ELSEVIER

Pattern Recognition Letters 21 (2000) 1149–1156

Pattern Recognition Letters

www.elsevier.nl/locate/patrec

Measuring homogeneity of planar point-patterns by using kurtosis

Jan-Olof Johansson [a,b]

[a] Mathematical Statistics, Centre for Mathematical Sciences, Lund University, Lund, Sweden
[b] Centre for Imaging Sciences and Technologies, Halmstad University, Sweden

Abstract

Kurtosis is generally associated with measurements of peakedness of a distribution. In this paper, we suggest a method where kurtosis can be used as a measure of homogeneity of any quantifiable property on a planar surface. A 2-dimensional, continuous and uniform distribution has kurtosis equal to 5.6. This value is also the limiting value for a discrete uniform distribution defined on a regular, rectangular grid when the number of grid points tend to infinity. Measurements of a planar surface, taken at regular grid points, are considered as realizations of random fields. These are associated with 2-dimensional random variables from which the value of kurtosis can be computed and used as a measure of the homogeneity of the field. A deviation from 5.6 indicates that the stochastic variable is not uniformly distributed and that the corresponding random field is not homogeneous. The model is applied on the spatial variation of the roughness on the surface of newsprint, an application where homogeneity is very important.

1. Introduction

Homogeneity is a demanded property in several manufacturing processes where substances are to be mixed together to form a product, for example the manufacturing of steel, medicine or newsprint to mention some applications.

Consequently, there is a need for measures of homogeneity. For practical use in industry, a single, fast computed figure is highly desirable. It can be used for comparisons and characterizations.

Often, measurements on the surface are sufficient for characterizing the homogeneity of the underlying distributions, for example the distribution of cellulose fibres in paper is closely related to the roughness of the surface.

We suggest the value of the 2-dimensional normalized kurtosis as a characterizing figure of homogeneity. The computation time is $O(N)$. Other methods, found in the literature, use computation times that are $O(N^2)$ or more, where N is the number of observations; see Thomas and Thomas, 1988, Constantine and Hall, 1994 and Davies and Hall, 1999 who use fractal index or dimensions to characterize the roughness of a surface. Also the use of co-occurrency matrices is possible but is very time consuming, cf. Haralick and Shapiro, 1992. A simple χ^2-test can be used but is restricted to binary data.

Kurtosis is the 4th normalized moment of a distribution. It is sensitive to peakedness and peaked distributions have high values of kurtosis.

E-mail address: jan_olof.johansson@ide.hh.se (J.-O. Johansson).

0167-8655/00/$ - see front matter
PII: S0167-8655(00)00076-3

In the 2-dimensional case, peakedness can be considered as a spatial clustering of values. This property is used for measuring the homogeneity of an attribute on a surface.

The method is applied on images, where the values of the pixels correspond to the measurements, see Fig. 1. These are quantized into a finite set of positive numbers, $\Gamma = \{0, 1, \ldots, G-1\}$.

In the present paper, we describe the method for $G = 2$, i.e. for binary images, mainly because it is more easily explained in that case. However, the method works well for other datasets too.

In Section 2, we give the mathematical definitions and formulas for the computations and discuss some numerical examples. In the subsequent section, we perform several simulations of 2-dimensional distributions and compute their value of kurtosis. In Section 4, we apply the method on measurements of the surface of newsprint and we finish with conclusions in Section 5.

2. Preliminaries and definitions

We define a p-dimensional, regular and rectangular grid, $\mathscr{G}$, as a finite or infinite number of points with equal spacing in the p-dimensions. The size of the grid is the number of points in the grid.

We focus on measurements that can be performed on planar, 2-dimensional surfaces, where the observations are taken at points belonging to regular, rectangular grids and we assume that the value at a grid point, u, is a realization of a random variable U. Continuous values are quantized into a finite set or thresholded into binary values. Fig. 2 shows the result from a thresholding operation on the image in Fig. 1.

The values of the points of the grid can thus be considered as a random field, $\boldsymbol{U} = \{U(x,y)\}$, where the arguments (x,y) refer to the location in a coordinate system. The thresholding operation is performed by mapping the random field $\boldsymbol{U}$ into another field $\boldsymbol{V}$, defined by

$$V(x,y) = \begin{cases} 1 & U(x,y) \geqslant z_0, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

A realization, $\boldsymbol{v}$, of this field is a binary image where the components $V(x,y)$ are stochastic variables with a common sample description space, $\Omega = \{0,1\}$. The value z_0 determines the probability of the outcomes 0 and 1. In order to compute kurtosis, we associate to every realization $\boldsymbol{v} = \{v(x,y)\}$, of $\boldsymbol{V}$ a 2-dimensional stochastic variable $\boldsymbol{Z} = (X, Y)$ defined by

$$p(x,y) = \begin{cases} 1/N & \text{if } v(x,y) = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where N is the number of elements of the grid where $v(x,y) = 1$. Spatial concentrations of outcomes of the same value are called clusters. For example, in the binary case, we speak about 0-valued and 1-valued clusters.

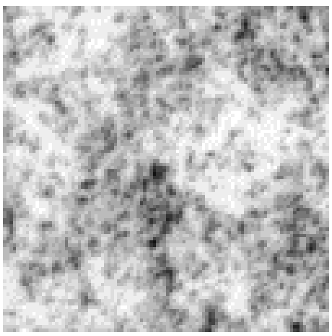

Fig. 1. A grey scale plot of 100×100 measurements of intensities of reflected light from a 2×2 cm^2 piece of paper illuminated with a laser.

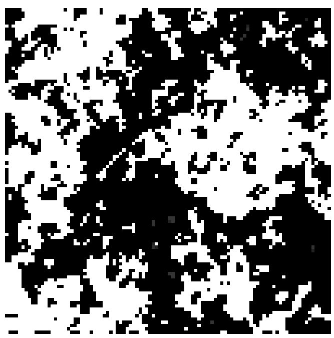

Fig. 2. A binary plot resulting from a thresholding operation on the image shown in Fig. 1.

2.1. Kurtosis

Let $\boldsymbol{Z}$ be a stochastic vector of dimension p with expectation $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The multi-dimensional kurtosis is then defined as

$$\beta_{2,p} = E\left[\left((\boldsymbol{X}-\boldsymbol{\mu})^t\ \boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right)\right]^2, \tag{3}$$

where

$$\Sigma = E\left[\left((\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^t\right)\right]^2 \tag{4}$$

is the covariance matrix.

Multi-dimensional kurtosis can be used to test multi-normality. The value of kurtosis of a p-dimensional normal distribution is $p(p+2)$. For instance, a sample of a 2-dimensional normal distribution should be close to 8, cf. Mardia, 1974.

2.2. 2-dimensional uniform distribution

Assume that $\boldsymbol{Z}$ is a 2-dimensional uniform distribution defined on a rectangle, R, $-a \leqslant x \leqslant a$; $-b \leqslant y \leqslant b$, which is symmetrically located around the origin. The density $f(x,y)$ is then

$$f(x,y) = \begin{cases} 1/4ab & -a \leqslant x \leqslant a, \quad -b \leqslant y \leqslant b, \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

and the kurtosis is

$$\beta_{2,2} = \frac{1}{4ab}\int\int\left\{\frac{3x^2}{a^2}+\frac{3y^2}{b^2}\right\}^2 \mathrm{d}x\,\mathrm{d}y = 5.6. \tag{6}$$

2.3. 2-dimensional discrete uniform distribution defined on a grid

Let $\boldsymbol{Z}$ have a 2-dimensional discrete uniform distribution, defined on a grid, $\mathscr{G}$, of size $(2n+1)\times(2m+1)$. The probability function is then

$$p(x,y) = \begin{cases} \frac{1}{(2n+1)(2m+1)} & \text{if } (x,y) \in \mathscr{G}, \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

and the kurtosis is computed to

$$\begin{aligned} \beta_{2,2} &= \frac{1}{(2n+1)(2m+1)} \\ &\quad \times \sum_{x=-n}^{n}\sum_{y=-m}^{m}\left((x \quad y)\ \boldsymbol{\Sigma}^{-1}\begin{pmatrix} x \\ y \end{pmatrix}\right)^2 \\ &= 5.6 - 0.6\left(\frac{1}{n(n+1)}+\frac{1}{m(m+1)}\right). \end{aligned} \tag{8}$$

Thus, the kurtosis tends to 5.6 when the size of the grid increases to infinity simultaneously in both directions.

2.4. 2-dimensional non-uniform distributions on a grid

In this section, we use some simple examples to illustrate the behaviour of kurtosis of non-uniform distributions. We define a stochastic variable, $\boldsymbol{Z}$, on a regular main grid, in such a way that only those grid points belonging to a set of sub-grids, $\mathscr{G}_i$, $i = \{1,2,\ldots,k\}$, can assume the value 1. Keeping the distribution on the sub-grids constant and letting the size of the main grid increase, the distribution of the 1-valued grid-points on the main grid becomes more and more peripheral, resulting in a non-uniform distribution on the main grid. We will show that the corresponding values of the kurtosis decrease to a limit.

If conversely, although not shown here, the distribution of 1-valued grid-points concentrate on an increasing grid, the values of the kurtosis will also increase.

Assume that an even number, k, of quadratic sub-grids, $\mathscr{G}_i$, $i = \{1,2,\ldots,k\}$, are superimposed, symmetrically around the origin, on a regular, rectangular grid, $\mathscr{G}$, with centre in the origin. Denote the horizontal and vertical distances between two adjacent grids with d_x and d_y, respectively. Fig. 3 illustrates a configuration consisting of four sub-grids.

Let each sub-grid consist of M^2 elements and define a 2-dimensional stochastic variable, $\boldsymbol{Z} = (I,J)$ with a discrete uniform distribution by

$$p(i,j) = \begin{cases} 1/kM^2 & \text{if } (i,j) \in \mathscr{G}_i, \ i = \{1,2,\ldots,k\}, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

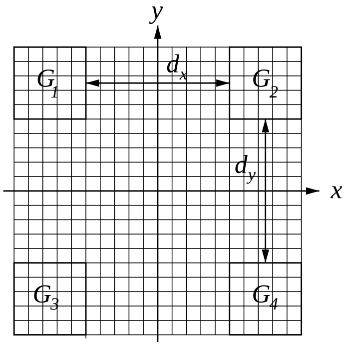

Fig. 3. Four sub-grids superimposed on a quadratic main grid $\mathcal{G}$.

The value of kurtosis is computed for different values of k and M. It can be proven that a limiting value exists for every arrangement of an even number of finite sub-grids, which are symmetrically arranged around the origin.

We consider, for example, two 3×3 and two 5×5 grids and compute the kurtosis and the variances, σ_x^2 and σ_y^2, in the two principal directions for different values of the distances, d_x and d_y, between the grids. The results are shown in Table 1.

These examples show that kurtosis decreases to a limit when the distances between the sub-grids tend to infinity. Kurtosis depends also on the size of the sub-grids. Other sizes and number of sub-grids confirm these results. The corresponding tables are not presented here.

Other configurations, where the center clusters are 1-valued yield values of kurtosis larger than 5.6. These examples are recorded in similar tables and are omitted here.

2.4.1. Conclusions

The calculations show that the kurtosis decreases for distributions, where the 1-valued pixels expand causing 0-valued clusters whereas it increases when the 1-valued pixels concentrate and cause 1-valued clusters. When the distribution is uniform the value of the kurtosis is close to 5.6.

3. Simulations

The objective with simulations is to generate binary images, where the 1-valued pixels are distributed randomly over a regular and rectangular grid. The simulated images are considered as realizations of random fields and using the connection to 2-dimensional random variables, it is possible to compute the corresponding value of the kurtosis.

We simulate homogeneous and non-homogeneous random fields. The size of the realizations are 100×100 pixels and the ratio between the 1-valued and the 0-valued pixels is 0.1.

The type of simulation used, is called sequential simulation and originates in the geostatistics, cf. Journel, 1989 and Deutch and Journel, 1992. It is based on the assumption that all involved distributions are Gaussian. The simulation is based on indicator kriging where the program computes the estimated probability of a pixel to belong to category 0 or 1 conditioned on a given set of data or observations, i.e. the program computes $P_k^*(x,y)$, defined as

$$P(\text{pixel } (x,y) = k \,|\, n \text{ fixed observations})^*; \qquad (10)$$

for $k = 0, 1$.

Based on the computed probabilities, $P_0^*(x,y)$ and $P_1^*(x,y)$, the value 0 or 1 is assigned to the pixels by sampling according to which is the larger of the two probabilities.

In the case of conditioned simulations n equals the number of a priori data when the first pixel is to be simulated. From the second point and onwards, n is the sum of a priori data and already

Table 1
Results from two 3×3 grids, left, and two 5×5 grids

d_x	d_y	σ_x^2	σ_y^2	Kurt.	d_x	d_y	σ_x^2	σ_y^2	Kurt.
0	0	0.67	2.92	5.23	0	0	2.0	8.25	5.48
0	3	0.67	9.67	4.76	0	5	2.0	27.0	4.98
0	9	0.67	36.67	4.57	0	25	2.0	227.0	4.73
0	$\to \infty$	0.67	$\to \infty$	4.50	0	$\to \infty$	2.0	$\to \infty$	4.70

simulated points. By assigning the value 0 or 1 to the set of these a priori data, the simulation will produce images with 0-valued or 1-valued clusters. In unconditioned simulations, no a priori data are given.

Totally, 48 simulations are performed, of which 16 are homogeneously distributed, 16 contain 1-valued clusters and 16 contain 0-valued clusters. We present the mean values and standard deviations from the experiments in Table 2. Two examples of the binary plots are shown in Fig. 4.

3.1. Results

The experiments show that kurtosis is close to 5.6 in the homogeneous case, less than 5.6 for 0-valued clusters and larger than 5.6 for 1-valued clusters. The dispersions around the means are, however, fairly large which complicate the classification. This is a consequence of the simulation method.

We also note that the absolute difference between the mean of the homogeneous images and those with 1-valued cluster are larger than that between the homogeneous images and the images with 0-valued cluster.

Table 2
Results from simulations

	Homogeneous	1-Cluster	0-Cluster
Mean	5.65	6.27	5.38
St.d.	0.19	0.29	0.21

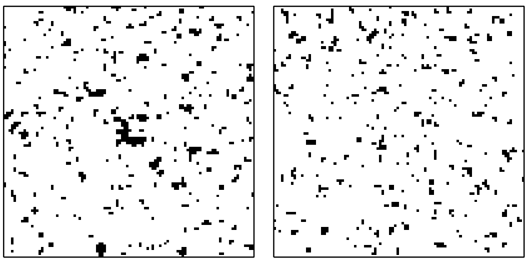

Fig. 4. Two examples from the simulations of binary images. To the left is an image of 1-valued clusters with kurtosis equal to 6.16. To the right is an example of a homogeneous image with kurtosis equal to 5.56. The 1-valued pixels are colored black and the 0-valued are colored white.

4. Applications

We use the method to measure the homogeneity of the surface roughness of newsprint. It is highly desirable that this property is homogeneously distributed on the surface due to its great influence on the quality of the print. The measures used today in the paper-makers and printers laboratories only use one-dimensional statistics and do not consider the spatial variation or homogeneity, cf. the ISO recommendations, ISO, 1988.

There are several methods developed to measure surface roughness, see Béland and Mangin, 1995 and Scott and Abbott, 1995. A simple method uses the reflected light. The light reflection from a surface, can be modelled by Phongs illumination model, which takes into account three types of reflections from a surface, ambient light, specular reflected light and diffusely reflected light, cf. Foley and Feiner, 1989.

The specular light is reflected only in the direction of reflection if the surface is a perfect mirror. If this is not the case, then there is a reflection in other directions as well. This reflection depends on the surface structure and can be used to get a measure of the surface roughness. A light emitter, for instance a laser, can be used to illuminate sample points and the intensity of the reflected light is measured at a direction V, with an angle ω to the normal, N, see Fig. 5. If the sample points are organized in regular grids, then the reflected light can be recorded into 2-dimensional arrays, or images, which then contain information of the 2-dimensional structure of the reflected intensities, or equivalently, the roughness of the surface.

In our experiments the images consist of 100×60 sample points with 0.4 mm^2 spacings and

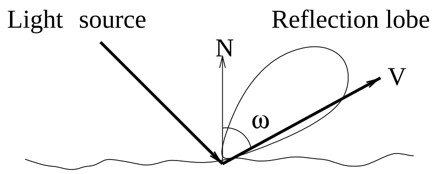

Fig. 5. The model of the reflection of specular light on a rough surface.

at each position we have observed the intensity of the reflected light. The observations, $u(x,y)$, are assumed to represent a realization of a stochastic variable $U(x,y)$ and the image a realization of a 2-dimensional stochastic field, $\{U(x,y)\}$.

The material, used in this investigation, originates from 8 different paper machines in Sweden and the measurements have been performed at Pappersbrukens Forskningslaboratorium, Djursholm, Sweden during 1994, cf. Sunnerberg, 1995.

4.1. Kurtosis of different percentiles

We have investigated 16 different datasets from each paper machines. For each dataset, 10 different percentiles, $u_{90}, u_{91}, \ldots, u_{99}$, have been calculated based on the univariate distribution of the observations. The images have been binarized using the percentiles as thresholds and the value 1 is assigned if $u(x,y) \geqslant u_p$, $90 \leqslant p \leqslant 99$. Fig. 7 shows an example of a binary plot of a sample of the u_{90} percentile from two different newsprints.

Our aim has been to study the distributions of the extreme values, which can be performed by choosing high values of u_p. The corresponding distributions reflect the behaviour of the fiber flocculation and the value of kurtosis constitutes a measure of the homogeneity of these distributions. Investigating the kurtosis at a number of different percentiles, also gives information of the variation of the distribution of the extreme values when the amplitudes, u_p, are changed. Fig. 6 illustrates the situation.

4.1.1. Results

The results from the 10 different percentiles and 8 different paper machines are presented in Table 3 and plotted in Fig. 9. We only present the mean value of the 16 samples from each paper machine. The standard deviations were of the same magnitude as for the simulation study, i.e. about 0.25.

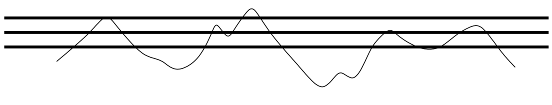

Fig. 6. A cross-section illustrating percentiles of a realization.

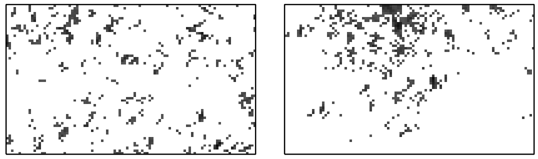

Fig. 7. Examples from the u_{90} percentile. To the left is a sample from paper-machine P1 with kurtosis equal to 5.54 and the images to the right is an example from paper-machine P8 with kurtosis equal to 7.87.

The numerical results agreed well with the visual appearance, cf. Fig. 7.

The 8 datasets were classified using the estimated values of the kurtosis. Due to the small difference between the class of homogeneous images and 0-valued only two classes were used for the classification, the homogeneous class and the 1-valued cluster class.

The samples which were taken from modern or modernized paper machines were all classified to the homogeneous class while those manufactured by older machines were classified as 1-valued clusters. This classification was also in accordance with traditional tests.

4.1.2. Remarks

The choice of size and resolution of the sample images are crucial for the results and must be considered with the distribution of the studied properties in mind. In this investigation, we have used 40×24 mm^2 images with resolution 100×60 pixels and the kurtosis has been computed over all the entire images. The size and resolution have been found adequate for characterizing the large-scale surface roughness.

A smaller sample image size could have been used, for example 50×50 pixels but the variation between the samples would then be greater. Larger sample images had caused a smoothing effect whereas the values of the kurtosis had been less diagnostic, cf. Table 3 and Fig. 9, where these circumstances are indicated.

5. Discussion

In the case of thresholded distributions, different distributions can produce the same value of

Table 3
Results of kurtosis from 10 different percentiles and 8 different papers

Dataset	90	91	92	93	94	95	96	97	98	99
P1	8.51	8.61	8.82	9.13	9.35	8.87	8.85	9.42	8.77	7.79
P2	8.67	9.17	9.18	9.36	9.55	10.23	10.53	11.77	11.10	12.71
P3	5.19	5.13	5.15	5.10	5.11	5.30	5.30	5.32	5.73	5.52
P4	5.07	5.06	5.09	5.13	5.35	5.28	5.28	5.73	6.22	6.50
P5	8.05	8.23	8.23	8.40	8.43	8.44	8.47	8.77	9.89	12.90
P6	5.65	5.71	5.75	5.84	5.89	5.81	5.85	5.98	5.71	5.90
P7	5.44	5.43	5.46	5.41	5.54	5.62	5.57	5.54	5.67	5.98
P8	7.87	8.03	8.05	8.14	8.36	8.21	8.66	8.96	8.83	9.25

kurtosis. This is simply demonstrated by a transformation of the original observations, $W_i = aU_i + b$, giving a new univariate distribution with another mean and variation but with the same value of the kurtosis because it is based on a percentile and is normalized.

In such cases another parameter can be taken into account, namely the coefficient of variation of the univariate distribution. It is defined as the ratio between the observed standard deviation and the observed mean.

Kurtosis and the coefficient of variation then give a 2-dimensional vector that well describes the fluctuations of a surface property.

When working with several grey levels, the 2-dimensional kurtosis is estimated using the formula

$$b_{2,2}^* = \frac{1}{\sum f_{(x,y)}} \sum f(x,y) \times \left(\begin{pmatrix} x - m_x & y - m_y \end{pmatrix}^t \boldsymbol{S}^{-1} \begin{pmatrix} x - m_x \\ y - m_y \end{pmatrix} \right)^2, \quad (11)$$

where $\boldsymbol{S}$ is the estimated covariance matrix, $(m_x \quad m_y)^t$ the estimated mean vector and $f_{(x,y)}$ denotes the value of a pixel in the position (x, y), which here is interpreted as a frequency of the pixel (x, y).

We have used a 100×100 image with 8 grey levels, simulated by using the Gibb's sampler with parameter values giving a peaked distribution, cf. Fig. 8. An inverted image is constructed from the original one by pixel-wise subtraction, $z' = 7 - z$. The value of the observed kurtosis was 5.75 for the peaked image and 5.52 for the inverted one. Thus, the deviation from 5.6 was 0.15 and -0.08, respectively. This result may indicate that the mapping of the deviation from homogeneity to the kurtosis is non-linear.

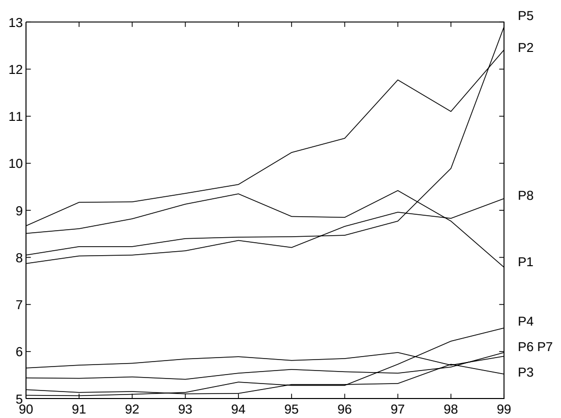

Fig. 9. Graphs of kurtosis from 10 percentiles and 8 datasets.

The distribution of the kurtosis is not known and it is therefore not possible to construct statistical tests. An interesting task for further research work could be to investigate this problem and possibly find an asymptotic distribution of the kurtosis.