Solar Flare Retrieval, Detection and Analysis

Tomáš Suk
Institute of Information Theory and Automation
The Czech Academy of Sciences
182 08 Praha 8, Czech Republic
Email: suk@utia.cas.cz

Stanislava Šimberová
Astronomical Institute
The Czech Academy of Sciences
251 65 Ondřejov, Czech Republic
Email: ssimbero@asu.cas.cz

***Abstract*—We propose methodology for the analysis of active regions on the Sun. It is based on high-order statistical moments of image histogram, particularly on its skewness. The methodology includes a new technique to select regions with possible formation of the flare. We track these areas in the time video sequences to search for triggers – turning point(s), when the pre-flare phase changes to the full developed event. The frequency analysis in the flare formation areas is based on Fourier analysis, Morlet wavelets, and principal component analysis. Results have been evaluated to achieve the periods of oscillations stimulating the flare origin.**

I. Introduction

While solar physics has been studied for many decades, many details about phenomena on the Sun are still not known. Explosive events defined by Dere et al. [1] originating from the outer layers of the Sun are embodied in solar activity. Monitoring of solar activity is one prerequisite for space weather prediction. The term "space weather" refers to conditions on the Sun, and its atmosphere - magnetosphere, ionosphere and thermosphere that can influence the performance and reliability of technological systems (space and ground based) and that can affect human life. An overview of these phenomena with all the consequences for the Earth system is given in [2].

Explosive events release energy in various forms and in various spectral ranges - from the shortest γ-rays, hard X-rays, EUV, visible, up to micro and radio waves. In our case we are interested in processing solar observations in the visible area of the spectra, namely in the wavelength 656.3 nm, called H_α line. The chromospheric flares (eruptions) are well observable in this spectral line by the ground based telescopes. We introduce the methodology of active region prediction, searching and location with emphasis on pre-flare phase. Analysis and determination of trigger mechanisms of the flares can significantly contribute to space weather prediction. The part of active region searching is based on statistical moments of high order computed in the observed sequence of images. Their unusual behavior in possible active regions we firstly noticed in [3] analyzing signs of flare arising in H_α high resolution observation and in X-ray range. Following steps of procedure target frequency analysis in the pre-flare phase and principal component analysis (PCA) using high-order moments.

The results of the frequency analysis can be compared with other observations: three- and five-minute oscillations in radio band 17 GHz [4], in various bands (optical, Hα, UV, microwave, soft X-rays and high-cadence longitudinal magnetograms) [5], in X-ray band [6] and in EUV (17.1 nm) [7]. Quasi-periodic pulsations are studied also in [8]. A detailed list of relevant references can be found in [9].

Our article is organized as follows: The section of flare searching covers the mathematical background of high order statistical moments and their application to the temporal image sequence of solar disk observation. In Section III we describe the method of image analysis regarding pre-flare time interval: searching for the flare trigger point and frequency analysis in this area based on Fourier transformation (FT), wavelets and PCA. Experiments on the real data and results are discussed in Conclusion.

II. Searching Solar Flares

In projection of the whole solar disk we can observe a few active regions at the same time. We proposed a new procedure based on statistical moments, how to search active regions in a sequence of images of the complete solar disk. Its brief flowchart can be seen in Fig. 1.

A. Statistical Moments

If we consider the brightness in an image to be a random variable, then the image histogram is an estimate of its density. The density of a random variable X can be described by its moments. The normalized moments are defined

$$m_n = E(X - EX)^n / S^n, \tag{1}$$

where n is the order of the moment and S is standard deviation. Moment m_3 is called skewness and m_4 kurtosis. Moment m_5 is sometimes called hyperskewness and m_6 hyperflatness.

B. Whole Solar Disk Data

We tested our operating procedures on the data from the Kanzelhöhe Observatory. They capture the whole solar disk in images of size 2048×2048 pixels with 12-bit data, i.e. we can distinguish up to 4096 gray levels in the images. Quiet state on the solar surface (no range is active) requires basic operational modus, only one image in ten minutes is recorded. When an active region appears, the cadence of the capturing grows up to one image per 6 seconds. One such interesting sequence was captured on October 26, 2015. The processed section of complete observation contains 1067 images, time interval from 8:18:02 UT to 13:30:39 UT (5h 12min 37s), the

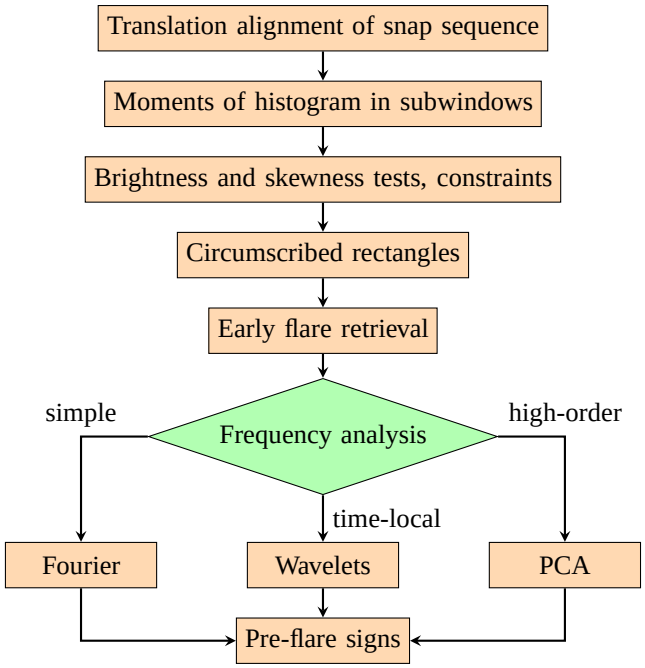


Fig. 1. Abridged flowchart of the proposed procedure.

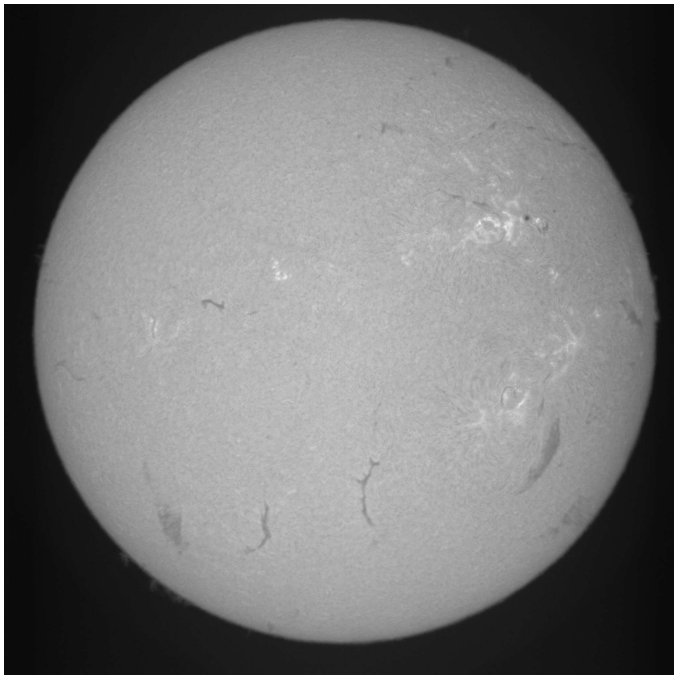


Fig. 2. The solar disk observed in H_α line (wavelength 656.3 nm), observatory Kanzelhöhe, Austria.

first snap is shown in Fig. 2. The minimum interval between the adjacent snaps is 5 s, the maximum interval is 147 s.

C. Skewness Method

The parts of an image outside the solar disk, i.e. the solar atmosphere (the dark background here), can violate the

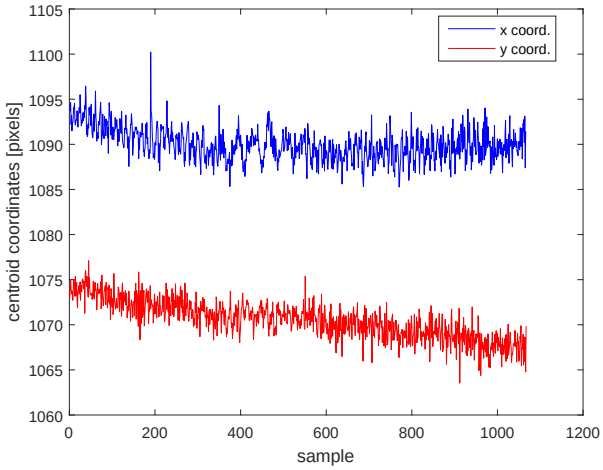


Fig. 3. The centroid coordinates through the sequence.

computation of the moments. For this reason we use the brightness threshold 256 as a mask. The less brightness parts of the image are excluded from the computation.

1) Translations among Images: There are translations among the individual images, we align them by means of centroid. We can look at a whole image as an instance of two-dimensional (2D) random field. Its moments M_{pq} (called geometric) are defined in [10]. The centroid coordinates are then $c_x = M_{10}/M_{00}$ and $c_y = M_{01}/M_{00}$. In the case of our sequence, the centroid coordinates are in Fig. 3. The mean values are $c_x = 1090.00$ and $c_y = 1070.54$. We translated each image so its centroid is mapped into this mean position. The maximum deviations were 10 pixels horizontally and 7 pixels vertically. The nearest-neighbor interpolation preserving the brightness values in the image has been applied.

2) Individual Images: The image is divided into non-overlapping windows; the skewness of histogram in each of them is computed. To optimize the window size, various values have been tested. Too small a window does not include the whole event and too big a window can connect more events together. The suitable sizes in pixels are between 64×64 and 192×192. In this sequence, the size 96×96 yielded the best results. We also tested overlapping windows, but connection of individual overlapping parts of the windows into one region is complicated and ultimately unreliable.

We distinguish three levels of activity in a window. If the minimum brightness is less than 369, we suppose there is a sunspot and we assign level 1 to the window. If the skewness m_3 is higher then 0.5, we suppose there is some pre-flare activity and assign the level 2. If $m_3 > 1$, we suppose there is a starting flare and assign the level 3. An example is in Fig. 4. The histograms are normalized by the maximum gray level 4096 and by the number of pixels. The pixels inside the solar disk boundary (not removed by the mask) are only calculated to provide the correct computation on the boundary of the disk.

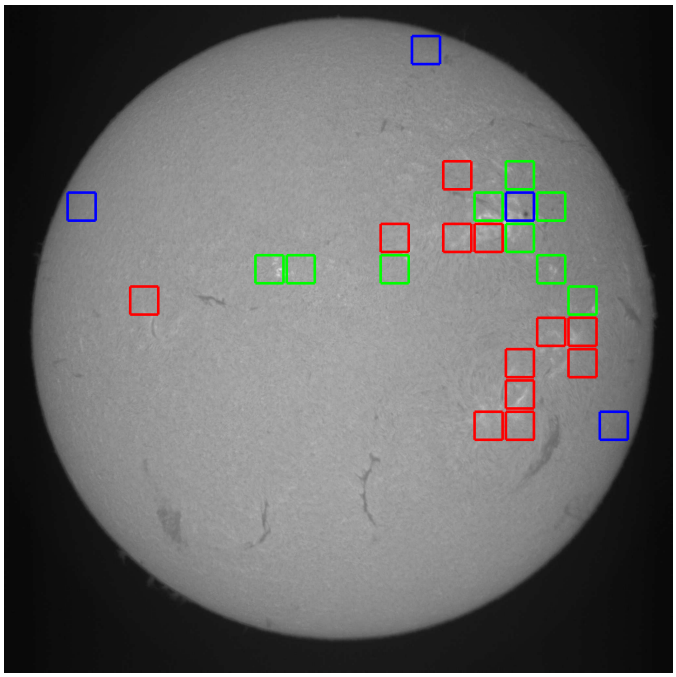


Fig. 4. The detected possible active particular windows. The windows with level 1 are marked blue, level 2 red and level 3 green. The snap number 196 from 09:23:01 UT.

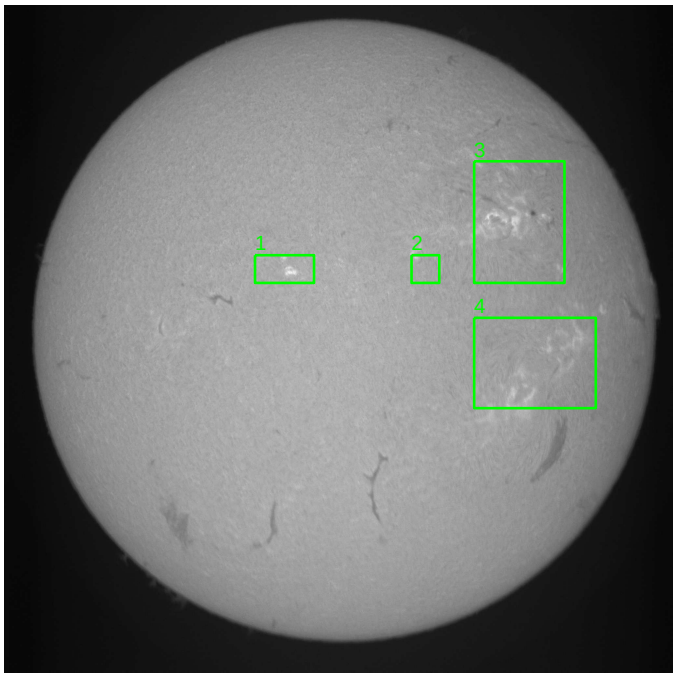


Fig. 5. The full-disk observation with discovered possible active regions. Observatory Kanzelhöhe, 20151026, resolution 1 pix = 1.02 × 1.02 arcsec, total length of record 12h 1min 38s.

3) Whole Sequence: We observe the score of the activity levels of the individual windows throughout the whole sequence. When their sum divided by three (maximum level) exceeds 50% of the number of images in the sequence, we consider the window as active. Then the adjacent active windows are connected and the circumscribed rectangle highlights the active region. Finally, we recompute the moments of these interconnected regions.

The regions found in the experimental sequence are in Fig. 5. In the central part of the region 3, a sunspot (dark stain) can be seen. The detail of the region 4 with maximum flare is in Fig. 6. The time curves of the skewness in the found regions are in Fig. 7.

III. Analysis of Solar Flares

We can distinguish a few phases of solar flare: inactive state, pre-flare, developing state with its maximum and post-flare. An interesting task is the search for the point when pre-flare phase changes to the developing state. Another interesting task is frequency analysis of oscillations, particularly in the pre-flare phase.

A. Search of Flare Beginning

The temporal evolution of a flare at the beginning manifests itself by fast accelerating brightness. Mathematically, it is typical for such points that the third derivative of a function crosses zero; we call it the turning point. We recall in this context that the point at which a function crosses zero is called zero-crossing, where its first derivative crosses zero is called extreme and where its second derivative crosses zero is called

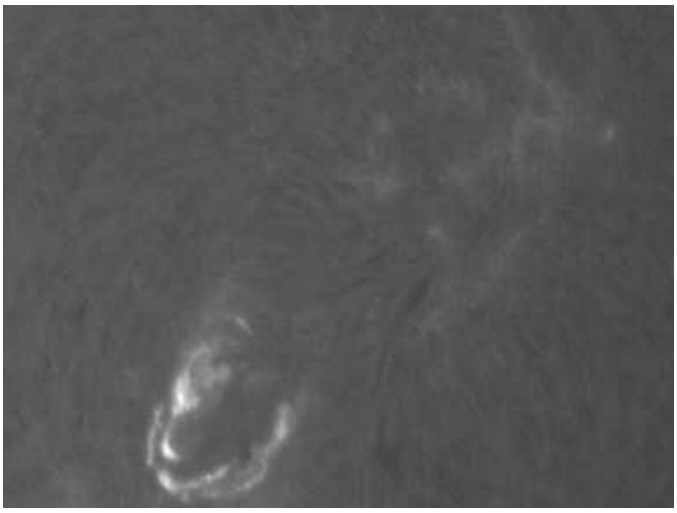


Fig. 6. The detail of active region 4 in the phase of flare maximum (skewness maximum at 10:28:50 UT).

inflection. If $f(x, y, t)$ is brightness at the point (x, y) at the time t, we search such time t_p that

$$f'''(x, y, t_p) = \left.\frac{\partial^3 f(x, y, t)}{\partial t^3}\right|_{t=t_p} = 0, \qquad (2)$$

$f'''(x, y, t) > 0$ for $t < t_p$ and $f'''(x, y, t) < 0$ for $t > t_p$, i.e. the second derivative $f''(x, y, t)$ has local maximum. The method based on these rules was published in [3]. Instead of the light curve at one point, we use the skewness in the detected region. Since it is very noisy and we search only

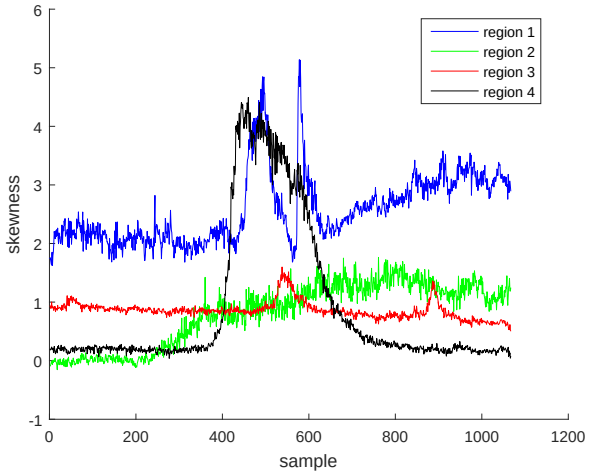


Fig. 7. The skewness in the four found regions.

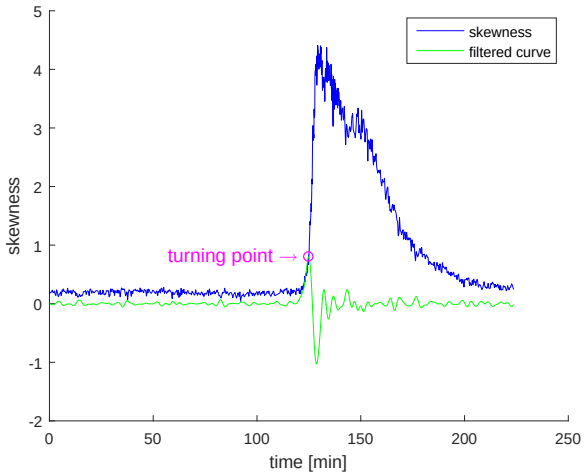


Fig. 8. The turning point found in region 4.

significant maxima, it is suitable to combine the computation of the second derivative with filtering by the Gaussian. The combined filter is

$$g_i = \left(\frac{i^2}{\sigma_g^2} - 1\right) \frac{1}{\sigma_g^3\sqrt{2\pi}} e^{-\frac{i^2}{2\sigma_g^2}}, \quad i = -n_g, -n_g + 1, \ldots n_g. \tag{3}$$

The filter half-length n_g should be four times longer than the standard deviation of the filter σ_g, we suggest $n_g = 60$ and $\sigma_g = 15$.

The sequence of the measurements should be re-sampled to equidistant time points prior the filtering. At this moment the post-flare phase is no longer relevant and we can shorten the tail of the sequence. The first 800 samples are interpolated to the sequence of 2684 samples by 5 seconds. In the case of region 4 (Fig. 8), the turning point is at the absolute maximum of the filtered signal, whereas region 1 with double flare (Fig. 9) evinces two turning points.

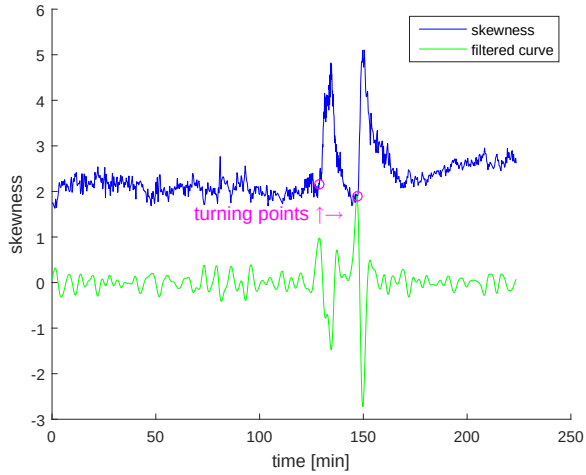


Fig. 9. The turning points found in region 1.

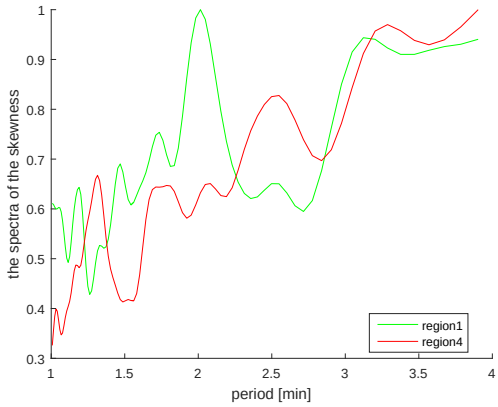


Fig. 10. The amplitude spectrum of the skewness plotted with respect to the periods.

B. Frequency Analysis

The trigger of the flares is incited by typical oscillations. The simplest way how to study the oscillations is FT of the signal.

1) Fourier Transformation of the Skewness: We computed amplitude Fourier spectrum of the skewness in the active regions. An appropriate filter is applied to remove noise. The spectrum filtered by Gaussian with the standard deviation 2.4 is in Fig. 10. From the theoretical point of view, the most interesting periods are between 1 minute and 4 minutes, therefore we display this interval.

We have found important oscillations that can be compared with theoretical results, see Tab. I. The maxima being too close to the others are not considered.

2) Wavelet Analysis: If we need to analyze the spectra of the pre-flare oscillations in more detail, the wavelets suggest themselves. They offer a more local investigation of the signal. Continuous wavelets are suitable for this purpose, e.g. Morlet wavelets [11], whose kernel function is product of Fourier and

TABLE I
SIGNIFICANT LOCAL MAXIMA OF THE AMPLITUDE FOURIER SPECTRUM OF THE SKEWNESS

region	1st maximum	2nd maximum	3rd maximum
1	3 min 7 s	2 min 1 s	1 min 28 s
4	3 min 17 s	2 min 33 s	1 min 19 s

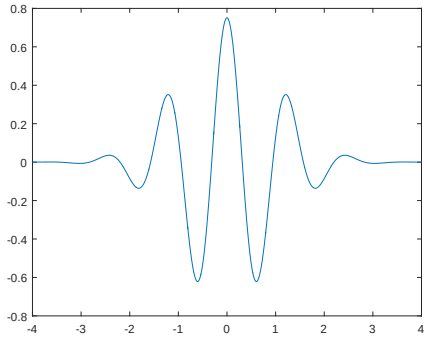


Fig. 11. The real part of the Morlet wavelet in its effective support [-4,4].

Gaussian components. They are defined

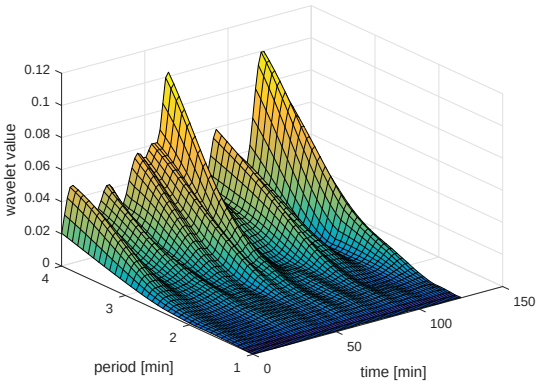
$$s_m(t,u) = \int_{-\infty}^{\infty} \frac{1}{\sqrt[4]{\pi}} e^{i\sigma_m u\tau} e^{-\frac{(u\tau)^2}{2}} f(t+\tau)\mathrm{d}\tau, \qquad (4)$$

where $i = \sqrt{-1}$, u is frequency (period is $1/u$), $s_m(t,u)$ is local Morlet spectrum at the time t, σ_m is ratio of the frequency and the standard deviation of the Gaussian component. The value $\sigma_m = 5$ yields good compromise between time and frequency resolution, see Fig. 11.

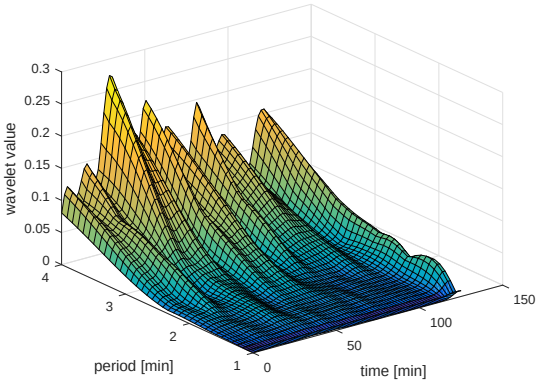
We analyzed the skewness curves from regions 1 and 4 by the Morlet wavelets by the method from [12]. We took only pre-flare phase (samples from 1 to 1500, i.e. the first 2 hours 5 minutes of the sequence). The parts of the local spectra for periods from 1 minute to 4 minutes can be seen in Fig. 12. The spectra are filtered by Gaussian with the standard deviation 1.8 samples. If we take a cross-section at the fixed time, we obtain the local spectrum. It can be compared with the global spectrum from Fig. 10.

3) Principal Component Analysis: High-order moments m_4, m_5, ... can also be used for the analysis of an event, but we must consider high correlations between them. We can reduce the correlations by PCA. It is based on such linear transformation of the measurement sequences that minimizes absolute value of their pair-wise correlations. The details can be found in [13]. FT is performed on the principal components instead of the moments.

In our case, we computed PCA of the moments from m_3 to m_8 separately in regions 1 and 4. The eigenvalues of the covariance matrices are in Tab. II. The eigenvalues drop quickly, therefore only the first few components are suitable for analysis. We did not use moment orders higher than eight, because now we can consider the covariance matrix to be positive definite. The spectra of the first three components are in Fig. 13, their significant maxima are in Tab. III.



(a)



(b)

Fig. 12. The real part of the amplitude Morlet spectrum of the skewness in region (a) 4, (b) 1.

TABLE II
EIGENVALUES OF THE COVARIANCE MATRICES IN PCA

1	2	3	4	5	6
5.85	0.141	0.00415	$1.36 \cdot 10^{-4}$	$9.17 \cdot 10^{-6}$	$3.75 \cdot 10^{-7}$
5.78	0.207	0.0108	$6.19 \cdot 10^{-4}$	$1.27 \cdot 10^{-5}$	$5.49 \cdot 10^{-7}$

TABLE III
SIGNIFICANT LOCAL MAXIMA OF THE AMPLITUDE FOURIER SPECTRUM OF THE FIRST PRINCIPAL COMPONENT

region	1st maximum	2nd maximum	3rd maximum
1	3 min 48 s	2 min 59 s	1 min 36 s
4	3 min 26 s	2 min 29 s	1 min 16 s

The PCA spectrum comparison with the skewness spectrum in Fig. 10 and Tab. I shows that both position and relevance are slightly moved. In the case of region 1, while the maximum at 2 min 1 s (now 1 min 58 s) is less relevant, there is new significant maximum at 3 min 48 s. In the case of region 4, the changes are less striking, the most significant one being higher relevance of the maxima at 1 min 19 s (now 1 min 16 s). The conclusions of the second component analysis are similar, while the spectrum of the third component is completely

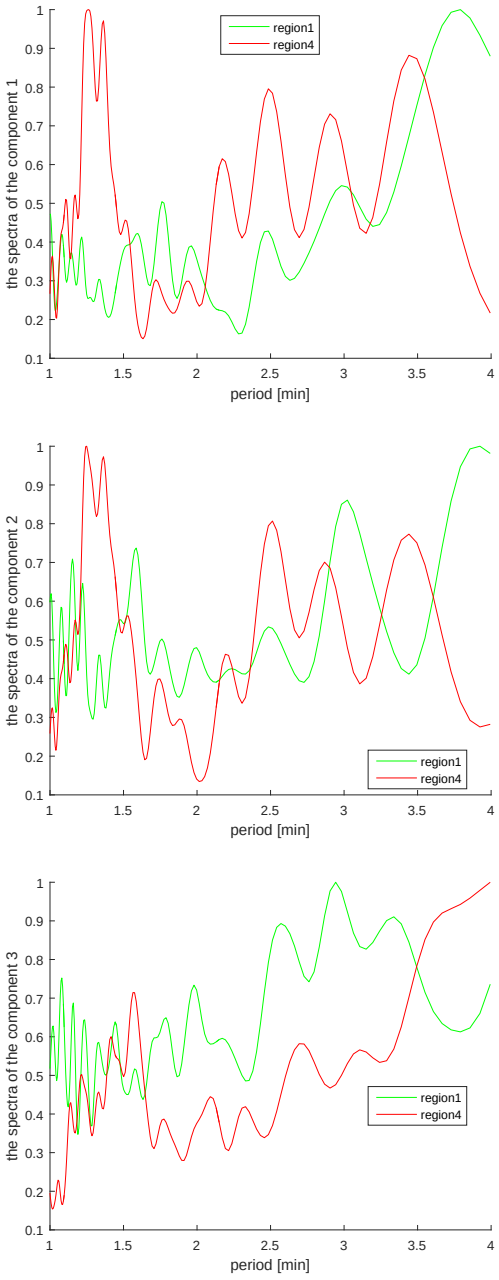


Fig. 13. The amplitude spectrum of the (a) first, (b) second and (c) third principal component.

different, these residual oscillations are not relevant.

IV. Conclusion

We suggested a new technique for processing and analyzing image data obtained from the ground based observation of the Sun. In our contribution we focused on active region searching and analysis of those events in terms of possible-flare-formation. Partial steps of the proposed procedures were experimentally verified on the real data. Data analysis is based on high-order statistical moments of subimage histogram, particularly on its skewness. It includes search for and identification of the turning point at which the pre-flare phase changes to the flare itself, or the frequency analysis of solar plasma oscillations that attend the flares.

The existing methods are based on the analysis of so-called "light curves", i.e., brightness changes in given points or areas over time. Unfortunately, light curves provide no information on pre-flare signs even if they are used for data analysis from various parts of the electromagnetic spectrum. We have analyzed the soft X-ray data from the GOES satellite [3] with the result that there is no turning point and the area for the study of oscillations leading to the flare formation cannot be delineated.

Our frequency analysis in the pre-flare area is based on Fourier analysis, Morlet wavelets or principal component analysis of different-order moments. The results are compared to achieve the main periods in the analyzed event. In our knowledge, nobody has used this method for analysis of solar flares. We expect applications in other fields that utilize analysis of long-term data sequences (medicine, chemistry, biology, remote sensing, etc.).

Acknowledgments

This research was supported by Czech Science Foundation (GAČR) Grant No. GA15-16928S and The Czech Academy of Sciences (AV ČR) Grant No. RVO:67985815.

H-alpha data were provided by the Kanzelhöhe Observatory, University of Graz, Austria.

References

[1] K. P. Dere, J.-D. F. Bartoe, and G. E. Brueckner, "Explosive events in the solar transition zone," *Solar Physics*, vol. 123, no. 1, pp. 41–68, 1989.

[2] R. Schwenn, "Space weather: The solar perspective," *Living Reviews in Solar Physics*, vol. 3, no. 2, pp. 1–72, 2006.

[3] S. Šimberová, M. Karlický, and T. Suk, "Statistical moments of active-region images during solar flares," *Solar Physics*, vol. 289, no. 1, pp. 193–209, 2014.

[4] V. E. Abramov-Maximov, G. B. Gelfreikh, and K. Shibasaki, "Quasi-periodic oscillations of solar active regions in connection with their flare activity NoRH observations," *Solar Physics*, vol. 273, no. 2, pp. 403–412, 2011.

[5] J. Li, D. L. Mickey, and B. J. LaBonte, "The X3 flare of 2002 July 15," *The Astrophysical Journal*, vol. 620, no. 2, pp. 1092–1100, Feb. 2005.

[6] I. De Moortel, A. W. Hood, J. Ireland, and R. W. Walsh, "Longitudinal intensity oscillations in coronal loops observed with TRACE," *Solar Physics*, vol. 209, no. 1, pp. 89–108, Sep. 2002.

[7] B. De Pontieu, R. Erdélyi, and I. De Moortel, "How to channel photospheric oscillations into the corona," *The Astrophysical Journal*, vol. 624, no. 1, pp. L61–L64, May 2005.

[8] V. M. Nakariakov, A. R. Inglis, I. V. Zimovets, C. Foullon, E. Verwichte, R. Sych, and I. N. Myagkova, "Oscillatory processes in solar flares," *Plasma Physics and Controlled Fusion*, vol. 52, p. 7, 2010.

[9] T. Wang, "Standing slow-modewaves in hot coronal loops: Observations, modeling, and coronal seismology," *Space Science Reviews*, vol. 158, no. 2–4, pp. 397–419, Jul. 2011.

[10] J. Flusser, T. Suk, and B. Zitová, *Moments and Moment Invariants in Pattern Recognition*. Chichester: Wiley, 2009.

[11] H. Li, "Complex Morlet wavelet amplitude and phase map based bearing fault diagnosis," in *WCICA*, 2010, pp. 6923–6926.

[12] S. Šimberová and T. Suk, "Analysis of dynamic processes by statistical moments of high orders," in *CIARP*, ser. LNCS, vol. 8258. Springer, 2013, pp. 221–228.

[13] ——, "Dynamic process analysis by moments of extreme orders," *Astronomy and Computing*, vol. 14, no. 1, pp. 43–51, 2016.