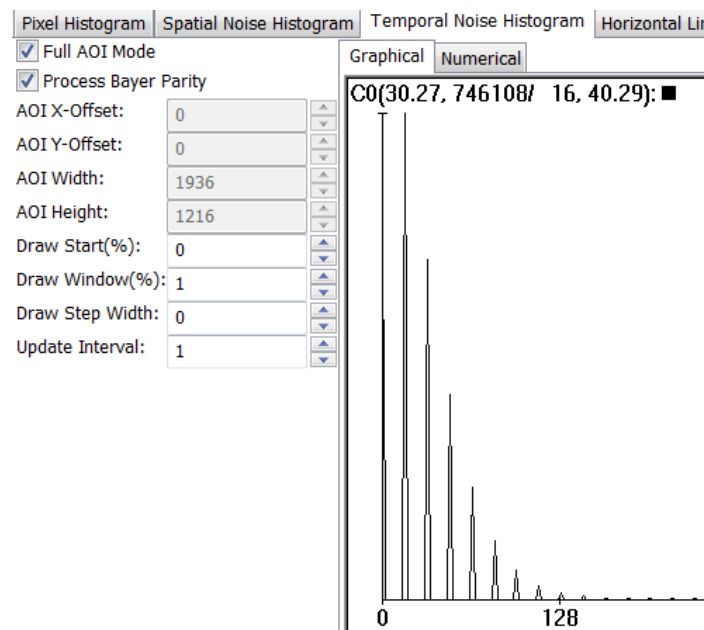
Die mvBlueFOX3-1xxx Serie kann bis zu 16 Bilder bei voller Bittiefe im Speicher der Kamera mitteln. Während die Kamera die Berechnung durchführt und die Bilder ausgibt, werden die neuen Bilder des Sensors. Der Datenfluss der gemittelten Bilder ist daher nicht kontinuierlich.

Die mvBlueFOX3-2xxx Serie kann die Bilder je nach Bildgröße und Bittiefe bis zu einer bestimmten Grenze summieren (z.B. mvBlueFOX3-2024: 56 Bilder@8 Bit oder 37 Bilder@12 Bit). Die Bildzufuhr ist kontinuierlich und kein eingehendes Bild wird verworfen.

Die mvBlueFOX3-2xxx Kameras unterstützen einen Modus für die summierte Ausgabe. D.h. das bspw. 56 Bilder bei 8 Bit als 16 Bit Summe ausgegeben werden.



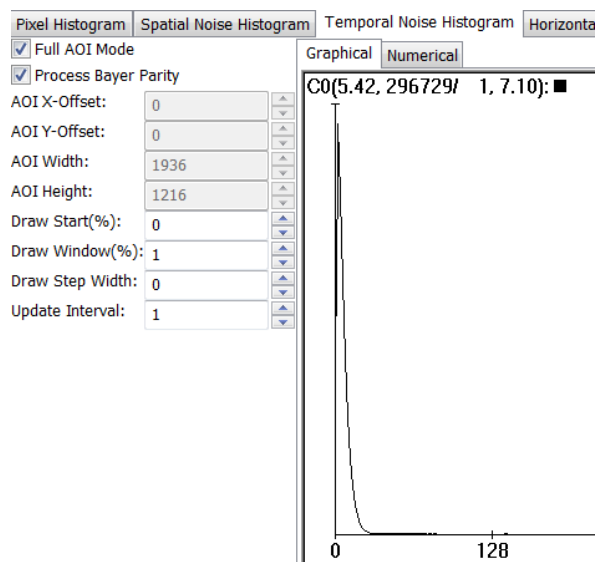
Der obige Screenshot zeigt ein Bild mit einer Verstärkung (Gain) von 42dB (mittels Bit-Shifting) ohne Averaging. Die Standardabweichung des temporären Rauschens (Pixelpositionen von einem Bild zum nächsten) des ganzen Bildes beträgt 40,29 Stellen (16 Bit).





Dieser Screenshot zeigt die gleiche Szene mit 32 gemittelten Bildern:

Das Rauschen wurde erheblich reduziert und keine Details gingen im Bild verloren. Das Bild hat immer noch einen Gain von 42dB! Dadurch konnte die Standardabweichung um den Faktor 5,6 = $\sqrt{32}$ verringert werden.



EMVA1288 Messungen mit der Kamera mvBlueFOX3-1013G, welche den CMOS-Sensor 76C560 von e2v mit 10 Bit ADC enthält, erzielt folgende Testergebnisse nach EMVA1288:

Averaging Modus	16x	4x	aus
Framerate [fps]	3,7	15,1	60,5
SNR,max [dB]	51,9	45,9	39,8
SNR,max [bit]	8,6	7,6	6,6
Dunkelrauschen [DN]	0,7	1,29	2,5
Dynamikumfang [dB]	62,7	57,3	51,4
Dynamikumfang [bit]	10,4	9,5	8,5

Sowohl SNR als auch DNR stimmten mit der dargelegten Theorie überein.

Was ist nun bei Anwendungen mit wenig Licht zu bevorzugen?

Bilder bei kurzer Belichtungszeit exp (erzeugt N_{exp} Photonen) mitteln avg oder ein Bild mit Belichtungszeit $avg * exp$ nehmen?

Dies kann mit der Formel für DNR beantwortet werden: In dunkleren Szenen fügt das Mitteln zu Beginn Photonen hinzu, während Dunkelstrom um die Quadratwurzeln hinzugefügt wird.

Formel 5: DNR-Erhöhung durch Averaging

$$DNR_{avg} \sim \frac{avg * N_{exp}}{\sqrt{avg * N_{dark}}} = \frac{avg * N_{exp}}{\sqrt{avg * n_{dark}^2}}$$

Formel 6: DNR-Erhöhung mittels Erhöhung der Belichtungszeit

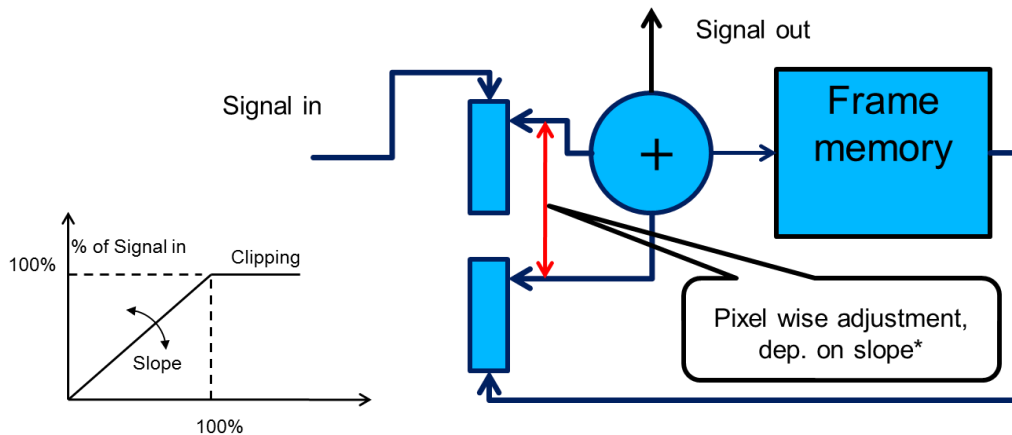
$$DNR_1 \sim \frac{avg * N_{exp}}{\sqrt{1 * N_{dark}}} = \frac{avg * N_{exp}}{\sqrt{1 * n_{dark}^2}} = \frac{avg * N_{exp}}{n_{dark}}$$

Die Aufnahme eines Bildes mit einer höheren Belichtungszeit führt zu einer höheren Dynamik, da der Dunkelstrom nur einmal einkalkuliert werden muss.

Während Averaging perfekt für statische Bilder funktioniert, ist es offensichtlich, dass es bei Bildern mit Bewegung zu Bewegungsunschärfen oder verschwommenen Bildern kommt.

Wie kann man der Bewegungsunschärfe Herr werden? – Durch Running Average (adaptive Rauschreduzierung)!

Das bekannte Prinzip der adaptiven Rauschreduzierung ist in den FPGA einiger MATRIX VISION CCD-Kameras des Typs mvBlueCOUGAR-X integriert.



Nehmen wir an, dass Pixel für Pixel die Abweichung des Grauwerts über Eingangssignal (*Signal in*) kontra Signal des Bildspeichers (Frame memory) ermittelt wird. Das eingehende Signal und das gemittelte Signal werden in einer umgekehrt proportionalen Art, dargestellt durch zwei Potenziometer, addiert. (D.h. umso mehr Eingangssignal desto weniger gemitteltes Signal)

Diese Eigenschaft wird durch einen „Slope“-Parameter gesteuert, sodass beispielsweise 10% Unterschied im Pixelwert zu 95% Eingangssignal und 5% gemitteltes Signal führen. Mit diesem Parameter kann die Rauschreduzierung auf Kosten von Bewegungsartefakten gesteuert werden. Dieser Ansatz eignet sich gut für Bewegung, denn es setzt auf den Fakt, dass das menschliche Auge auf Rauschen in bewegten Bildflächen empfindlicher reagiert, als bei statischen Bildflächen. Und genau das macht die Kamera. Statische Bildflächen werden entrauscht, da die Schleife „geschlossener“ ist. Bewegte Teile im Bild werden weniger entrauscht, da die Schleife für diese Pixel „geöffneter“ ist.

Weder wird durch das Entrauschen die Bildfrequenz reduziert noch kommt es zu einer Verzögerung. Der Grund ist einfach: Die Bildausgabe wird im FPGA noch vor dem Speicher erzeugt. Jedoch reduziert sich natürlich die Bildrate, da es zum Entrauschen einige Bilder benötigt.

Die folgenden Screenshots zeigen, dass sowohl in Bildern mit statischem Inhalt Entrauschung möglich ist, als auch in Bildern mit bewegtem Inhalt gleichzeitig wenig Rauschen und Bewegungsunschärfe möglich sind.

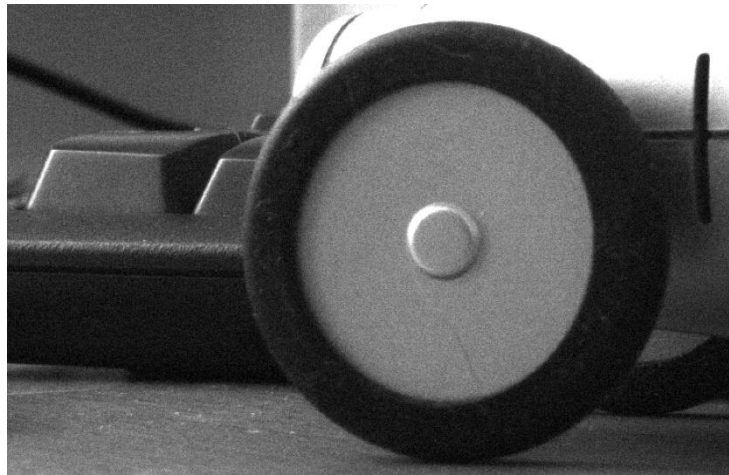
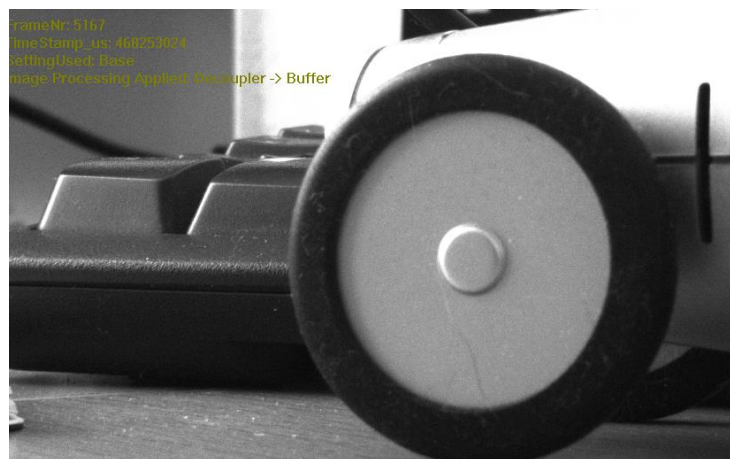
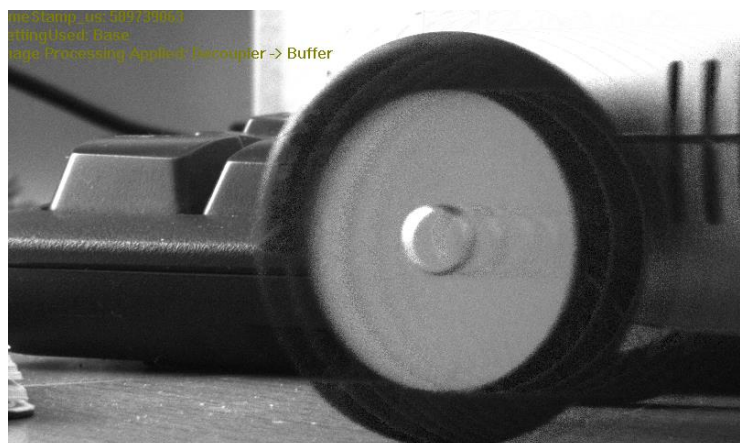


Bild mit statischem Inhalt und hoher Verstärkung (Gain); ohne Rauschreduzierung



Gleiche, entrauschte Szene (Slope 1000)



Szene mit sich bewegendem Reifen

Addendum (2021-06-01)

Belichtungszeit vs. Averaging / Summierung

Was ist der Unterschied zwischen eine N-mal längeren Belichtungszeit und N-Bild Averaging und N-fachen Verstärkung (Gain) eines Bildes (falls das aktuelle Bild zu dunkel ist)?

Beide Methoden führen zu einem gleichen Helligkeitsniveau beim finalen Bild.

- Der erste offensichtliche Unterschied ist die Framerate. Beim Averaging wird die Framerate um die Anzahl der Bilder reduziert, die beim Averaging verwendet werden.
 - Aus Sicht der Bildqualität haben wir bereits gelernt, dass sich das temporäre Rauschen um die Wurzel der Anzahl der gemittelten Bilder reduziert. Das führt zu einem Signal-Rausch-Verhältnis, was annähernd dem N-mal Belichtung vs. N-mal Averaging und N-mal Gain entspricht. Der Hauptunterschied ist, dass das Averaging auch beim Ausleserauschen des Sensor Anwendung findet und unabhängig von der Belichtungszeit ist. Das bedeutet, dass die Reizschwelle (Sensitivity Threshold) auf jeden Fall reduziert wird.
- ⇒ **Averaging erhöht die Empfindlichkeit der Kamera, während eine längere Belichtungszeit zu einem höheren Dynamikumfang führt. Für lichtschwache Anwendungen ist letzteres meist weniger wichtig, Averaging dagegen schon.**

Bild ist N-mal zu dunkel:

- Bei einer N-mal längeren Belichtung erhalten wir:

$$SNR_{NxExp}, DNR_{NxExp} \text{ und } SensitivityThreshold_{NxExp}$$

- Bei N-mal Gain und N-mal Averaging erhalten wir:

$$SNR_{NxAvg} \approx SNR_{NxExp},$$

$$DNR_{NxAvg} \approx \frac{1}{\sqrt{N}} DNR_{NxExp},$$

$$SensitivityThreshold_{NxAvg} \approx \frac{1}{\sqrt{N}} SensitivityThreshold_{NxExp}$$

Ableitung

SNR

$SNR = \frac{N_{sat}}{\sigma_{sat}} \approx \frac{N_{sat}}{\sqrt{N_{sat}}} = \sqrt{N_{sat}}$, aufgrund der Poisson-Verteilung entspricht die Abweichung den Mittelwert.

Wenn wir annehmen, dass wir nur $\frac{1}{N}$ der möglichen Sättigungsniveaus (saturation) N_{sat} erreichen, können wir N-mal Belichtung verwenden, um N_{sat} zu erhalten:

$$SNR' \approx \sqrt{N'_{sat}}, \text{ with } N'_{sat} = \frac{1}{N} \cdot N_{sat}$$

$$SNR \approx \frac{N \cdot N'_{sat}}{\sqrt{N \cdot N'_{sat}}} = \sqrt{N \cdot N'_{sat}} = \sqrt{N} \cdot SNR'$$

jetzt können wir Bilder hinzufügen (für Averaging gilt das gleiche):

$$SNR = \frac{\sum_{n=1}^N N'_{sat}}{\sqrt{\sum_{n=1}^N N'_{sat}}} = \frac{N \cdot N'_{sat}}{\sqrt{N \cdot N'_{sat}}} = \sqrt{N} \cdot SNR'$$

Auf diese Weise erhalten wir für N-mal Belichtung oder N-mal Averaging/Summierung den gleichen SNR-Wert

DNR

$$DNR = \frac{N_{sat}}{\sigma_{dark}}$$

Wird die Belichtungszeit verlängert, erhalten wir:

$$DNR = \frac{N \cdot N'_{sat}}{\sigma_{dark}} = N \cdot DNR'$$

Hinzufügen von Bildern:

$$DNR = \frac{\sum_{n=1}^N N'_{sat}}{\sqrt{\sum_{n=1}^N \sigma_{dark}^2}} = \frac{N \cdot N'_{sat}}{\sqrt{N \cdot \sigma_{dark}^2}} = \frac{N \cdot N'_{sat}}{\sqrt{N} \cdot \sigma_{dark}} = \sqrt{N} \cdot \frac{N'_{sat}}{\sigma_{dark}} = \sqrt{N} \cdot DNR'$$

Reizschwelle (Sensitivity Threshold)

$$\mu_{e.min} \approx \sigma_{dark} + \frac{1}{2}$$

Bei unterschiedlichen Belichtungszeiten bleibt der Wert konstant, da keine Abhängigkeit zur Belichtungszeit besteht.

Sobald Bilder hinzugefügt werden, erhalten wir:

$$\sigma_{dark} = \frac{1}{N} \cdot \sqrt{\sum_{n=1}^N \sigma'^2_{dark}} = \frac{1}{N} \cdot \sqrt{N \cdot \sigma'^2_{dark}} = \frac{1}{N} \cdot \sqrt{N} \cdot \sigma'_{dark} = \frac{1}{\sqrt{N}} \cdot \sigma'_{dark}$$

Und folglich:

$$\mu_{e.min} \approx \frac{1}{\sqrt{N}} \cdot \sigma'_{dark} + \frac{1}{2}$$

Simulationswerte:

Max Saturation: 10000 e-

Read-Noise: 2 e-

original	: Bright Noise: 27.95 e-, Dark Noise: 2.02 e-, SNR: 93.52 39.42 dB, DNR: 1294.49 62.24 dB
4 x exposure	: Bright Noise: 58.26 e-, Dark Noise: 2.21 e-, SNR: 180.01 45.11 dB, DNR: 4742.76 73.52 dB
4 x average	: Bright Noise: 13.96 e-, Dark Noise: 1.01 e-, SNR: 187.28 45.45 dB, DNR: 2588.27 68.26 dB