Analysis of Dynamic Processes
by Statistical Moments of High Orders

Stanislava Šimberová[1] and Tomáš Suk[2,⋆]

[1] Astronomical Institute, Academy of Sciences of the Czech Republic,
Ondřejov, Czech Republic
[2] Institute of Information Theory and Automation, Academy of Sciences
of the Czech Republic, Prague 8, Czech Republic
ssimbero@asu.cas.cz, suk@utia.cas.cz
http://asu.cas.cz, http://utia.cas.cz

Abstract. We present a new approach to image analysis in temporal sequence of images (data cube). Our method is based on high-order statistical moments (skewness and kurtosis) giving interesting information about a dynamic event in the temporal sequence. The moments enable precise determination of the "turning points" in the temporal sequence of images. The moment's curves are analyzed by continuous complex Morlet wavelet that leads to the description of quasi-periodic processes in the investigated event as a time sequence of local spectra. These local spectra are compared with Fourier spectrum. We experimentally illustrate the performance on the real data from astronomical observations.

Keywords: Statistical moments, Frequency analysis, Fourier and wavelet transformations, Dynamic processes.

1   Introduction

Random variables can be characterized for application purposes by consideration of quantities called "moments". Since simple and widely known statistical moments about the origin - EX (mean value μ, the first order moment) via the central moments of second order $E(X - EX)^2$ (variance σ^2) and its square root $S = \sqrt{E(X - EX)^2}$ (standard deviation σ), we ascend to the third and higher orders. In pattern recognition these moments are used as the regional descriptors for structural shape of regions, boundary determination, texture analysis, etc. The practical use of moments in statistics is e.g. in [1].

The third order moment m_3 is called the "skewness" of the distribution of random variable X. It is defined: $m_3 = E(X - EX)^3/S^3$. The skewness is a measure of reflection symmetry, i.e. if the distribution of X is symmetric, then $m_3 = 0$. If the distribution is steeped in left (right) of μ, it is denoted skewed to the right (left), respectively. The fourth moment called "kurtosis" is defined

similarly: $m_4 = E(X - EX)^4/S^4$. It expresses the peakedness, in other words it is a measure of how "heavy" the tails of distribution are. The distribution is said to be leptokurtic (narrower than Gaussian), platykurtic (broader than Gaussian), or mesokurtic (as Gaussian). The Gaussian distribution has $m_4 = 3$, therefore the value $\hat{m}_4 = m_4 - 3$ is sometimes used.

Each plane of data cube in the temporal sequence of images consists of pixels of different brightnesses. It can be statistically described by their distribution as well as by the high-order statistical moments. These moments have applications in signal and image processing. A detailed study of the statistical moments in pattern recognition is in [2]. In astronomical applications they are used for example in the investigation of statistical system behavior, mass-density field distribution, statistical studies of the interstellar medium [3], [4]. Pattern analysis of cosmic structure formation is in [5], statistical modeling of lines in atomic spectra [6], detection of non-Gaussianity deviations [7], etc. The generalized spectral-kurtosis estimator and its statistics is in [8] and [9]. Another type of statistical moments applied in the UV spectral range was described in [10].

Our contribution deals with an analysis of the dynamical temporal sequences obtained by the ground-based astronomical observations in optical range. By our methodology we reliably identify the "turning point" where the dynamic event starts. Determination of this point leads to the specification of temporal intervals for further analysis. In these selected sections the periodicity of signals has been searched and results by Fourier and wavelet analysis have been compared. The next section introduces the typical behavior of moments during an observed temporal sequence where a dynamical event appears. The following sections present results of frequency analysis and conclusions.

2   Dynamical Event Diagnostic by Statistical Moments

An example of the observed temporal sequence is in Fig. 1. It is monospectral observation of the Sun surface - solar chromosphere ($\lambda = 656.3$ nm).
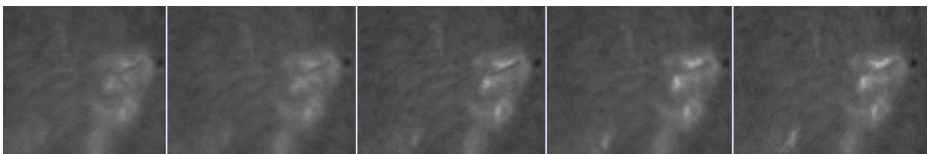

Fig. 1. Patterns of the data cube planes: the light parts in the image represent an active region with arising flare. From left to the right: the beginning of the sequence, the second pattern is from the "trigger area" and gradually up to the fully developed flare.

The dimensionless third and fourth moments are computed from the image histogram. In this sense the skewness of one image is estimated as the third moment normalized to scaling by the standard deviation

$$s = \frac{1}{\sigma^3} \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \mu)^3 \tag{1}$$

and similarly the normalized fourth moment - kurtosis

$$k = \frac{1}{\sigma^4} \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \mu)^4, \tag{2}$$

where x_i is the ith realization of the random variable X. In our case x_i is the brightness of the ith pixel of a region of interest, where N is the total number of pixels in the area. The simplified flowchart of this part of processing follows:

1. Observational sequence (data cube) of active-region images covering the time of a dynamical event (e.g. solar flares).
2. Computation of the high-order moments $[s_i, k_i]$ of each image in the whole data sequence to obtain their temporal evolution.
3. Determination of the starting point of the flare (time or corresponding plane number) and selection of the time interval for consecutive frequency analysis.

The typical evolution of m_3 and m_4 during the flare development is in Fig. 2. It is very interesting to see a fast increase of moments, the temporal curves of m_3 bears resemblance to the m_4 and both unambiguously enable determination of the starting point. As a matter of interest the temporal curve of μ (called the light curve) of the same sequence does not provide any relevant information about the position of "trigger area", see Fig. 2b. The dynamic phenomenon causes both lighter and darker regions with respect to the quiet state, the densities are averaged and the light curve cannot intercept any change of trend.
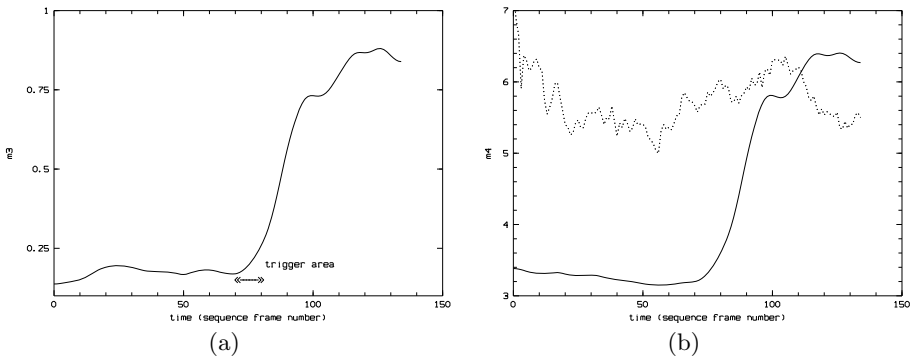

Fig. 2. Temporal characteristic of (a) skewness and (b) kurtosis during solar flare development. Temporal evolution of mean μ has been drawn for comparison (b, dotted).

3   Frequency Analysis

For an automatic searching of significant points (times) in the temporal moment's curves we decided for filtering by the Laplacian, the significant points = maxima of the Laplacian. Since the observation is often distorted by high-frequency noise, a combination of appropriate filters would be suitable. The Gaussian and Laplacian can be combined into one filter, proposed in 2D by [11]. The 1D version is

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

Application of convolution filter (3) enables an automatic identification of changes in the noisy moment curve. The absolute maximum of the Marr-filtered curve then determines the precise position of the main "fault" assigned as the flare starting time (ST). Likewise the subsidiary maxima also give useful information, especially the period of maxima. See Fig. 3a.
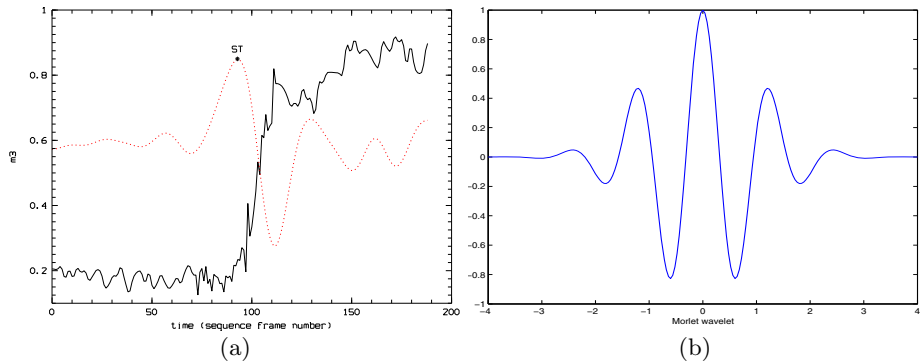

Fig. 3. (a) Significant and start (ST) point searching in the moment curve (black) and Marr filter convolution, (red dotted, mask size 59), (b) The Morlet wavelet in its effective support [-4,4].

For the frequency analysis the moment's curve can be basically divided into two parts: pre-flare time interval, i.e. the time interval before the flare start time ST (located in the trigger area and determined by an analysis of the moment evolution), and the time interval after this start time.

To get an information about the pre-flare time interval we need to analyze the frequencies of a quasi-periodic sequence. The analysis is usually done by comparison with some pattern wave that is used as the kernel function of the integral transformation. We can use either a global wave passing through the whole sequence, typically the sinusoidal signal $\exp(-2\pi\mathrm{i}x)$ of Fourier transformation, or some local wave, typically wavelet. There are several wavelet families, some continuous wavelet is suitable for this type of frequency analysis. In our

experiments we used the real part of Morlet wavelet $\cos(\sigma x)\exp(-x^2/2)$ with the most usual parameter $\sigma = 5$ that yields good compromise between time and frequency resolution, see Fig. 3b and e.g. [12].

The real data experiments consist of seven various events from the two different ground-based telescopes. We introduce two cases illustrative for each type of data, see Fig. 4a (the sampling period is 5 s) and Fig. 5a (the sampling period is 6 s). The left part of the second signal was interpolated from the sampling period 60 s to 6 s. The results of the frequency analysis by Morlet wavelets are in Figs. 4b and 5b. They show absolute value of its real part: in a row, there is the significance of a specific frequency; in a column, there is the local spectrum.

The period of the oscillation is related to the length of the sequence. If there are 210 samples in the first sequence, then the part from -0.5 to 0.5 of the wavelet with length 1 from Fig. 3b was mapped onto the whole sequence in the first row of Fig. 4b. If we are interested in the response of the whole wavelet from Fig. 3b with the length 8, we have to look at the row $210/8 \approx 26$ of Fig. 4b and $225/8 \approx 28$ of Fig. 5b.

The Fourier spectra are in Figs. 4c and 5c. To be comparable as most as possible, we use the real parts of the spectra and the same frequencies as in the case of the wavelets. The most significant maxima (i.e. those with the highest absolute value) are summarized in Tab. 1. Both times and periods are expressed in sample numbers, i.e. if the first sequence has 5 s per sample, then 140 samples represent $140 \times 5 = 700$ s. The Fourier transformation yields frequencies prevailing in the whole sequence, while the wavelet transformation yields an idea about the significant frequencies in the individual samples.

Table 1. Wavelet and Fourier analysis – sample numbers, periods (in the samples) and sizes of the most significant maxima

Sequence from Fig. 4a					Sequence from Fig. 5a				
Wavelets			Fourier		Wavelets			Fourier	
Sample	Period	Size	Period	Size	Sample	Period	Size	Period	Size
140	210	1.76	169	12.56	119	225	7.62	186	79.6
1	210	1.28	94	7.88	116	124	7.55	130	58.1
174	117	1.05	121	7.01	193	124	7.51	101	44.2
101	117	0.77	58	6.76	38	127	7.11	82	40.2
202	37	0.73	50	5.16	202	78	5.06	60	27.9
178	38	0.61	45	4.44	26	75	4.13	69	27.6
58	61	0.59	37	4.07	158	73	3.67	53	20.8
22	60	0.57	76	3.92	70	72	3.08	43	19.86

The comparison of the lowest frequencies is difficult, because the difference of the wave form over whole sequence is too significant. So, the wavelet periods 210 and 225 samples does not correspond to the Fourier maxima 169 and 186 samples. The precise wave form is less important in the higher frequencies, we can see the oscillations with period 117 samples detected by wavelets have good counterpart in the 121 samples of the Fourier spectrum in the case of the first
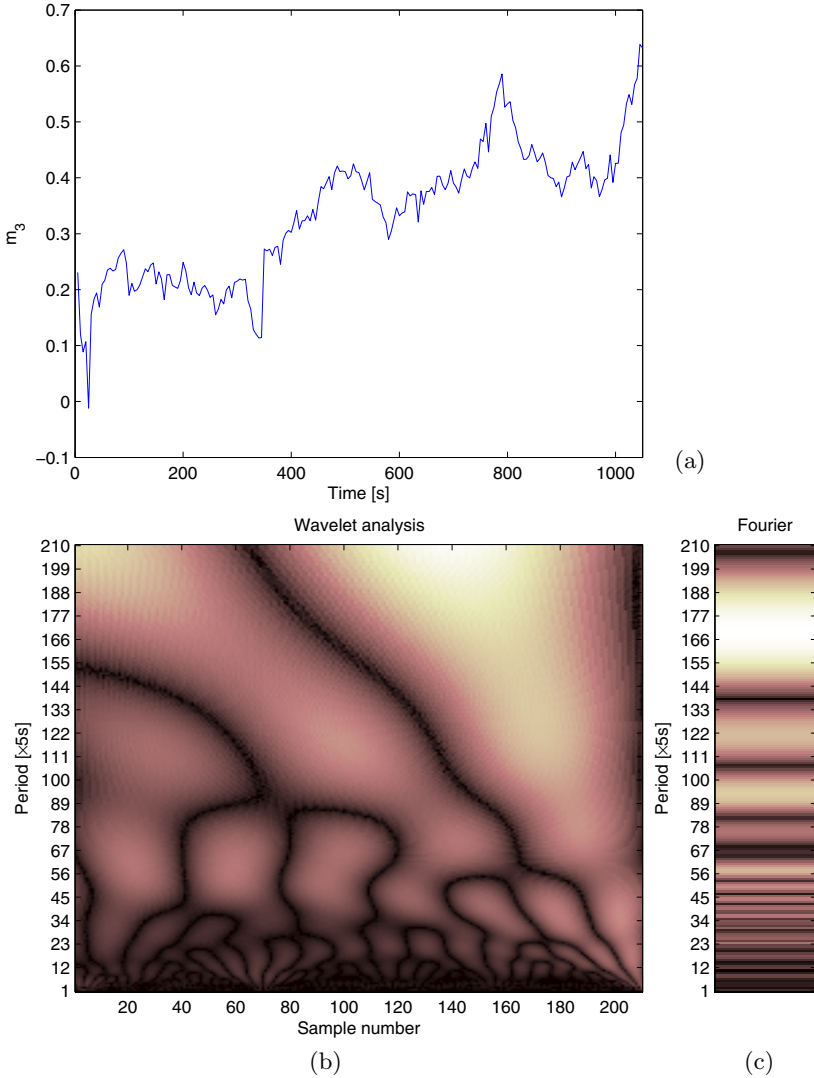
Fig. 4. The analyzed data, (a) the sequence with the sampling period 5 s, (b) the wavelet analysis and (c) Fourier amplitude spectrum

sequence. Similarly 37-38 correspond to 37 samples and 60-61 correspond to 58 samples. In the case of the second sequence, 124-127 samples correspond to 130 samples and 72-78 correspond to 69-82 samples. In both Fourier spectra, there are local maxima without direct counterpart in the wavelet analysis (94, 76, 50 and 45 samples in the first case and 101 and 60 samples in the second case, 53 and 43 samples have weak counterparts). They are not significant in any local time, while their sum over the whole sequence is significant.
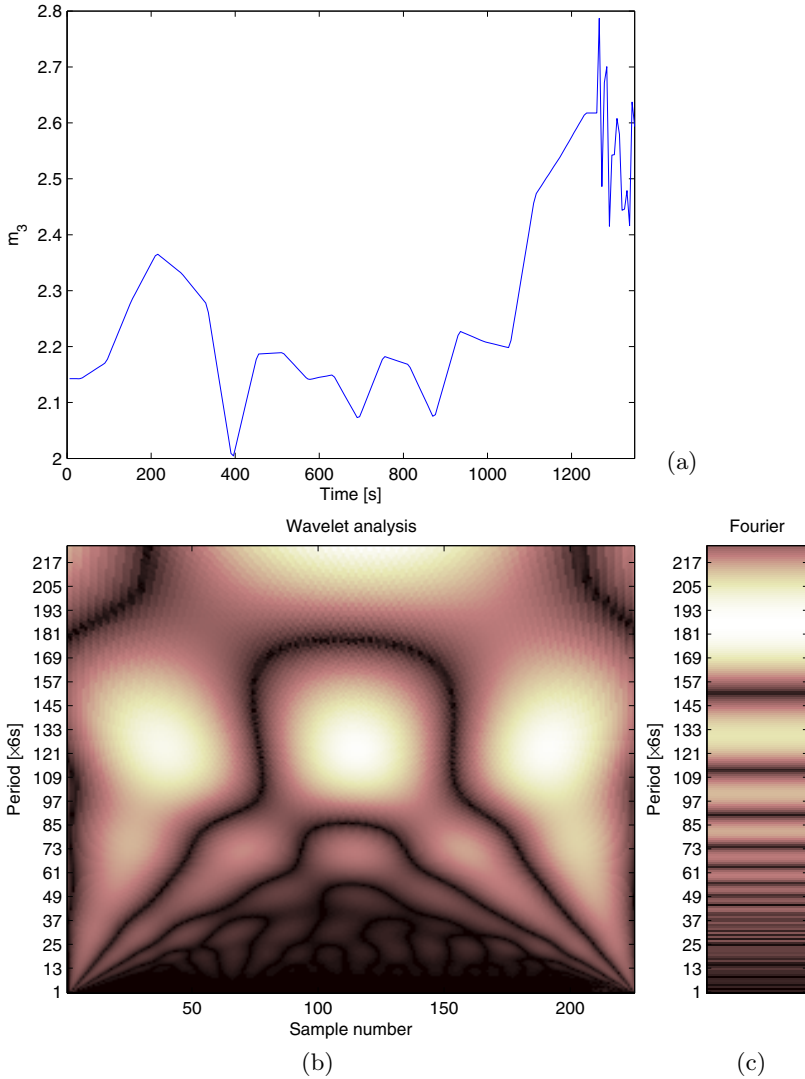
Fig. 5. The analyzed data, (a) the sequence with the sampling period 6 s, (b) the wavelet analysis and (c) Fourier amplitude spectrum

4   Conclusion

The moment curves express evolution of a dynamic process in a new way. Our experiments proved the moments of high orders are sensitive to changes in the image brightness during the initial phase. Determination of the starting point is of great importance for astrophysical interpretation as well as the oscillation analysis in the pre-flare times. It can give more information about the

mechanism of flare formation, magnetic field configurations, cosmic weather prediction, etc. Particularly, the comparison of the found frequencies from the oscillation analysis (both wavelet and Fourier) with the model of arising flare leads to its improvement.

The advantage of suggested method is its unusual robustness, it is possible to apply the method to the raw data files, neither preprocessing nor calibration is needed. For the future we intend to design special algorithms to be included into the astronomical observation pipeline. We are collecting data for other experiments in the future: temporal sequences for the meteoritic swarms searching, automatic determination of flashes of gamma lighting and applications in remote sensing.

References

1. Dudewicz, E.J., Mishra, S.N.: Modern Mathematical Statistics. Wiley and Sons, New York (1988)
2. Flusser, J., Suk, T., Zitová, B.: Moments and Moment Invariants in Pattern Recognition. Wiley, Chichester (2009)
3. Coles, P., Jones, B.: A lognormal model for the cosmological mass distribution. Monthly Notices of the Royal Astronomical Society 248, 1–13 (1991)
4. Burkhart, B., Falceta-Gonçalves, D., Kowal, G., Lazarian, A.: Density Studies of MHD Interstellar Turbulence: Statistical Moments, Correlations and Bispectrum. Astronomical Journal 693, 250–266 (2009)
5. Takada, M., Jain, B.: The kurtosis of the cosmic shear field. Monthly Notices of the Royal Astronomical Society 337, 875–894 (2002)
6. Pain, J.C., Gilleron, F., Bauche, J., Bauche-Arnoult, C.: Effect of third- and fourth-order moments on the modeling of unresolved transition arrays. High Energy Density Physics 5, 294–301 (2009)
7. Grossi, M., Branchini, E., Dolag, K., Matarrese, S., Moscardini, L.: The mass density field in simulated non-Gaussian scenarios. Monthly Notices of the Royal Astronomical Society 390, 438–446 (2008)
8. Nita, G.M., Gary, D.E.: The generalized spectral kurtosis estimator. Monthly Notices of the Royal Astronomical Society 406, L60–L64 (2010)
9. Nita, G.M., Gary, D.E.: Statistics of the Spectral Kurtosis Estimator. Publications of the Astronomical Society 122, 595–607 (2010)
10. Alipour, N., Safari, H., Innes, D.E.: An Automatic Detection Method for Extreme-ultraviolet Dimmings Associated with Small-scale Eruption. Astronomical Journal 746, 12 (2012)
11. Marr, D., Hildreth, E.: Theory of edge detection. Proceedings of the Royal Society of London. Series B, Biological Sciences 207(1167), 187–217 (1980)
12. Li, H.: Complex Morlet wavelet amplitude and phase map based bearing fault diagnosis. In: Proceedings of the 8th World Congress on Intelligent Control and Automation, pp. 6923–6926. IEEE (July 2010)