Advanced Camera and Image Sensor Technology

Dr. Joachim Linkemann

Senior Product Manager

Basler AG
• Physical model of a camera
• Definition of various parameters for EMVA1288
• EMVA1288 and image quality
• Noise in cameras
• Spectral response and penetration depth
• Area scan and line scan
• Spatial trigger
• Monochrome and color
• Single-, dual-, tri-linear sensors, TDI
• Mechanical and optical pixel size limitations
• Camera mounting standards
Physical Model of a Camera

A number of photons ...
... hitting a pixel during exposure time ...
... creating a number of electrons ...
... forming a charge which is converted by a capacitor to a voltage ...
... being amplified ...
... and digitized ...
... resulting in the digital gray value.
When light is emitted from a light source and passes through the optics and hits the Silicon, electrons will be generated.

The probability of how many electrons will be generated per 100 photons is called Quantum Efficiency (QE).

\[ QE = \frac{\text{number of electrons}}{\text{number of photons}} \]

Typical values are 30 to 60 %.
What Are The Important Things?
Full Well and Saturation Capacity

- Each pixel has a maximum capacity to collect electrons.
- But we will limit that to a lower maximum, because non-linear effects occur.
- A pixel has a saturation capacity of 50000 electrons.
- A glass of wheat beer contains approx. 50000 drops of beer.
- A glass of wine or champagne might have 10000 drops.
When refilling the pixel there is a slight jitter, although you try to get the same amount of electrons in. The "noise" is equivalent to the square root of the number of electrons.

Noise = sqrt (number of electrons)

Max. Signal to Noise Ratio (SNR) = (max number of elec.)/(sqrt(max number of elec.)) = Saturation Capacity/(sqrt(Saturation Capacity)) = sqrt(Saturation Capacity)

Sqrt(50000) = approx. 224

Refilled beer glasses do not have the same content.
• There is some remaining noise, although no light hits the sensor.

• Depending on the sensor, this is between 8 to 110 electrons.

• Compare this to the remaining drops of beer in an “empty” glass.
• If the signal is the same as the dark noise, we call this detection limit.

• CCD sensors have a detection limit of 8 to 25 electrons. This is according to 15 to 70 photons.

• CMOS sensors will start at 14 to 110 electrons, corresponds to 33 to a few hundreds.
Dynamic range is the ratio between a full and an empty glass of beer.

Dynamic range
= Full / Empty
= Saturation Cap. / Dark Noise
= 50000 / 25
= 200

A glass of beer will always have dynamics!
• We are using a logarithmic scale to cover from small to large numbers.
• On a linear scale you will not see the small numbers (a total scale of 1 meter has to show the detection limit in the range of less than a millimeter).

Set Quantum Efficiency to 50%,
100 photons will generate 50 electrons.
We are using a logarithmic scale to cover from small to large numbers.

A linear scale always adds the same unit. A logarithmic scale always multiplies with the same unit.

Bits ($2^n$) and decibel (dB) are logarithmic scales.

Multiplying by 2 is 1 bit or 6 dB.

$1 \text{ dB} = 20 \times \log x$

1 bit = 6.02 dB
An excellent image is \( \text{SNR} = 40 \) or better.

A good image quality is \( \text{SNR} = 10 \).
Saturation capacity: 18000 electrons
Dark noise: 9 electrons
QE @ 545 nm: 56 % (electrons per photons)

Saturation with $18000 / 0.56 \approx 32140$ photons
Detection limit $9 / 0.56 \approx 16$ photons

Max SNR: $\sqrt{18000} \approx 134$ (this is close to $128 = 2^7$)
7.1 bits or 43 dB

Dynamic range: saturation cap./dark noise =
$18000 / 9 = 2000$
(this is close to $2048 = 2^{11}$)
11 bits or 66 dB

Examples: Sony ICX285
Saturation capacity: 8000 electrons
Dark noise: 8 electrons
QE @ 545 nm: 50 % (electrons per photons)

Saturation with $\frac{8000}{0.50} \approx 16000$ photons
Detection limit $\frac{8}{0.50} \approx 16$ photons

Max SNR: $\sqrt{8000} \approx 89$ (this is between $64 = 2^6$ and $128 = 2^7$ )
6.5 bits or 39 dB

Dynamic range: saturation cap./dark noise $= \frac{16000}{16} = 1000$
(this is close to $1024 = 2^{10}$ )
10 bits or 60 dB

Examples: Sony ICX274
Image quality depends on the signal to noise ratio (SNR).

Total noise consists of:
- Temporal Noise ($\sim \sqrt{\text{number of photons}}$)
- Dark Noise ($\sim \text{number of photons}$)
Noise in Cameras I

There are three main noise sources on a sensor:

- **ADC noise**
  - When using two (like on a CCD) or multiple (like on a CMOS sensor) the ADCs might behave a little different from one device to the next.
  - It can be corrected / aligned if there is no light to the sensor and no voltage to the pixels.

- **Fixed Pattern Noise (FPN) or Dark Signal Non Uniformity (DSNU):**
  - Every single pixel has a different threshold when starting to convert photons to electrons. Especially on CMOS sensors the FPN is an issue.
  - It can be corrected / aligned with all voltages on, but no light to the sensor. The worst pixel is the threshold.
• Photo Response Non-Uniformity (PRNU)
  – Every single pixel has a slightly different conversion factor. This depends on geometric factors, material differences, etc.
  – It can be corrected / aligned if there is a uniform illumination to the sensor. Variations to the smoothed average are aligned with an individual pixel gain correction.

• Defect Pixels
  – Defect pixels, like dead pixels or hot pixels can be identified by similar measurements. A marked defect pixel can be interpolated by neighbor pixels.
The spectral response depends on wavelength and sensor type. CMOS sensors have often a wavy curve due to interferometric issues.
Cameras cover a spectral range of the visible (VIS) and near infrared (NIR), wavelengths from 400 to 1100 nm.

Penetration depth:

\[ \lambda \quad \text{Wavelength [nm]} \]

Wavelength: 1/100 of a human hair

Penetration depth: 1/100000 of a human hair
Area Scan and Line Scan

- Scan types can be separated by area scan and line scan.

- Area scan is known from a digital still camera.
  - 1 shot and the image is taken.
  - As an example: resolution 1300 x 1000 pixels. After an exposure time of 10 ms, everything is captured. All pixels have an exposure time of 10 ms.

- Line scan is known from a Xerox machine.
  - One line after the other is taken to get the total image.
  - As an example: resolution 1000 pixels, 1300 lines. With a total exposure time of 10 ms the image is taken, BUT: every single line (or each pixel) has only an exposure time of 7.7 µs!
  - This is a very short exposure time → You need much more light!
The Need of Spatial Trigger

- A trigger by time squeezes the object for different speeds (e.g. acceleration after a traffic light).
- Only a spatial trigger gives the right information.
- It does not depend on the speed of the object.
• Color on area scan can be taken either with a
  – 3 CCD setup: a beamsplitter separates the colors to three different CCDs.
  – Advantage: Every pixel has the full color information.
  – Disadvantages: Expensive, special lenses, alignment, color shades.

  ![Image of 3 CCD setup]

  – In most cameras a Bayer pattern is used.

  ![Image of Bayer pattern]
Color: Line Scan I

Single Line

Dual Line
Color: Line Scan II

Triple Line

3 CCD Line
The object on the conveyor belt is moved beneath the camera.
• Take 1, 8, and 15, but there is a color shade within one group RGB.

<table>
<thead>
<tr>
<th>Lines 1, 8, 15:</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lines 2, 9, 16:</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lines 3, 10, 17:</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lines 4, 11, 18:</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What is Possible with the Camera?

- Raw Image
- Corrected Image
TDI Line Scan Sensor

- TDI stands for Time Delay and Integration.
- The object is exposed several times, charges are accumulated and shifted simultaneously with the trigger.
- Signal is taken $N$ times, Noise reduces by $\sqrt{N}$. 

![Diagram showing Time Delay and Integration process](image)
The depth of focus (DOF) depends on the pixel size, the diameter of the iris and the focal length of the lens.
The depth of focus (DOF) depends on the pixel size, the diameter of the iris and the focal length of the lens.

\[
\text{DOF} = 2 \times \text{Pixelsize} \times \frac{f}{d} = 2 \times \text{Pixelsize} \times F/\# 
\]
Tilt of the Sensor

The sensor has to be aligned perpendicular to the optical axis.

Given max Depth of Focus

The whole sensor has to be taken into account
Examples of Geometrical DOF

Real numbers:
- Pixelsize = 5 µm
- F/# = 4

Depth of Focus: ±20 µm (40 µm)

Real numbers:
- Pixelsize = 4 µm
- F/# = 2

Depth of Focus: ±8 µm (16 µm)
Due to the physical structure of light as an electro-magnetical wave, the rays are blurred by diffraction. Diffraction depends on the F-number and wavelength $\lambda$.

$$\varnothing_{\text{Airy}} = 2.44 \times \lambda \times F/#$$

As a rule of thumb the diameter of the Airy disc is $F/#$ in microns, like $F/#$ is 4, the diameter of the Airy disc is approx. 4 $\mu$m.
The real spotsize and DOF is the geometrics folded with the diffraction.

To make life a little easier we will treat the diameter as independent errors (deviations):

$$\varnothing_{\text{total}} = \sqrt{\varnothing_{\text{geom}}^2 + \varnothing_{\text{diff}}^2}$$

Total DOF

Geometric DOF
In case of an 8k line scan sensor with 10 µm pitch we might have further issue:

The sensor might be bent by $\Delta z = 50$ to $80$ µm (still normal)

The lens (e.g. 5.6/90mm) will cause a field curvature of $\Delta z = 50$ to $100$ µm, whereas the astigmatism might occur and the meridional and sagittal focal plane might run the opposite direction.
In case of an 8k line scan sensor with 10 µm pitch we might have a difference of about 100 µm between both focal planes (meridional and sagittal).
Pixel Size

- Pixels are mainly between 10 to 3.5 µm.
- People are going to smaller pixels, because higher resolution with less silicon. More sensors from a single wafer.
- “Reasonable“ limit: 5 µm for monochrome, 2.5 µm for color.
- Full-Well capacity (saturation capacity, resp.) is lower for smaller pixels. Therefore the max SNR is not as good as on a larger pixel.
### Recommended Mechanical Interfaces (Mounts)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0</td>
<td>4</td>
<td>≈ ¼</td>
<td>C-, CS-, NF-, S-Mount</td>
</tr>
<tr>
<td>II</td>
<td>4</td>
<td>16</td>
<td>≈ 1</td>
<td>C-, CS-, NF-Mount</td>
</tr>
<tr>
<td>III</td>
<td>16</td>
<td>31.5</td>
<td>≈ 2</td>
<td>F-Mount, 48 mm Ring, M42 x 1, M48 x 0.75</td>
</tr>
<tr>
<td>IV</td>
<td>31.5</td>
<td>50</td>
<td>≈ 3</td>
<td>M58 x 0.75 (and F-Mount if possible)</td>
</tr>
<tr>
<td>V</td>
<td>50</td>
<td>63</td>
<td>≈ 4</td>
<td>M72 x 0.75</td>
</tr>
<tr>
<td>VI</td>
<td>63</td>
<td>80</td>
<td>≈ 5</td>
<td>M95 x 1</td>
</tr>
<tr>
<td>VII</td>
<td>80</td>
<td>100</td>
<td>≈ 6</td>
<td>M105 x 1</td>
</tr>
</tbody>
</table>

Please see: JIIA LER-004-2010 (Draft 0.20)
Contact Information

Dr. Joachim Linkemann
Senior Product Manager

Basler AG
An der Strusbek 60-62
22926 Ahrensburg
Germany

Basler Inc.
855 Springdale Drive
Suite 203
Exton, PA 19341 USA

Phone: +49 4102 463 236
Email: Joachim.Linkemann@baslerweb.com
Web: www.baslerweb.com
Get all the Manufacturing News - Check Manufacturing Newsletter daily.
http://www.manufacturingnewsletter.com/