Some simple algorithms for detecting anomalous bright pixels

Alessandro Gentilini

March 2, 2014

Abstract

I describe some simple algorithms I am using for detecting anomalous bright pixels in images taken by a laptop webcam. The webcam lens is covered with black tape and so the visible light is not collected by the sensor; the grey levels different from zero in the image are usually due to the so called *dark current*; sometimes there are bright pixels (and I classify the image as an *event*) and an hypothesis is that the bright pixels are the result of the interaction of cosmic ray muons with the semiconductor sensor of the webcam.

1 Algorithms

I am experimenting various algorithms in order to classify an image as an *event*.

I have a sequence of images, let I_i be the i-th image in the sequence. The image has R rows and C columns. Let $I_i(r,c)$ be the grey value of the pixel at $row = r$ and $column = c$ in the image I_i. The program computes M_i as

$$M_i = \max_{\substack{0 \le r \le R-1 \\ 0 \le c \le C-1}} I_i(r,c)$$

So M_i is the maximum grey level in the image I_i.

The simplest algorithm uses a fixed threshold t and the image I_i is classified as an event if

$$M_i > t \tag{1}$$

For example formula 1 is used in `dkirkby/cosmic`[1] with an additional filter stage with the aim to filter out the so-called *hot pixels*. If the maximum grey level M_i happens at the same pixel position (r_H, c_H) more than a certain number of times then the pixel at (r_H, c_H) is classified as an hot pixel and any following images with maximum at (r_H, c_H) is discarded[2].

A different algorithm considers the average grey level avg_i of the image I_i and the standard deviation sd_i of the grey levels of the image I_i, the image I_i is then classified as an event if

$$M_i > avg_i + n \cdot sd_i \tag{2}$$

In a third algorithm, the program keeps running statistics for M_i, in particular $\overline{M}_i$ is the mean of the maximum grey levels and it is computed as

$$\overline{M}_i = \frac{1}{i}\sum_{k=1}^{i} M_k$$

and σ_{Mi} is the standard deviation of the maximum grey level and it is computed as

$$\sigma_{Mi} = \sqrt{\frac{1}{i(i-1)}\left(i\sum_{k=1}^{i} {M_k}^2 - \left(\sum_{k=1}^{i} M_k\right)^2\right)}$$

The image I_i is then considered an event if

$$M_i > \overline{M}_i + n \cdot \sigma_{Mi} \tag{3}$$

2 Data

Figures 2 to 14 show data collected in various days. The time between two images acquisition has a mean value of 0.13s and a standard deviation of 8ms. The exposure time is unknown. The images were collected with a laptop webcam, the location is in Northern Italy, the time is local.

[1] *Cosmic ray detector for iOS* available at `https://github.com/dkirkby/cosmic`

[2] See `https://github.com/dkirkby/cosmic/blob/master/Cosmic/CosmicBrain.m` accessed March 2, 2014, where the number of time is `MAX_REPEATS` and the threshold t is `MIN_INTENSITY`.

Figure 1: A crop from an image classified as an event using formula 3 with $n = 10$.

The graphs should be considered as a work in progress, in particular I have some issues with them:

1. The range for M_i in Figures 2 and 3 seems quite different with respect to the remaining Figures.
2. The curve for $\overline{M}_i$ seems too low in Figure 2 and 3.
3. The curve for $\overline{M}_i$ does not properly follow the trend of M_i in Figures 6 and 14.

The above issues could have been caused by some errors in the program collecting the data, I should investigate further.

The summary for the hourly rate event is the following:

```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.700   1.355   1.840   1.991   2.468   4.460
```

This table summarize the data collected so far.

events	elapsed time	events per hour
9	03h14m18s	2.78
12	08h03m03s	1.49
23	09h27m57s	2.43
15	06h37m57s	2.26
6	08h37m23s	0.70
10	08h06m30s	1.23
8	09h05m33s	0.88
22	08h44m34s	2.52
17	08h44m33s	1.94
7	05h21m33s	1.31
12	07h16m38s	1.65
11	06h19m07s	1.74
20	08h04m50s	2.48

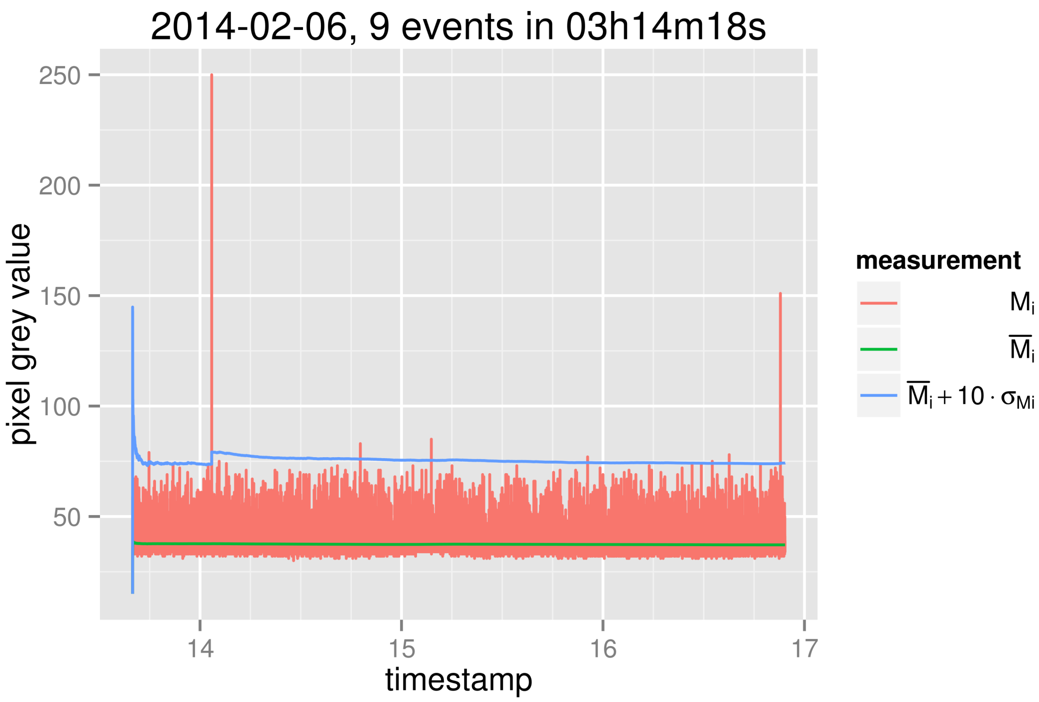

Figure 2: Data collected using formula 3 with $n = 10$.

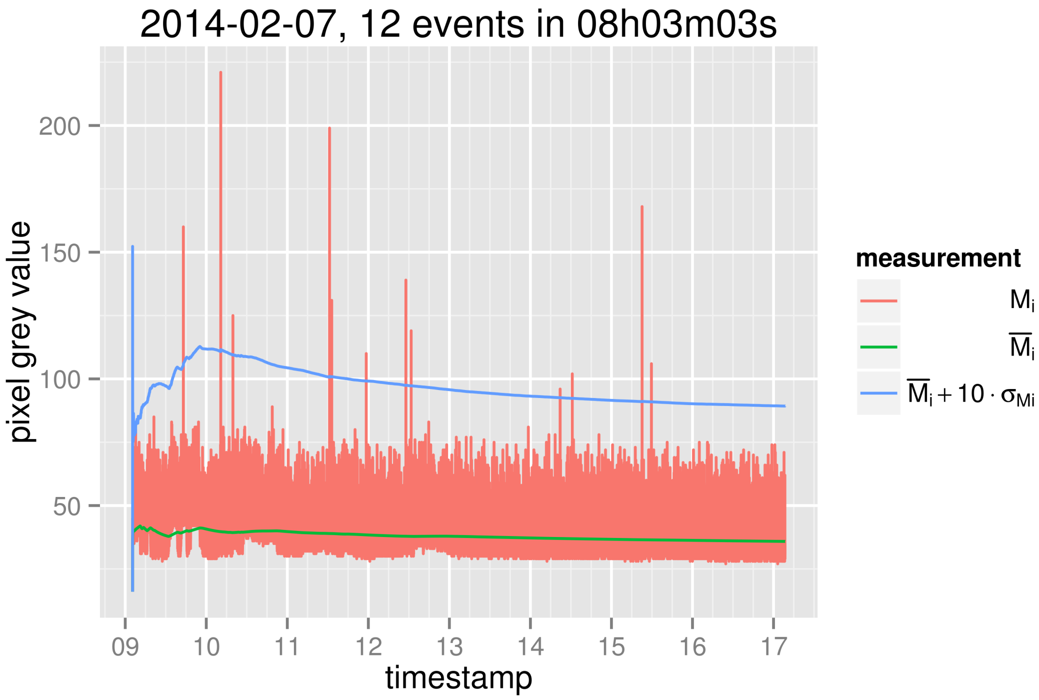

Figure 3: Data collected using formula 3 with $n = 10$.

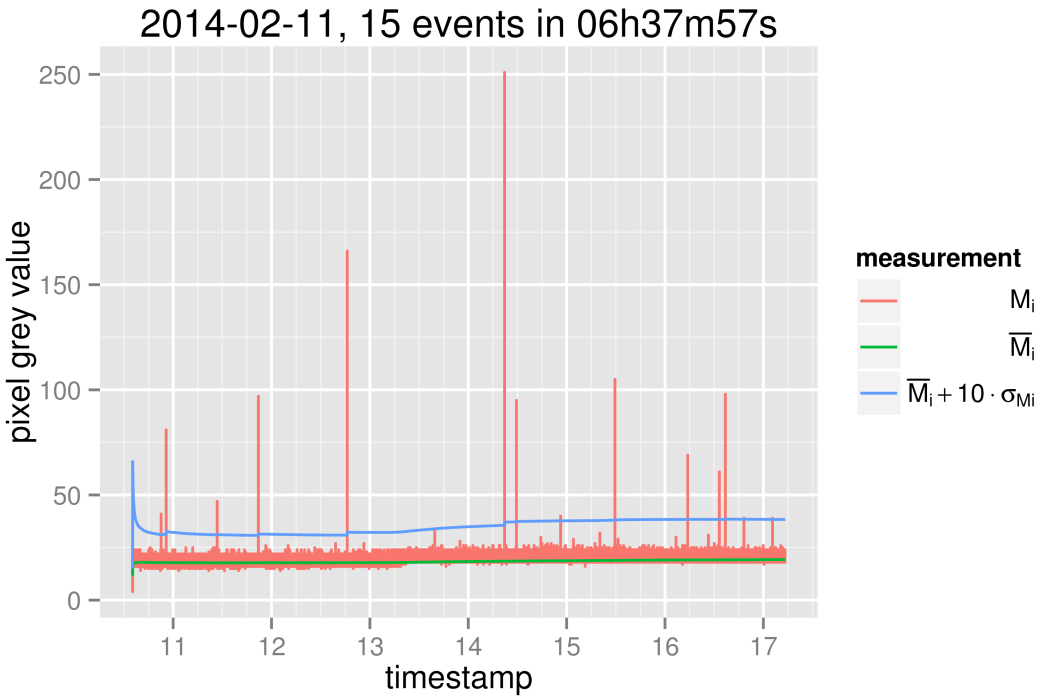

Figure 4: Data collected using formula 3 with $n = 10$.

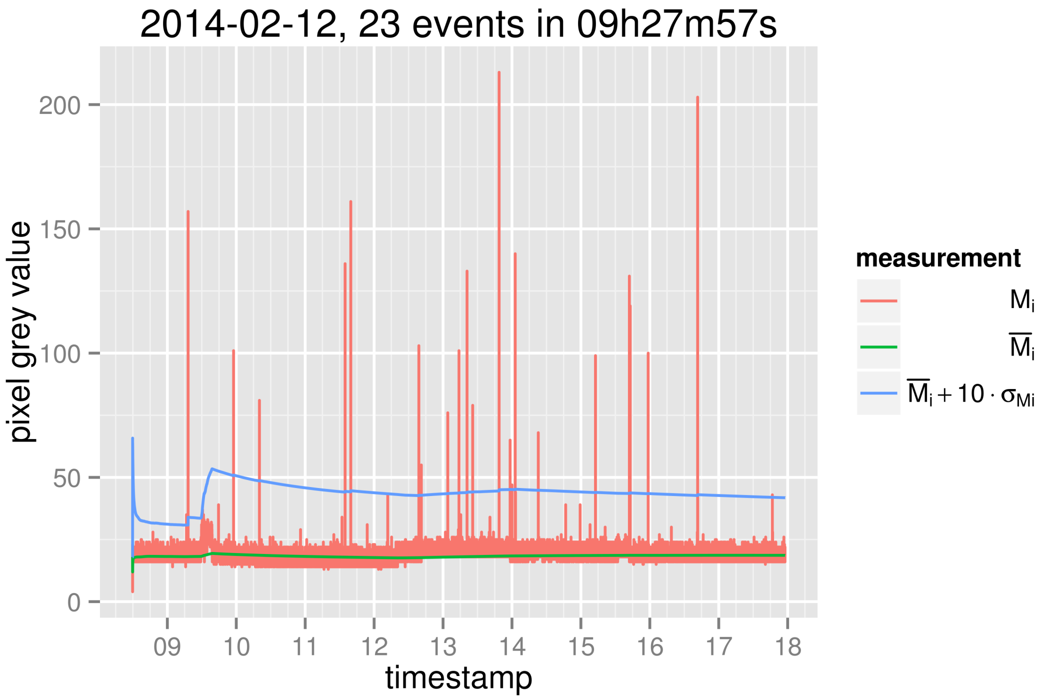

Figure 5: Data collected using formula 3 with $n = 10$.

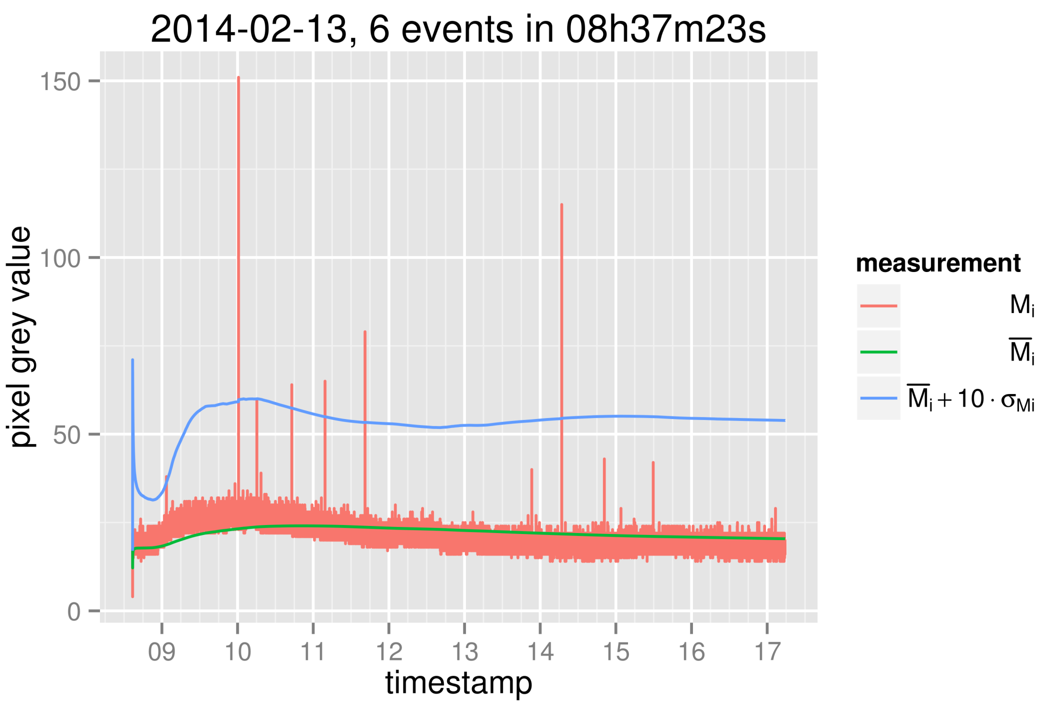

Figure 6: Data collected using formula 3 with $n = 10$.

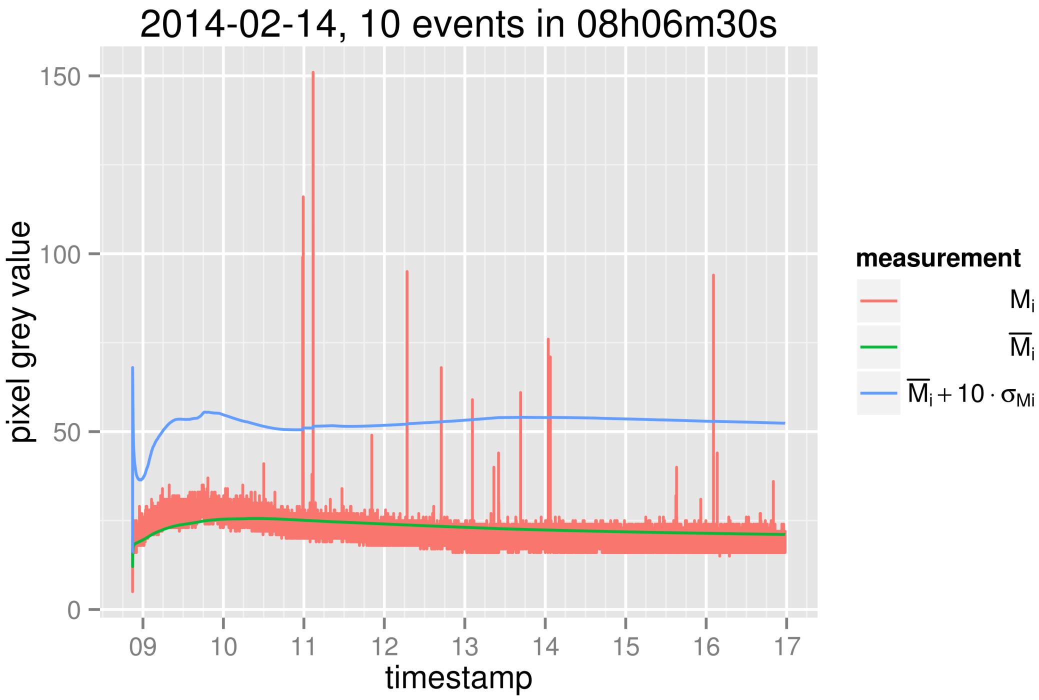

Figure 7: Data collected using formula 3 with $n = 10$.

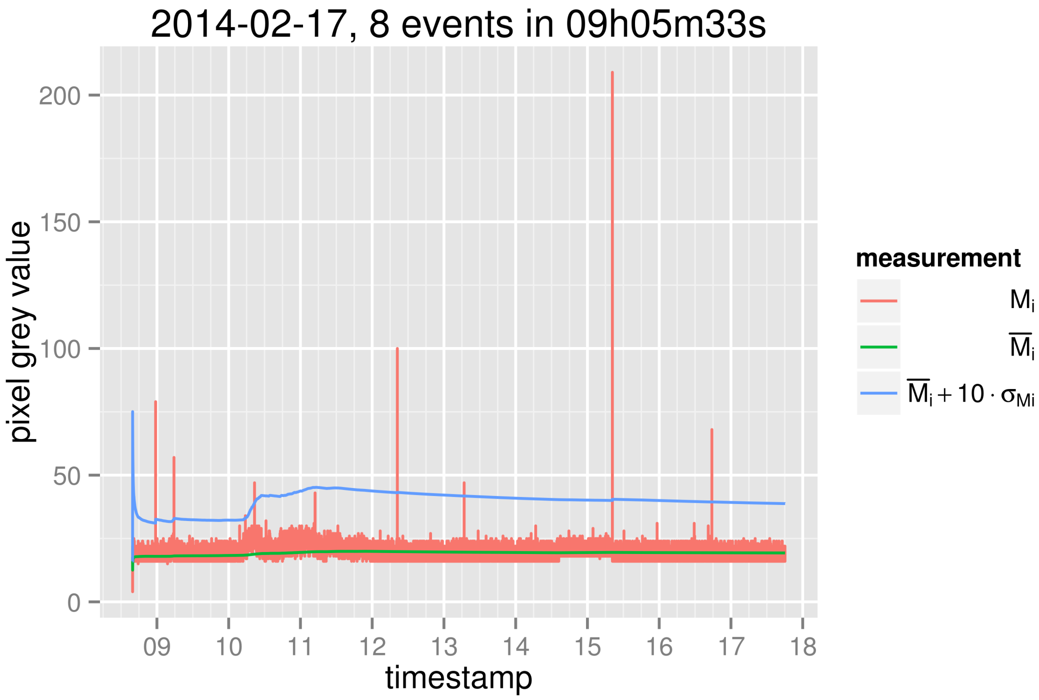

Figure 8: Data collected using formula 3 with $n = 10$.

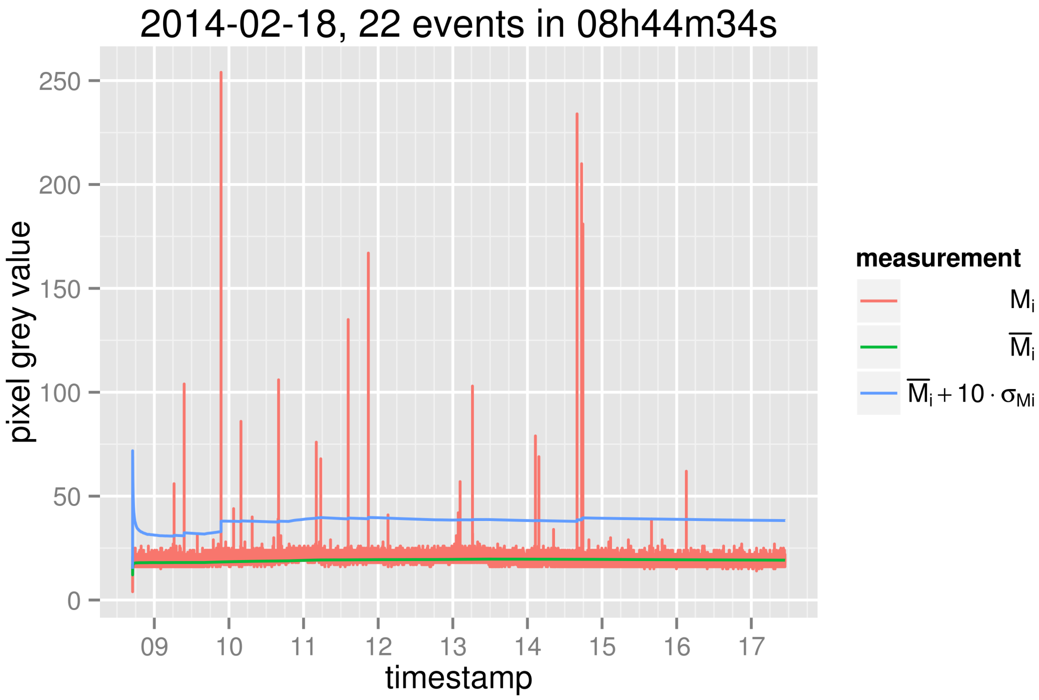

Figure 9: Data collected using formula 3 with $n = 10$.

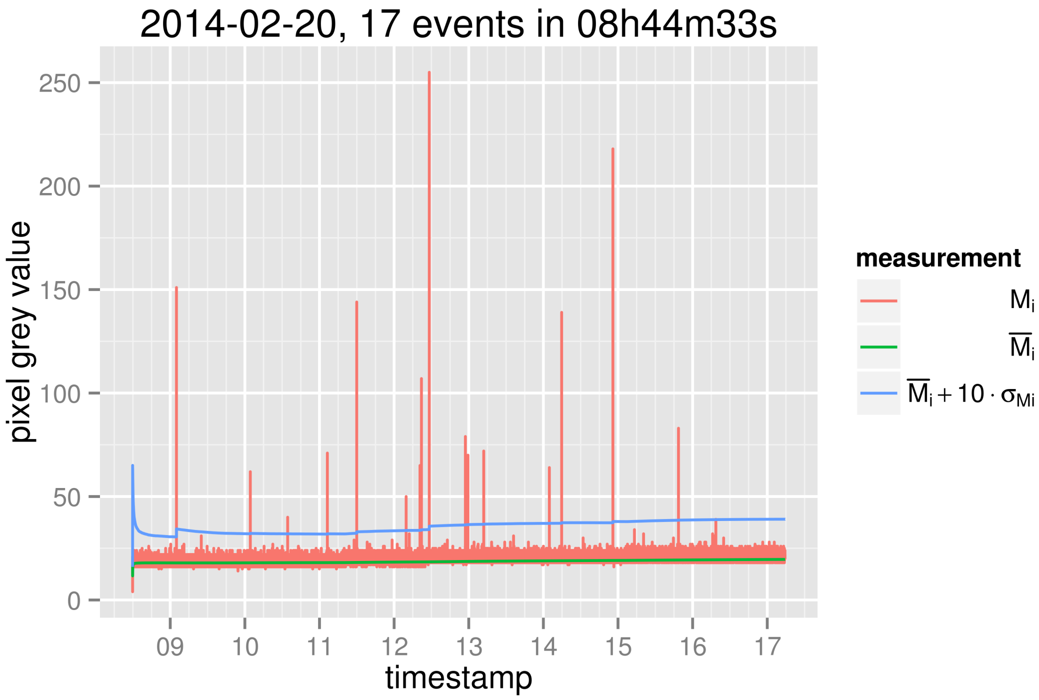

Figure 10: Data collected using formula 3 with $n = 10$.

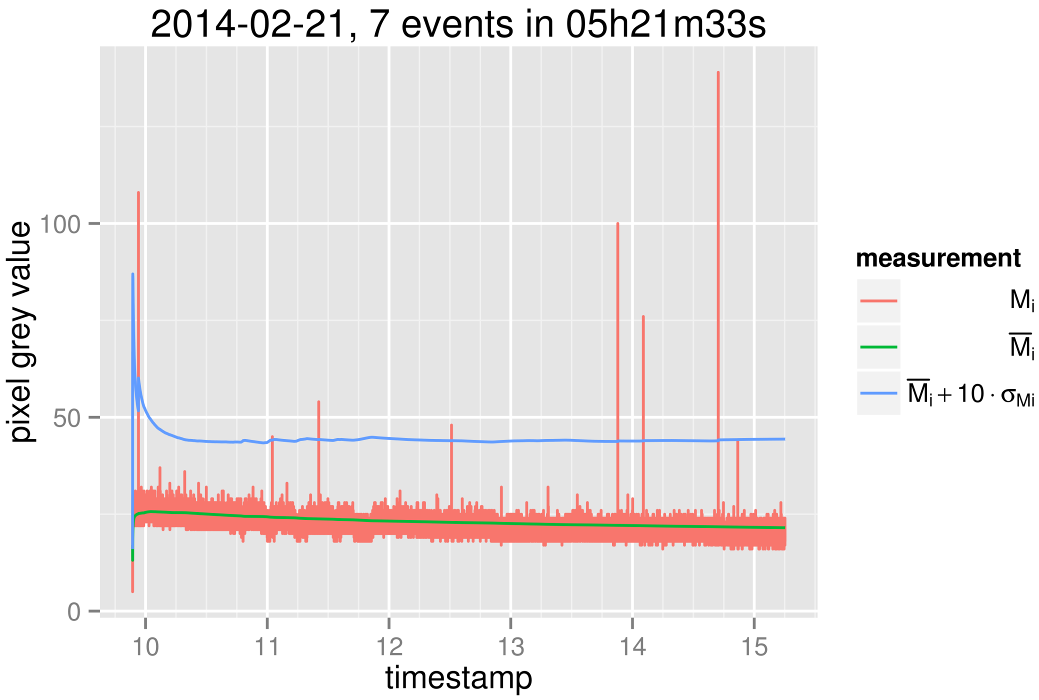

Figure 11: Data collected using formula 3 with $n = 10$.

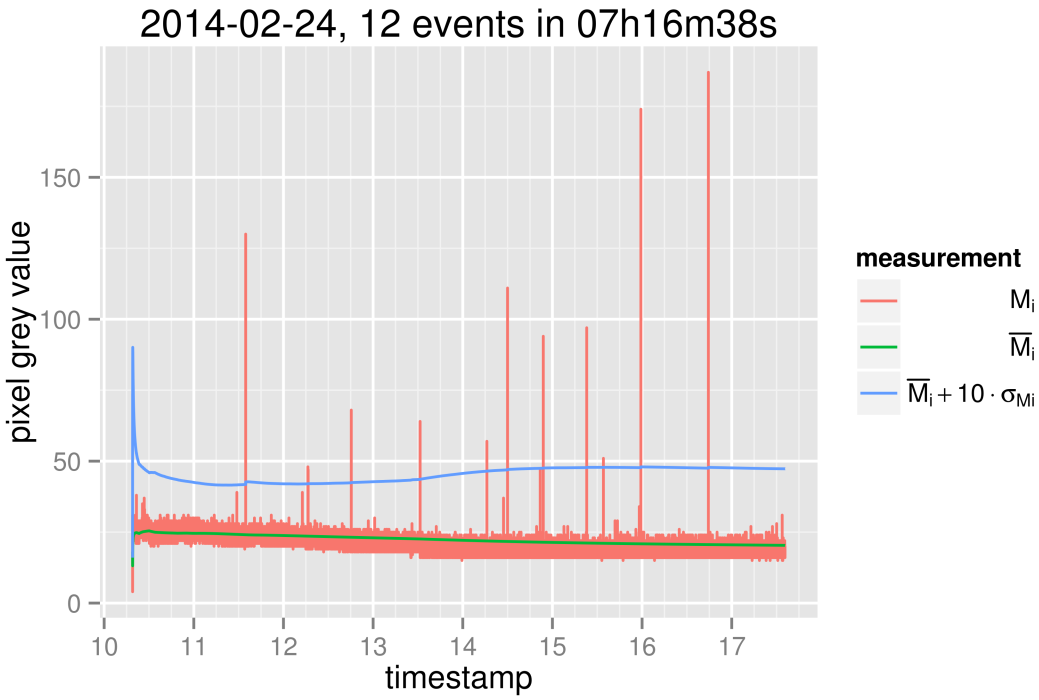

Figure 12: Data collected using formula 3 with $n = 10$.

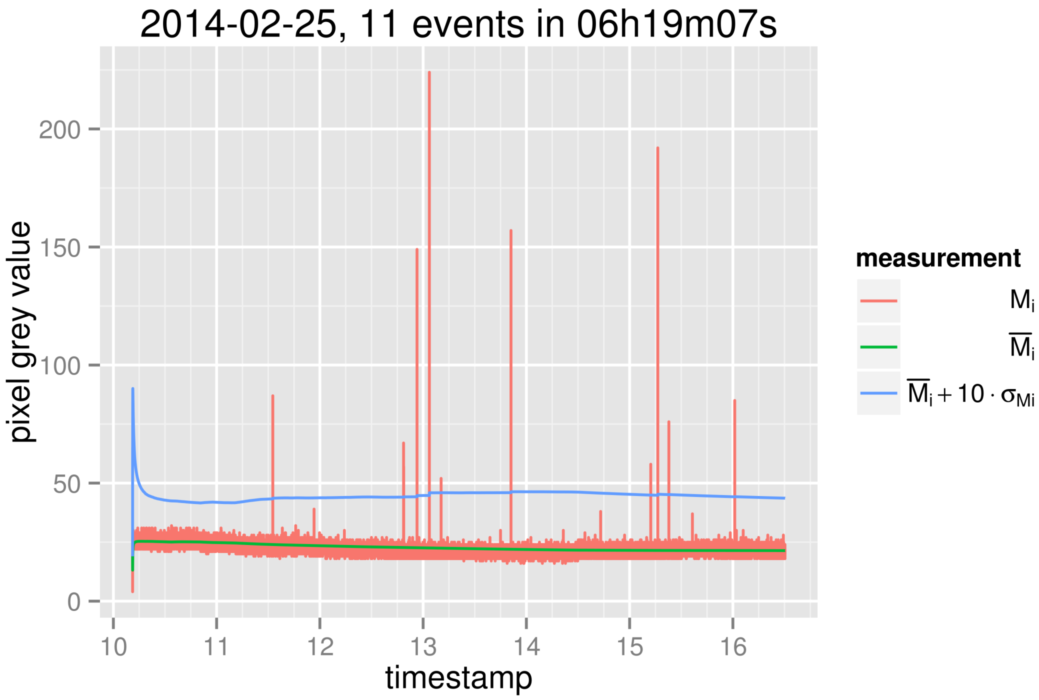

Figure 13: Data collected using formula 3 with $n = 10$.

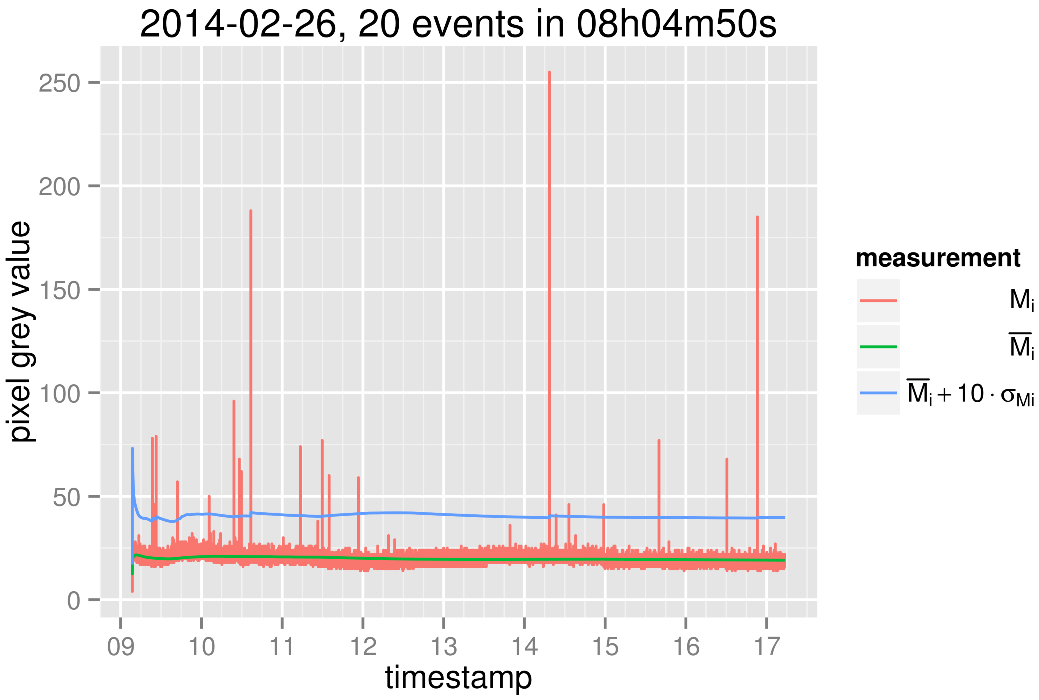

Figure 14: Data collected using formula 3 with $n = 10$.