

Article

On Structural Entropy and Spatial Filling Factor Analysis of Colonoscopy Pictures

Szilvia Nagy^{1,*} , Brigita Sziová¹  and János Pipek² ¹ Széchenyi István University, Egyetem tér 1, H-9026 Győr, Hungary; szilvia.nagy@sze.hu² Budapest University of Technology and Economics, Budafoki út 8, H-1111 Budapest, Hungary; pipek@phy.bme.hu* Correspondence: nagysz@sze.hu; Tel.: +36-96-613-773

Received: 31 December 2018; Accepted: 27 February 2019; Published: 6 March 2019



Abstract: Colonoscopy is the standard device for diagnosing colorectal cancer, which develops from little lesions on the bowel wall called polyps. The Rényi entropies-based structural entropy and spatial filling factor are two scale- and resolution-independent quantities that characterize the shape of a probability distribution with the help of characteristic curves of the structural entropy–spatial filling factor map. This alternative definition of structural entropy is easy to calculate, independent of the image resolution, and does not require the calculation of neighbor statistics, unlike the other graph-based structural entropies. The distant goal of this study was to help computer aided diagnosis in finding colorectal polyps by making the Rényi entropy based structural entropy more understood. The direct goal was to determine characteristic curves that can differentiate between polyps and other structure on the picture. After analyzing the distribution of colonoscopy picture color channels, the typical structures were modeled with simple geometrical functions and the structural entropy–spatial filling factor characteristic curves were determined for these model structures for various parameter sets. A colonoscopy image analyzing method, i.e., the line- or column-wise scanning of the picture, was also tested, with satisfactory matching of the characteristic curve and the image.

Keywords: computer aided diagnostics; colonoscopy; Rényi entropies; structural entropy; spatial filling factor

1. Introduction

Colorectal cancer develops from colorectal polyps. The detection of the colorectal polyps is mostly carried out by special endoscopes, called colonoscopes [1,2]. These devices possess not only the image acquiring equipment with light source, but also forceps, needle, laser scalpel, or loop instrument for removing polyps or tissue samples for biopsy. Beside normal, white light pictures, some of the endoscopes can take narrow band images (NBI), which emphasize the blood vessels and the shadows, as can be seen in [3–5], which help find unusual vein patterns that are typical in the case of malignant polyps. In many cases, indigo carmine [4] or other food dyes of bluish hue can be sprayed beside the usual cleansing water to make the pits and valleys of the bowel wall more visible (chromoendoscopy). In addition, magnifying endoscopy is becoming more and more common to detect the fine scale patterns of the surface [4]. Virtual endoscopy [6,7] is a computer tomography based alternative for the optical endoscopy. Capsule endoscopy was developed with the goal of decreasing the discomfort of the patients. It is a small capsule with two cameras and light sources at both ends. It can be swallowed, and travels through the bowel [8]. Unfortunately, it is generally less effective in finding polyps than the classical endoscopy, and cannot perform any operations, as it is a passive device.

Although colonoscopy is considered to be the most effective way of cancer screening [1], it still has a non-negligible miss rate [9], which has not decreased much, even though the equipment has

improved over time [10]. A colorectal polyp can be missed for the following reasons. First, if the polyps are small, then they are much harder to find [9], even though curvelet-based methods exist that can improve the diagnosis probability [11]. In addition, usually, the better quality the picture is, the lower the risk of missing a small polyp. Moreover, although before endoscopy sections, the bowel is theoretically purified, impurities often occur [12], mostly in the form of yellowish liquid or solid pieces. These impurities can be removed by spraying water on the given surface segment from the endoscope. The next reason is that the bowel has continuously moving, shiny, pink walls, which sometimes (despite the inflation) fold over polyps. The last factor is fatigue: after a couple of seconds of watching the screen of the endoscopes, the eyes and brain of the gastroenterologist gets used to the environment, and after a longer time of watching they often get tired; this is true not only in the case of inexperienced medical staff [9,10]. This last factor was the reason for thinking about developing a computer aided diagnosis protocol. Computer aided diagnosis is not meant to substitute the human medical expert, only to draw their attentions to certain points, in a way that does not disturb other aspects of the diagnosis.

Image processing tools can improve the diagnosis performance; however, colonoscopes are not developed for machine processing, no matter if it is a conventional or capsule version. Usually, live video signal arrives from the endoscope, which has quite low bit rate, hence the small resolution and/or large compression ratio of the pictures. This means that, even though the video signal seems to be of good or at least acceptable quality for the human eye, the individual pictures have large distortion and many compression artifacts. Since the bowel walls move continuously, usually the pictures are blurred. Impurities are often present, and, even if they are removed by water, this water together with the native liquids makes the bowel walls shiny, and the thus arising reflections make image processing more complicated. Even though chromoendoscopy or NBI can make the color spectrum a bit more stretched, in a common endoscopy picture, the colors are mostly only shades of pink.

In medicine, to determine whether a lesion is benign or potentially malignant, a sample is taken from the tissue for biopsy. As fatigue might also influence the working efficiency of a pathologist, there are possibilities to introduce computer aided diagnosis to this point, too. This branch is investigated by multiple research groups. Kayser and his group applied the neighborhood relations based structural entropy for this purpose [13]; dos Santos and his group used sample entropy [14]; Ribeiro and his fellow researchers proposed curvelet and fractal analysis combined with Haralick structure descriptors and various classification methods [15]; Chaddad and his coworkers combined multiple texture characterizing entropy-related quantities and studied their multiresolution behavior [16,17]; and Wang trained neural networks for the picture components derived for both of the applied stains [18]. However, even though these methods might work well for tissue samples or other illnesses [19], they are unfortunately not applicable directly to endoscopy pictures.

Another image classification task related to colorectal polyps or cancer arose after magnifying endoscopy appeared [20]. Kudo and his coworkers, based on biopsy results, found that the pit pattern of the polyp surface can classify the polyps without performing the biopsy, or removing the benign polyps [21]. The classification of polyps based on the Kudo classes [22] is the second branch of image processing based computer aided diagnosis for colonoscopy [23]. However, the task for finding the polyps is also of great interest.

The colorectal polyps can have a wide range of shapes, from lesions depressed into the bowel wall through flat and slightly protruding sessile to pedunculated polyps with expressed stalk. Generally, the roundish polyps—either sessile or pedunculated—are the target for automatic image processing methods. The MICCAI Endoscopic Vision Challenge made several databases and methods available [24–26]. Its results are summarized in [27]. The methods collected in [27] include evolutionary algorithms, neural networks, and shape- and lighting-based classical similarity recognition algorithms [25]. For the detailed description of the methods and the results achieved by the research groups that participated in the MICCAI Endoscopic Vision Challenge, we refer to [27]. There

are also methods for computer aided classification of lesions from capsule endoscopy images [28,29], but most of the studies work with conventional endoscopes, as they are more widespread.

A fuzzy classification scheme based on the method detailed in [30,31] and summarized in [32] also appeared [33,34]. This proposal uses edge density and statistical parameters, such as the mean, standard deviation or the Rényi entropy based structural entropy, for determining whether a segment of the colonoscopy picture contains polyp.

Entropies are often used for image analysis [19,35,36]. There are two approaches for determining the structural entropy. First, a graph theory based definition was given in [37]. Later, independently a Rényi entropy based structural entropy was also introduced. Although this later structural entropy is first applied in electron structure analysis in [38], its use in image processing is first presented in [39] for characterizing microstructures of the metal electrode materials on semiconductor surfaces. The idea to use it in medical image processing are presented in in [33]. The results are good, well above 90% hit rate for some types of pictures (especially the ones where either the color or the pattern is very different for the polyp and the background, or the polyp has strong, visible contour), while for some other types (e.g., when the polyps are lit too strongly from the side, or when there are image distortions on the pictures due to low resolution or some dark impurities), the miss rate is around 50%, which is of course unacceptable. The false positive rate is always very low, except for some extremely impure cases.

The aim of this paper was not to provide or improve a method for classification of images, but to study a tool that can be used in classification algorithms as one of the parameters related to the shape of the pixel intensity distribution of the picture.

The Rényi entropy based structural entropy is a very simple quantity, easy to calculate, and, together with the spatial filling factor, gives visible information about the shape of the studied distribution. In the following, we discuss some of the most important properties of the roundish colorectal polyps and model their pixel intensity distributions and structural entropy behavior to get characteristic curves that help to understand the reason of differences between image segments with and without polyps and image types.

The remainder of this article is organized as follows. The properties of structural entropy and its use in image analysis are summarized in Section 2. Next, in Section 3, the distribution of bowel picture segments with and without polyps are studied and model structures for reproducing certain aspects of the pictures are introduced, which are required to generate structural entropy–spatial filling factor characteristic curves for the picture segments with polyps, as presented in Section 4. In this section, the dependence of the structural entropy–spatial filling factor curves of both the model surfaces and the real images are studied according to several parameters. Finally, the conclusion is summarized in Section 5. The characteristic curves are collected in Appendices A and B.

2. The Rényi Entropy-Based Structural Entropy

One of the first attempts to describe information was by Hartley, who used a number of yes/no questions to identify an element of a set of possible strings as the information revealed by identifying the string [40]. A couple of years earlier, Nyquist used a very similar formula [41]: both used the logarithm of the number of possible choices to define information content. Shannon wrote his article about the theory of communication in 1948 [42]. He first defined entropy in the information content sense, referring to statistical mechanics and the Gibbs entropy [43] when introducing this quantity. In addition, in quantum mechanics, the entropy of a density distribution was introduced [44], as it was based on the notations used by von Neumann [45]; later, it was named after him. In both cases, the entropy of a probability distribution $\{p_1, p_2, \dots, p_N\}$ is defined as

$$S(p_1, p_2, \dots, p_N) = K \sum_{i=1}^N p_i \log_a \frac{1}{p_i}. \quad (1)$$

where, both the constant K and the basis of the logarithm are freely chosen; however, both in quantum mechanics and in information theory, the constant is generally selected to be one. The basis of the

logarithm is 2 in the case of the information theory applications (in this case, the unit of the entropy is Shannon or simply bit), and e in the case of physics.

The entropies used in image processing, as well as the quantities originated from the sample or structural entropy usually define the probability distribution corresponding to an image in a rather complex way. Stantchev based the probabilities on the number of connections of a given node in a graph. The entropy from [13,46] calculates probabilities from distances between neighboring vertexes and connections weights; Humeau-Heurtier and her co-workers generalized the sample entropy, which uses probabilities consisting of the ratio of the number of cases when two sample vectors froming a series have sufficiently small distance, and a similar number for their shortened versions [36,47–49]. All these methods introduce quite complicated concepts, which are scale-dependent. Of course, scale dependency gives valuable information on the structure, such as in the case of the multiscale entropy [36].

However, there is another concept for characterizing the structure the shape of a picture by entropies. Images have native distributions, their pixel intensities, which can be easily normalized to fulfill the conditions for being a probability distribution, i.e.,

$$\sum_{i=1}^N p_i = 1 \quad (2)$$

$$p_i \geq 0 \quad \text{for } i = 1, 2, \dots, N, \quad (3)$$

if the already non-negative pixel intensities I_i are divided by their sum as $p_i = I_i / \sum_i I_i$.

In electron structure calculations, instead of the probabilities, the electron density is used; it is also normalized similarly to the probabilities. Although the electron density is usually a continuous function, it can be approximated, or modeled as a distribution over a regular grid, thus the similarity of the electron states and picture pixel intensities can be seen. For measuring how localized is an electron state of a solid, a participation ratio, or delocalization measure, was introduced [50,51] the following way,

$$D = \frac{1}{\sum_{i=1}^N p_i^2}. \quad (4)$$

This quantity tells approximately the number of the higher probability grid points, i.e., the number of grid points the electron density extends to, or, in the case of the pictures, it is the number of the light pixels.

If in the entropy we substitute the real probability distribution by a step distribution that extends to the D pixels, we exclude the shape information and keep only the information related to the extension of the distribution. The entropy thus becomes the extension entropy [38]

$$S_{ext} = \log D. \quad (5)$$

This means that, if we subtract S_{ext} from the total Shannon entropy, the remaining part has the information about the shape or the structure of the distribution. Structural entropy was introduced as

$$S_{str} = S_1 - S_{ext} = S_1 - \log D. \quad (6)$$

In Ref. [38], Equation (6) uses natural logarithm, and we apply this convention (even though any basis for the logarithm could be used).

Using Shannon's entropy definition, Alfréd Rényi generalized [52,53] the quantity characterizing the amount of information based on Faddeev's postulates [54]. His zeroth entropy was Hartley's information content; the first one was Shannon's entropy; and the next members of this series are

$$S_n = \frac{1}{1-n} \log \sum_{i=1}^N p_i^n. \quad (7)$$

If we study the extension entropy in Equation (5), knowing the Rényi entropy series in Equation (7), we can find that the second Rényi entropy

$$S_2 = \frac{1}{1-2} \log \sum_{i=1}^N p_i^2 = \log \frac{1}{\sum_{i=1}^N p_i^2}. \quad (8)$$

is the extension entropy itself [55].

Pipek and Varga introduced another quantity that describes the structure of the distribution. If the participation ratio D is compared to the total number of grid points (pixels), i.e.,

$$q = \frac{D}{N} \quad (9)$$

is defined, we receive the so-called spatial filling factor, which is a quantity between 0 and 1. Pipek and Varga [38] showed that, if for a distribution of a given shape its structural entropy S_{str} is plotted versus its spatial filling factor q , then the point is along a curve that is characteristic for the shape of the distribution. Each type of shape, e.g. Gaussian, exponential or power law distribution has its separate characteristic curve (which is of course different for one-, two-, or three-dimensional distributions). Moreover, in [55], the relation

$$\log q = \log \frac{D}{N} = \log D - \log N = S_2 - S_0 \quad (10)$$

is also derived, and it is usual to plot the $S_{str}(\ln q)$ curves, instead of the originally proposed structural entropy–filling factor plots. Some characteristic curves for the two-dimensional exponential, Gaussian and second-order power law distributions are shown in Figure 1. In addition, the theoretical limit of the structural entropy

$$S_{str} \leq -\ln q \quad (11)$$

is plotted. For the proof of this formula, $S_{str} \geq 0$, the completeness of the allowed domain, as well as the shape of the characteristic curves, we refer to the appendices of [38].

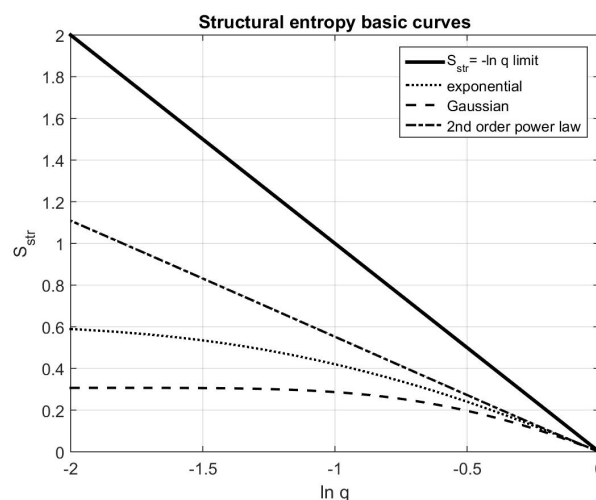


Figure 1. A structural entropy $\ln q$ plot showing the limiting line $S_{str} = -\ln q$ in thick continuous line, and the trend lines for the second-order over law, exponential and Gaussian type distributions for two-dimensional case.

The Rényi entropy based structural entropy and the filling factor is introduced in scanning electron microscope image characterization in [39], and for determining superstructures within a nanostructure in [56]. For characterizing surfaces of electrodes, Bonyár and his coworkers used the structural entropy based localization factor with success [57,58]. Based on these results, we surmised that also colorectal polyps can be identified using their structural entropy versus filling factor plots. The classification of a distribution needs characteristic curves, to which the structural entropy and filling factor point of the distribution can be related. To make the Rényi entropy-based structural entropy applicable for characterizing images or image segments of colonoscopy origin, we need to find possible structures present in such an image, as well as their characteristic lines on the $S_{str}(\ln q)$ map. The purpose of this study was to determine if there are differences between characteristic curves of images with and without polyps, and if there is a way they might be used for distinguishing the two types of images.

3. Results: Simplifying and Modeling the Structures Present on Colonoscopy Images

3.1. Across Real Pictures

We used the database of Etis Larib from the MICCAI Endoscopy Vision Challenge [24] for this study, as their pictures have very high resolution (1225 by 966 pixels), only small black frame, and only very few compression artifacts. The three color channels of two selected images are plotted in Figure 2. (The first one belongs to the well, but not extremely well classifiable group in [33,34], the second to the not too badly classifiable group.) It can be seen that the different color channels emphasize different features of the image: the veins are visible in the green color channel, the shadows can be seen in the blue and red channels and yellowish liquids show in the blue channel.

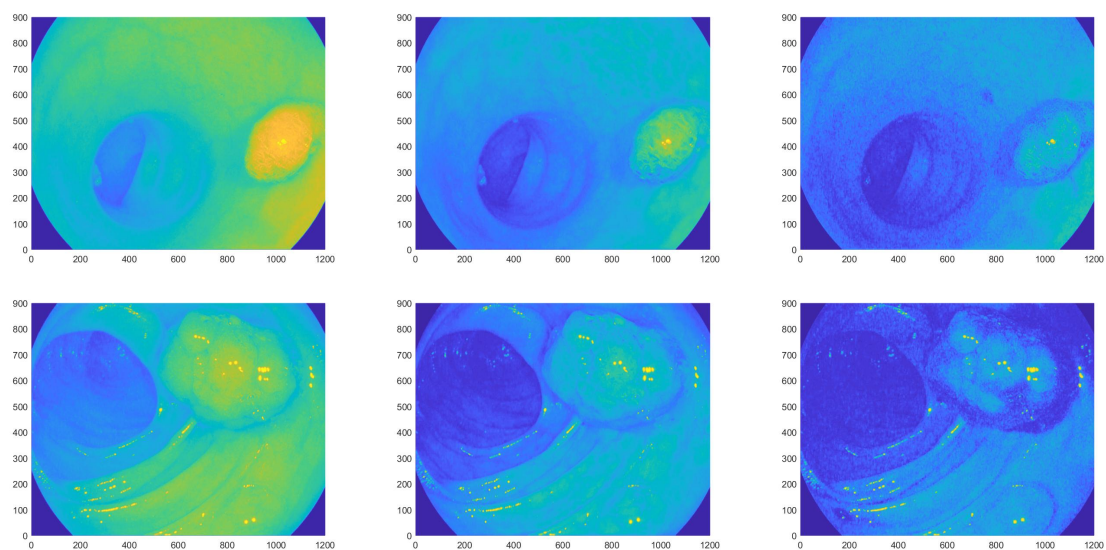


Figure 2. The red, green and blue color channels of pictures 83.tif and 114.tif from database ETIS Larib [24].

In these pictures, the elementary structures seem to be waves and sphere or ellipsoid segments. To understand the behavior of the structural entropy of the different image segments, structural entropy versus filling factor plots of waves with straight or curved wave fronts, as well as of hemispheres are determined and the characteristic lines are given for these structures as a first step. As in these pictures the sphere segments are sitting on the wavy background, the next step would be to plot structural entropy characteristic curves of these superposed structures. To determine the more detailed structure around the polyps, we prepared cross section cuts through the polyps both in both dimensions. Some examples are shown in Figure 3.

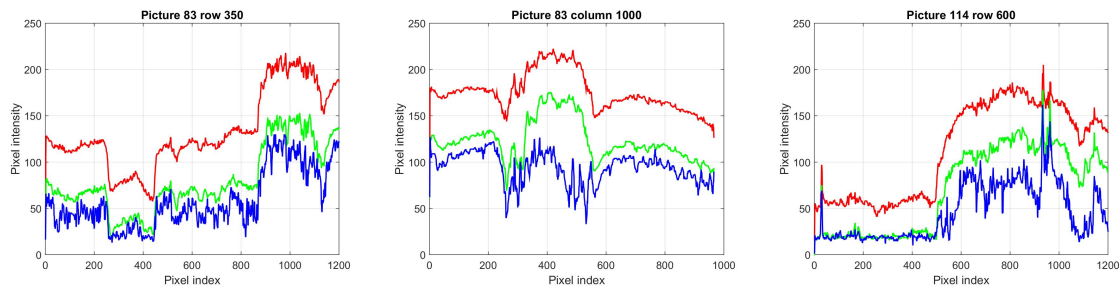


Figure 3. Cross section plots through polyps of Figure 2. The red, green and blue color channels are plotted with red, green and blue, respectively.

According to the cross sections, the environment of the polyps can be modeled as if a hard hemisphere would be pressed into an elastic surface, i.e., almost all the polyps had some kind of ditches around them, similar to the function shown in Figure 4. This is of course the shadow around the polyp. After testing some functions to reproduce this structures, we found that, if we subtract Gaussian function of the same standard deviation as the radius of the sphere, the behavior is rather well modeled.

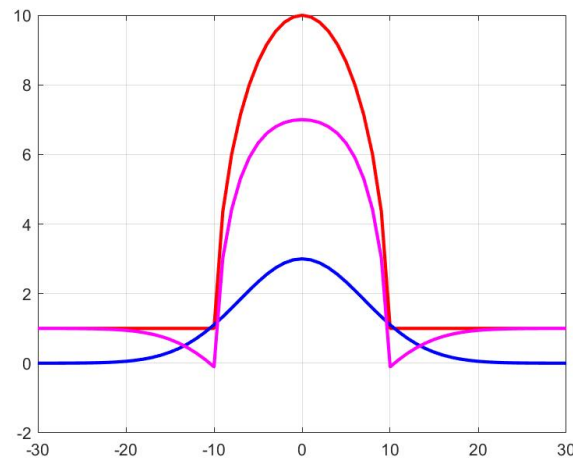


Figure 4. A semicircle (red line), a Gaussian function (blue line) of the same parameters $R_{circle} = \sigma_{Gaussian}$ and their difference (magenta line), which is reminiscent of the shape around the polyp in Figure 3. This kind of function can be used as a simple model the pixel intensity around the polyps.

3.2. Model Structures for Waves of the Bowel Wall

The bowel wall without polyp forms waves. As a first step, these waves can be modeled as sinusoidal function with straight wave profile over a grid of size $N \times N$, as can be seen in Figure 5. A sinusoidal distribution

$$f(x, y) = A \cdot \sin\left(\frac{2\pi}{T}x + \varphi\right) \quad (12)$$

has three parameters: the amplitude A , the wavelength T (frequency $1/T$) and the phase φ . In our case, as the distributions are normalized to be a probability distribution, changing the amplitude is out of the question, thus the remaining parameters are frequency and phase shift. An offset or DC term can also be introduced as a parameter to study, and the angle of the wave front, is the distribution, thus becomes:

$$f_W(x, y) = \sin\left(\frac{2\pi}{T_x}x + \frac{2\pi}{T_y}y + \varphi\right) + B + 0.5, \quad (13)$$

with B being the offset, and T_x and T_y the two components of the wavelength. A default offset 0.5 was introduced to fulfill Equation (3). The distribution is normalized according to Equation (2) before

calculating the structural entropies in all the cases, thus the normalization step is not mentioned in the further models.

As in most cases only part of a whole period is visible in the studied image segments, the parameter set was selected to cover the cases when the period of the wave is between 0.1 and 10 times the size of the tile size. There is no point going below 0.1 as the surface is practically a plane with gradient of $2\pi/T$. The offset was studied to be between 10^{-10} and 10^{-1} . The zero offset was not used, as in the $p_i = 0$ case (Equation (1)) is not computable by machine; of course, its limit can be derived by l'Hospital's rule, however for running time reasons the conditional branching for calculating the $p_i = 0$ entropies was not implemented. The parameter set for φ was from 0° to 180° , and the T_y/T_x ratio from 0.1 to 10.

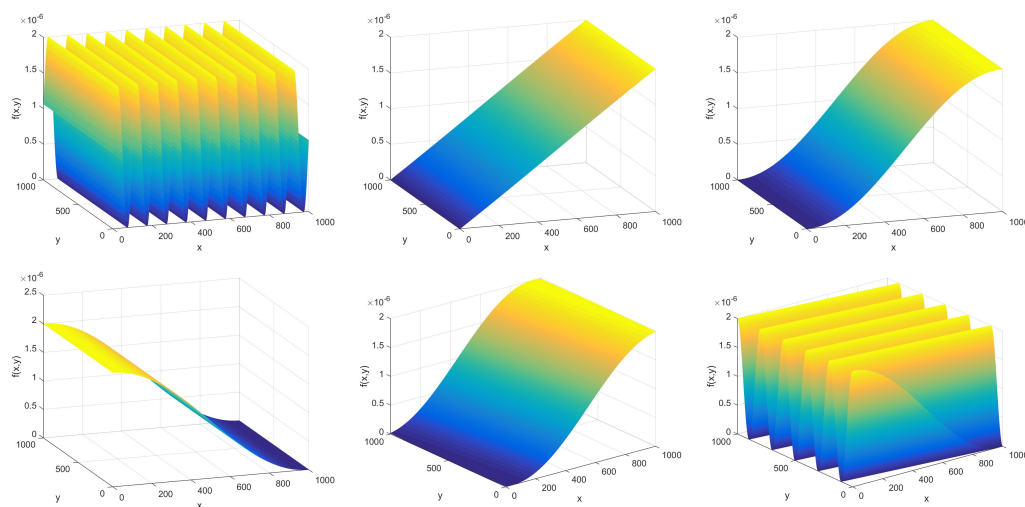


Figure 5. Pictures of straight waves over 1000 by 1000 grid. The limits of the parameter sets are plotted: (1st and 2nd subplots) the largest wavelength and the smallest wavelength used; (3rd and 4th subplots) the minimum and the maximum of the phase shifts; and (5th and 6th subplots) the minimum and maximum wavefront direction angles.

As the bowel is a tube, and in perspective the waves of the wall might seem to be concentric, the circular, or elliptic waves are also of interest. The distribution for these waves is modeled as

$$f_C(x, y) = \sin \left(\sqrt{\left(\frac{2\pi}{T_x}(x - x_0) \right)^2 + \left(\frac{2\pi}{T_y}(y - y_0) \right)^2} + \varphi \right) + 0.5 + B. \quad (14)$$

The six parameters are the two wavelengths T_x and T_y , the two coordinates of the center x_0 and y_0 , the phase φ and the offset B . The studied parameter domains for the ratio T_x/T_y are from 1/5 to 5, for the center coordinates (x_0, y_0) from the center of the tile, i.e., from $(0, 0)$ to $(N, 0)$, and for the phase φ from 0° to 180° , as can be seen in Figure 6.

Tilted Waves

As in the pictures the further parts of the bowel are darker, waves with a tilt were also studied. In this case, instead of the constant offset B , a plane with a slight slope was also applied. The direction of the slope was perpendicular to the wave front, as mostly the wave fronts are perpendicular to the bowel axis. Plane tilt was given to the waves with straight front and conical tilt to the circular fronts. The parameter was the ratio of the wavelength and the gradient, which was between 2^0 and 2^{-6} .

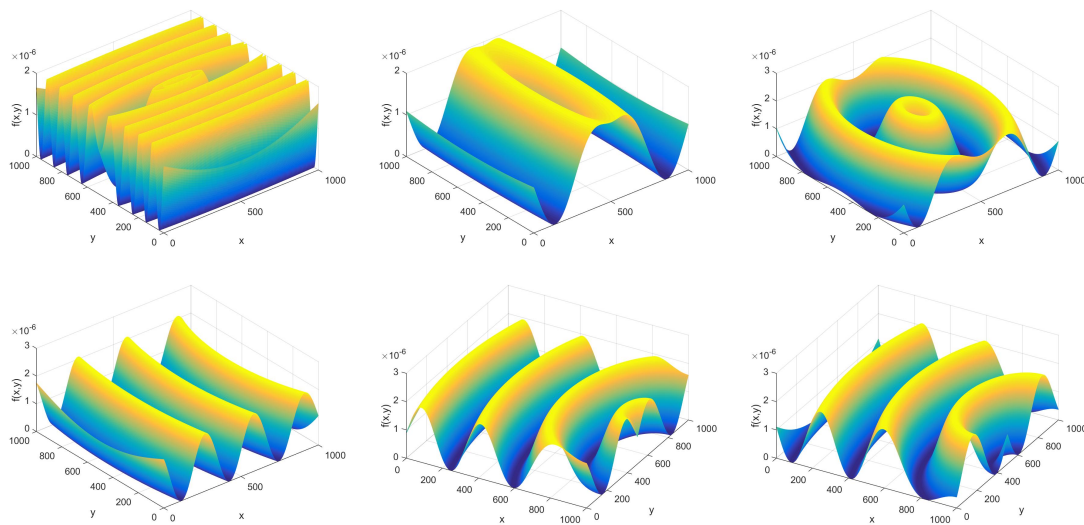


Figure 6. Picture of elliptical waves with minimum and maximum wavelength ratio in the first two subplots, minimum and maximum off-center positions in the fourth and fifth subplots, and minimum and maximum phases in the last two subplots.

3.3. Model Structures for the Polyps and their Shadows

The polyp can be quite well modeled as hemispheres, ellipsoid or sphere segments. The studied distribution was

$$f_S(x, y) = \max \left(\sqrt{R^2 - \left(\left(\frac{(x - r_x)}{R_x} \right)^2 + \left(\frac{(y - r_y)}{R_y} \right)^2 \right)}, B \right), \quad (15)$$

with R being the radius of a sphere, R_x and R_y the parameters distorting the hemisphere to a half ellipsoid, (r_x, r_y) the coordinates of the center of the object, and B the background height around the ellipsoid or sphere segment, which was usually set to 10^{-10} . The analysis went on in two directions: first the size and the position of a hemisphere was varied, and then the positions remained at the center and at the edge of the picture, but the shape was distorted to ellipsoid. The the distributions corresponding to the limits of the parameter sets are given in Figure 7.

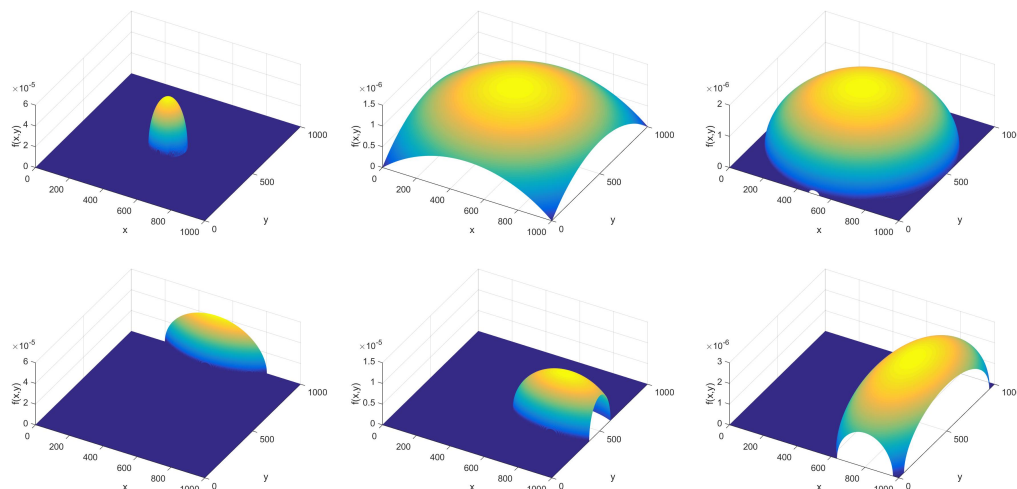


Figure 7. Picture of hemispheres. In the first two subplots the radius is at the two extremum of the parameter set, in the third and fourth subplots the offset, and in the fifth and sixth subplots the deformation toward an ellipsoid.

Model Structures for the Shadows around the Polyps

Only Gaussian functions were used for generating the valley representing the shadow around the spheres; however, distributions of type

$$f_G(x, y) = \exp \left(- \sqrt{ \left(\frac{(x - r_x)^2}{\sigma_x^2} + \frac{(y - r_y)^2}{\sigma_y^2} \right) }^{\alpha} \right), \quad (16)$$

were also studied. Here, the center was always set as the same position as the center of the half ellipsoid, and the variances σ_x and σ_y as the same as the radial parameters of the ellipsoid. As functions that use higher power α in Equation (16) have wider and flatter central part and quicker decrease, they were also tested for reproducing the shadows around the hemispheres, with less distortion in the spheres. However promising this idea was, the results were usually less similar to the real polyps than the $\alpha = 2$ case, as can be seen in Figure 8.

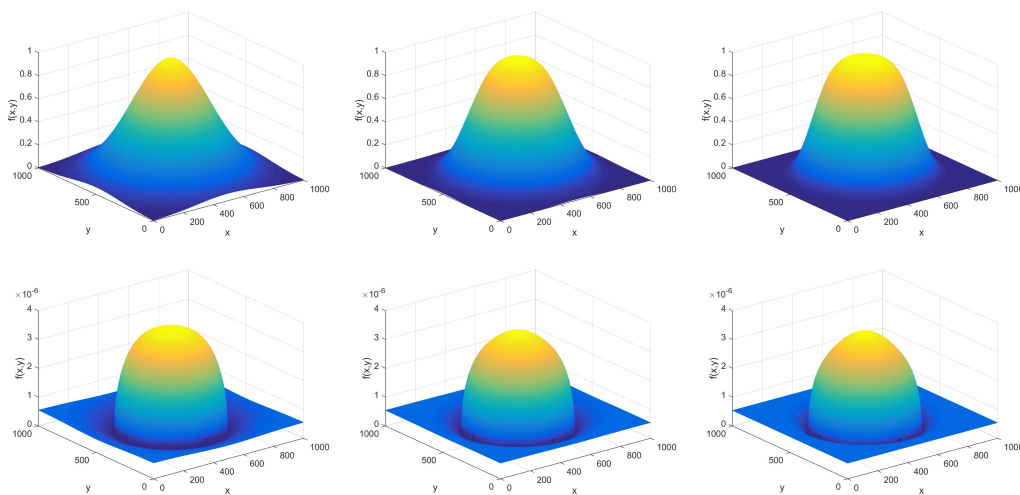


Figure 8. Picture of a Gaussian function together with its superposition with a hemisphere, as well as the $\alpha = 3$ (first row) and $\alpha = 4$ (second row). The hemisphere has the same radius as the σ of the exponential functions, but its height is triple.

4. Discussion of the Structural Entropy Characteristic Curves

4.1. Characteristic Lines from Artificial Model Systems

After deciding the possible models and their parameter sets, the structural entropy versus spatial filling factor plots were studied. Two parameters were changed in one plot series: the first given as the third axis of the plot, and the second as the color and marker of the plotted points. Even though the parameter sets consist of discrete values, the points corresponding to the second parameter value and varying along the first parameter were connected as a guide of the eye. Most of the characteristic curves are presented in Appendices A and B. The reason for this is manifold. First, with only the title containing the information about the model type, it is easier to see the result. A similar statement is true for the text about the characteristic curves in this section: not breaking the text with images helps keep the focus. Second, there are many parameter combinations that do not seem to be very important at this point, and their results can be summarized in one sentence. Third, usually three or four plots are given for one result, which is too many; however, the 3D plots with parameter–filling factor–structural entropy axes are usually interesting not only from one point of view, but from the three projections and one perspective plot as well.

4.2. Dependence on the Image or Tile Size

The dependence on the tile size can be excluded as a parameter if all the other parameters are given in the relation to the tile size N . The only exception is the offset parameter B . In the cases when B was used only for technical reasons, namely for treating the $p_i = 0$ cases without having to use if-then conditional branchings in the program, the offset should be small enough to be negligible compared to the rest of the intensity values. To set a suitable default offset, the tile size dependence of the structural entropy and the spatial filling factor of all our model surfaces were determined. In all cases, we found that, between the realistic limits of $N = 20$ (smallest applied tile size in the case of lower resolution images in [33]) and $N = 1000$ (magnitude of the full image size of database [24]), the tile size- $\ln q$ - S_{str} curves are practically the same if B is smaller than 10^{-5} .

An example (of waves with straight wavefront) is plotted in Figure 9. This serves as a demonstration of how the three projections of the 3D plot look similar. It can also be seen in Figure 9 that, although above the tile size $N = 200$ the values $\ln q$ and S_{str} are almost completely independent of the tile size (even of the tile size to wavelength ratio), in the region of smaller N , the tile size plays not negligible role, thus we can conclude that using larger tile sizes in the evaluation process makes the results more stable and reliable. However, fulfilling the condition of using at least 200 by 200 pixel sized tiles is not always possible, especially if the images are of 384×288 size, such as in the case of the CVC Clinic database [27].

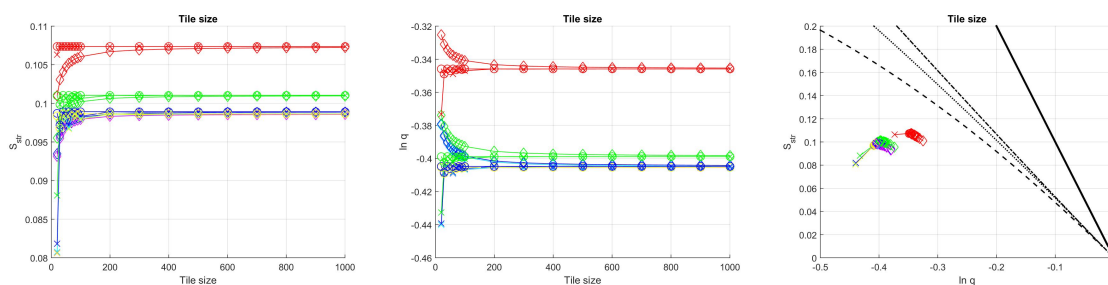


Figure 9. The structural entropy S_{str} and the logarithm of the spatial filling factor q versus the tile size for straight waves. The offsets are the following: red, 0.1; green, 0.01; blue, 0.001; cyan, 0.0001; magenta, 0.00001; yellow, 0.000001. The last three are barely distinguishable. The wavelength to tile size ratios are the following: \diamond : 2, \circ : 1, \times : 0.5.

4.3. Dependence on the Wavelength Compared to the Image Size

If the wavelength T of the wave in the picture segment is chosen as the first parameter, it can be either larger than the tile size N , or at most one third of it, thus the tile size to wavelength ratio was selected to be between $1/4$ and 3 . In the case of N to T being $1:4$, the resulting wave distribution starts to resemble a plane; this was the reason the parameter space was extended to $1:10$ limit, as in that case the limiting behavior could also be studied. In addition, the other limit was extended compared to the realistic case to see whether there is a limiting behavior in the small wavelength domain as well.

The offsets can also play important role if they are larger than 0.001 , thus, as a first step, the second parameter was selected to be B . The characteristic curves for both the waves and the limiting planes are given in Figures A1 and A2 in the Appendix, both for straight and for circular wavefronts. Figure A1 gives the three projections of the $S_{str}(\ln q, T/N)$ plots, while Figure A2 shows 3D perspective. It can be seen that in the large $T:N$ domain the curves follow their corresponding limits' characteristics (i.e., the planes for the straight wave front and the cones for the circular wavefront). The small $T:N$ ratio part of the curve oscillates around a value with decreasing amplitude in both $\ln q$ and S_{str} , resulting in ribbon-bow-like, eight-shaped loops in the $S_{str}(\ln q)$ plots.

The second parameter can also be the phase. With a very small offset $B = 10^{-10}$; the characteristic curves can be seen in Figures A3 and A4. It is clearly visible that phase influences the structural entropy, and spatial filling factor values, especially in the lower $T:N$ domains.

Another possibility for the other parameter beside the wavelength is the tilt slope to wavelength ratio. The results are summarized in Figures A5 and A6 in the Appendix. It can be seen that, for wavelengths larger than the tile size, the tilt does not have real influence, however, in the small wavelength direction, the characteristic curve oscillates much more vehemently if tilt is present than in the tiltless cases, and also some points with very high curvature—turning points—arise in the case of the waves with tilt.

4.4. Dependence on the Phase and Center-Shift

In the case of the colonoscopy image categorization, besides dividing the pictures into fixed tiles, applying sliding tiles and analyzing the characteristics of the arising $S_{str}(q)$ or $S_{str}(\ln q)$ point set is another option. This can be represented as a phase scanning in the case of the waves with straight wavefront, and as moving the center in the case of the circular waves.

For the studies of the phase, the fourth dimension can be either the wavelength, or the offset (constant or linear). For all three cases (i.e., wavelength, offset, and tilt), only the 3D plots are given in Figure A7 of Appendix A. The non-varying parameters were set the following way. The offset in the first image, where the wavelength varied, was set to be negligible (10^{-10}). The wavelength in the second column, where the offset varied, was set to $2N$, as for wavelength values smaller than N neither the structural entropy nor the spatial filling factor had dependence on wavelength, in the case of the straight wavefront, and very simple sinusoid wave-like characteristic curves arise in the case of the circular wavefronts. In both cases, if the wavelength is larger than the tile size, the loops formed on the $S_{str}(\ln q)$ plot are turned back at a point, resulting in hook-like lines, which seem to have derivative singularities, or at least very rapid variation in their gradients. As can be seen in the plot with varying offset, if the offset becomes negligible, this turning point becomes a simple inflection on the characteristic line. In the case of the linear offset (tilt) of the third column, the hook-like behavior becomes rather loop-like.

However, as both the offset and the tilt can easily be removed from a picture by image processing means (the offset by a counter-offset, i.e., by setting the minimum of the pixel intensities as 0, whereas the tilt by removing a mean-filtered version of sufficiently large filter size from the image), it is more advisable to remove these unnecessary information sources from the picture.

Moving along a diameter of a circular wave results in the characteristic lines given in Figures A8 and A9. Both the $\ln q$ and the S_{str} curves are periodic at the higher center shift domains, and, similar to the straight waves, they have hook-like characteristics, if the wavelength is larger than the tile size. In the case of smaller wavelengths, the oscillations are of much smaller magnitude.

If the center is moved in the other direction as well, the upper hook becomes more and more asymmetric, and a shift also appears, as is demonstrated in the second row subplots of Figures A10 and A11 of Appendix A.

As in the case of scanning a row of a picture the tile size is usually smaller than the wavelength, the large wavelength curves are of greater interest from the point of view of polyp detection. In addition, as the center of the elliptical waves are generally in the more distant parts of the image, i.e., practically never in the same frame as the polyp, the offsets larger than the tile size are of more interest. In these cases, as can be seen in Figures A10 and A11, the straight waves model the behavior of the circular waves very well.

The dependence directions of the straight waves and the axis ratio of the elliptic waves can be seen in Figure A12, and in its 3D version in Figure A13. We can conclude that the direction does not influence the characteristic curves of straight waves if the wavelength is below the tile size. The hook-like characteristic curves with smaller or larger asymmetry remain for both the straight and the circular waves, and for the elliptical waves the ratio of the axes becomes negligible if the center is shifted out of the tile.

To summarize this subsection, scanning the picture with a moderately large window along a line or column can be of greater interest from structure detection point of view. In this case, for larger

distance of the center of the elliptical waves, they behave similarly to the waves with straight wavefront: periodic, hook-like characteristic curves are usual, which can be distorted by other parameters.

4.5. Hemispheres

In the case of the sphere or ellipsoid segments, the parameters we selected are the radius to tile size relation, the ratio of the axes of the ellipsoid and the center shift. The characteristic curves can be seen in Figures A14 and A15 of Appendix B. It can be seen that the hemispheres or half ellipsoids have very low structural entropy because, in the case of a sphere with radius to tile size ratio 0.3, a very large part of the picture is completely flat and dark, with 0 entropy (and thus 0 structural entropy). In addition, the radius and axis ratios influence only the spatial filling factor; the structural entropy does not change as long as the whole ellipsoid is within the tile.

If the shadow part of the picture is also included into the model, i.e., the Gaussian like functions (Equation (16)) are subtracted from the hemispheres, the structural entropy of course becomes much larger, as the part of the image with zero pixel intensity becomes very small. The results are given in Figures A16 and A17 of Appendix B. The characteristic curves of the Gaussian-like structures and their negative counterpart are also given, but only in the 3D plot form.

The Gaussian distributions are on their theoretical characteristic line for that central region, where the $S_{str}(\ln q)$ points are constant, and deviate from their theoretical value if significant part of the distribution is outside of the tile (the deviation starts to be visible at the shift of about $5R$ and in very small variation cases, if the shift is larger than about eight times the radius, the structural entropy's deviation starts to grow, and then its value sinks to the origin of the plot).

The structural entropy plots of the hemispheres with shadow have big loops if their radius is small, and hooks start to form with the increasing of the size of the polyp model. The Gaussian-like structures with higher power α were also tested, however, their result did not differ much from the Gaussian case, only the loop area became a little bit smaller, as the power increased.

In addition, the depth of the shadow, or the shadow to ellipsoid height ratio, is interesting. If the shadow is much deeper than the polyp, we arrive at the distant part of the bowel, the tunnel, which almost always has a darkening part and a turning, which often appears as hemisphere or similar object in the cross section of the distribution. The results are plotted in the Appendix in Figures A18 and A19. The pictures show that, as the shadow deepens, the hooks at the sides of the hemisphere decrease, moving inward, toward the point with 0 center shift. In addition, more smaller hooks appear in the inner domain.

Tilt is important in this case, too. The effects of introducing and increasing tilt are shown in Figures A20 and A21 of Appendix B. It is clearly visible that the distance of the received $S_{str}(\ln q)$ points decrease from the origin, and other little hooks emerge in the middle region of the plot. This is of course not always this visible: if the radius is too small, the hooks disappear here as well, such as in the case of the blue curve in Figure A16. This means that, if the polyp is much smaller than the window used for scanning, it behaves completely differently from the ones with radii more similar to the tile size.

The components of the hemisphere with Gaussian shadow were also studied and the effect of the tilt to their properties are given in Figures A22–A27. In the case of the hemispheres, the tilt increases the structural entropy and decreases the spatial filling factor, thus elevating the hemisphere's curve from the $\ln q$ axis. In the case of the Gaussians, the magnitude decreases with increasing tilt. In those center shift values, where the sphere dominates in the tile (i.e., when the tile center is around the center of the sphere), the movement toward the origin is less than those parts that contain picture domains with 0 value.

4.6. Superpositions

The superpositions of the semi-ellipsoids and the waves have rather complex behavior, depending on which component is dominant according to the magnitude and size. If the wavelength is larger than the tile size, and the hemisphere diameter is smaller, then the setup is very similar to a roundish polyp largely protruding into the bowel volume. If the wavelength is smaller, the arising picture is similar to those flatter polyps, which are sitting at the bends of the bowel wall, making these bends only slightly thicker at a given region.

Characteristic curves for such superstructures can be seen in Figures A28 and A29 of Appendix B.

The resulting characteristic curves are also of two types: for the larger wavelengths, the periodic behavior dominates, hooks similar to the ones in Figure A8 appear, and the sphere segment and its shadow causes only slight asymmetries. If the wavelength is smaller than the tile, the two components can decrease each other's structural entropy and filling factor.

The Rényi entropy based structural entropy and the spatial filling factor is able to distinguish parts of a superstructure, if they are multiplied and not added. In the case of a multiplicative superstructure, the S_{str} and $\ln q$ values of the components are simply added together. Unfortunately, in the case of additive superstructures, the S_{str} and $\ln q$ values of the component structures can only be detected, if one of the structures is dominant. It might also be possible to detect components of the superstructures, if they are of different characteristic lengths, and wavelet analysis or other filtering method is used to separate the different characteristic lengths [59].

4.7. Summary of the Artificial Surface Characteristic Curve Properties

For a better visibility, we summarize the previous results in Table 1, concentrating on how other parameters influence the center-shift curves.

Table 1. The effect of the increasing of various parameters to the characteristic curves for the waves and the hemispheres in the case of off-center shift being the first parameter.

Parameter	Wave	Sphere with Shadow
General shape	periodic loops	2 hooks, M-shape
Radius	-	different position
Wavelength	different period	-
Offset	increased magnitudes	Decreased magnitudes
Tilt	broken symmetry of the shapes	smaller curves
Height ratio	-	decreased hook size
-	-	more hooks
Direction	shallower hooks	no systematic effect
Tile size	no effect for large tiles	no effect for large tiles

4.8. Typical Characteristic Curves of Real Images

In the case of real images, instead of fixed tiles, we applied the sliding tile method suggested in Section 4.4. In Figure 10, the two cuts in Figure 3 are scanned with tiles of size 50 by 50. The characteristic curves of these cuts are very similar to the ones given in Figure A16, however, as the distant, dark part of the bowel is also similar to the hemisphere with shadow model profile in some cases (see pixels 200-600 in picture 83, row 350), such occasions may cause misinterpretation of the $S_{str}(\ln q, i)$ curves, and thus false positive categorization.

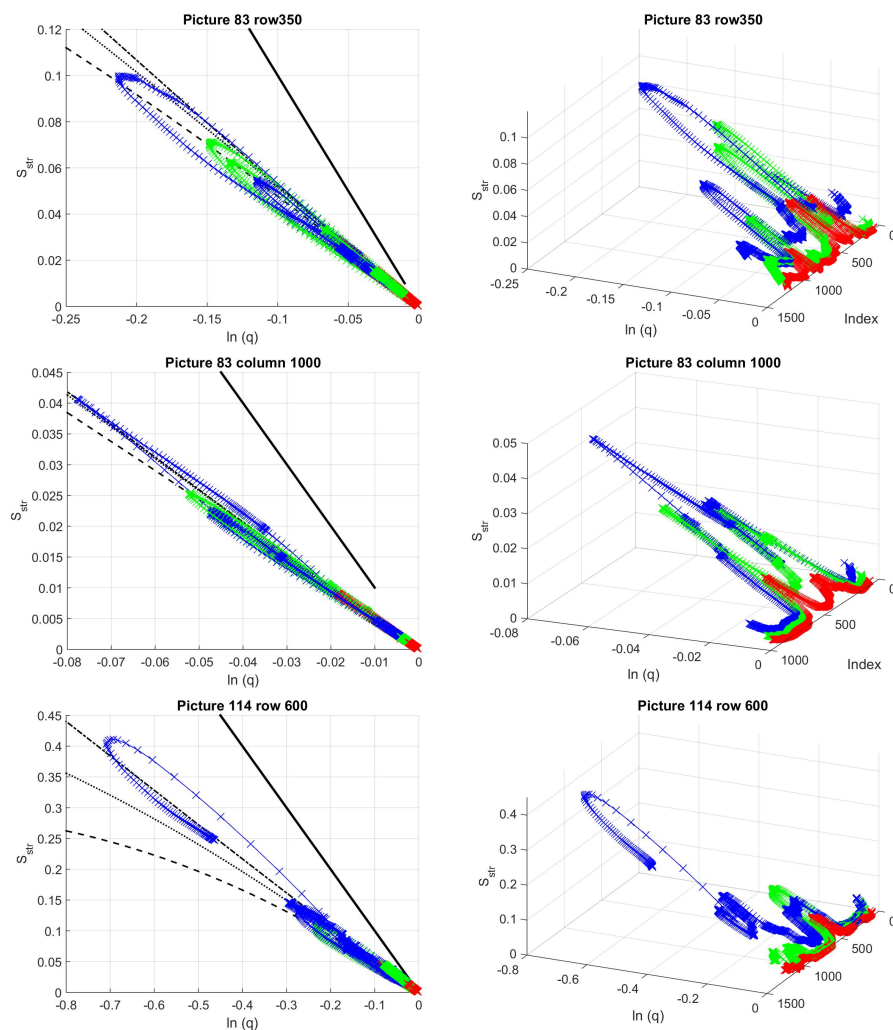


Figure 10. Structural entropy across real polyps. The center of the sliding tile is at the curves plotted in Figure 3. Fixed tile size of 50 by 50 pixels. The polyps are located approximately between pixel indices 900 and 1100 for picture 83, row 350, between 300 and 500 for the same picture's 1000th column, and between 600 and 1100 for picture 114, row 600.

In Figure 11, two scans without polyp are given as an example, one with clearly distinguishable waves, and the other across a polyp-like appearing curvature of the bowel, where the shadows are much larger, and the spherical characteristics are much weaker than in real polyps. These polyp mimicking parts with much more emphasized shadows generally have larger loops than the real polyps. The waves do not have such expressed, curvy hook-like behavior as can be seen in Figures A10 and A11.

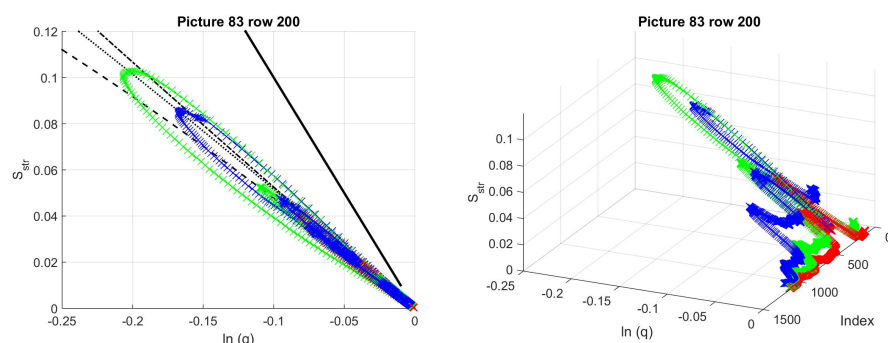


Figure 11. Cont.

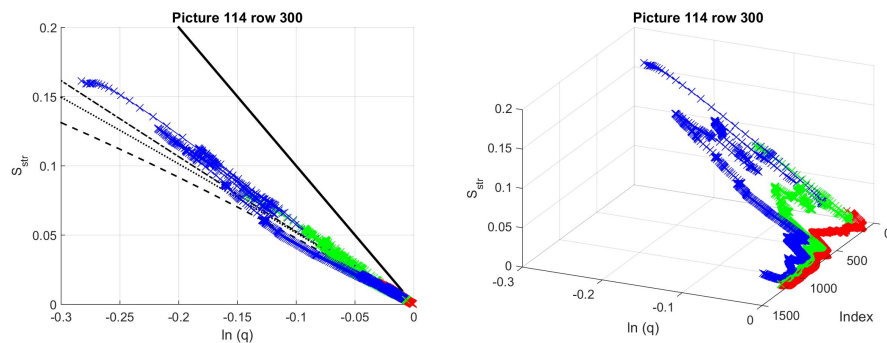


Figure 11. Structural entropy across pictures in scans without polyps. The center of the sliding tile is at the rows 200 and 300 of the pictures in Figure 2. Fixed tile size of 50 by 50 pixels. The dark part mimicking polyp is approximately located between pixel indices 300 and 700 in picture 83, row 200, while the waves are between pixels 200 and 800 in picture 114, row 300.

4.9. Real Picture versus Model: The Applicability

As an example, part of a real picture was studied. As we suggest removing the offset and the tilts from the image by shifting the 0 level and applying a larger scale mean filtered version of the image, first we show their effects on the structural entropy–filling factor plots. We used the same 1000th column of image 83 from database [24], as in the previous section. The tile size remained 50 by 50.

The image preprocessing algorithm consists of only the following steps: reflection removing, histogram stretching and removing of the mean-filtered background pattern [60]. As the image size is around 1000 in both directions, we applied filter sizes of 100 by 100 and 200 by 200. The pictures, the cross section cuts and the structural entropy–filling factor–scanning window center position plots are given in Figure 12. The average diameter of the polyp is also around 200–250 pixels, thus the background generated with the smaller filter size suppressed the sphere-like characteristics of the polyp, as can be seen in the first and second subplots of Figure 12. The symmetrizing effect of removing the tilt from the background can also be seen, even with such rough background subtracting algorithm.

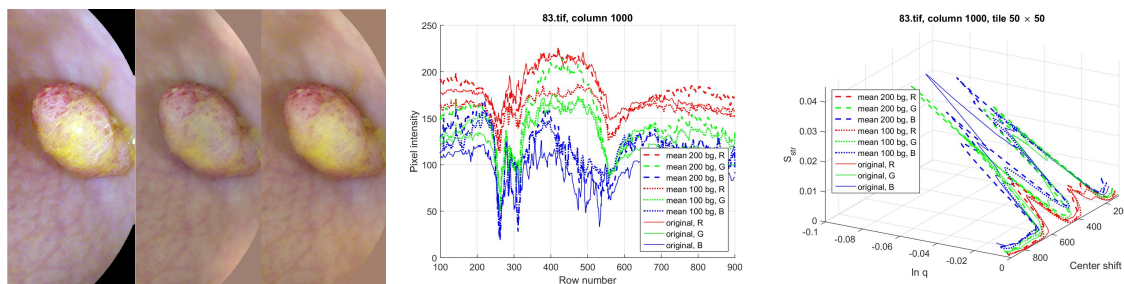


Figure 12. (First plot) A cut from the not processed, and background subtracted versions of picture 83.tif from database of [24]. The second picture segment's background is generated by a 100×100 sized mean filter, the third slices by a 200×200 sized one. (Second plot) The cross sections at the studied row before and after background subtraction. (Third plot) Structural entropy across the 1000th column of the original picture, approximately at the middle of the cuts. Tiles of 50 by 50 pixels.

As the image preprocessing method was applied to the whole picture, not only to the shown segment, there is still an offset in the color channels, about 150 in the red channel, 80 in the green, and 50 in the blue. We used a simplified model to demonstrate that the characteristic curves of model systems are similar to those of the real images, even though many aspects of the real picture, such as the fine patterns, the details of the background, or the yellowish spot that causes a depression in the middle of the polyp in the blue channel of the original picture, are neglected.

In the model system, a hemisphere was used with Gaussian, or higher-order exponential shadow, and flat, constant offset. The offset values were chosen according to the picture color channels' offsets.

The polyp diameter was selected to be 120 pixels, and the height of the hemisphere to be 80 for the red, 120 for the green and 100 for the red channel. The shadows were Gaussian in both the red and the green picture parts, and third-order exponential for the blue part. The depth of the shadows were adjusted to be 30, 60 and 80 for the R, G, and B channels, respectively.

The resulting characteristic curves for the three color channels can be seen in Figure 13. From the $S_{str}(\ln q)$ plot, one might conclude that the models fit the real image very well, however, from the 3D curves, it can be seen that the fine structure of the structural entropy and filling factor around the shadow-polyp transition is not too well represented. As the fine-scale behavior of the pixel intensity distribution is not studied, these deviations may be attributed to the smaller sized patterns, however, this aspects needs further investigation. As the fine-scale pattern is useful in the case of pit pattern based classification, we decided not to study this problem in this article.

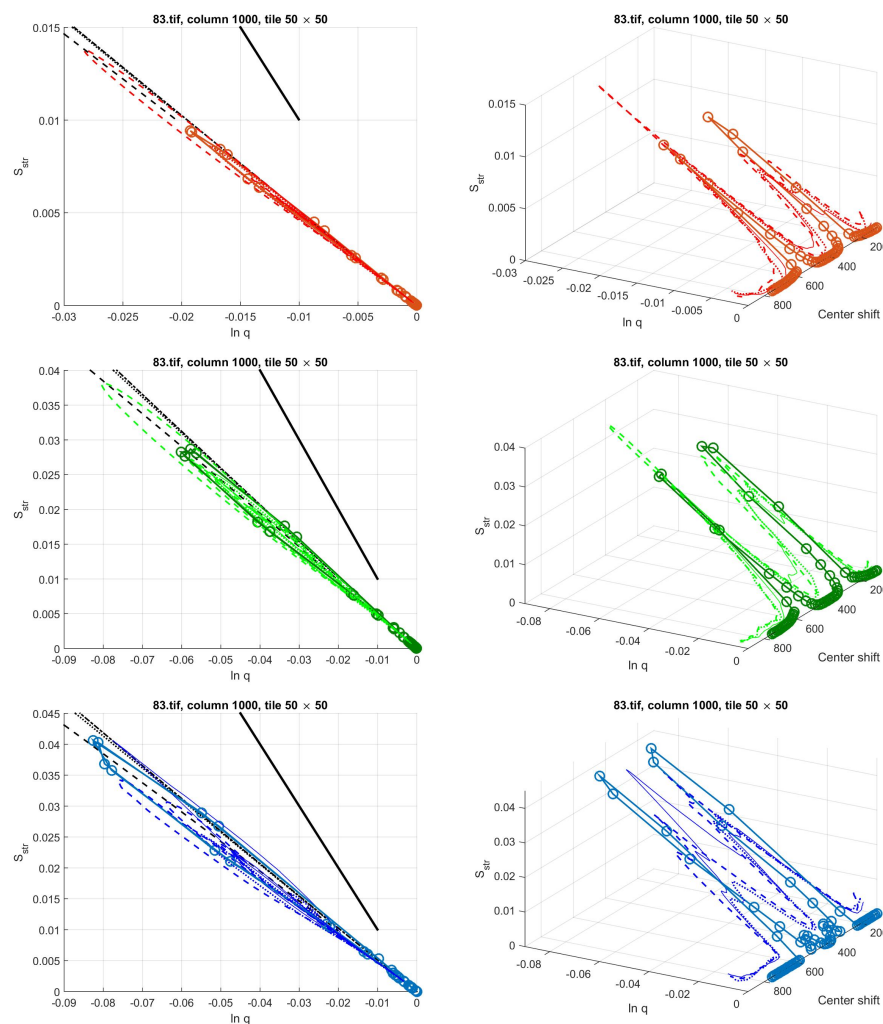


Figure 13. Rényi entropy based structural entropy–spatial filling factor–scanning window center curves for picture 83.tif from database of [24] in continuous lines. The preprocessed-image segments are given in Figure 12 in dashed and dotted lines. The structural entropy and filling factor values of the model system of hemisphere with Gaussian shadow and constant offset model systems. **First row:** red channel, **Second row:** green channel, **Third row:** blue channel. The shadows in the red and green channels are modeled with Gaussian functions of $\alpha = 2$, while for the blue channel, a higher, $\alpha = 3$ rd order polynomial was used in the exponential function.

5. Conclusions

The shape of a probability distribution can be characterized by Rényi entropy based quantities, which are called spatial filling factor and structural entropy. Although the name is similar to the graph neighborhood relation based structural entropy that appeared earlier in the literature, these quantities of the same name are significantly different.

The Rényi entropy based structural entropy uses the native probability distribution of an image, i.e., its pixel intensity distribution, simply normalized in a manner that it would form a probability distribution. This structural entropy is from one point of view more complex than the graph theory based one, as it uses generalized, Rényi entropy differences instead of Shannon entropy. The main advantage of the method, however, is its simpleness. The probability distribution used in the entropies is straightforward, easy to generate, and does not need topological knowledge and neighborhood statistics, which might change with the resolution. These Rényi entropy differences can possibly be used as input parameters for fuzzy, support vector machine, or other metaheuristic or learning algorithms.

The application of the Rényi entropy based structure parameters requires plotting the structural entropy as the function of the filling factor, and comparing the result with existing characteristic curves. This process is easy to be visualized, but might be rather hard to understand and apply. Simply using the two quantities as input parameters for classification methods might lose a lot of information, which lies in the position on the $S_{str}(q)$ map related to characteristic lines. If this information is also to be included into the analysis, previous knowledge about the possible shapes are necessary to know which characteristic lines should be used as references, as these characteristic lines might overlap. This is the main disadvantage of the method.

Characteristic curves of simple distributions such as the Gaussian or exponentially decreasing probability distributions were known for a long time; however, distributions related to structures present on colonoscopy images were never mapped before this study. Here, besides roundish colorectal polyps, different types of waves were also investigated using a rather broad set of possible parameters. The collection of the characteristic curves in the appendices could be used as references or extended and refined if the application deems it necessary.

For some real images, some aspects can be found by using the characteristic structures listed in this contribution. However, we did not pay attention to the fine-scale behavior, as they are not as important in the process of finding a polyp. Superpositions of different types of distributions are rather complicated, if additive and not multiplicative superstructures are studied. We suggest using wavelet-analysis or other, scale sensitive methods to separate the components of an additive superstructure.

To summarize the other findings, the following can be suggested for using structural entropy in image classification methods, especially in colorectal polyp searching cases. Instead of static tiling of the images, and using structural entropy and filling factor as two parameters of the image segment classification—although they provide valuable information about the shape of the distribution—it is more advisable to use sliding tiles and study the thus arising curves on the $S_{str}(\ln q)$ plots. We also suggest removing tilts and offsets from the image segments using simple image processing tools. We demonstrated that simple background subtraction techniques can change the characteristic curves very much without introducing extra information or losing valuable information.

Regarding the characteristic curves, we could conclude the following. The directions of the patterns do not influence the types of the characteristic curves. The size of the tiles also do not influence the results, provided that sufficiently large tile sizes are used. We suggest using larger tile sizes to achieve more stable results. In addition, for larger wavelengths, in realistic cases, the elliptical waves produce similar characteristic curves to those of the waves with straight wavefront, and they are both suitable to model the waves on the bowel walls.

Author Contributions: Conceptualization, J.P.; methodology, J.P.; software, S.N.; validation, S.N.; resources, B.S.; and writing—original draft preparation, S.N. and B.S; correction, S.N. and J.P.

Funding: This research was funded by the New National Excellence Programme of Hungary grant number ÚNKP-18-3.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Characteristic Curves for Waves

Appendix A.1. Waves, Wavelength to Tile Size Ratio, Offset, Limiting Behaviors

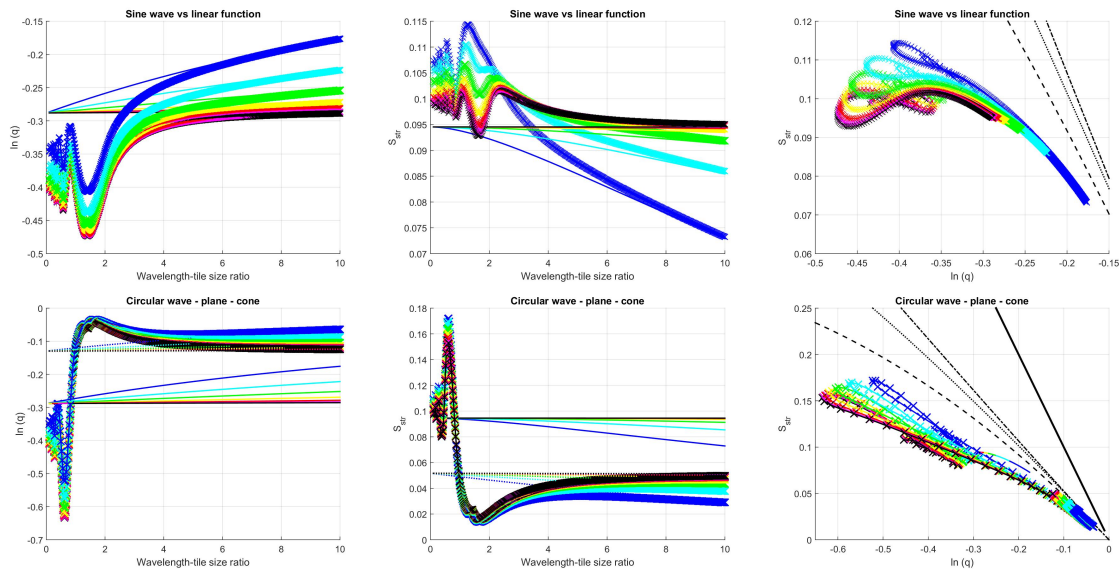


Figure A1. Characteristic curves of waves with straight (**upper row**) and circular (**lower row**) wavefronts, if the parameters are the wavelength to tile size ratio and the constant offset. The colors denote the following constant offsets: $B = 0.1$, blue; $B = 0.05$, cyan; $B = 0.025$, green; $B = 0.0125$, yellow; $B = 0.00625$, red; $B = 0.003125$, magenta; $B = 0.0015625$, black. As a reference, the characteristic curves for planes with the same gradient as the $x \rightarrow 0$ gradient of the waves are plotted with solid lines without markers, whereas the cones corresponding to the initial slope of the circular waves are plotted in dotted lines.

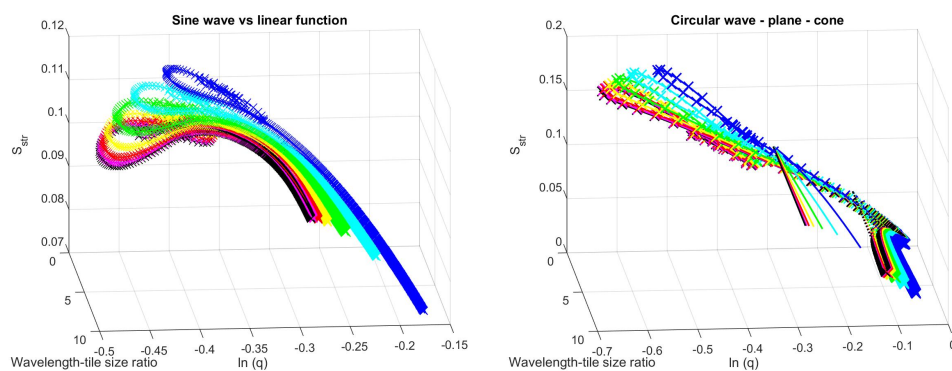


Figure A2. 3D view of Figure A1.

Appendix A.2. Waves, Wavelength to Tile Size, Phase

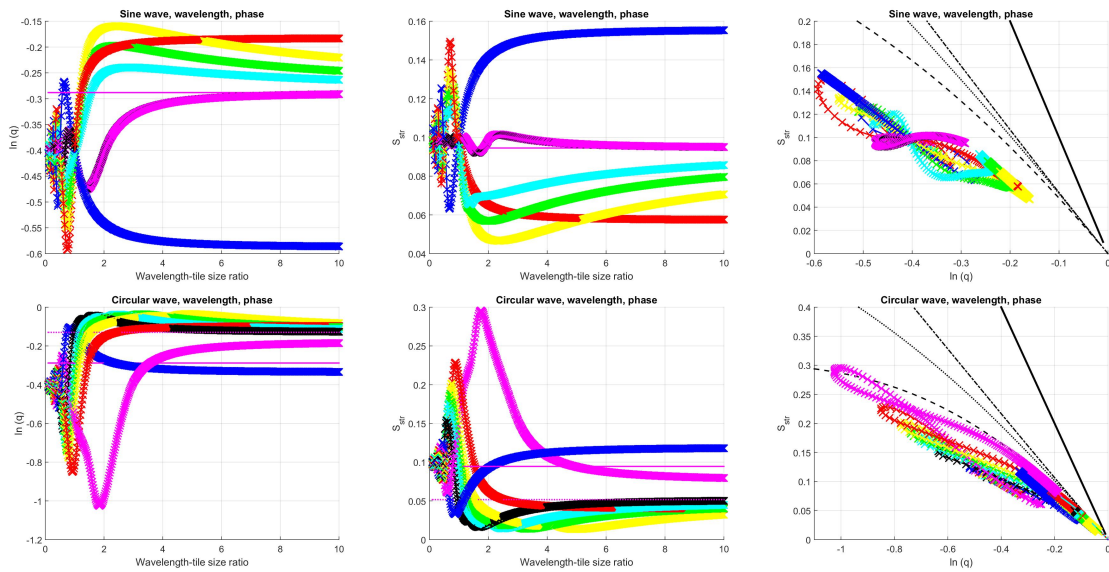


Figure A3. Characteristic curves of waves with straight (**upper row**) and circular (**lower row**) wavefronts, if the parameters are the wavelength to tile size ratio and the phase φ . The colors denote the following: $\varphi = 0^\circ$, black; $\varphi = 30^\circ$, cyan; $\varphi = 45^\circ$, green; $\varphi = 60^\circ$, yellow; $\varphi = 90^\circ$, red; $\varphi = 180^\circ$, magenta; $\varphi = 270^\circ$, blue. As a reference, the characteristic curves for planes with the same gradient as the $x \rightarrow 0$ gradient of the waves are plotted with solid lines without markers, whereas the cones corresponding to the initial slope of the circular waves are plotted in dotted lines.

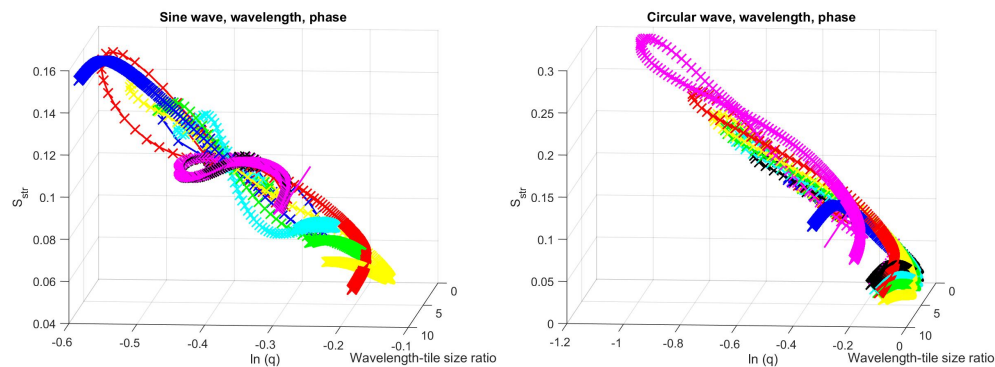


Figure A4. 3D view of Figure A3.

Appendix A.3. Waves, Wavelength to Tile Size, Tilt Gradient

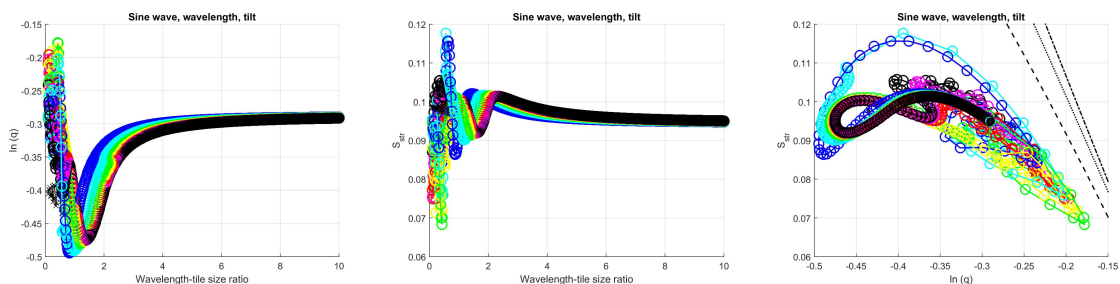


Figure A5. Cont.

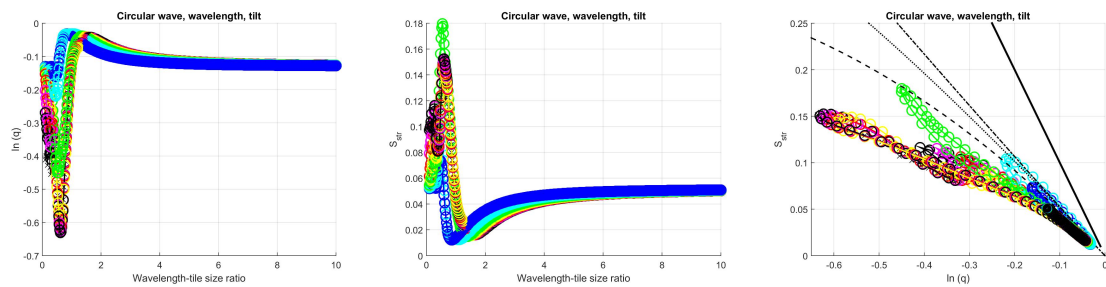


Figure A5. Characteristic curves of waves with straight (**upper row**) and circular (**lower row**) wavefronts, if the parameter is the wavelength to tile size ratio and the tilt gradient. The colors denote the following linear offset slopes to wavelength ratios: $B/T = 1$, blue; $B/T = 0.5$, cyan; $B/T = 0.25$, green; $B/T = 0.125$, yellow; $B/T = 0.0625$, red; $B/T = 0.03125$, magenta; $B/T = 0.015625$, black. As a reference, the characteristic curves waves without tilt are plotted in thinner lines with different \times markers.

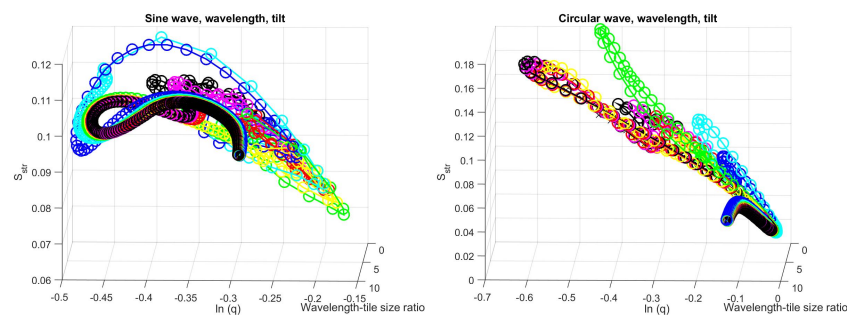


Figure A6. 3D view of Figure A5.

Appendix A.4. Waves, Phase and Wavelength, Offset, Tilt

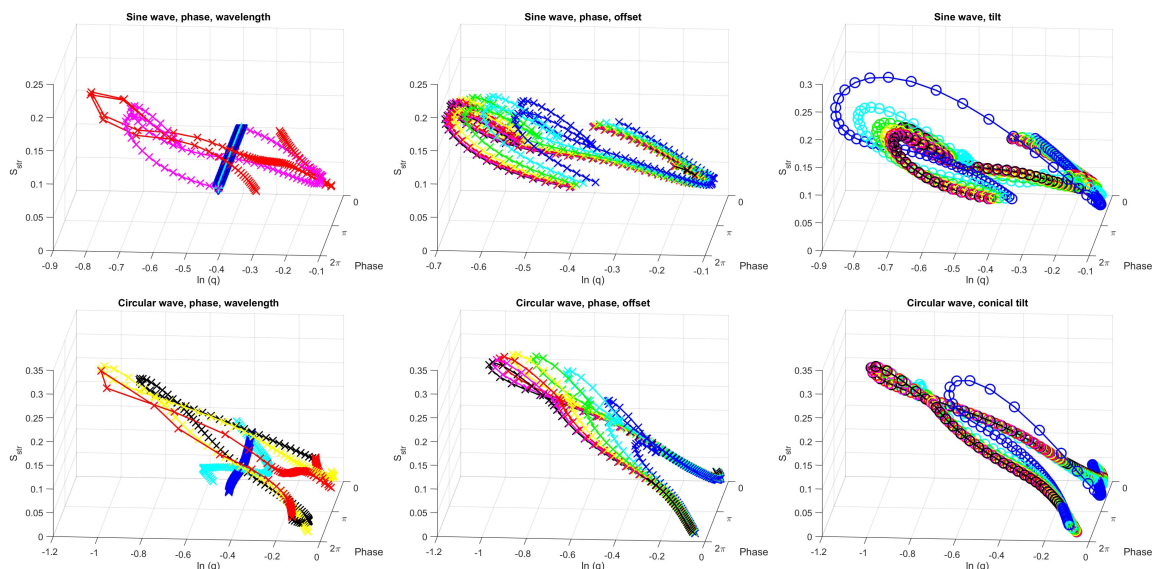


Figure A7. Characteristic curves of waves with straight (**upper row**) and circular (**lower row**) wavefronts, if the parameter is phase. The other parameters are the following. **First column:** wavelength. The colors denote the following wavelength to tile size ratio: $T/N = 0.1$, blue; $T/N = 0.5$, cyan; $T/N = 1$, black; $T/N = 2$, yellow; $T/N = 10$, red. **Second column:** constant offset. The colors mean the following: $B = 0.1$, blue; $B = 0.05$, cyan; $B = 0.025$, green; $B = 0.0125$, yellow; $B = 0.00625$, red; $B = 0.003125$, magenta; $B = 0.0015625$, black. **Third column:** linear offset. The colors mean the following: $B/T = 1$, blue; $B/T = 0.5$, cyan; $B/T = 0.25$, green; $B/T = 0.125$, yellow; $B/T = 0.0625$, red; $B/T = 0.03125$, magenta; $B/T = 0.015625$, black. The markers \times denote the zero-tilt limit, while the circles show the results for the tilted waves.

Appendix A.5. Wave, Center Shift (One Direction), Wavelength

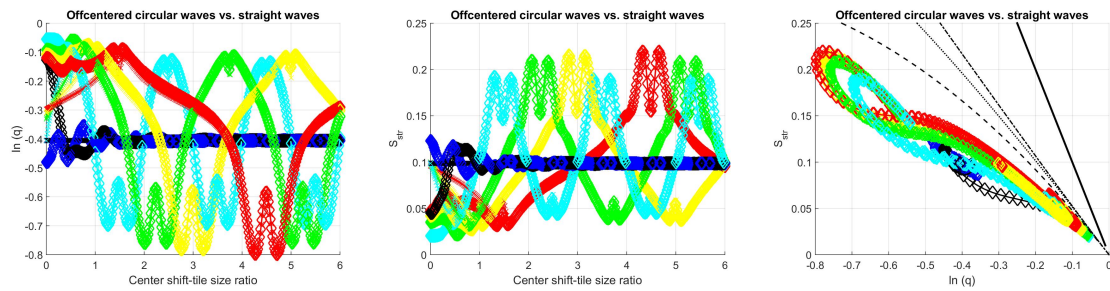


Figure A8. Characteristic curves of circular wavefronts, if the parameter is center shift x_0 and wavelength to tile size ratio. In the first row, the other coordinate of the center shift was set to $y_0 = 0$ and the colors denote the following wavelength to tile size ratios: $T/N = 0.5$, blue; $T/N = 1$, black; $T/N = 2$, cyan; $T/N = 3$, green; $T/N = 4$, yellow; $T/N = 6$, red. The markers are \diamond . As a reference, the characteristic curves of straight waves of same parameters were plotted with thinner lines and smaller markers of shape \times (the phase had to be set to π to match the curves). As a reference, the characteristic curve of a straight wave with the same parameters was plotted with thinner black line and smaller markers of shape \times (the phase was again π).

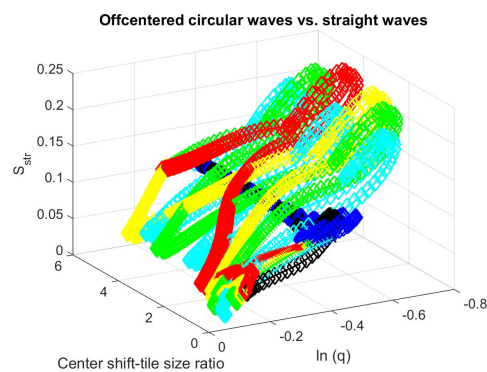


Figure A9. 3D view of Figure A8.

Appendix A.6. Wave, Center Shift, both Directions

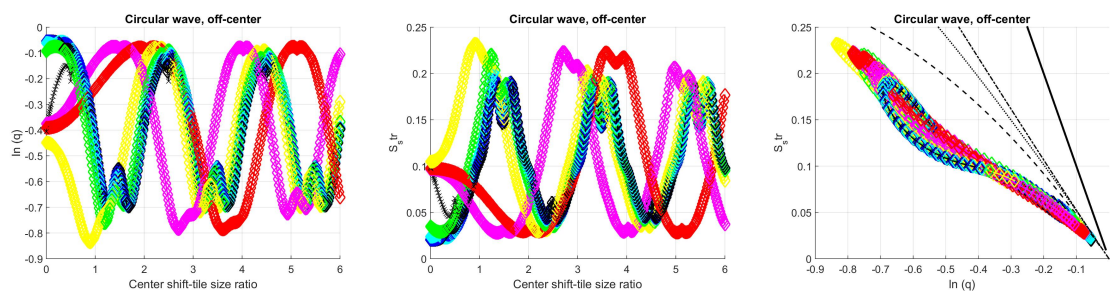


Figure A10. Characteristic curves of circular wavefronts, if the parameter is center shift x_0 and center shift in the other direction. The wavelength to tile size ratio was set to $T/N = 2$, and the colors mean the following other coordinate of the center shift to tile size ratios, $y_0 = 0$ and the colors denote the following: $y_0/N = 0$, black; $y_0/N = 1/8$, blue; $y_0/N = 1/4$, cyan; $y_0/N = 1/2$, green; $y_0/N = 1$, yellow; $y_0/N = 2$, magenta; $y_0/N = 4$, red. The markers are \diamond . As a reference, the characteristic curve of a straight wave with the same parameters was plotted with thinner black line and smaller markers of shape \times (the phase was again π).

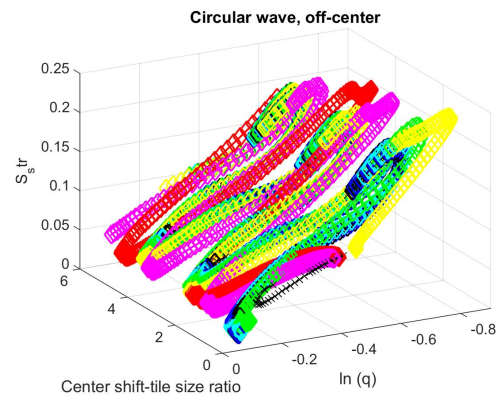


Figure A11. 3D view of Figure A10.

Appendix A.7. Waves, Center Shift, Direction

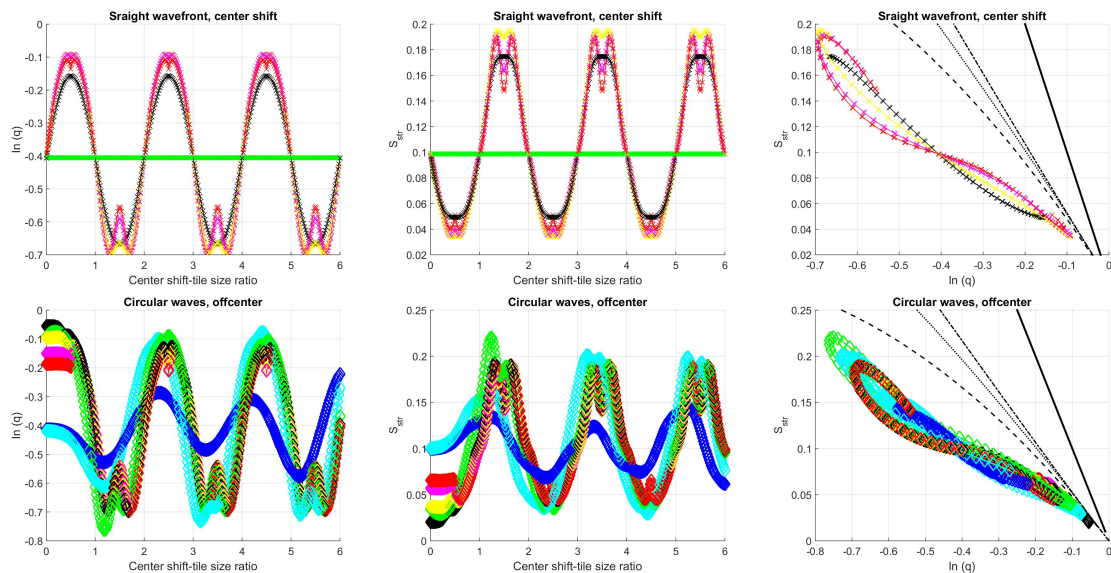


Figure A12. Characteristic curves of straight (upper row) and circular (lower row) wavefronts, if the parameters are the center shift and the direction of the wavefront. The colors denote the following ratios: $T_y/T_x = 1/8$, blue; $T_y/T_x = 1/4$, cyan; $T_y/T_x = 1/2$, green; $T_y/T_x = 1$, black; $T_y/T_x = 2$, yellow; $T_y/T_x = 4$, magenta; $T_y/T_x = 8$, red.

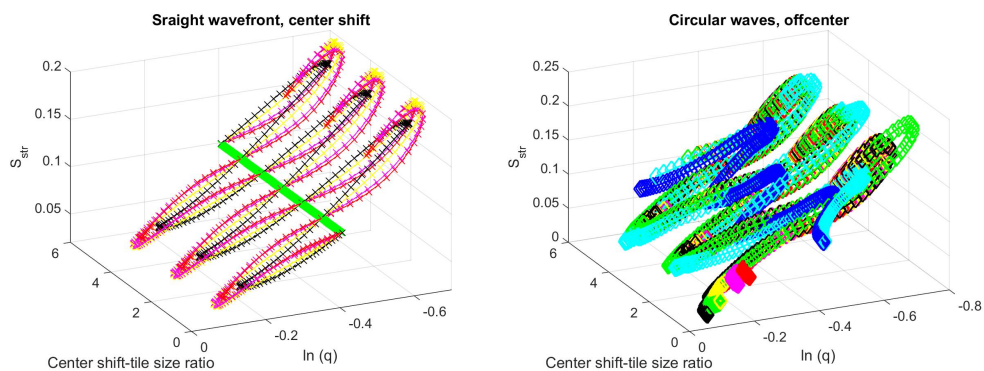


Figure A13. 3D view of Figure A12.

Appendix B. Characteristic Curves for Semi-Ellipsoids, Hemispheres

Appendix B.1. Semi-Ellipsoid, Center Shift, Axis Ratio

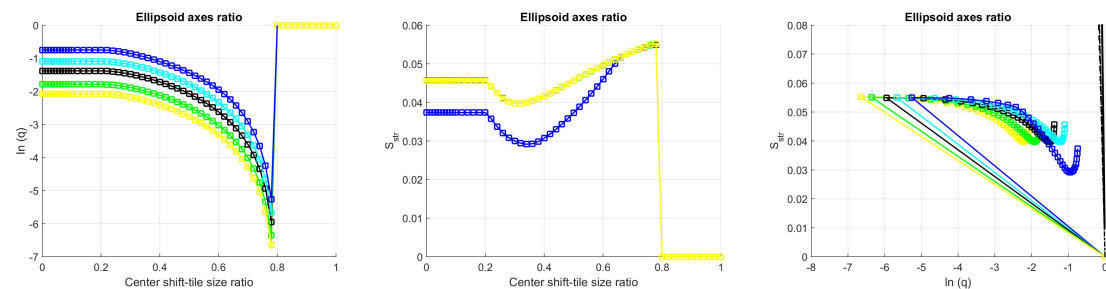


Figure A14. Characteristic curves of hemispheres, if the parameter is center shift x_0 . The other coordinate of the center shift was set to $y_0 = 0$. The colors denote the following axes ratios: $R_y/R_x = 0.5$, blue; $R_y/R_x = 0.75$, cyan; $R_y/R_x = 1$, black; $R_y/R_x = 1.5$, green; $R_y/R_x = 2$, yellow. R was set to $0.3N$, while R_y was set to 1.

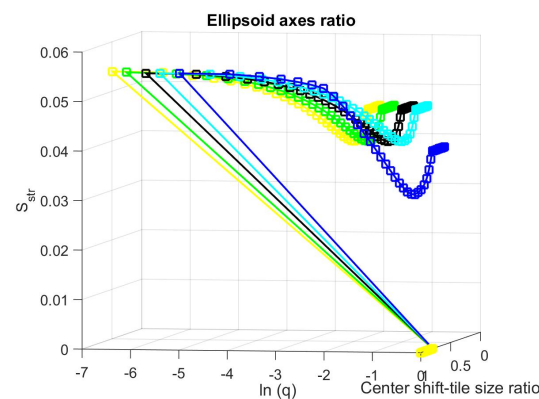


Figure A15. 3D view of Figure A14.

Appendix B.2. Semi-Ellipsoids with Gaussian Subtraction (Shadow), Center Shift, Radius

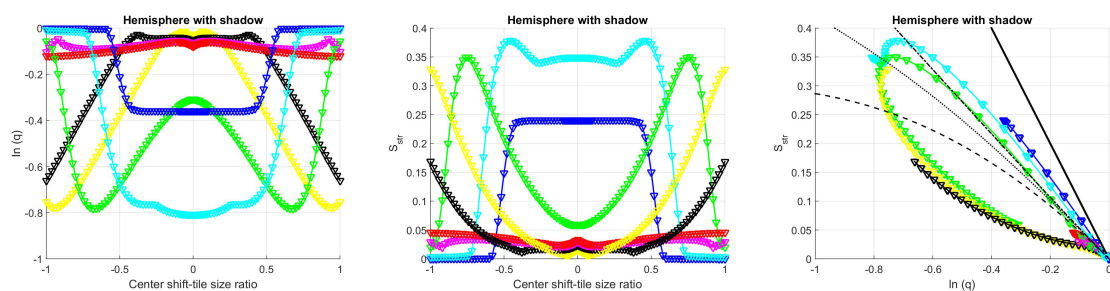


Figure A16. Characteristic curves of hemispheres with Gaussian shadows, if the parameter is center shift x_0 and radius to tile size ratio. The other coordinate of the offset was set to $y_0 = 0$. The colors denote the following radius to tile size ratios: $R/N = 0.1$, blue; $R/N = 0.25$, cyan; $R/N = 0.5$, green; $R/N = 0.75$, yellow; $R/N = 1$, black; $R/N = 1.5$, magenta; $R/N = 2$, red.

Appendix B.3. Components of the Semi-Ellipsoids with Gaussian Shadow, Center Shift, Radius

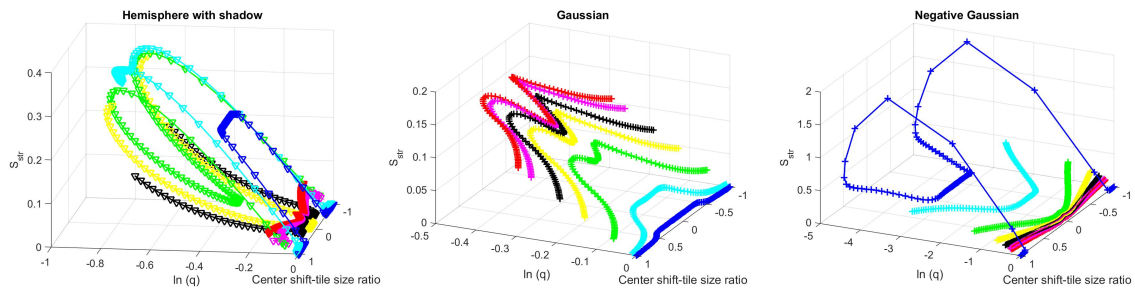


Figure A17. (First subplot) 3D view of Figure A16; (Second subplot) Gaussian functions; and (Third subplot) negative Gaussian functions. The variances of the Gaussians are the same as the square of radii of the hemispheres. The color notation is the same as in Figure A16. The ratio of the Gaussian's height and the radius is 1:3.

Appendix B.4. Semi Ellipsoids with Gaussian Shadow, Center Shift, Ratio of the Two Components: Shadow Depth

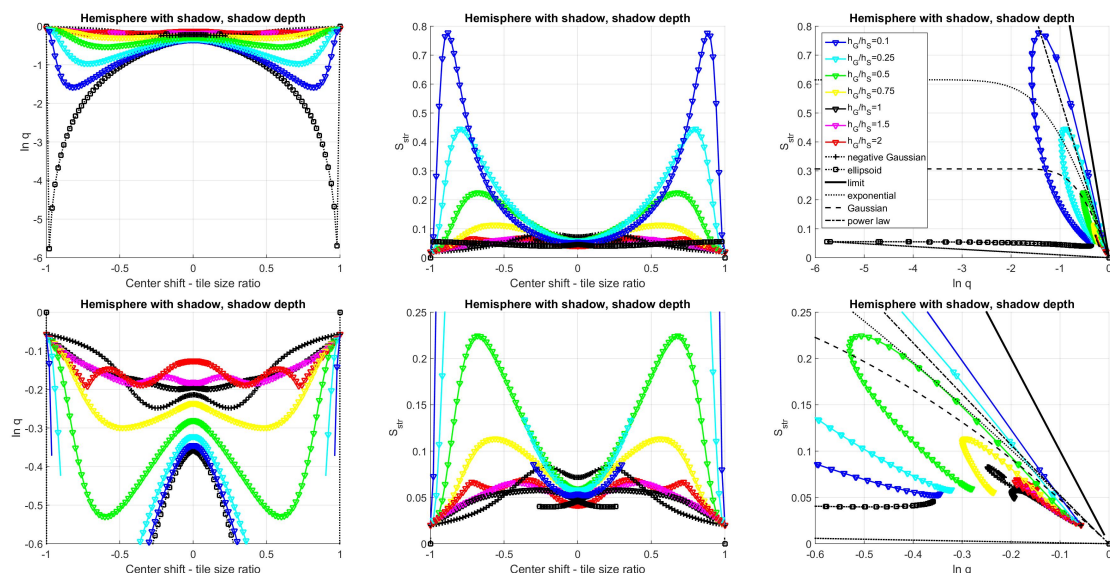


Figure A18. Characteristic curves of hemispheres with Gaussian shadows, if the parameters are center shift x_0 and ellipsoid to Gaussian height ratio h_G/h_S . The lower row is the zoomed version of the upper one. All pictures use the same markers with the same meaning as given in the upper right hand side subplot.

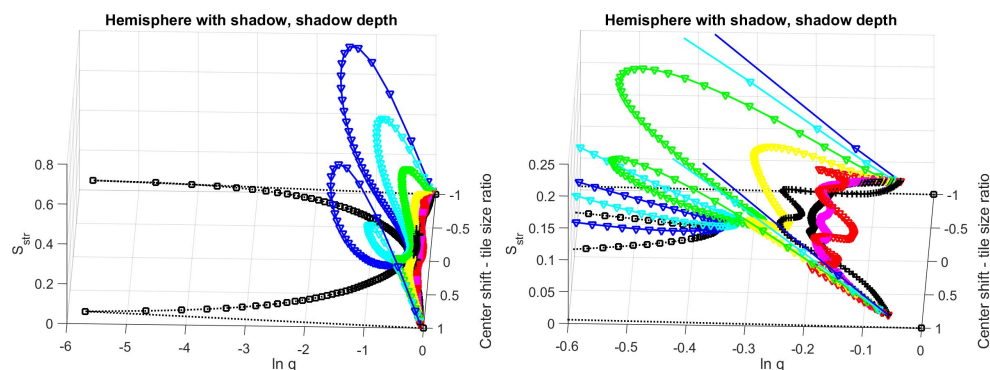


Figure A19. 3D version of Figure A18 both for the full and the zoomed image.

Appendix B.5. Semi-Ellipsoids with Gaussian Shadow, Center Shift, Tilt

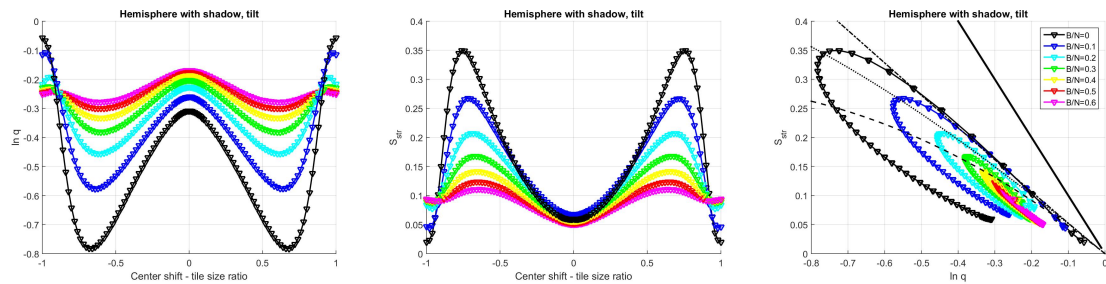


Figure A20. Characteristic curves of hemispheres with Gaussian shadows, if the parameters are center shift x_0 and tilt gradient B .

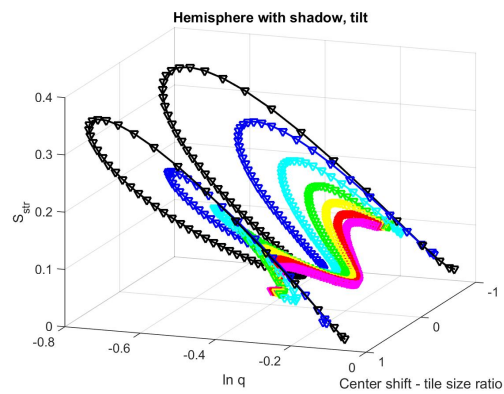


Figure A21. 3D version of Figure A20.

Appendix B.6. Semiellipsoids, Center Shift, Tilt

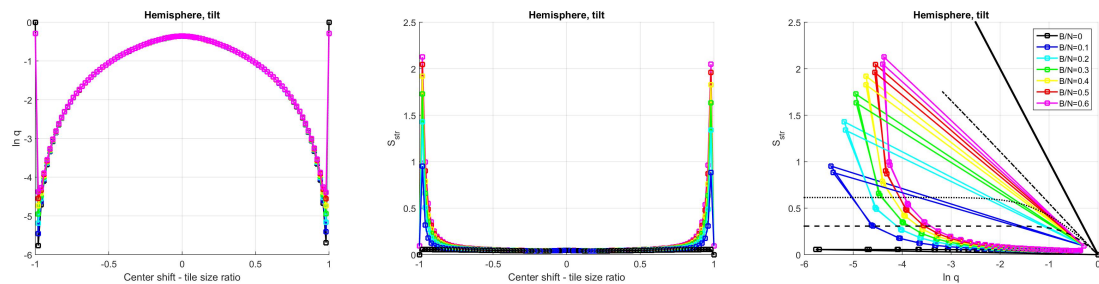


Figure A22. Characteristic curves of hemispheres, if the parameters are center shift x_0 and tilt gradient B .

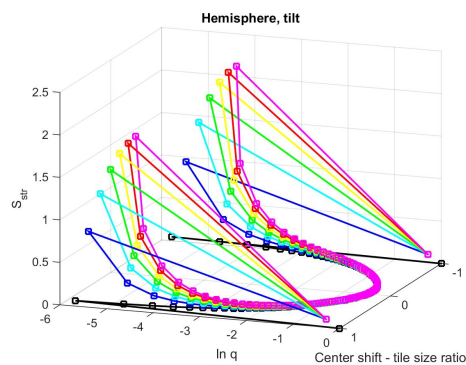


Figure A23. 3D version of Figure A22.

Appendix B.7. Gaussian Distribution, Center Shift, Tilt

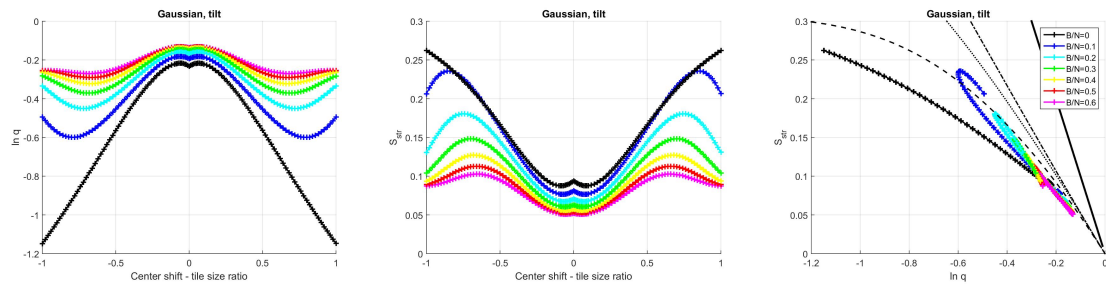


Figure A24. Characteristic curves of Gaussian distribution segments, if the parameters are center shift x_0 and tilt gradient B .

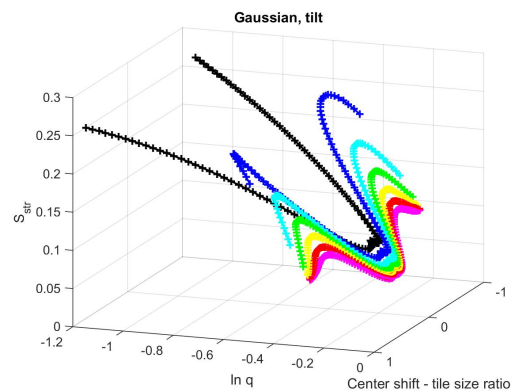


Figure A25. 3D version of Figure A24.

Appendix B.8. the Negative Gaussian Shadow, Center Shift, Tilt

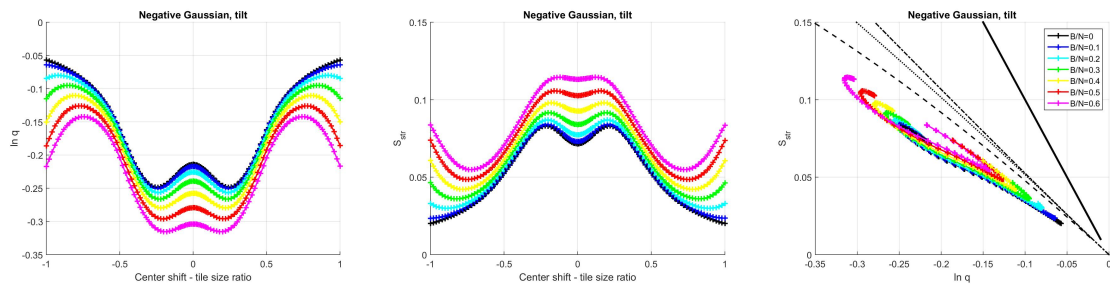


Figure A26. Characteristic curves of the negative Gaussian shadows, if the parameters are center shift x_0 and tilt gradient B .

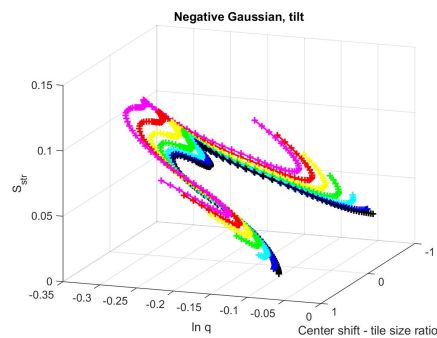


Figure A27. 3D version of Figure A26.

Appendix B.9. Superposition of Wave and Hemisphere with Gaussian Shadow

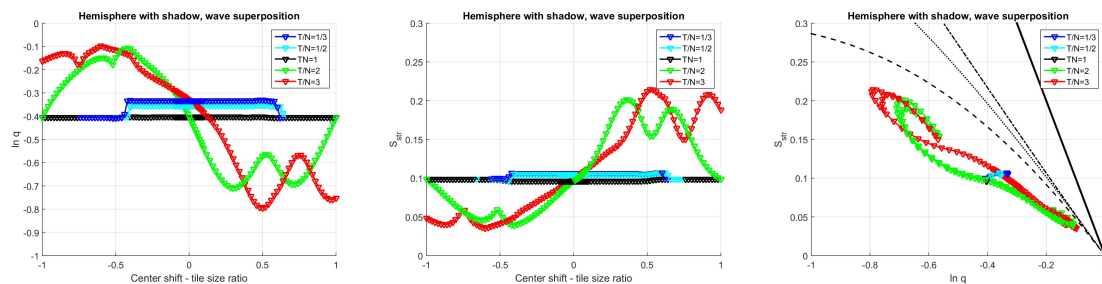


Figure A28. Hemisphere with gaussian shadow superposed to a wave of wavelength T with picture segment size $N \times N$. The radius was set to $0.1N$. The small wavelength curves are similar to polyps sitting on a bending of the bowel wall making it slightly thicker. The large wavelength lines are similar to roundish polyps sitting on slightly curved bowel walls.

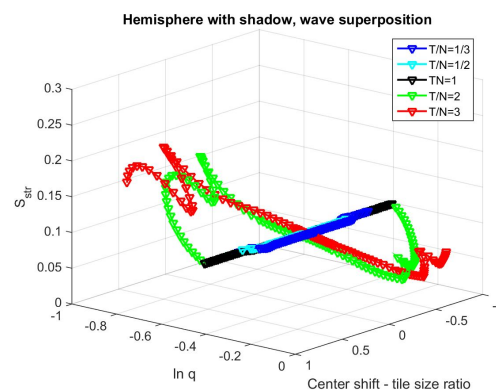


Figure A29. 3D version of Figure A28.

References

1. Burnand B.; Harris J.K.; Wietlisbach, V.; Froelich, F.; Vader, J.P.; Gonvers, J.J. Use, appropriateness, and diagnostic yield of screening colonoscopy: An international observational study (EPAGE). *Gastrointest Endosc.* **2006**, *63*, 1018–1026, doi:10.1016/j.gie.2006.01.051.
2. Levin, B.; Lieberman, D.A.; McFarland, B.; Smith, R.A.; Brooks, D.; Andrews, K.S.; Dash, C.; Giardiello, F.M.; Glick, S.; Levin, T.R.; et al. Screening and surveillance of the early detection of colorectal cancer and adenomatous polyps, 2008: A joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology* **2008**, *134*, 1570–1595, doi:10.3322/CA.2007.0018.
3. Gono, K.; Obi, T.; Yamaguchi, M.; Machida, H.; Sano, Y.; Yoshida, S.; Hamamoto, Y.; Endo, T. Appearance of enhanced tissue features in narrow-band endoscopic imaging. *J. Biomed. Opt.* **2004**, *9*, 568–577, doi:10.1117/1.1695563.
4. Tischendorf, J.J.W.; Wasmuth, H.E.; Koch, A.; Hecker, H.; Trautwein, C.; Winograd, R. Value of magnifying chromoendoscopy and narrow band imaging (NBI) in classifying colorectal polyps: A prospective controlled study. *Endoscopy* **2007**, *39*, 1092–1096, doi:10.1055/s-2007-966781.
5. Stehle, T.; Auer, R.; Gross, S.; Behrens, A.; Wulff, J.; Aach, T.; Winograd, R.; Tautwein, C.; Tischendorf, J. Classification of colon polyps in NBI endoscopy using vascularization features. In Proceedings of the Medical Imaging 2009: Computer-Aided Diagnosis (SPIE 7260), Lake Buena Vista (Orlando Area), FL, USA, 7–12 February 2009; Volume 7260, doi:10.1117/12.808103.
6. Pickhardt, P.J.; Choi, J.R.; Hwang, I.; Butler, J.A.; Puckett, M.L.; Hildebrandt, H.A.; Wong, R.K.; Nugent, P.A.; Mysliwiec, P.A.; Schindler, W.R. Computed Tomographic Virtual Colonoscopy to Screen for Colorectal Neoplasia in Asymptomatic Adults. *N. Eng. J. Med.* **2003**, *349*, 2191–2200, doi:10.1056/NEJMoa031618.

7. Dachman, A.H. Diagnostic performance of virtual colonoscopy. *Abdom. Imaging* **2002**, *27*, 260–267, doi:10.1107/s00261-001-0167-8.
8. Van Gossum, A.; Munoz-Navas, M.; Fernandez-Urien, I.; Carretero, C.; Gay, G.; Delvaux, M.; Lapalus, M.G.; Ponchon, T.; Neuhaus, H.; Philipper, M.; et al. Capsule Endoscopy versus Colonoscopy for the Detection of Polyps and Cancer. *N. Engl. J. Med.* **2009**, *361*, 264–270, doi:10.1056/NEJMoa0806347.
9. Rex, D.K.; Cutler, C.S.; Lemmel, G.T.; Rahmani, E.Y.; Clark, D.M.; Helper, D.J.; Lehman, G.A.; Mark, D.G. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology* **1997**, *112*, 24–28, doi:10.1016/S0016-5085(97)70214-2.
10. Pullens, H.J.M.; Leenders, M.; Schipper, M.E.I.; van Oijen, M.G.H.; Siersema, P.D. No Decrease in the Rate of Early or Missed Colorectal Cancers After Colonoscopy With Polypectomy Over a 10-Year Period: A Population-Based Analysis. *Clin. Gastroenterol. Hepatol.* **2015**, *13*, 140–147, doi:10.1016/j.cgh.2014.04.032.
11. Liu, G.; Yan, G.; Kuang, S.; Wang, Y. Detection of small bowel tumor based on multi-scale curvelet analysis and fractal technology in capsule endoscopy. *Comput. Biol. Med.* **2016**, *70*, 131–138, doi:10.1016/j.combiomed.2016.01.021.
12. Clark, B.T.; Protiva, P.; Nagar, A.; Imaeda, A.; Ciarleglio, M.M.; Deng, Y.H.; Laine, L. Quantification of Adequate Bowel Preparation for Screening or Surveillance Colonoscopy in Men. *Gastroenterology* **2016**, *150*, 396–405, doi:10.1053/j.gastro.2015.09.041.
13. Kayser, G.; Görtler, J.; Weis, C.A.; Borkenfeld, S.; Kayser, K. The application of structural entropy in tissue based diagnosis. *Diagnos. Pathol.* **2017**, *3*, 251, doi:10.17629/www.diagnosticpathology.eu-2017-3:251.
14. dos Santos, L.F.S.; Neves, L.A.; Rozendo, G.B.; Ribeiro, M.G.; do Nascimento, M.Z.; Azevedo Tosta, T.A. Multidimensional and fuzzy sample entropy (SampEnMF) for quantifying H&E histological images of colorectal cancer. *Comput. Biol. Med.* **2018**, *103*, 148–160, doi:10.1016/j.combiomed.2018.10.013.
15. Ribeiro, M.G.; Neves, L.A.; do Nascimento, M.Z.; Roberto, G.F.; Martins, A.S.; Tosta, T.A.A. Classification of colorectal cancer based on the association of multidimensional and multiresolution features. *Expert Sys. Appl.* **2019**, *120*, 262–278, doi:10.1016/j.eswa.2018.11.034.
16. Chaddad, A.; Daniel, P.; Niazi, T. Radiomics Evaluation of Histological Heterogeneity Using Multiscale Textures Derived From 3D Wavelet Transformation of Multispectral Images. *Front. Oncol.* **2018**, *8*, 96, doi:10.3389/fonc.2018.00096.
17. Chaddad, A.; Tanougast, C. Texture Analysis of Abnormal Cell Images for Predicting the Continuum of Colorectal Cancer. *Anal. Cell. Pathol.* **2017**, *2017*, 8428102, doi:10.1155/2017/8428102.
18. Wang, C.; Shi, J.; Zhang, Q.; Ying, S. Histopathological image classification with bilinear convolutional neural networks. In Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, Korea, 11–15 July 2017; pp. 4050–4053, doi:10.1109/EMBC.2017.8037745.
19. Chaddad, A.; Desrosiers, C.; Niazi, T.M. Deep radiomic analysis of MRI related to Alzheimer’s disease. *IEEE Access* **2018**, *6*, 58213–58221, doi:10.1109/ACCESS.2018.2871977.
20. Buchner, A.M.; Shahid, M.W.; Heckman, M.G.; McNeilm, R.B.; Cleveland, P.; Gill, K.R.; Schore, A.; Ghabril, M.; Raimondo, M.; Gross, S.A.; et al. High-Definition Colonoscopy Detects Colorectal Polyps at a Higher Rate Than Standard White-Light Colonoscopy. *Clin. Gastroenterol. Hepatol.* **2009**, *8*, 364–370, doi:10.1016/j.cgh.2009.11.009.
21. Kudo, S.; Hirota, S.; Nakajima, T.; Hosobe, S.; Kusaka, H.; Kobayashi, T.; Himori, M.; Yagyuu, A. Colorectal tumours and pit pattern. *J. Clin. Pathol.* **1994**, *47*, 880–885, doi:10.1136/jcp.47.10.880.
22. Kudo, S.; Tamura, S.; Nakajima, T.; Yamano, H.; Kusaka, H.; Watanabe, H. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointest. Endosc.* **1996**, *44*, 8–14, doi:10.1016/S0016-5107(96)70222-5.
23. Rácz, I.; Horváth, A.; Szalai, M.; Spindler, Sz.; Kiss, Gy.; Regőczy, H.; Horváth, Z. Digital Image Processing Software for Predicting the Histology of Small Colorectal Polyps by Using Narrow-Band Imaging Magnifying Colonoscopy. *Gastrointest. Endosc.* **2015**, *81*, AB259, doi:10.1016/j.gie.2015.03.1345.
24. Silva, J.S.; Histace, A.; Romain, O.; Dray, X.; Grando, B. Towards embedded detection of polyps in WCE images for early diagnosis of colorectal cancer. *Int. J. Comput. Assist. Radiol. Surg.* **2014**, *9*, 283–293, doi:10.1107/s00261-001-0167-8.
25. Bernal, J.; Sánchez, F.J.; Vilariño, F. Towards Automatic Polyp Detection with a Polyp Appearance Model. *Pattern Recognit.* **2012**, *45*, 3166–3182, doi:10.1016/j.patcog.2012.03.002.

26. Bernal, J.; Sánchez, F.J.; Fernández-Esparrach, G.; Gil, D.; Rodríguez, C.; Vilariño, F. WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Comput. Med. Imaging Graph.* **2015**, *43*, 99–111, doi:10.1016/j.compmedimag.2015.02.007.
27. Bernal, J.; Tajkbaksh, N.; Sánchez, F.J.; Matuszewski, B.; Chen, H.; Yu, L.; Angermann, Q.; Romain, O.; Rustad, B.; Balasingham, I.; et al. Comparative Validation of Polyp Detection Methods in Video Colonoscopy: Results from the MICCAI 2015 Endoscopic Vision Challenge. *IEEE Trans. Med. Imaging* **2017**, *36*, 1231–1249, doi:10.1109/TMI.2017.2664042.
28. Charisis, V.S.; Hadjileontiadis, L.J. Potential of hybrid adaptive filtering in inflammatory lesion detection from capsule endoscopy images. *World J. Gastroenterol.* **2016**, *22*, 8641–8657, doi:10.3748/wjg.v22.i39.8641.
29. Charfi, S.; El Ansari, M. Computer-aided diagnosis system for colon abnormalities detection in wireless capsule endoscopy images. *Multimed. Tools Appl.* **2018**, *77*, 4047–4064, doi:10.1007/s11042-017-4555-7.
30. Lilik, F.; Botzheim, J. Fuzzy based Prequalification Methods for EoSHDSL Technology. *Acta. Tech. Jaurinensis* **2011**, *4*, 135–144.
31. Lilik, F.; Kóczy, L.T. Performance Evaluation of Wire Pairs in Telecommunications Networks by Fuzzy and Evolutionary Models. In Proceedings of the IEEE Africon, Pointe-Aux-Piments, Mauritius, 9–12 September 2013; pp. 712–716, doi:10.1109/AFRCON.2013.6757602.
32. Lilik, F.; Nagy, Sz.; Kóczy, L.T. Improved Method for Predicting the Performance of the Physical Links in Telecommunications Access Networks. *Complexity* **2018**, *2018*, doi:10.1155/2018/3685927.
33. Nagy, Sz.; Lilik, F.; Kóczy, L.T. Entropy based fuzzy classification and detection aid for colorectal polyps. In Proceedings of the IEEE Africon, Cape Town, South Africa, 18–20 September 2017; pp. 78–82, doi:10.1109/AFRCON.2017.8095459.
34. Nagy, Sz.; Sziová, B.; Kóczy, L.T. The effect of image feature qualifiers on fuzzy colorectal polyp detection schemes using KH interpolation - towards hierarchical fuzzy classification of coloscopic still images. In Proceedings of the FuzzIEEE, Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–7, doi:10.1109/FUZZ-IEEE.2018.8491479.
35. Lin, Y.-H.; Liao, Y.-Y.; Yeh, C.-K.; Yang, K.-C.; Tsui, P.-H. Ultrasound Entropy Imaging of Nonalcoholic Fatty Liver Disease: Association with Metabolic Syndrome. *Entropy* **2018**, *20*, 893, doi:10.3390/e20120893.
36. Humeau-Heurtier, A. The multiscale entropy algorithm and its variants: A review. *Entropy* **2015**, *17*, 3110–3123, doi:10.3390/e17053110.
37. Stantchev, I. Structural Entropy: A New Approach for Systems Structure's Analysis. In *Cybernetics and Systems '86*; Trappl, R., Ed.; Springer: Dordrecht, The Netherlands, 1986; pp. 139–186, doi:10.1007/978-94-009-4634-7_19.
38. Pipek, J.; Varga, I. Universal classification scheme for the spatial-localization properties of one-particle states in finite, d -dimensional systems. *Phys. Rev. A* **1992**, *46*, 3148–3163, doi:10.1103/PhysRevA.46.3148.
39. Mojzes, I.; Dominkovics, C.; Harsányi, G.; Nagy, Sz.; Pipek, J.; Dobos, L. Heat treatment parameters effecting the fractal dimensions of AuGe metallization on GaAs. *Appl. Phys. Lett.* **2007**, *91*, 073107, doi:10.1063/1.2768911.
40. Hartley, R.V.L. Transmission of Information. *Bell Syst. Tech. J.* **1928**, *7*, 535–563.
41. Nyquist, H. Certain Factors Affecting Telegraph Speed. *Bell Syst. Tech. J.* **1924**, *3*, 324–346.
42. Shannon, C.E. A mathematic theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
43. Gibbs, J.W. *Elementary Principles in Statistical Mechanics, Developed with Especial Reference to the Rational Foundation of Thermodynamics*; Charles Scribner's Sons: New York, USA, 1902; p. 179.
44. Von Neumann, J. Thermodynamik quantenmechanischer Gesamtheiten. *Nachricht. Ges. Wiss. Gött. Math.-Phys. Kl.* **1927**, *102*, 273–291. (In German)
45. von Neumann, J. *Mathematische Grundlagen der Quantenmechanik*; Springer: Berlin, Germany, 1932; p. 26. (In German)
46. Craciunescu, T.; Murari, A.; Gelfusa, M. Improving Entropy Estimates of Complex Network Topology for the Characterization of Coupling in Dynamical Systems. *Entropy* **2018**, *20*, 891, doi:10.3390/e20110891.
47. Humeau-Heurtier, A.; Omoto, A.C.M.; Silva, L.E. Bi-dimensional multiscale entropy: Relation with discrete Fourier transform and biomedical application. *Comput. Biol. Med.* **2018**, *100*, 36–40.
48. Silva, L.E.V.; Senra Filho, A.C.S.; Fazan, V.P.S.; Felipe, J.C.; Murta Junior, L.O. Two-dimensional sample entropy: Assessing image texture through irregularity. *Biomed. Phys. Eng. Express* **2016**, *2*, 045002, doi:10.1088/2057-1976/2/4/045002.

49. Azami, H.; Escudero, J.; Humeau-Heurtier, A. Bidimensional Distribution Entropy to Analyze the Irregularity of Small-Sized Textures. *IEEE Signal. Proc. Lett.* **2017**, *24*, 1338–1342, doi:10.1109/LSP.2017.2723505.
50. Bell, R.J.; Dean, P. Atomic vibrations in vitreous silica. *Discuss. Faraday. Soc.* **1970**, *50*, 55–61, doi:10.1039/DF9705000055.
51. Pipek, J. Localization measure and maximum delocalization in molecular systems. *Int. J. Quantum Chem.* **1989**, *36*, 487–501, doi:10.1002/qua.560360405.
52. Rényi, A. On measures of information and entropy. In Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960; pp. 547–561.
53. Amigó, J.M.; Balogh, S.G.; Hernández, S. A Brief Review of Generalized Entropies. *Entropy* **2018**, *20*, 813, doi:10.3390/e20110813.
54. Faddeev, D.K. Zum Begriff der Entropie Einer endlichen Wahrscheinlichkeitsschenmas. In *Arbeiten zu Informationstheorie*; Deutschen Verlag der Wissenschaften: Berlin, Germany, 1957; pp. 85–90. (In German)
55. Varga, I.; Pipek, J. Rényi entropies characterizing the shape and the extension of the phase space representation of quantum wave functions in disordered systems. *Phys. Rev. E* **2003**, *68*, 026202, doi:10.1103/PhysRevE.68.026202.
56. Molnár, L.M.; Nagy, S.; Mojzes, I. Structural entropy in detecting background patterns of AFM images. *Vacuum* **2010**, *84*, 179–183, doi:10.1016/j.vacuum.2009.04.025.
57. Bonyár, A.; Molnár, L.M.; Harsányi, G. Localization factor: A new parameter for the quantitative characterization of surface structure with atomic force microscopy (AFM). *Micron* **2012**, *43*, 305–310, doi:10.1016/j.micron.2011.09.005.
58. Bonyár, A. AFM characterization of the shape of surface structures with localization factor. *Micron* **2016**, *87*, 1–9, doi:10.1016/j.micron.2016.05.002.
59. Nagy, Sz.; Fehér, A. Topology analysis of scanning microscope images with structural entropy and discrete wavelet transform. In Proceedings of the 18th International Conference on Systems, Signals and Image Processing IWSIP2011, Sarajevo, Bosnia Herzegovina, 16–18 June 2011; pp. 101–104.
60. Nagy, Sz.; Sziová, B.; Solecki, L. The effect of background and outlier subtraction on the structural entropy of two-dimensional measured data. *Int. J. Reason. Intell. Syst.* **2019**, submitted for publication.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).