pixel size & sensitivity

The relationship between pixel size of an image sensor and its sensitivity is discussed in detail to illuminate the reality behind the myth that "larger pixel image sensors are always more sensitive than small pixel sensors".
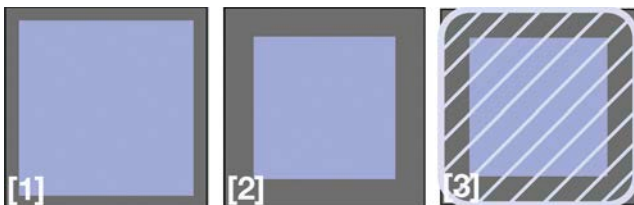
1 pixel size – fill factor

Figure 1
Pixels with different fill factors (blue area corresponds to light sensitive area): [1] pixel with 75% fill factor, [2] pixel with 50% fill factor and [3] pixel with 50% fill factor plus micro lens on top.

The fill factor of a pixel describes the ratio of light sensitive area versus total area of a pixel, since a part of the area of an image sensor pixel is always used for transistors, electrodes or registers, which belong to the structure of the pixel of the corresponding image sensor (CCD, CMOS, sCMOS). Only the light sensitive part might contribute to the light signal, which the pixel detects.
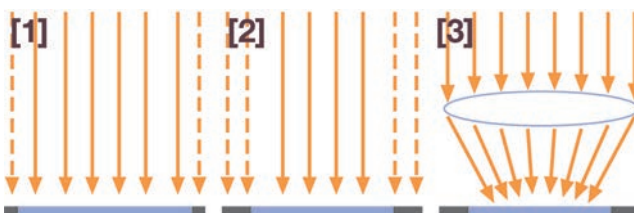

Figure 2
Cross sectional view of pixels with different fill factors (blue area corresponds to light sensitive area) and the light rays which impinge on them (orange arrows). Dashed light rays indicate, that this light does not contribute to the signal:
[1] pixel with 75% fill factor, [2] pixel with 50% fill factor and [3] pixel with 50% fill factor plus micro lens on top.

In case the fill factor is too small[i], this fill factor usually is improved by the addition of micro lenses, where the lens collects the light impinging onto the pixel and focuses the light to the light sensitive area of the pixel (see figure 2).

table 1: correspondences & relations

fill factor	light sensitive area	quantum efficiency	signal
large	large	high	large
small	small	small	small
small + micro lens	larger effective area	high	large

Although the application of micro lenses is always beneficial for pixel with fill factors below 100%, there are some physical and technological limitations to consider:

table 2: some limitations

limitations	
size limitation for micro lenses	For pixels larger than 12µm pixel pitch the stack height of the process to make a good micro lens can't be made high enough to generate a good lens. This stack height is proportional to the area, which should be covered[ii]
CMOS pixels have limited fill factor	due to semi-conductor processing there has always to be 25% of the pixel area covered with metal => maximum fill factor = 75%
light sensitive areas larger than 25µm x 25µm	are difficult to realize because diffusion length and therefore the probability for recombination increase

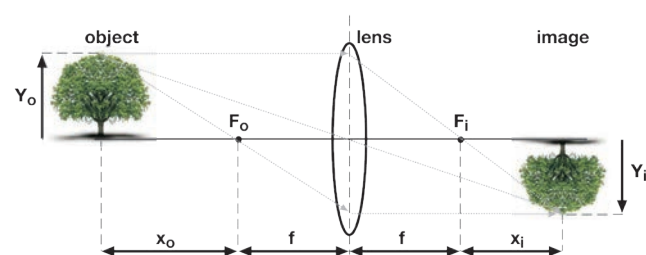## 2 pixel size – optical imaging

Figure 3
Optical imaging with a simple optical system based on a thin lens and characterized by some geometrical parameters: f - focal length of lens, F_o - focal point of lens on object side, F_i - focal point of lens on image side, x_o - distance between F_o and object = object distance, x_i - distance between F_i and image = image distance, Yo - object size, Y_i - image size

Figure 2 shows the situation of an optical imaging with

pixel size & sensitivity

a simple optical system based on a thin lens. For this situation the Newtonian imaging equation is valid:

$$x_0 \cdot x_i = f^2$$

or the Gaussian lens equation:

$$\frac{1}{f} = \frac{1}{(x_0 + f)} + \frac{1}{(x_i + f)}$$

and the magnification M is given by the ratio of the image size Y_i to the object size Y_o:

$$M = \left| \frac{Y_i}{Y_0} \right| = \left| \frac{f}{x_0} \right| = \left| \frac{x_i}{f} \right|$$

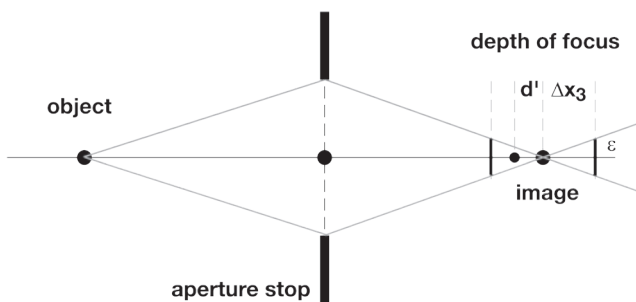### 3 pixel size – depth of focus / depth of field[iii]

Figure 4
Illustration of the concept of depth of focus, which is the range of image distances in which the blurring of the image remains below a certain threshold.

The image equation (chapter 2) determines the relation between object and image distances. If the image plane is slightly shifted or the object is closer to the lens system, the image is not rendered useless. It rather gets more and more blurred, the larger the deviation from the distances become, given by the image equation.

The concepts of depth of focus and depth of field are based on the fact that for a given application only a certain degree of sharpness is required. For digital image processing it is naturally given by the size of the pixels of an image sensor. It makes no sense to resolve smaller structures but to allow a certain blurring. The blurring can be described using the image of a point object as illustrated in figure 3. At the image plane, the point object is imaged to a point. It smears to a disk with the radius ε (see fig. 4) with increasing distance from the image plane.

Introducing the f-number $f_{/\#}$ of an optical system as

the ratio of the focal length f_0 and diameter of lens aperture D_0:

$$f_{/\#} = \frac{f_0}{D_0}$$

the radius of the blur disk can be expressed:

$$\varepsilon = \frac{1}{f_{/\#}} \cdot \frac{f_0}{(f_0 + d')} \cdot \Delta x_3$$

where $\Delta x3$ is the distance from the (focused) image plane. The range of positions of the image plane, [d' - $\Delta x3$, d' + $\Delta x3$], for which the radius of the blur disk is lower than ε, is known as depth of focus. The above equation cam be solved for $\Delta x3$ and yields:

$$\Delta x_3 = f_{/\#} \cdot \left(1 + \frac{d'}{f_0} \right) \cdot \varepsilon = f_{/\#} \cdot (1 + |M|) \cdot \varepsilon$$

where M is the magnification from chapter 2. This equation shows the critical role of the f-number for the depth of focus.
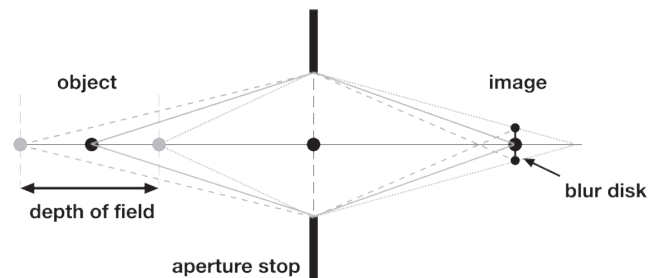

Figure 5
Illustration of the concept of depth of field, which gives the range of distances of the object in which the blurring of its image at a fixed image distance remains below a given threshold.

Of even more importance for practical usage than the depth of focus is the depth of field. The depth of field is the range of object positions for which the radius of the blur disk remains below a threshold ε at a fixed image plane.

$$d \pm \Delta x_3 = \frac{-f^2}{d' \mp f_{/\#} \cdot (1 + |M|) \cdot \varepsilon}$$

In the limit of $|\Delta x_3| \ll d$ we obtain:

$$|\Delta x_3| \approx f_{/\#} \cdot \frac{1 + |M|}{M^2} \cdot \varepsilon$$

If the depth of field includes infinite distance, the depth of field is given by:

$$d_{min} \approx \frac{f_0^2}{2 \cdot f_{/\#} \cdot \varepsilon}$$

[iii] This chapter comes from the book "Practical Handbook on Image Processing for Scientific and Technical Applications" by Prof. B. Jähne, CRC Press, chapter 4. Image Formation Concepts, pp. 131 - 132.

pixel size & sensitivity

Generally the whole concept of depth of field and depth of focus is only valid for a perfect optical system. If the optical system shows any aberrations, the depth of field can only be used for blurring significantly larger than those caused by aberrations of the system.

4 pixel size – comparison total area / resolution

If the influence of the pixel size on the sensitivity, dynamic, image quality of a camera should be investigated, there are various parameters that could be changed or kept constant. In the following different pathways are followed to try to answer this question.

4.1 constant area

constant => image circle, aperture, focal length, object distance & irradiance,
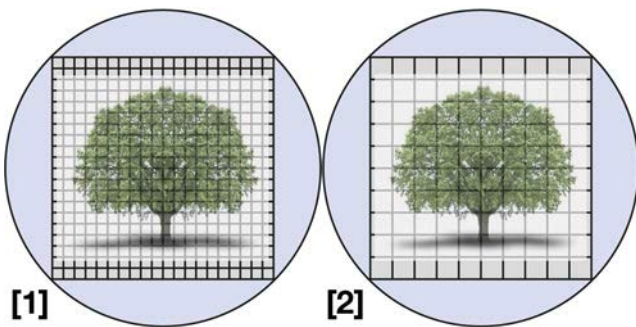**variable** => resolution & pixel size

Figure 6
Illustration of two square shaped image sensors within the image circle of the same lens, which generates the image of a tree on the image sensors. Image sensor [1] has pixels with the quarter size area of the pixels of sensor [2].

For the ease of comparison square shaped image sensors are assumed, which fit into the image circle or imaging area of a lens. For comparison a homogeneous illumination is assumed therefore the radiant flux per unit area also called irradiance E [W/m^2] is constant over the area of the image circle. Let us assume that the left image sensor (fig.6, sensor [1]) in figure 6 has smaller pixels but higher resolution e.g. 2000 x 2000 pixels at 10µm pixel pitch, while the right image sensor (fig. 6, sensor [2]) subsequently has 1000 x 1000 pixels at 20µm pixel pitch. The question would be which sensor has the brighter signal and which sensor has the better signal-to-noise-ratio (SNR)?

To answer this question it is either possible to look at a single pixel, but this neglects the different resolution, or to compare the same resolution with the same lens but this corresponds to compare a single pixel with 4 pixels.

Generally for the considerations it is also assumed that both image sensors have the same fill factor of their pixels. The small pixel measures the signal m, it has its own readout noise r_0 and therefore a signal-to-noise-ratio s could be determined for two important imaging situations: low light, where the readout noise is dominant and bright light, where the photon noise is dominant.

table 3: consideration on signal and SNR for different pixel sizes same total area

pixel type	signal	readout noise	SNR low light	SNR bright light
small pixel	m	r_0	s_0	s_1
large pixel	4 x m	> r_0	> s_0	2 x s_1
4 small pixels	4 x m	2 x r_0[iv]	2 x s_0	2 x s_1

Still the proportionality of SNR to pixel area at a constant irradiance is valid, meaning the larger the pixel size and therefore the area, the better the SNR will be. However, ultimately this means that one pixel with a total area which fits into the image circle has the best SNR:
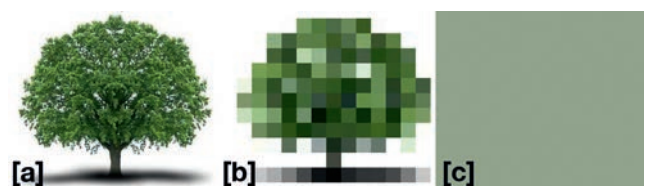

Figure 7
Illustration of three resulting images which where recorded at different resolutions: a) high resolution, b) low resolution and c) 1 pixel resolution

Assuming three image sensors with the same total area, which fits into the image circle of a lens, but they have different resolutions. As can be seen in figure 7 a) shows the nicest image but has the worst SNR per pixel, in figure 7 b) the SNR / pixel is better but due to the smaller resolution the image quality is worse compared to a). Finally figure 7 c) with a super large single pixel shows the maximum SNR per pixel but unfortunately the image content is lost.

[iv]since it is allowed to add the power of the readout noise, the resulting noise equals the square root of (4 x r_0^2)

pixel size & sensitivity

4.2 constant resolution

constant => aperture & object distance,
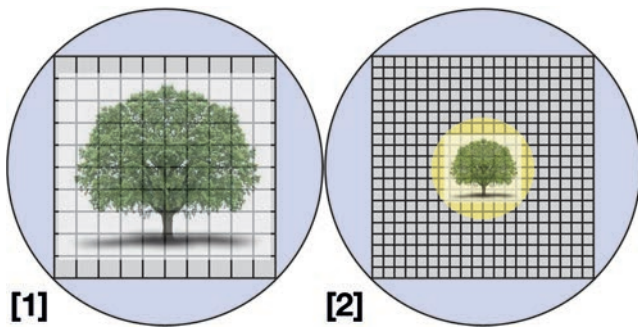*variable* => pixel size, focal length, area & irradiance

Figure 8
Illustration of two square shaped image sensors within the image circle of the same lens type (e.g. f-mount). The tree is imaged to the same number of pixels which require lenses with different focal lengths: Image sensor [1] has pixels with 4 times the area of the pixels of sensor [2].

Another possibility to look at a comparison is if the number of pixels should be the same at the same object distance. To achieve that condition it would be necessary to change the focal length of the lens in front of the image sensor with the small pixels. Since the information (energy), which before has been spread to a larger image area, now has to be "concentrated" to a smaller area (see figure 8 and 9).
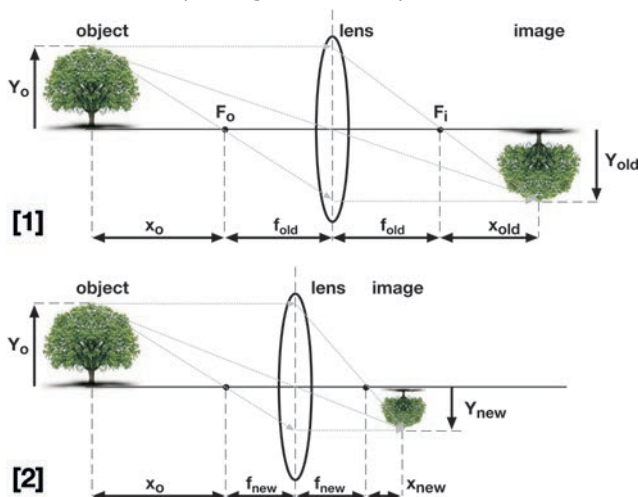

Figure 9
Illustration of two imaging situations, which are required for comparison:
[1] this is the original lens which images completely to the large pixel sensor - [2] this is the lens with the changed focal length, which images to the smaller pixel size sensor

If for example the resolution of the image sensor is 1000 x 1000 pixels, and the sensor with the smaller pixels has a pixel pitch of half the dimension of the larger pixel sensor, the image circle diameter d_{old} of the larger pixel sensor amounts to $\sqrt{2} \cdot c$ with c = width or height of image sensor. Since the smaller pixel sensor has half the pitch, it also has half the width and height and half the diagonal of the "old" image circle, which gives:

$$d_{new} = \frac{\sqrt{2}}{2} \cdot c = \frac{1}{\sqrt{2}} \cdot c$$

To get an idea of the required focal length, it is possible to determine the new magnification M_{new} and subsequently the required focal length f_{new} of the lens (see chapter 2 and figure 9), since the object size Y_0 remains the same:

$$M_{new} = \frac{|Y_{new}|}{|Y_0|} = \frac{\left|\frac{Y_{old}}{2}\right|}{|Y_0|} = \frac{1}{2} \cdot M_{old} \text{ with } \frac{Y_{new}}{Y_{old}} = \frac{\frac{1}{\sqrt{2}} \cdot c}{\sqrt{2} \cdot c}$$

Since the photon flux Φ_0 coming from the object or a light source is constant due to the same aperture of the lens (NOT the same f-number!), the new lens with the changed focal length achieves to spread the same energy over a smaller area, which results in a higher irradiance. To get an idea on the new irradiance, we need to know the new area A_{new}:

$$\frac{A_{new}}{A_{old}} = \frac{\pi \cdot \left(\frac{\frac{1}{\sqrt{2}} \cdot c}{2}\right)^2}{\pi \cdot \left(\frac{\sqrt{2} \cdot c}{2}\right)^2} = \frac{1}{4}$$

With this new area it is possible to calculate how much higher the new irradiance I_{new} will be.

$$I_{new} = \frac{\Phi_0}{A_{new}} = \frac{\Phi_0}{\frac{1}{4} \cdot A_{old}} = 4 \cdot \frac{\Phi_0}{A_{old}} = 4 \cdot I_{old}$$

The reason for that discrepancy between the argumentation and the results is within the definition of the f-number $f_{/\#}$[iii]:

$$f_{/\#} = \frac{f}{D}$$

pixel size & sensitivity

Figure 10
Images of the same test chart with same resolution, same object distance, same fill factor but different pixel size and different lens settings (all images have the same scaling for display):
a) pixel size 14.8 µm x 14.8 µm, f = 100mm, f-number = 16
b) pixel size 7.4 µm x 7.4 µm, f = 50mm, f-number = 16
c) pixel size 7.4 µm x 7.4 µm, f = 50mm, f-number = 8

The lenses are constructed such that the same f-number always generates the same irradiance in the same image circle area. Therefore the attempt to focus the energy on a smaller area for the comparison is not accomplished by keeping the f-number constant. It is the real diameter of the aperture D, which has to be kept constant. Therefore the new f-number would be, if the old f-number was $f_{/\#}$=16 (see example in figure 10):

$$f_{/\#new} = \frac{f_{new}}{D_0} = \frac{f_{new}}{\left(f_{old} \middle/ f_{/\#old}\right)} = \frac{f_{new}}{f_{old}} \cdot f_{/\#old} = \frac{50}{100} \cdot 16 = 8$$

Again the question would be which sensor has the brighter signal and which sensor has the better signal-to-noise-ratio (SNR)? This time we look at the same number of pixels but with different size, but due to the different lens at the same aperture the irradiance for the smaller pixels is higher. Still it is assumed that both image sensors have the same fill factor of their pixels.

table 4: consideration on signal and SNR for different pixel sizes same resolution

pixel type	signal	readout noise	SNR low light	SNR bright light
small pixel	m	r_0	s_0	s_1
large pixel	m	$> r_0$	$< s_0$	s_1

If now the argument would be, that the larger f-number will cause a different depth of field, which will in turn change the sharpness of the image and therefore the quality, the equation from chapter 3 can be used to look at the consequences.

From the above example we have the focal lengths f_{old} = 100mm and f_{new} = 50mm, and we are looking for the right f-number to have the same depth of field. Therefore:

$$\frac{f_{old}^2}{2 \cdot f_{/\#old} \cdot \varepsilon} = \frac{f_{new}^2}{2 \cdot f_{/\#new} \cdot \varepsilon}$$
$$\Leftrightarrow$$
$$f_{/\#new} = \frac{f_{new}^2}{f_{old}^2} \cdot f_{/\#old} = \frac{50^2}{100^2} \cdot 16 = 4$$

Since we have calculated a f-number = 8 for the correct comparison, the depth of field for the image sensor with the smaller pixels is even better.

5 pixel size – fullwell capacity, readout noise, dark current

The fullwell capacity of a pixel of an image sensor is one important parameter, which determines the general dynamic of the image sensor and therefore also for the camera system. Although this fullwell capacity is influenced by various parameters like pixel architecture, layer structure and well depth, there is a general correspondence also to the light sensitive area. This is also true for the electrical capacity of the pixel and the dark current, which is thermally generated. Both, dark current and capacity[v] add to the noise behaviour, and therefore larger pixels also show larger readout noise.

table 5: consideration on fullwell capacity, readout noise, dark current

pixel type	light sensitive area	fullwell capacity	dark current	capacity	readout noise
small pixel	small	small	small	small	small
large pixel	large	large	large	large	large

6 pixel size – dynamic (intra scene dynamic, pixel dynamic)

The intra scene dynamic dyn_{max} of an image sensor is the expression, which generally is used to determine the dynamic of an image sensor or a camera system. It describes the maximum contrast or light level, which an image sensor can measure or detect within one image.

[v] the correspondence between pixel area, capacity and therefore kTC-noise is a little bit simplified, because there are newer techniques from CMOS manufacturers like Cypress, which overcome that problem, nevertheless the increasing dark current is correct, whereas the sum still follows the area size. Further in CCDs and newer CMOS image sensors most of the noise comes from the transfer nodes, which are cancelled out by correlated double sampling (CDS).

pixel size & sensitivity

It is defined as ratio of the maximum number of electrons, that a pixel can take - the fullwell capacity divided by the smallest light (signal) level that can be used - the readout noise.

$$dyn_{max} = \frac{fullwell\ capacity\ [e^-]}{readout\ noise\ [e^-]}$$

If now pixels with different sizes have to be compared, the fullwell capacity and the readout noise roughly follow the pixel size. For some very large pixel the increase in fullwell is more pronounced, therefore they have a larger dynamic dyn_{max}. On the other hand the very small pixels have still decreasing fullwell but persistent readout noise, which decreases the dynamic. Hence, the maximum pixel dynamic corresponds to the signal-to-noise-ratio of the light, because solely the signal is then photon noise limited:

$$pixel\ dynamic_{max} = \frac{fullwell\ capacity\ [e^-]}{\sqrt{fullwell\ capacity\ [e^-]}}$$

Therefore for normal image sensors (normal = no image sensors with incredible high readout noise), the larger pixels have the best maximum pixel dynamic.

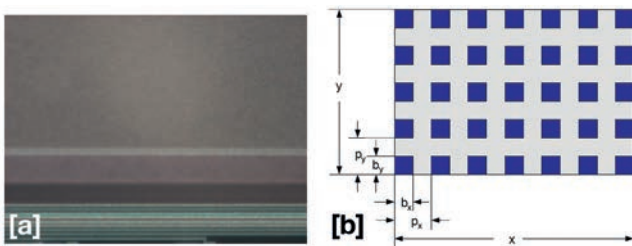## 7 pixel size – resolution, contrast & modulation transfer function (MTF)

Figure 11
a) a macro image of an image sensor showing the pixels at the edge of the imaging area, b) illustration of an image sensor with characteristic geometrical parameters: x, y - horizontal, vertical dimensions, p_x, p_y - horizontal, vertical pixel pitch, b_x, b_y - horizontal, vertical pixel dimensions

The modulation transfer function, MTF, is used to describe the ability of the camera system to resolve fine structures, resolution, and the contrast, which finally defines the quality of the image.

The resolution ability depends on the geometrical parameters of the image sensor, and the used lenses for imaging (including micro lenses).
The maximum spatial resolution is therefore given by:
- line pair [lp] needs two pixels to be resolved
- assumption: square pixels with bx = by = b and px = py = p
- the Nyquist frequencies correspond to spatial frequencies which are usually expressed in line pairs per mm [lp/mm]

This results in the maximum possible axial R_{axial} and diagonal $R_{diagonal}$ resolution ability, that is given by the dimensions of the pixel:

$$R_{axial} = \frac{1}{2 \cdot p} \text{ and } R_{diagonal} = \frac{1}{\sqrt{2} \cdot 2 \cdot p}$$

In the following table there are the maximum resolution ability values for image sensors with different pixel sizes given and for comparison some corresponding lens values.

table 6: some MTF data of image sensors & lenses

item	image sensor / lens type	pitch [µm]	focal length [mm]	R_{axial} [lp/mm]	$R_{diagonal}$ [lp/mm]
ICX285AL	CCD	6.45	-	77.5	54.8
MT9M413	CMOS	12	-	41.7	29.5
ICX205AK	CCD	4.65	-	107.5	76
APO-Xenoplan 1.4/23	c-mount	-	23	75	-
Cinegon 2.1/6.0	c-mount	-	6	30	-

The next important parameter is the contrast or modulation M, which is defined with the intensity I [count] or [DN[vi]] in an image:

$$M = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

The modulation depends on the spatial frequencies, which means that M is a function of the resolution R: M = M(R). The quality of the imaging process is described by the modulation transfer function, MTF.

[vi] DN = digital number, like count

pixel size & sensitivity

So both parameters, the resolution and the contrast, define the quality of an image, as is illustrated in figure 12. Increasing resolution improves the sharpness of an image while increasing contrast adds to the "brilliance".
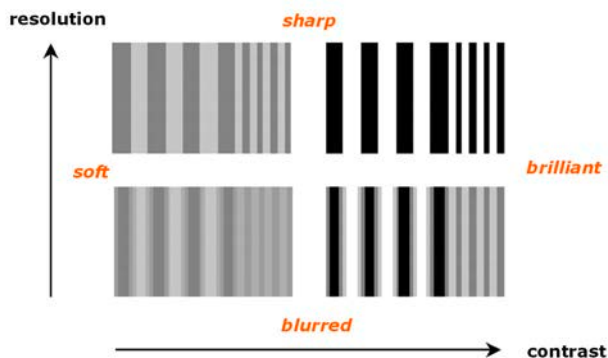

Figure 12
Illustration of the influence of resolution and contrast on image quality

8 Compare cameras with respect to pixel size & sensitivity

Generally it is a good idea to image the same scene to each camera, which means:

- *Keep the object distance and the illumination constant!*

If the cameras should be compared, they should use the same resolution, which either means analogue binning or mathematical summation or averaging, or usage of a region/area of interest.

- *Keep or adjust the same resolution for all cameras!*

Then select a proper focal length for each camera, whereas each camera should see the same image on the active image sensor area.

- *Select corresponding lens with the appropriate focal length for each camera!*
- *Adjust the lens with the largest focal length with the largest useful f-number!*

For a proper comparison use the equation for the f-number on page 8, keep the aperture D constant, and calculate the corresponding f-number for the other lenses and adjust them as good as possible - then compare the images!

- *Adjust the f-numbers of the other lenses in a way, such that the aperture of all lenses is equal (similar)!*
- *Compare for sensitivity!*