Filtering Digital Profile Observations

Mohsen Mostafa Hassan
Saray el Kobba, P. O. Box 85, Cairo 11712, Egypt

ABSTRACT: The elimination of noise in digital terrain profile data requires the analysis of these data in the frequency domain. Data processing in the frequency domain involves inferences regarding the periodogram interpretation. A periodogram is an estimate of the spectral density function of a given set of data. This estimate describes the frequency properties of these data. Knowing the frequency characteristics of random errors, it is possible to discriminate them from the "true" elevations of a terrain profile. The presented method of filtering depends on the fact that random errors have a nearly constant periodogram over all the frequency range. On the other hand, the periodogram of smooth data is concentrated at low frequencies and decreases rapidly when moving towards the high frequency region. The periodogram of actual terrain profiles represents a combination of both properties; however, it is always possible to distinguish between the zone that represents noise and the region representing true values. It was found, through this investigation, that it is possible to modify the periodogram of a given set of data such that the zone representing noise is eliminated. The reconstruction of data from the modified periodogram is then performed to obtain a new set of data which is believed to represent more accurately the true terrain profile.

INTRODUCTION

An estimate of the spectral density function of a given set of discrete data can be determined by the transformation of these data from the spatial domain to the frequency domain using the discrete Fourier transform. Although this estimate is a function of the frequency rather than the period, it was called a periodogram by statisticians who first used it to look for periodicities or seasonal trends in data.

Periodograms provide the necessary information about the frequency content of the given data, and have been widely used in the past few years with the advent of many powerful Fast Fourier Transform algorithms.

In Photogrammetry and more specifically in digital elevation models (DEM) processing, spectral analysis has been used in some different applications. Fredriksen (1980) has used this technique for the statistical description of the different landscapes and for the estimation of the standard deviation between the terrain surface and the DEM. He also used it for predicting the accuracy of a digital elevation model (Fredriksen *et al.*, 1984). Papo and Gelbman (1984) have used the same concept to extend the area of digital terrain models to include digital representation of slopes and curvatures. Hassan (1986) has also used the spectral analysis concept for the estimation of the optimum sampling density of digital elevation models.

In the present paper the same concept is employed to filter DEM data that are contaminated by noise. The approach is based on the idea of modifying the computed periodogram of the DEM data so that the contribution of random errors is eliminated. This idea has not been previously applied in DEM processing. The method can be depicted as a transformation of the spatial data to the frequency domain where their periodogram is modified, then a new set of filtered data is reconstructed from the modified periodogram using the inverse Fourier transform. The modification of the computed periodogram is performed in accordance with the fact that noise has noticeable values in the periodogram at high frequencies.

ANALYSIS IN THE FREQUENCY DOMAIN

In order to reduce noise in naturally occuring signals, as in the case of DEM data, it is most convenient to describe them in the frequency domain. The analysis of terrain profile data in the frequency domain serves to reduce masses of data and reveal hidden patterns which can be used for random errors detection and elimination (Dudgeon *et al.*, 1978).

Discrete spatial data are described in the frequency domain by applying a Fourier transformation to obtain the power spectral density function which is known as the Spectrum $S(\psi)$. There are several different procedures for computing power spectral density estimates. One of these methods is to define the spectrum as the discrete Fourier transform of the autocorrelation function $R(k)$ as follows:

$$S(\Psi) = \sum_{k=-\infty}^{\infty} R(k)\, e^{-jk\Psi} \quad (1)$$

where

Ψ is the frequency at which the power spectral density is computed and
k is the lag of the autocorrelation function.

An estimate of the power spectral density function can be related directly to the observed data Z_N (nd) as follows:

$$S_N(m\omega) = \sum_{n=0}^{N-1} Z_N(nd)\, e^{-jndm\omega} \quad (2)$$

where

ω is the frequency spacing in the spectrum plot,
d is the sampling interval,
m = 0, 1, 2, . . . , $M-1$,
n = 0, 1, 2, . . . , $N-1$,
M is the number of spectrum points, and
N is the number of data points.

The discrete Fourier transform of the N data samples is a complex periodic function with a period of 2π. This function can be written as

$$S_N(m\omega) = A_N(m\omega) - j\, B_N(m\omega) \quad (3)$$

where

$$A_N(m\omega) = \sum_{n=0}^{N-1} Z_N(nd) \cos ndm\omega \quad (4)$$

$$B_N(m\omega) = \sum_{n=0}^{N-1} Z_N(nd) \sin ndm\omega \quad (5)$$

The magnitude and angle of the complex discrete Fourier transform can be calculated as follows:

$$|S_N(m\omega)| = [A_N^2(m\omega) + B_N^2(m\omega)]^{1/2} \quad (6)$$

$$< S_N(m\omega) = \tan^{-1}[-B_N(m\omega) / A_N(m\omega)] \quad (7)$$

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 54, No. 10, October 1988, pp. 1391–1394.

0099-1112/88/5410–1391$02.25/0

The magnitude function $|S_N(m\omega)|$ has even symmetry about the $m\omega = 0$ axis, while the angle function $< S_N(m\omega)$ has an odd symmetry about the same axis. The symmetry and periodicity of the discrete spectrum imply that all the frequency information is represented in the finite band of frequencies $0 \le m\omega \le \pi$. The periodogram $P_N(m\omega)$ is defined as the magnitude squared of the discrete Fourier transform of the data divided by the number of data points (Chatfield, 1975). The periodogram is, therefore, unrelated to the angle of the complex spectrum.

$$P_N(m\omega) = \frac{1}{N}|S_N(m\omega)|^2 \quad (8)$$

The reason for introducing the $1/N$ factor is to help the value of $P_N(m\omega)$ converge as $N \to \infty$. As the peridogram is always nonnegative, as a squared quantity, it shares the properties of a true spectrum. The periodogram as defined in Equation 8 has come into prominence in the past few years and is widely used now with the development of very powerful fast Fourier transform programs.

Referring to sampling theory, it can be stated that the value of the periodogram must be computed at intervals not greater than $2\pi/(N-1)d$ (Schwartz *et al.*, 1975); therefore, it is found convenient in the present research to chose an interval of $2\pi/Nd$ between periodogram points. As the discrete periodogram is computed at frequencies spaced by this interval, it is expected that a smooth curve through the resulting values should be a good representation of the corresponding spectrum. The choice of the number of points at which $P_N(m\omega)$ should be determined depends on the frequency properties of the data and the purpose of determining the periodogram. In this paper the number of points is taken to be $N/2$.

Evaluation of $P_N(m\omega)$ requires many calculations of sines and cosines which are very time consuming processes. Fourier transform programs make use of the fact that the used angles are all integer multiples of $2\pi/N$. Therefore, once $\sin(2\pi/N)$ and $\cos(2\pi/N)$ have been calculated, it is faster to compute other sines and cosines as functions of them. Moreover, it is noted that certain functions are computed more than once; for example, $\cos(12\pi/N)$ occurs six times for different values of ω and n. The fast Fourier transform programs make use of such properties to compute the values of the periodogram of large data sets in a very short time (Chatfield, 1975).

INTERPRETATION OF PERIODOGRAMS

Spectral density estimates are typically ragged curves requiring considerable interpretation and analysis. It is important, however, to know the various characteristics of such curves in order to be able to distinguish between portions representing noise and those indicating true elevations of the profile.

From the stand point of Fourier analysis, noise satisfies all the conditions of a function and allows for Fourier representation. However, as noise is highly irregular, its Fourier series will converge very slowly and its periodogram will not diminish at high frequencies. This behavior of the noise periodograms yields a very effective method for the separation of signal (true values) from noise (random errors).

The periodogram is sometimes interpreted as the output of a narrowband filter with adjustable passband center frequency. A narrow band filter at frequency Ψ will pass a sine wave only if its frequency is near Ψ. A set of filters centered at different frequencies can measure relative amounts of power in different frequency bands and therefore result in a periodogram.

The following are some properties of periodograms which were the bases of contriving the method presented in this paper for eliminating noise from terrain profile data:

- The periodogram of random data has essentially small and nearly constant amplitude. The purely random data have a constant spectrum over all frequencies. This type of data is called "white noise" which indicates that it has the properties of white light that contains all wavelengths.
- For smooth data, the periodogram is concentrated at low frequencies with a peak at zero frequency. The values of the periodogram decrease rapidly as the higher frequency zone is approached.
- Periodic variations result in peaks at the values of their frequencies, while trend produces a peak at frequency zero.

Using these properties and observing Figures 1, 2, and 3, it is possible to reach the following conclusions:

- At low frequencies the value of $P_N(m\omega)$ is mainly due to the true elevations of the terrain profile. The effect of noise in this zone is overwhelmed by the much greater effect of signal.
- At some specific frequency the effect of noise and signal may be equal. This is the so called cut-off frequency.
- For frequencies higher than the cut-off frequency, the contribution of noise is greater than that of true values. The value of $P_N(m\omega)$ in this zone is small and the periodogram curve is irregular.

In the present investigation, fictitious data generated to represent terrain profiles are employed to show the efficiency of the contrived filtering technique. Sampling is performed in profiles with unit spacing. Data are generated by trigonometric polynomials of sine and cosine terms. The presence of these terms is the reason behind the noticeable peaks in the periodograms of some profiles. Fictitious noise values are generated as an additional set of data that varies randomly within specified limits. In some cases this set of noise values was exaggerated in order to emphasize the efficiency of the presented method. The degree of smoothing can be controlled by specifying the

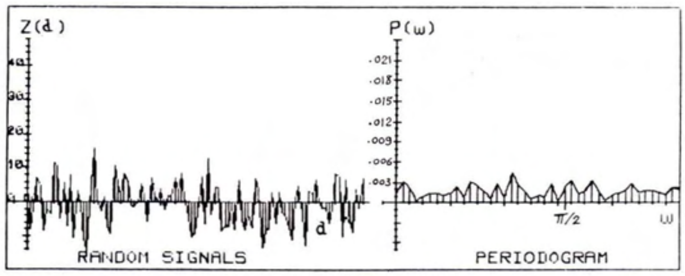

FIG. 1. Periodogram of random signals.

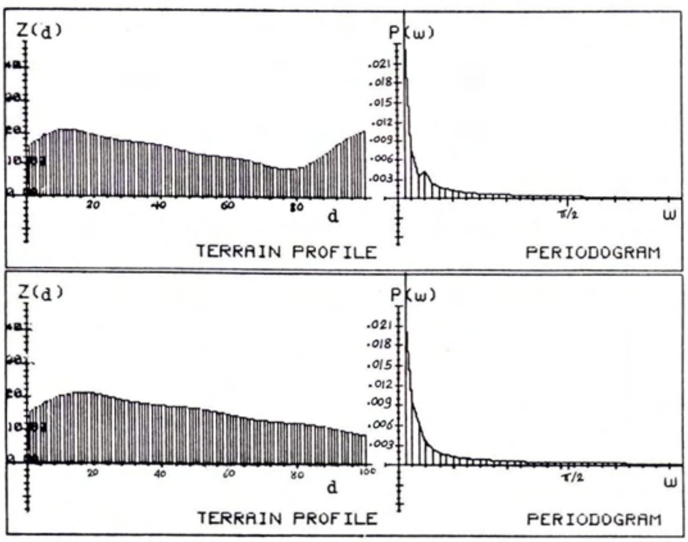

FIG. 2. Periodogram of smooth profiles.

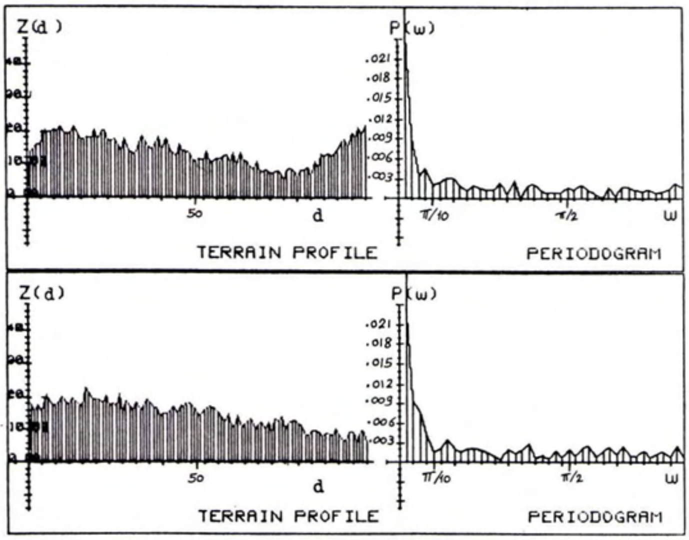

FIG. 3. Periodogram of data contaminated by noise.

level of elimination of the high frequency values of the periodogram.

Fictitious data have many advantages for testing new algorithms. In the present paper this type of data provides the following benefits:

- It is possible to employ data that represent different terrain types, which can be objectively classified. These data types are useful in testing the efficiency of the contrived algorithm in different cases. The use of real data may limit the testing process to areas where DEM data are available. These areas may not represent the different types of terrain relief that are required to perform a comprehensive testing of the suggested method.
- The level of random error values can be controlled so that the filtering effect can be easily judged.

The presented algorithm can readily be used for filtering photogrammetrically acquired DEMs. Terrain surfaces may be sampled photogrammetrically in regular or irregular patterns. In regular patterns, the data points may be taken in homogenious grids or in profiles. Homogenious grids can be considered as a special case of profiles in which the distance between profiles is equal to the interval between points in each profile. It is a common practice in regular sampling to measure the elevations of data points while scanning the photogrammetric models in profiles. Therefore, it can be argued that the method of filtering data in profiles is suitable for DEMs acquired by photogrammetric means in regular patterns. The filtering process in this case follows the same discipline of the data acquisition process.

PERIODOGRAM MODIFICATION AND NOISE ELIMINATION

It is possible to reconstruct the original profile data from the discrete spectral density function using the inverse Fourier transform summation:

$$Z(nd) = \frac{1}{N} \sum_{m=0}^{M-1} S(m\omega)\, e^{jndm\omega} \qquad (9)$$

This process of restoring data from their spectra is sometimes called Fourier synthesis. The term $\frac{1}{N} S(m\omega)$ can be interpreted as the weight with which the periodic component is represented in $Z(nd)$.

Equation 9 can be written in the form

$$Z(nd) = C(nd) + j\, D(nd) \qquad (10)$$

where $C(nd)$ and $D(nd)$ can be expressed as functions of $A(m\omega)$ and $B(m\omega)$ as follows:

$$C(nd) = \frac{1}{N} \sum_{m=0}^{M-1} A(m\omega) \cos ndm\omega - B(m\omega) \sin ndm\omega \qquad (11)$$

$$D(nd) = \frac{1}{N} \sum_{m=0}^{M-1} A(m\omega) \sin ndm\omega - B(m\omega) \cos ndm\omega \qquad (12)$$

In the case of digital elevation model data, the original elevations can be restored using Equation 11, while $D(nd)$ will be zero for all points. However, if the values of $A(m\omega)$ and $B(m\omega)$ in Equations 4 and 5 are modified so that the effects of random errors are suppressed, it will be possible to reconstruct a set of data values which are free from noise. In this way, profile data can be filtered digitally and the resulting smoothing profiles may be considered as better representation of the true terrain surfaces.

Figure 4 shows a typical example of signal contaminated by noise. The signal here represents the true elevations of a terrain profile, while the noise represents the random errors added to these signals to produce the observed data. Figure 5 shows the signal and its periodograms after eliminating the random errors.

Examination of Figure 6 shows that the high frequency zone of the periodogram plot can be regarded as the result of noise contamination. Therefore, truncating this zone results in a periodogram which represents more truly the expected smooth terrain surface.

Figure 7 shows another example of a terrain profile with obvious periodic components which result in peaks in the frequency domain. It should be noted here that the smoothed profile will have the same periodic characteristics. Applying the inverse Fourier transform to the modified periodogram, a filtered value for each elevation point can be obtained. It is ob-

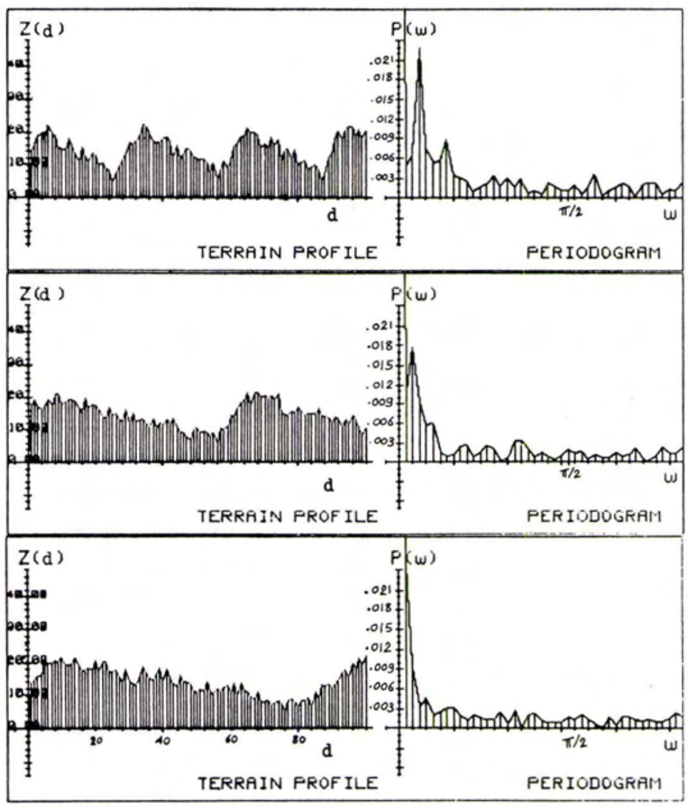

FIG. 4. Periodograms of different terrain profiles.

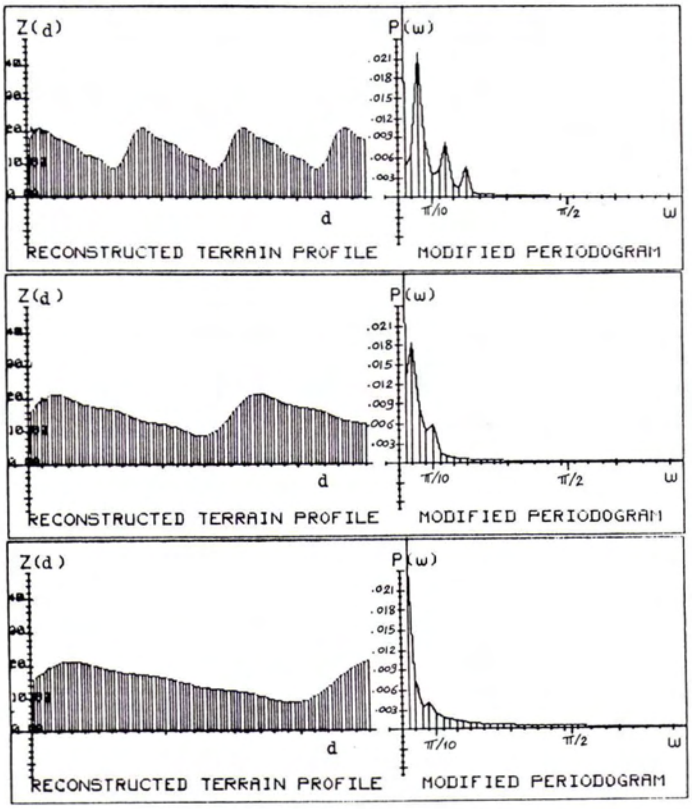

FIG. 5. Periodograms of smoothed terrain profiles.

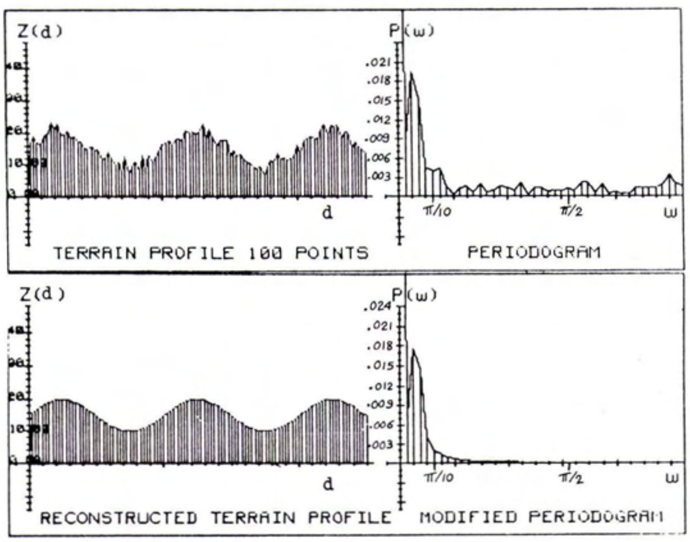

FIG. 6. Filtering terrain profiles using the periodogram modification method.

vious, however, that it is not possible to eliminate random errors completely from a given set of data, because the remaining part of the periodogram after truncation is also influenced by noise and there is no way to purify the data by omitting the effects of noise in this zone.

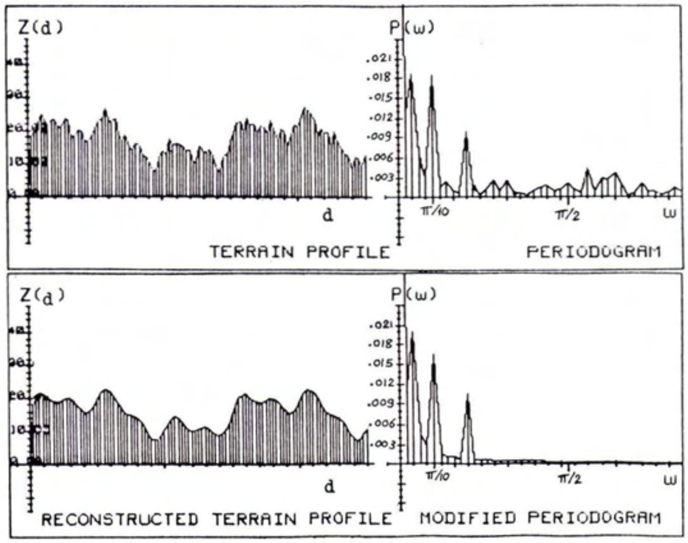

FIG. 7. Filtering a terrain profile with periodic characteristics using the periodogram method.

RESULTS AND CONCLUSION

Frequency domain analysis of terrain profile data can be a very useful tool in eliminating some of the random errors caused by observational imperfections.

The algorithm used in this paper has been tested using different types of terrain profiles. The results show that the same technique can be employed to smooth any irregular curve that is sampled at equal intervals. The smoothing of digitized contour lines is a typical example of a straightforward application of the method.

The results show also that periodogram analysis of digital profile data is a useful diagnostic tool in the filtering of digital elevation model data. The filtering effects of the method are illustrated in Figures 4 and 5 which show different types of terrain surfaces ranging from very rough to moderately smooth profiles.

It should be mentioned that the presented method of filtering has the advantage of restoring the data with minimum distortion in the original signal due to filtering.

REFERENCES

Chatfield, C., 1975. *The Analysis of Time Series*, Chapman and Hall, London, 263 p.

Dudgeon, D. E. and R. M. Mersereau, 1978. *Multidimensional Digital Signal Processing*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 399 p.

Fredriksen, P., 1980. Terrain Analysis and Accuracy prediction by means of the Fourier Transformation, *Proceedings ISP Congress*, Comm. IV, Vol. XXIII, Part B4, pp. 284–293.

Fredriksen, P., O. Jacobi, and K. Kubik, 1984. Accuracy prediction for Digital Elevation Models, *Annals, ISPRS XVth Congress*, Vol. XXV, 3A, pp. 246–255.

Hassan, M. M., 1986. A spectral analysis method for estimating the sampling density of Digital Elevation Models, *ISPRS, proceedings of the Symposium "From Analytical to Digital"*, Comm. III, Int. Archives Vol. 26, pp. 306–316.

Papo, H. B., and E. Gelbman, 1984. Digital Terrain Models for Slopes and Curvatures, Photogrammetric Engineering and Remote Sensing, Vol. 50, No. 6, pp. 695–701.

Schwartz, M., and L. Shawl, 1975. *Signal Processing*, McGraw Hill Kogakusha, 396 p.

(Received 31 December 1987; revised and accepted 17 June 1988)