ROBUST ETIMATES IN DARK CURRENT SUPPRESSION

[1]*František Mojžíš*, [1]*Jan Švihlík*, [2]*Jaromír Kukal*

[1]Department of Computing and Control Engineering, ICT Prague
[2]Faculty of Nuclear Sciences and Physical Engineering, Department of Software Engineering in Economics, CTU in Prague

Abstract

This paper is devoted to noise suppression in astronomical images. These are contamined by thermally generated noise, called the Dark Current, which occurs in Charge Coupled Devices that are used as image sensors in astronomical science. Many basic statistical and filtering methods have been studied for Dark Current elimination. Applied algorithms must be sensitive to processed data, i.e., must not remove important objects and information. This paper introduces denoising methods based on the robust statistics, mainly on L and M-estimates. Proposed methods are applied to the real astronomical data. Results of used filters are compared from vizual point of view and also by Signal to Noise Ratio characteristic.

1 Introduction

Noise suppression is one of the most fundamental topics in digital image processing [1, 2]. This is especially important in astronomical science, where we want to suppress unwanted information added to the original pattern, but on the other hand we want to preserve information concerned in the acquired data.

Astronomical images present special kind of digital images. These are acquired with long exposure times and during night, when the light conditions are poor. Astronomical cameras [3] used for image acqusition, contain Charge Coupled Devices [3, 4] (CCD), that are used as image sensors. Useful signal is generated by photons incident onto photosensitive area of CCD. Except of useful information, thermal noise is also generated. This noise is called Dark Current and can be described by the following relation

$$I_d = A\mathrm{e}^{-\frac{\beta}{kT}} \tag{1}$$

where A and B are material constants, T is thermodynamic temperature in Kelvin and β is a Boltzman constant, $\beta = 1.38 \times 10^{-23}$ J K^{-1}.

From Eq. (1) can be seen, that this type of noise is temperature dependent and it is also proportional to the exposure time. Dark Current can be eliminated by cooling of CCDs and can be almost removed when sensor temperature decrease under -213.15 K. This condition is not always satisfied.

From mathematical point of view, Dark Current can not be described as a statistical random variable with normal probability distribution. Thus lot of denoising methods can not be used, but some useful algorithms exist.

The simpliest method that can be used, is in having information about used system. In astronomical science are known three type of images. The first one is a light image that represents data of the the night sky, etc. Second one is dark image, that is acquired by astronomical camera, when the shutter is closed. This image maps Dark Current position. The third one is called flat filed and presents vizualization of system optical path. When we have flat and the dark image, then we are able to remove Dark Current from the light image.

Previous presumptions are nice, but are not usually satisfied. In practice we have only light and the dark image, and in the worst case only the light one. In this situation we have to use different filtering methods. Commonly used algorithm for Dark Current suppression is median filter and Wiener filter can be also used [5].

The main aim of this paper is to test methods based on the robust statistical approaches [6] and their comparison with mentioned median filter. These methods flow from methods of robust estimation like L and M-estimates [7, 8, 9]. Excpet of this methods, there were tested methods of neigbourhood selection of the processed pixel.

The rest of this paper is organized as follows. Robust estimates in noise suppression introduces statistical L and M-estimates and methods of processed area selection. Results of used denoising algorithms are presented in the Result section and further discussed in the Conclusion.

2 Robust estimates in noise suppression

Robust [6, 10] estimation is approach, which is insensitive to small departures from the idealized assumptions which have been used to optimize the algorithm. These techniques concerns:

- L-estimates - based on linear combination of order statistics,
- M-estimates - use maximum likelihood methods,
- R-estimates - use statistical rank tests.

These estimtes found its application in suppression of various noises like Gaussian, Cauchy, shot noise, etc. An example of mentioned estimates used in shot noise removing is median filter, that belongs to L-estimates group.

In this paper were used only two first classes of mentioned robust estimates, beacuse of their simple realization. All of the algorithms were written and tested in Matlab.

2.1 L-estimates

L-estimates [7, 10] are based on statistical sample sorting and linear combination of sorted values, it means that weighted average is computed. Each value of sorted list $\boldsymbol{x}$ has its specific weight w_i and $\sum_{i=1}^{n} w_i = 1$. Result of L-estimate is then expressed as

$$y = \sum_{i=1}^{n} x_i w_i. \tag{2}$$

2.1.1 Median filter

Median filter [1, 2] is well known as order statistic filter and belongs to nonlinear digital filtering techniques. Let us have ordered list $\boldsymbol{x}$, which number of elements is n, then median [13] is defined as

$$y = \begin{cases} \frac{x_{n/2} + x_{n/2+1}}{2} & \text{if } n \text{ is even} \\ x_{n/2} & \text{if } n \text{ is odd} \end{cases} \tag{3}$$

2.1.2 Trimmed mean

Trimmed mean [10, 11] or truncated mean is one of the most simpliest form of L-estimates. From both ends of the list is truncated same number of values and a new vector of weights is computed. This new vector of weights (in original vector had all values same weight) has the same length as the original vector of weights but trucanted values have their weights equal to zero. Trimmed mean can be analytically expressed as

$$y = \frac{1}{n-2l} \sum_{i=l+1}^{n-l} x_i \tag{4}$$

where n is total number of elements in $\boldsymbol{x}$ and l is number of trimmed values.

2.1.3 Windsorized mean

As in the case of trimmed mean, winsorized mean [10, 11] removes given number of values from the original vector and new vector of weights is computed. First non-zero weights are higher, because of taking over the weights of truncated values. This can be written as

$$y = \frac{1}{n} \left((l+1)x_{l+1} + \sum_{i=l+2}^{n-l-1} x_i + (l+1)x_{n-l} \right) \tag{5}$$

where n is total number of elements in $\boldsymbol{x}$ and l is number of replaced values.

2.1.4 Triangular distribution

Triangular distribution [12] is also combined together with trimmed mean but weights are uniformly increasing. Maximum is in the middle of new weight vector. Triangular distribution of random variable k is defined as

$$w(k; a, b, c) = \begin{cases} 0 & \text{for } k < a \\ \frac{2(k-a)}{(b-a)(c-a)} & \text{for } a \leq k \leq c \\ \frac{2(b-k)}{(b-a)(b-c)} & \text{for } c \leq k \leq b \\ 0 & \text{for } b < k \end{cases} \tag{6}$$

where a and b are lower and upper limits and c is a mode. Result value using this L-estimate is not analytically expressed and has to be evaluated using Eq. (2) and (6).

2.1.5 Binomial distribution

Binomial distribution [13] is almost the same as triangle distribution but the weights are given by the following relation

$$w(k, n, p) = \binom{n}{k} p^k (1-p)^{n-k} \tag{7}$$

where k is number of successes, $k = 0, 1, \cdots, n$, n is number of trials and p is success probability in each trial. Results value of this L-estimate has to be evaluted by Eq. (2) and (7), because there is no analytical expression too.

2.1.6 Tukey BES

Tukey Best Easy Systematic Estimation [11], known as trimean, is an average of three values (1^{st}, 2^{nd} and the 3^{rd} quartile [13]). Tukey BES is expressed by

$$\text{BES}(\boldsymbol{x}) = \frac{1}{4} \left(x_{(\lceil \frac{n}{4} \rceil)} + x_{(\lfloor \frac{n+1}{2} \rfloor)} + x_{(\lceil \frac{n+1}{2} \rceil)} + x_{(\lfloor \frac{3n+4}{4} \rfloor)} \right) \tag{8}$$

where $\boldsymbol{x}^{\mathrm{T}} = (x_1, x_2, ..., x_n)$ is a list and $n \in \mathbb{N}$.

Mentioned algorithm is possible to modify using BES estimation of Walsh list. This modification is called WBES. Walsh list is defined as a list containing elements $\frac{x_i+x_j}{2}$, where $i < j$.

2.2 M-estimates

In statistics, M-estimators are a broad class of estimators, which are obtained as the minima of sums of functions of the data. Least-squares estimators and many maximum-likelihood estimators are M-estimators. M-estimators are used in robust statistics. The statistical procedure of evaluating an M-estimator on a data set is called M-estimation. In practice, they are based on influence function ψ [8, 9], which gives lower weights to the outliers and higher weights to the more important data. Thanks to this weights w, we are able to evaluate weighted average.

Generally, result value using M-estimates is evaluated as the sequence of following steps. At firts the data are normalized [14]

$$u = \frac{x - \mu}{\sigma} \tag{9}$$

where x are original data, μ and σ respectively are mean value and standard deviation of x. After this, weights w_i are evaluated. This is followed by denormalization and summation [14]

$$y = \frac{1}{n}\left(\sum_{i=1}^{n}(u_i\sigma + \mu)w_i\right). \tag{10}$$

where n is total number of elements in $\boldsymbol{x}$ Graphical forms of used influence functions are figured in Fig. 1 and their definitions are listed bellow:

- Andrews weight function

$$w = \begin{cases} \sin(\frac{u}{c})\frac{c}{u} & \text{if } |u| \leq \pi c \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

- Bisquare weight function

$$w = \begin{cases} \left(1 - \left(\frac{u}{c}\right)^2\right)^2 & \text{if } u < c \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

- Cauchy weight function

$$w = \frac{1}{1 + \left(\frac{u}{c}\right)^2} \tag{13}$$

- Fair weight function

$$w = \frac{1}{\left(1 + \left|\frac{u}{c}\right|\right)^2} \tag{14}$$

- Huber weight function

$$w = \begin{cases} 1 & \text{if } |u| < c \\ \frac{c}{|u|} & \text{otherwise} \end{cases} \tag{15}$$

- Logistic weight function

$$w = \tanh\left(\frac{u}{c}\right)\frac{c}{u} \tag{16}$$

- Welsch weight function

$$w = \exp\left(-2\left|\frac{u}{2c}\right|^2\right). \tag{17}$$

Constant c used in Eq. (11) - (17) are specific for each M-estimate and can be found in the following table.

Table 1: Constants c applied in used influence functions.

influence function	Andrews	Bisquare	Cauchy	Fair	Huber	Logistic	Welsch
constant c	1.339	4.685	2.385	1.400	1.345	1.205	2.985

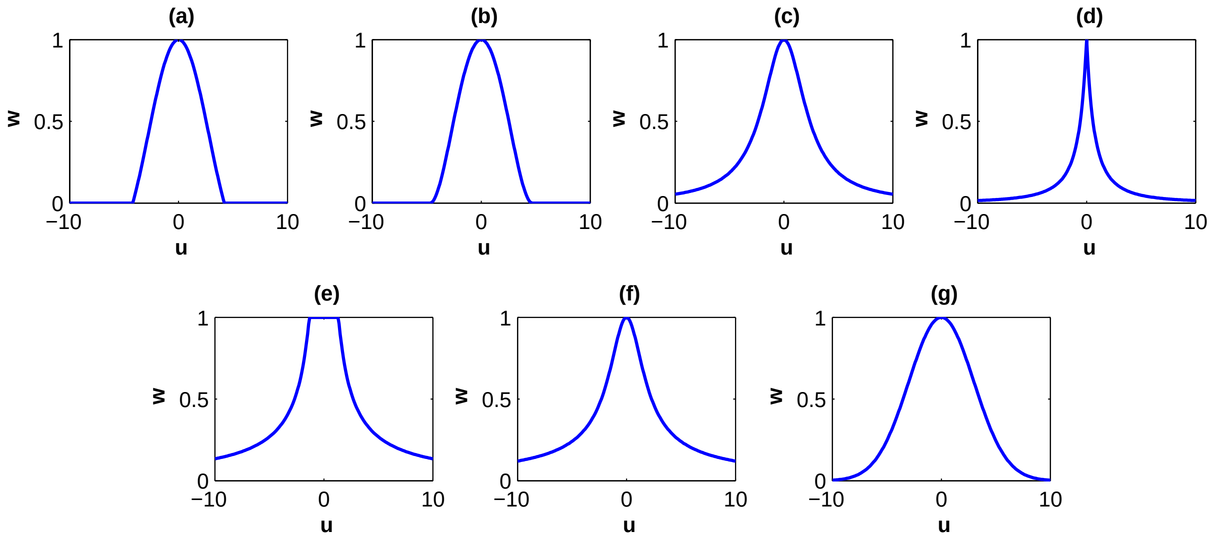

Figure 1: Weight functions shapes of used M estimates, (a) Andrews, (b)Bisquare, (c) Cauchy, (d) Fair, (e) Huber, (f) Logistic, (g) Welsch.

2.3 Masks definition

Mask is a binary matrix, which is applied to processed matrix (image) and selects defined region that is further processed. The most important is size of the mask, because of variance grow up in the neighbourhood of processed pixel.

At first it is necessary to define radius of the mask, which realizes a selection from neighbourhood of (x, y), thus the size of rectangle (usually square) region $S_{x,y}$ must be odd. If the middle row and column of the region $S_{x,y}$ are intersected in point (x, y), then the neighborhood of that point is called radius of the mask, marked r.

Let us have a square region $S_{x,y}$, then the square mask defined on this region is given by

$$\mathrm{m}(s,t) = 1, \tag{18}$$

where $(s, t) \in S_{x,y}$. Next type of the mask is called rhomb mask and it is defined as

$$\mathrm{m}(s,t) = \Big|\mathrm{sign}\big(r - (|s| - |t|)\big)\Big|. \tag{19}$$

Annulus mask is given by the following equation

$$\mathrm{m}(s,t) = \Big|\mathrm{sign}\big((r_1 - \sqrt{s^2+t^2})(\sqrt{s^2+t^2} - r_2)\big)\Big| \tag{20}$$

where r_1 and r_2 are radius of the mask and inner radius of the annulus. Circle mask is a special kind of annulus mask, where the inner radius r_2 is equal to zero. After modifying Eq. (20) it can be expressed as

$$\mathrm{m}(s,t) = \Big|\mathrm{sign}\big(r - \sqrt{s^2+t^2}\big)\Big|. \tag{21}$$

3 Results

Methods introduced in previous section were applied to astronomical images, which are described in the Tab. 2. Results are compared from vizual point of view and also by Signal to Noise Ratio (SNR).

Table 2: Analyzed astronomical images information

image information	image name	
	20050724233135-0301.fits	2g980831.006.fits
bit depth	16	16
dimensions (px)	1530×1020	1536×1024
exposure time (s)	100	60
CCD's temperature (K)	267.92	277.36
exposure date (dd-mm-yyyy)	24-07-2005	01-09-1998

For better vizual results of denoising algorithms, there are figured only cuts of analyzed images, Figs. 2(a) and 2(b).

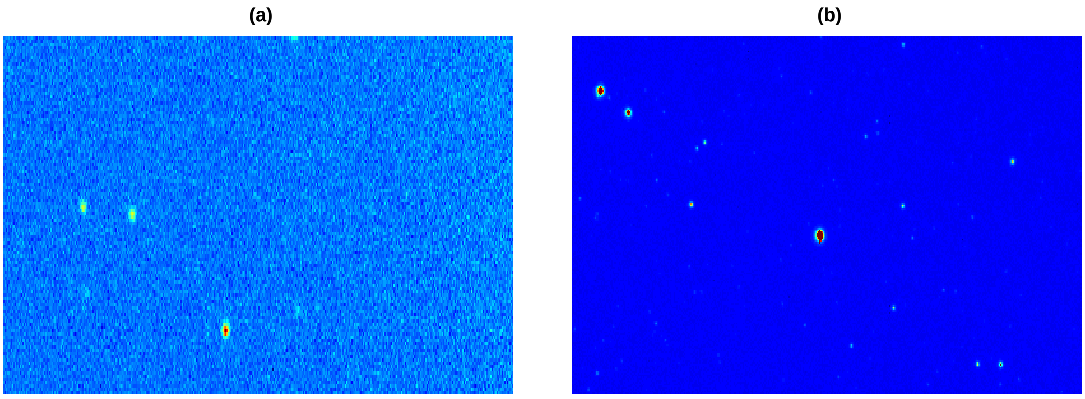

Figure 2: Cuts of original images (a) 20050724233135-0301.fits, (b) 2g980831.006.fits.

Tables 3 and 4 present results SNR of applied L-estimates on both processed images.

Table 3: SNR results - application of L-estimates with values trimming on analyzed astronomical images.

filter	mask	20050724233135-0301.fits					2g980831.006.fits				
		radius = 1 trimmed values		radius = 2 trimmed values			radius = 1 trimmed values		radius = 2 trimmed values		
		1	2	1	2	3	1	2	1	2	3
trimmed mean	square	2.1678	2.4257	1.8376	1.8705	1.9075	2.5621	2.5277	2.2020	2.1921	2.1791
	rhomb	3.1262	3.2816	2.1626	2.3037	2.5451	2.7547	2.7353	2.4099	2.3882	2.3619
winsorized mean	square	2.0457	2.0553	1.8087	1.8102	1.8132	2.5664	2.5297	2.2038	2.2033	2.2007
	rhomb	3.0520	0.7942	2.0718	2.0792	2.0913	2.7455	0.7885	2.4154	2.4024	2.3894
triangle distribution	square	2.4693	2.7608	2.0276	2.0635	2.1040	2.5420	2.5162	2.1675	2.1571	2.1473
	rhomb	3.2090	3.2816	2.4741	2.6283	2.8076	2.7601	2.7353	2.3809	2.3615	2.3428
binomial distribution	square	2.7650	2.9499	2.4040	2.4274	2.4520	2.5225	2.5102	2.1151	2.1135	2.1119
	rhomb	3.2090	3.2816	2.8276	2.8739	2.9085	2.7601	2.7353	2.3423	2.3357	2.3290

Table 4: SNR results - application of L-estimates without values trimming on analyzed astronomical images.

filter	20050724233135-0301.fits				2g980831.006.fits			
	square mask		rhomb mask		square mask		rhomb mask	
	radius = 1	radius = 2	radius = 1	radius = 2	radius = 1	radius = 2	radius = 1	radius = 2
median	3.0746	3.2816	2.5827	2.9621	2.4457	2.7353	2.0959	2.3099
BES	2.0629	3.1262	1.9889	2.0485	2.5158	2.7547	2.1644	2.3899
WBES	1.9649	2.7006	1.8907	2.2241	2.5607	2.7957	2.1939	2.4077

Chosen vizual results are figured in Figs. 3(a) and 3(b).

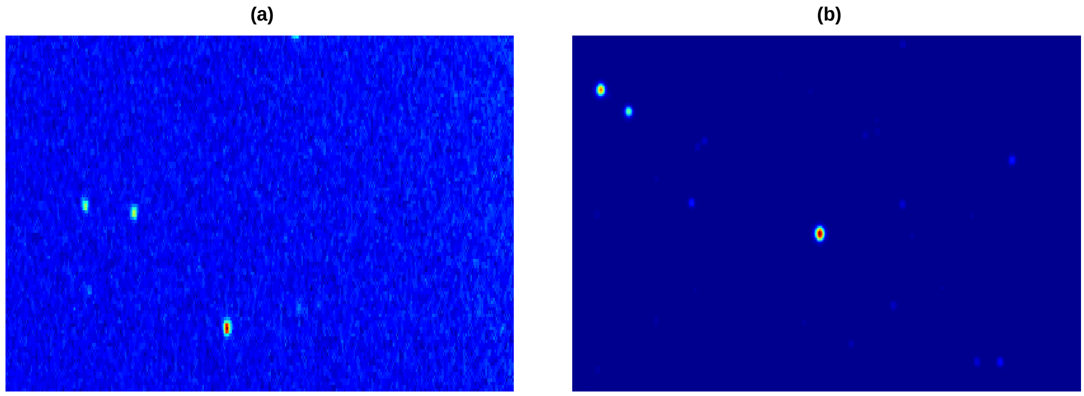

Figure 3: Cuts of filtered images by L-estimates (a) 20050724233135-0301.fits - binomial distributed weights with rhomb mask, radius r = 1 and two trimmed values, (b) 2g980831.006.fits - WBES with rhomb mask, radius r = 1

From the SNR results can be seen that the best results were obtained by triangle or binomial distributed weights or by median. From the vizual point of view does not matter which L-estimate was used, because vizual results are almost comparable.

Table 5 presents SNR results of M-estimates application. Figured results are shown in Figs. 4(a) and 4(b).

Table 5: SNR results - application of M-estimates on analyzed astronomical images.

filter	20050724233135-0301.fits				2g980831.006.fits			
	radius = 1, mask		radius = 2, mask		radius = 1, mask		radius = 2, mask	
	square	rhomb	square	rhomb	square	rhomb	square	rhomb
Andrews	1.8748	1.7372	2.0626	1.8815	2.0022	1.8514	2.1000	1.9433
Bisquare	1.8785	1.7400	2.0682	1.8853	2.0072	1.8554	2.1063	1.9475
Cauchy	1.7003	1.6021	1.8210	1.6998	1.7747	1.6987	1.8329	1.7461
Fair	0.6640	0.6167	0.7271	0.6444	0.6715	0.6532	0.7211	0.6601
Huber	2.0638	1.8822	2.3253	2.0882	2.2898	2.0588	2.4077	2.2016
Logistic	0.9900	1.1225	0.8253	1.0576	1.1543	1.3925	0.9193	1.2652
Welsch	1.9998	1.8153	2.2757	2.0155	2.2661	2.0348	2.3991	2.1743

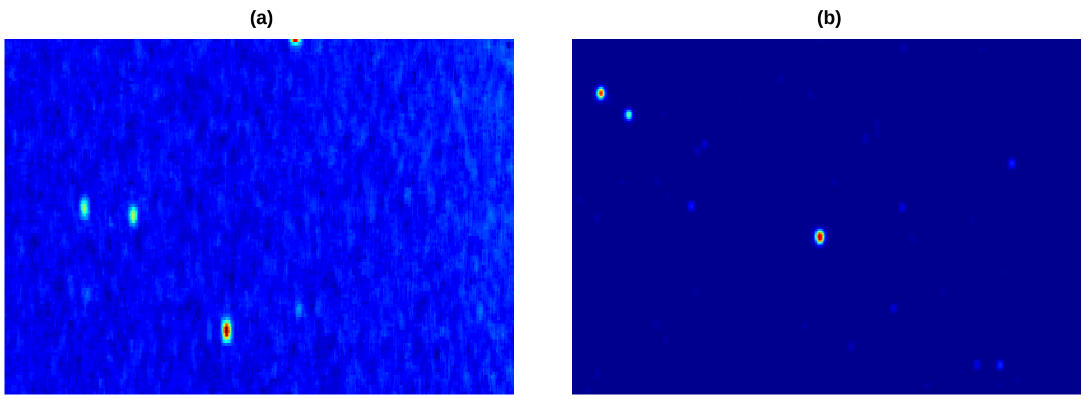

Figure 4: Cuts of filtered images by M-estimates (a) 20050724233135-0301.fits - values adjusted by Welsch weight function with square mask and radius r = 2, (b) 2g980831.006.fits - values adjusted by Huber weight function with square mask and radius r = 2.

Best results from the point of SNR were given by Huber weight function applied to the processed pixels in selected region of analyzed images with square mask. In the case of M-estimates, we can see that values of SNR are lower than in the case of L estimates. M-estimates have to use masks with higher radius. Worse SNR may be given by the fact, that the outliers are not trimmed (have non-zero weights) as L-estimates do.

4 Conclusion

The main goal of this paper was suppression of Dark Current, which is thermally generated noise occuring in CCD sensors. For this purpose were used methods of mathematical statistics, based on robust estimation. There were applied L and M-estimates to the real astronomical data. In processing of analyzed images was also discussed selection of pixels in the neigbourhood of the processed one. This was realized by two types of masks with different radiuses. Results of noise suppression were then compared from vizual point of view and also by SNR characteristic.

From the SNR results can be said that L-estimates are more suitable than M-estimates. When we compare chosen vizual results of applied filters, there can not be seen such a big different between these two estimation methods. The higher value of SNR in the case of L-estimates may explained that L-estimates trim specific number of values from processed vector, but M-estimates have usually non-zero weights given by the influence function. If we compare influence of used masks and its radius, then we can say that the shape of used does not influence results so musch as the radius. Radius may be limiting factor in loss of information, because the higher the radius is the higher the loss of information can be. This is unwanted effect for further data processing, so the used radius should be the smallest one.

Acknowledgement

This work has been supported by the financial support from specific university research MSMT No 21/2012, research project MSM 6046137306 of the Ministry of Education, Youth and Sports of the Czech Republic and by the Czech Grant Agency under grants No. 205/09/1302 "Study of sporadic meteors and weak meteor showers using automatic video intensifiers cameras", No. P102/10/1320 "Research and modeling of advanced methods of image quality evaluation" and by the research program MSM 6840770014 "Research in the Field of Information and Communication Technology" and SGS11/165/OHK4/3T/14 CTU in Prague.

References

[1] Acharya T., Ajoy K. R.: *Image Processing: Principles and Applications.* Jonh Wiley & Sons Inc., U.S.A., 2005.

[2] Gonzalez C. R., Woods E. R.: *Digital Image Processing.* Prentice Hall, U.S.A., second edition, 2002.

[3] Nakamura J.: *Image sensors and signal processing for digital still cameras.* CRC Press (Taylor & Francis), U.S.A, first edition, 2005.

[4] Buil C.: *CCD Astronomy: Construction and Use of an Astronomical CCD Camera.* Willmann-Bell, U.S.A, first edition, 1991.

[5] Kňazovická L., Švihlík J. Dark Current Elimination in Charged Coupled Devices, *International Conference Technical Computing*, November 2010, pp. 067/1-7.

[6] Wikipedia EN: *Robust statistics*, 2006 - 2012 - [cite 20th August 2012]. Online: http://en.wikipedia.org/wiki/Robust_statistics.

[7] Wikipedia EN: *L-estimator*, 2006 - 2012 - [cite 20th August 2012]. Online: http://en.wikipedia.org/wiki/L-estimator.

[8] Wikipedia EN: *M-estimator*, 2006 - 2012 - [cite 20th August 2012]. Online: http://en.wikipedia.org/wiki/M-estimator.

[9] Zhang Z: *M-estimators*, 1996 - 2012 - [cite 20th August 2012]. Online: http://research.microsoft.com/en-us/um/people/zhang/inria/publis/tutorial-estim/node24.html.

[10] Vetterling W. T., Flannery B. P., Press W. H., Teukolsky S. A.: *Numerical recepies in C.* Cambrigde University Press, Cambridge, U.K., third edition, 2007.

[11] Tukey J. W.: *Exploratory Data Analysis.* Addison Wesley, U.S.A., first edition, 1977.

[12] Weisstein, E. W.: *Triangular Distribution*, 1999 - 2012 - [cite 21st August 2012]. Online: http://mathworld.wolfram.com/TriangularDistribution.html.

[13] Pavlík J. a kol.: *Aplikovaná Statistika.* ICT, Prague, CZ, first edition, 2005.

[14] Rejta M., Matouš O.: *Nelineární filtrace 2D obrazu v Matlabu.* semestral project, ICT, Prague, CZ, 2008.

František Mojžíš
Department of Computing and Control Engineering
Institute of Chemical Technology in Prague
Technická 5, 166 28 Prague 6, Czech republic
E-mail: frantisek.mojzis@vscht.cz

Jan Švihlík
Department of Computing and Control Engineering
Institute of Chemical Technology in Prague
Technická 5, 166 28 Prague 6, Czech republic
E-mail: jan.svihlik@vscht.cz

Jaromír Kukal
Faculty of Nuclear Sciences and Physical Engineering, CTU in Prague
Department of Software Engineering in Economics
Trojanova 13, 120 00 Prague 2, Czech Republic
E-mail: jaromir.kukal@fjfi.cvut.cz