Celestre, Rafael; Rosenberger, Maik; Notni, Gunther

A novel algorithm for bad pixel detection and correction to improve quality and stability of geometric measurements

A novel algorithm for bad pixel detection and correction to improve quality and stability of geometric measurements

R Celestre, M Rosenberger and G Notni

Ilmenau University of Technology
Department of Quality Assurance and Industrial Image Processing
Gustav-Kirchhoff-Platz 2, 98693 Ilmenau, Germany

E-mail: `rafael.celestre@tu-ilmenau.de`

Abstract. An algorithm for detection and individually substitution of bad pixels for further restoration of an image in the presence of such outliers without altering overall image texture is presented. This work presents three phases concerning image processing: bad pixel identification and mapping by means of linear regression and the coefficient of determination of the pixel output as a function of exposure time, local correction of the linear and angular coefficients of the outlier pixels based on their neighbourhood and, finally, image restoration. Simulation and experimental data were used as means of code benchmarking, showing satisfactory results.

1. Introduction
Image based measuring techniques rely on the most precise estimation of dark to bright transitions, the so called edge detection, to extract geometric features from the examinee. Going towards the subpixel resolution, values delivered by pixels, as means of their grey values, play a much important role for precise geometric measurements [1]. For accurate, improved and stable measurements it is vital that bad pixels are not only identified, but also corrected locally - not to alter overall image texture, before the image is used for metrology or quality control applications.

2. Bad pixel definition
A bad pixel can be defined as a pixel that does not behave as expected, producing anomalous values and therefore, no valuable information can be extracted from them. The data provided by them is not only less relevant but also less reliable than the produced by its neighbourhood. Bad pixels occur either isolated or in clusters, the latter being more difficult to handle.

2.1. Classification of outliers
There are plenty of ways in which a pixel may not deliver reliable information: from defects on electronics to reading and accessing the data generated by the sensors. It is out of the scope of this work to discuss the origins of bad pixels, since for this application it is enough to group outlier pixels by their behaviour on the resulting image, i.e. resulting grey value:

- **Linear pixels with false bias**: a category where the pixel output behaves linearly with illumination intensity or exposure time, but due to either a low or high bias compared to its neighbours, values displayed by this pixel differs considerably from the ones around it, given an homogeneous illumination;

- **Nonlinear pixels**: a subset of defect pixels, where the output signal displayed on an image does not relate to the illumination intensity or exposure time in a linear fashion;

- **Dead pixels**: are those who have very low sensitivity to illumination intensity variations, consistently presenting low grey values;

- **Hot pixels**: this category is highly sensible to illumination intensity variations, presenting persistent high grey values.

3. Bad pixel mapping

For applications where subpixel precision is aimed, it is important to be able to perform pixel correction without disturbing healthy parts of the image, i.e. altering the texture or smoothing edge transitions. This approach can be done by mapping unhealthy pixels and performing localised correction operations. The criteria used on this work for tagging a pixel as an outlier is its linearity as a function of sensor integration time, measured by the coefficient of determination [2] calculated from the linear regression done pixel-wise by the method of the least squares [3] to each pixel of this image stack. The procedure for generation of a so called bad pixel map is as follows:

Firstly, multiple images are taken at a fixed integration time (t_{int}). The exposure time is incrementally increased until saturation of the detector is reached, so that at the end of this calibration procedure a set of multiple images is obtained for each integration time - see [4] for further guidance. This set is then reduced by averaging the images obtained at each integration time. Doing so is important in order to suppress temporal noise, which can occasionally disturb bad pixel mapping. This happens because the variance of the temporal noise ($\sigma_{temp.} \approx 1\%$) is normally larger than the variance of spacial inhomogeneities of the image acquisition device ($\sigma_{spac.} \approx 0.3 - 0.5\%$) [4]. From this reduced set, *averaged image stack*, a set of three matrices is generated based on the assumption that for zero intensity until 90% of the saturation intensity, the grey values are a linear function of the integration time, as shown in fig-1(right). Linear coefficient (LC), angular coefficient (AC) and coefficient of determination (R^2) are 2D pixel maps of the quantities they are named after.

A binary pixel map for unhealthy pixels, *bad pixel map*, is generated based on a threshold for R^2. The closer this parameter is to one, the better the linear model is for the data set in question [2]. For each matched pixel, i.e. pixel whose R^2 is lower than the threshold, its coordinates are saved on a coordinate array, called *bad pixel location vector*.

4. Image restoration in the presence of bad pixels

The recovery of the lost information due to unhealthy pixels is done in two steps: generation of the corrected matrices and the image restoration in the presence of bad pixels. Firstly a map of irrecoverable pixels is generated from the *bad pixel map* by visiting each location from the *bad pixel location* array and substituting the outlier with the median of the values of its Moore neighbourhood:

$$\text{corrected pixel value} = \text{median}_{nxn}\{\text{bad pixel neighbourhood}\}, \tag{1}$$

where n is related to the Moore neighbourhood range r by: $r = (n-1)/2$ [1].

After this stage, all remaining low values on the *irrecoverable pixel map* cannot be corrected by the presented method. Generally, this matrix highlights pixel clusters that are too big for correction. The procedure of local substitution of bad pixels with the aid of the *bad pixel*

location array is also applied for the linear and angular coefficient matrices, as well as for the acquired image. Those resulting matrices are called *corrected* matrices and will be used for image restoration.

The second step is done under the following assumptions: grey values assumed by a pixel are a linear function of the integration time of the imaging sensor:

$$\text{grey value} = \text{linear coefficient} + \text{angular coefficient} \times t_{int}, \tag{2}$$

and that the average number of photons per squared unity of area at the detector (A), e.g. pixel surface, is a multivariable function of exposure time, wavelength (λ) and irradiance (E) on the sensor surface, the latter, being given in power per area unity [4]:

$$\frac{\text{number of photons}}{A} = \frac{t_{int}\lambda E}{hc}, \tag{3}$$

where h is the Planck's constant and c is the speed of light. By letting both wavelength (ideally monochromatic) and irradiance constant, the number of photons per squared unity of area is a linear function of the exposure time. One can, therefore, relate the pixel grey value directly and linearly to the number of incident photons. This is the core idea of the correction algorithm: pixel-wise recover the number of photons as a function of the grey value. From equations 2 and 3:

$$\text{number of photons} = \frac{\text{grey value} - \text{linear coefficient}}{\text{angular coefficient}} \cdot \frac{\lambda E}{hc}, \tag{4}$$

without losing generality, the constant term can be dropped, hence:

$$\text{number of photons} \propto \frac{\text{grey value} - \text{linear coefficient}}{\text{angular coefficient}}, \tag{5}$$

The image recovery, performed on the *corrected image* from the previous step, is done as follows: the *corrected linear coefficient* matrix is extracted from the *corrected image* and the resulting matrix is pointwise divided by the *corrected angular coefficient* matrix. Those operations lead to the restored image, that is, a mapping of the number of photons per pixel - a more realistic description of the scene measured:

$$\text{recovered image} = \frac{\text{corrected image} - \text{corrected linear coefficient matrix}}{\text{corrected angular coefficient matrix}}. \tag{6}$$

This image can be pointwise multiplied by the *irrecoverable pixel map* to make visually clear where bad pixels that are unable to be corrected are located.

5. Results

The benchmarking of the algorithm was performed in three phases: bad pixel mapping and recovery of lost information with simulated data, in order to explore the code sensitivity in a controlled environment; bad pixel mapping and recovery of lost information with measured data in accordance with the EMVA standard 1288 guidelines; and a probe measurement to evaluate image restoration and its effectiveness.

5.1. Bad pixel mapping and recovery of lost information with simulated data
Were done as a proof of concept and to investigate the algorithm sensitivity. A set of images created to mimic the experimental process as described in sec. 3. Four types or pixel misbehaves, namely the ones presented at subsec. 2.1, were introduced at random positions on the images, as shown in fig. 1(left). The methodology described in sec. 3 was applied to the artificial images.
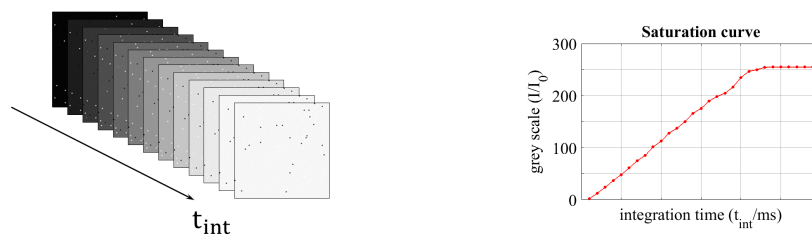
Figure 1. (left) Shows stack of averaged images for increasing exposure times (from dark to saturated images). (right) Integration time is increased until saturation is reached, so that a suitable integration time can be chosen as a sound operation point.

The aforementioned routine for bad pixel mapping is sensitive to most types of pixel misbehaves (non-linear, hot and dead pixels). Nonetheless, false bias and false angular coefficients - i.e. angular coefficients that differ significantly from those of its neighbourhood, but nonetheless are still linear - are often not ruled out as bad pixels. The only conditions where the latter misbehaves are tagged as bad pixels, is when the grey values reach saturation prematurely, that is, below the operation point and have a behave similarly to a hot pixel.

5.2. Bad pixel mapping and recovery of lost information with measured data
In order to test the algorithm on real imaging sensors, a calibration setup was assembled following guidelines displayed at [4]. A schematic can be seen on fig. 2, where the main components are displayed. The sensor under test is a EV76C661 with CMOS technology from the company *e2v*. The procedures followed here are detailed at sec. 3. The light intensity was chosen so that the saturation of the sensor was reached at 1s of integration time. This working point is particular interesting for low light applications and high dynamic range imaging. The light required for the experiment is monochromatic, which was obtained by means of optical filters. Although a multispectral characterisation of the sensor is out of scope, measurements were performed from 400nm to 950nm in 50nm steps. The results described next were taken at 650nm - fig. 3. The Moore neighbourhood range used was of a unity.
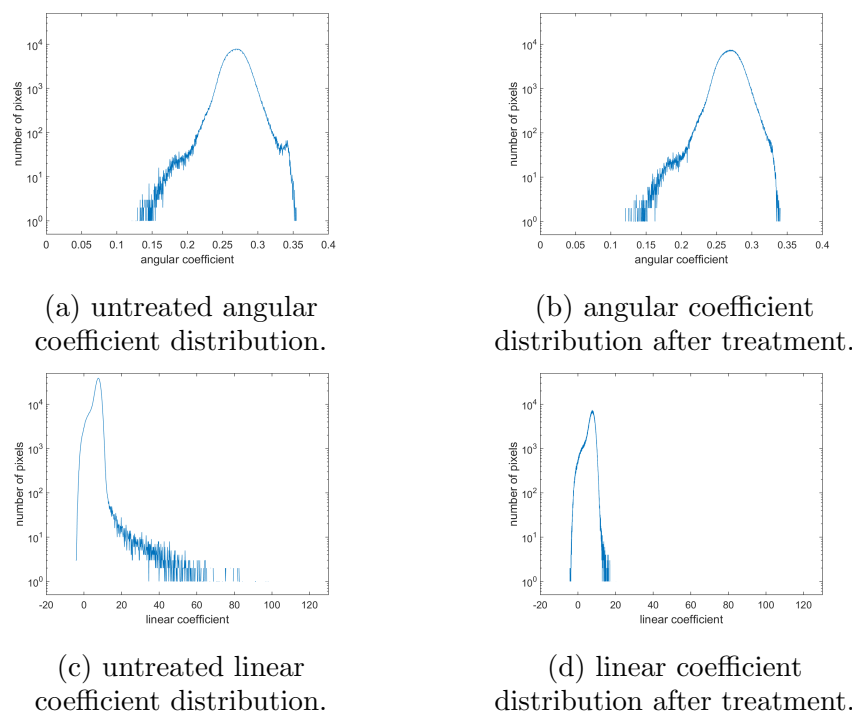
The angular coefficient (fig. 3(a)) gives an idea of how *fast* a pixel increases its grey value, in other words, how fast it reaches saturation. The linear coefficient (fig. 3(c)) can be understood as the pixel value without an illumination source or as an offset. To get a glimpse of how bad pixels behave, one must take a look at the angular and linear coefficient distribution after the proposed treatment (figs. 3(b) and (d), respectively). The former has a *lump* right before 0.35 (grey value/s), which is smoothed after treatment (figs. 3(a) and (b)). Low angular coefficients seem to be unaffected by the algorithm. The latter shows alarmingly high offsets: from 20 to over 100 grey values. After correction, all linear coefficients are under 20 - figs. 3(c) and (d). It is clear that the *lump* found on the angular coefficient histogram and the spikes over 20 grey values at the linear coefficient histogram are as bad pixels - they correspond to less than 0.35% of the sensor surface. As for the generation of the *irrecoverable pixel map*, no bad pixels were identified. Since there are no clusters of foul pixels on this sensor, one is lead to believe in a complete elimination of bad pixels.

5.3. Evaluation of image restoration in the presence of bad pixels
After the successful benchmarking of the bad pixel mapping and recovery of lost information, it was necessary to study the performance of the image restoration algorithm - described at sec. 4. The experimental setup shown in fig. 2 was kept unaltered, except by the introduction of an objective lens between (6) and (7). The results can be seen at fig. 4. The reader might point out
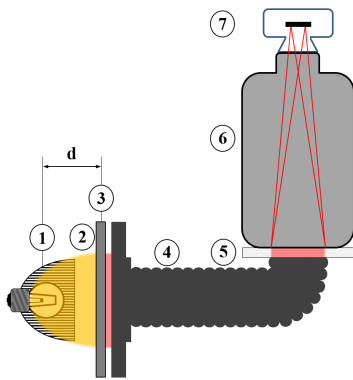
Figure 2. Shows a schematic of the measurement setup for sensor characterisation, where: 1) is a halogen lamp, 2) broadband light, 3) metallic filter for wavelength selection (bandwidth of ± 10nm), 4) optical fibre, 5) diffuser plate to ensure homogeneity of the illumination field, 6) tube to prevent contamination from ambient line and 7) is the sensor being probed. The d stands for the variable distance between light source and filter, used to adequate the light intensity to the detector, which was chosen to make the sensor reach saturation at around 1s of acquisition time.

(a) untreated angular coefficient distribution.

(b) angular coefficient distribution after treatment.

(c) untreated linear coefficient distribution.

(d) linear coefficient distribution after treatment.

Figure 3. Shows real data of a sensor characterisation with the proposed algorithm. Histograms from angular and linear coefficients calculated with the least squares method are displayed.

that the treated image, fig. 4(c), appears to be noisier than its untreated counterpart, fig. 4(a), although seemingly true, it can be accounted as different pixel sensitivities - figs. 3(a) and (b) show how broadly distributed they are. On the other hand, the contrast is increased by a factor of 3 - from a contrast of 7 times bright to dark values (untreated image) to 22 times bright to dark values (treated image), which is very positive for geometric measurements. Nonetheless, new studies are meant to tackle better algorithms for image restoration, in order to suppress the aforementioned inhomogeneities.

6. Conclusion
The presented work can be seen as two modular and independent methods of image processing: bad pixel identification, value correction and image restoration. The first block is able to identify outliers and locally substitute their values (*in situ* action), without altering overall image texture:

(a) untreated image.

(b) grey value histogram.

(c) treated image.

(d) grey value equiv. histo.
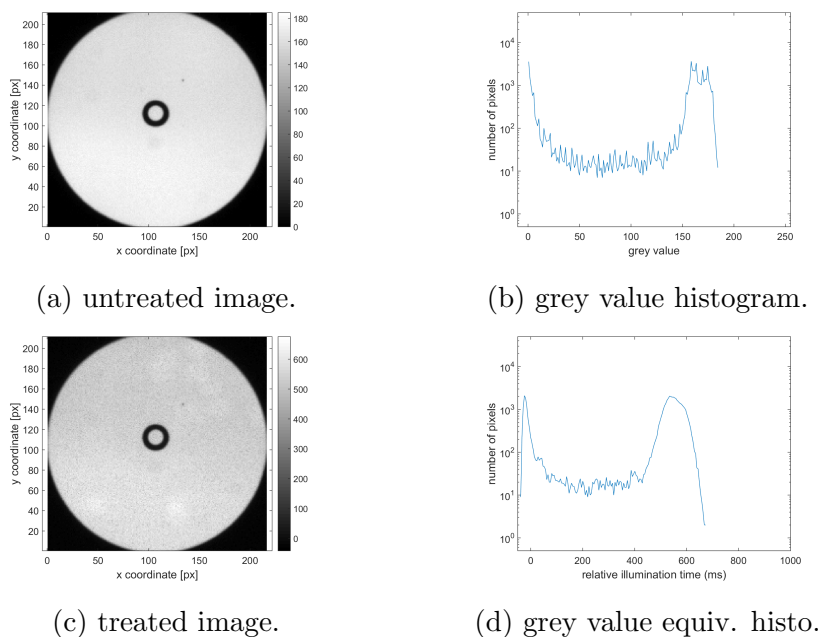
Figure 4. Shows (a) untreated image and (a) the distribution of grey values in a histogram. (c) image treated with the proposed algorithm and the (d) distribution of relative illumination time histogram.

its counterparts not only do not generate a list with outliers position (eg. EMVA1288 only gives the number of outliers), but also often disturb healthy pixels (see filtering operations [1]). There seems to be a correlation between high offset-pixels and *fast* ones - fig. 3, which leads to an interesting conclusion for pixel misbehaves: false bias and unusually high angular coefficients are often linked. The second method described is somehow similar to the shading correction (compensation for illumination), widely used for image processing [1], but done in a more reliable way for being able to account for bad pixels, whereas the latter is done mainly to account for illumination inhomogeneity. The recovered image seems to be more noisy, but contrast between bright and dark spots is increased. The resulting image is theoretically a better description of the measured scene. The current method is appealing to low light applications and HDR imaging, where integration times are often high and bad pixels have a more prominent role on image quality. Although multispectral analysis is out of scope, it could be seen - in preliminary analysis - that for the different colour channels, the behaviour of this algorithm is quite similar.

7. Acknowledgments

8. References

[1] W. Abmayr. *Einfuehrung in die digitale Bildverarbeitung.* Teubner, 1994.
[2] Richard Anderson-Sprecher. Model comparisons and r^2. *The American Statistician*, 48(2):113–117, 1994.
[3] I. Kreyszig. *Advanced Engineering Mathematics.* John Wiley & Sons, Inc., 2006.
[4] European machine vision association. EMVA Standard 1288 A3.0 - Standard for Measurement and Presentation of Specifications for Machine Vision Sensors and Cameras, 2010.