Spearman's rho modification in digital image processing

B Nalepa and A Gwiazda
Silesian University of Technology, Faculty Mechanical Engineering, Institute of Engineering Processes Automation and Integrated Manufacturing Systems, Konarskiego 18A, 44-100 Gliwice, Poland
E-mail: bartlomiej.nalepa@polsl.pl

Abstract. The article presents an alternative Spearman's rank correlation pattern as an innovative method in detecting the edges of an image by defining characteristic points of the image. Edge detection methods can be counted as the image processing stage following the implementation of the "initial" image processing, which includes: use of filters such as low-pass or high-pass, wavelet transform, etc. In this article hitherto known edge detection algorithms will be replaced with image ranking method. The first chapter of this work contains an analysis of the current knowledge on the ranking of the image. The operation of the Spearman rho algorithm together with its various configurations and the comparison of the two methods studied are presented in turn. Next, the modification of the Spearman rho pattern is presented along with the justification for the use of such a transformation. The same chapter also presents the effects of the algorithm, comparing it with known image ranking methods. The next stage of works consisted in developing a method for the interpretation of the results obtained. In the last chapter of the work, the obtained test results were summarized and analysed.

1. Introduction

Image processing by defining the algorithm is used both by vision systems that are environment recorders and in a virtual environment (applications) based on images created by computer applications. In this article, only images recorded by vision systems originating from the natural environment will be analysed.

Digital image processing algorithms are used to modify the image in a way that identifies specific features. Several stages of image processing can be distinguished. Initial operations are mainly used to improve the image quality, i.e. to use different types of filters or to change image parameters such as gradient, brightness. The second phase is usually algorithms aimed at highlighting a previously defined characteristic in the image, such as change of colour palette, various transformations, algorithms of edge detection (e.g. machining simulation programs [1-4]), image thresholding, etc. The last stage of image processing is usually the right algorithm for recognizing and determining the position of an object, tracking a specific item.

In this article, the methods of object edge detection were mainly analyzed. So far known methods of edge detection, e.g. Canny [5-7] are based on complex mathematical algorithms using, among others, determination of the gradient from the data being processed. Edge detection is not performed in real time, which makes it impossible to use it to track objects, i.e. in autonomous vehicles. Common usage of algorithms for edge detection as well as defects of these methods, among others Difficult interpretation of results with a variable gradation of the image is an argument for trying to define new
relationships. This article proposes the correlation of two variables to object edge detection. The Spearman rho algorithm is presented, which is used, among others, in the MOS (Mean Opinion Scores) method, for supporting medical image analysis [8, 9] and as an example of applying the correlation of two variables [10]. The main application of the Spearman rho method is the use of an algorithm to support the image filtration process (salt-and-pepper noise) [11]. The project used Visual Studio 2017 software together with the OpenCV library supporting the image processing process.

2. Spearman's Rho in digital image processing
Chapter 1 presents the use of the Spearman rho algorithm, which is mainly used as a support element of digital imaging procedures. The image processing algorithm proposed in this article based on Spearman's rho statistics will be the main element of the system consisting of the following procedures:
- determination of the sample size,
- ranking two samples,
- calculation of Spearman's rho coefficient,
- multiplication of the obtained factor by the maximum range of the pixel value - 255,
- creating a new image.

Figure 1(a) shows the image before applying the Canny algorithm, while figure 1(b) shows the image after the edge detection operation.

![Image before applying the algorithm](a)

![Picture after applying the Canny algorithm](b)

**Figure 1.** (a) Image before applying the algorithm, (b) picture after applying the Canny algorithm.

The basis of the considered algorithm is ranking of pixels and then calculating the Spearman rank correlation coefficient. The Spearman rho coefficient is calculated from the following dependence [7, 12-16]:

\[
\rho_s = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2-1)}
\]  

(1)
According to many scientists, the equation (1) is insufficient due to the repetition of particular ranks in the set. The following dependence takes into account the following:

\[ r_s = \frac{\frac{1}{6}(n^3-n)-(\sum_{i=1}^{n} d_i^2)^2-T_x^{-}T_y}{\sqrt{\frac{1}{6}(n^3-n)-2T_x\left(\frac{1}{6}(n^3-n)-2T_y\right)}} \]  

(2)

In image processing equations (1) and (2) can be transformed into the form:

\[ \Lambda_{s=1}^{S} \Lambda_{p=1}^{D} r_{sp} = 1 - \frac{6}{n(n^2-1)} \sum_{i=1}^{n} (R(x_{i+y_i+p})-R(x_{i+y_i+p}))^2 \]  

(3)

\[ \Lambda_{s=1}^{S} \Lambda_{p=1}^{D} r_{sp} = \frac{\frac{1}{6}(n^3-n)-(\sum_{i=1}^{n} (R(x_{i+y_i+p})-R(x_{i+y_i+p}))^2)^2-T_x^{-}T_y}{\sqrt{\frac{1}{6}(n^3-n)-2T_x\left(\frac{1}{6}(n^3-n)-2T_y\right)}} \]  

(4)

where:
- \(RI(x_iy_i)\) – the rank of the pixel value in a given subset,
- \(n\) – number of ranks,
- \(X_c\) – image size along the X axis,
- \(Y_c\) – image size along the Y axis,
- \(S = X_c - 1\),
- \(D = Y_c - 1\).

The image processing algorithm contains the following elements:
- determination of the size of the extracted image area according to figure 2 and figure 3,
- ranking pixel values,
- calculation of dependencies (3) and (4) for a given subset,
- multiplying the obtained Spearman rank correlation coefficient with the maximum pixel value - 255 (white colour) - figure 2 and figure 3.

![Figure 2. Spearman's rho image processing algorithm.](image)

![Figure 3. The result of the Spearman rho algorithm.](image)
Figure 4 shows the application of the equations (3) and (4) ranking individual fragments from 1 to 2 for the image from figure 1(a). Next, in figures 4, 5, 6 the application of these dependencies is presented depending on the scope of the area.

![Figure 4](image1)

**Figure 4.** (a) Image with the use of Spearman's rho according to the equation (3), (b) Image with the use of Spearman's rho according to the equation (4).

![Figure 5](image2)

**Figure 5.** Images using Spearman's rho according to the dependence (5): (a) for \( n = 2 \), (b) for \( n = 3 \), (c) for \( n = 4 \).

![Figure 6](image3)

**Figure 6.** Images using Spearman's rho according to the dependence(4): (a) for \( n = 2 \), (b) for \( n = 3 \), (c) for \( n = 4 \).
The use of equations (3) and (4) returns an image consisting of black boxes (zeros) defined by
the algorithm as a background and white fields being the potential edges of the area. Analysing the
results presented in figures 4, 5 and 6, it can be concluded that increasing the area to be cut reduces
accuracy and smears the image. The best solution, therefore, is to cut the 2x2 area, although in
statistical terms calculations on such a small sample of data are not recommended.

3. Spearman rho modification

The presented solution in chapter 2 is not accurate because the rank correlation coefficient returns
values in the range [-1; 1]. The value -1 indicates the negative correlation of results and 1 the positive
correlation. In addition, the algorithm in image processing is not accurate because the positive
(table 1(a)) and negative (table 1(b)) correlating is not accurate as shown in figure 7.

Table 1. Rank for the area \( n = 5 \):

(a) Spearman's rank correlation coefficient equals 1,
(b) Spearman's rank correlation coefficient equals -1.

|   |   |   |   |
|---|---|---|---|
|1  |1  |0  |0  |
|2  |2  |0  |0  |
|3  |3  |0  |0  |
|4  |4  |0  |0  |
|5  |5  |0  |0  |

\[ \sum_{i=1}^{n} d_i^2 = 0 \quad \sum_{i=1}^{n} d_i^2 = 40 \]

In order to define the Spearman rho modifications, you first need to consider all the 2x2 image rank
configurations that are shown in table 2.

Table 2. Possible configurations for ranking the 2x2 area.

|   |   |   |   |
|---|---|---|---|
|1  |2  |   |   |
|2  |1  |   |   |
|1.5| 1.5|   |   |
|1.5| 1.5|   |   |

Figure 7 shows the possible configurations of the couples shown in table 2.

The values for one variable are the same - the potential edge

Figure 7. The possible configuration pairs are shown in figure 8.

The Spearman's rank correlation coefficient presented in equations (3) and (4) in relation to the
pairs presented in figure 7 does not give the same priority weights. The best solution, therefore,
is to modify the Spearman rho in such a way that it primarily takes into account the cases from figure 7.
This article proposes the following dependence for pre-processing of images based on metric spaces [17]:

\[ r_m = \frac{12 \sum_{i=1}^{n} (R_x^2 - R_y^2)}{n(n^2-1)} \]  

(5)

The above equation (5) can also be presented in the form intended for digital image processing:

\[ \Lambda_x^S \Lambda_y^B r_{sp} = \frac{12 \sum_{i=1}^{n} (R_I(x_{i+p}y_{i+p})^2 - R_I(x_{i+p}y_{i+p})^2)}{n(n^2-1)} \]  

(6)

By carrying out the proof of the effectiveness of the above equation (6), you can make an analysis based on the pairs presented in figure 7.

**Table 3.** Proof of the effectiveness of Spearman's rho modifications.

| \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | 1  | 1  | 0  | 1.5 | 1.5 | 0  |
|---|---|---|---|---|---|---|---|---|
| 1  | 2  | -3 | 2  | 1  | 0  | 1.5 | 1.5 | 0  |
| 2  | 1  | 3  | \( r_{sp} \) | 0  | \( r_{sp} \) | 0  | \( r_{sp} \) | 0  |

| \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1  | 1.5| -1.25| 1.5| 1  | 1.25| 1.5| 2  | -1.75| 1.5| 3  | 16  |
| 2  | 1.5| 1.75| \( r_{sp} \)| 1  | \( r_{sp} \)| -1 | \( r_{sp} \)| -1 | \( r_{sp} \)| -1 | \( r_{sp} \)| -1 |

Based on the data analysed in table 3, it can be concluded that the algorithm specifically considers cases that contain the same rank. It is also possible to provide further proof of the theoretical effectiveness of the method by checking the case for a sample ranked from 1 to 5 as shown in table 4.

**Table 4.** Examples of the effectiveness of the Spearman. Rho modifications for the ranking test ranging from 1 to 5.

| \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3  | 1  | 8  | 1  | 3  | -8 | 1  | 3  | -8 | 3  | 1  | 8  |
| 3  | 2  | 5  | 2  | 3  | -5 | 2  | 3  | -5 | 3  | 2  | 5  |
| 3  | 3  | 0  | 3  | 3  | 0  | 3  | 3  | 0  | 4  | 3  | 7  |
| 3  | 4  | -7 | 4  | 3  | 7  | 4  | 3  | 7  | 5  | 3  | 16 |
| 3  | 5  | -16| 5  | 3  | 16 | 5  | 3  | 16 |
| \( r_{sp} \)| -1 | \( r_{sp} \)| -1 | \( r_{sp} \)| -1 | \( r_{sp} \)| -1 | \( r_{sp} \)| -1 |

| \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) | \( x_i \) | \( y_i \) | \( R_x^2 - R_y^2 \) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3  | 2.5| 2.75| 3  | 2  | 5  | 3  | 2  | 5  | 3  | 2  | 5  |
| 3  | 2.5| 2.75| 3  | 4  | -7 | 3  | 4  | -7 | 3  | 4  | -7 |
| 3  | 2.5| 2.75| 3  | 4  | -7 | 3  | 4  | -7 | 3  | 4  | -7 |
| 3  | 5  | -16| 3  | 5  | -16| 3  | 5  | -16| 3  | 5  | -16|
| \( r_{sp} \)| -0.5 | \( r_{sp} \)| -0.5 | \( r_{sp} \)| -0.5 | \( r_{sp} \)| -0.5 | \( r_{sp} \)| -0.5 | \( r_{sp} \)| -0.5 | \( r_{sp} \)| -0.5 | \( r_{sp} \)| -0.5 |
On the basis of the presented examples, it can be concluded that, as in the classical Spearman rho, the correlation coefficient of ranks is in the range [-1; 1]. However, the modified version of Spearman's rho includes mainly the repeatability of ranks. Figure 8 presents the operation of the modified Spearman rho version together with its comparison with the methods of equations (3) and (4). Figure 9 provides a comparison of the modified Spearman rho version (5) and (6) depending on the size of the n-slice image.

![Figure 8](image1.png)

**Figure 8.** Images using Spearman's rho for \( n = 2 \); (a) according to dependence (3) (b) according to dependence (4), c) according to the modified version.

![Figure 9](image2.png)

**Figure 9.** Images using the modified Spearman rho version: (a) forn = 2, (b) forn = 3, c) forn = 4.

Analysing figure 9, it can be concluded that the larger the sample, the more the resulting picture is blurred. As a result of the application of the algorithm, a set of points can be obtained, which according to known mathematical laws can be used to determine the characteristic parameters of the image.

4. The method of least squares and central moments of the image as applied to the modified Spearman rho method

The modified Spearman rho version analysed in chapter 3 returns a set of points to which various mathematical operations can be applied. The first of these is the method of least squares. The function defining the set of points can be a circle:

\[
(x - a)^2 + (y - b)^2 = r^2 
\]  

(7)

Converting dependence to the form:

\[
x^2 - 2ax + a^2 + y^2 - 2by + b^2 - r^2 = 0
\]

(8)

and simplifying to dependencies:

\[
x^2 + Xx + y^2 + Yy + Z = 0
\]

(9)

Then perform the function minimization:
\[
\chi^2(X,Y,Z) = \sum_{i=1}^{n} (x_i^2 + Xx_i + y_i^2 + Yy_i + Z)^2
\]

where:

- \(X = -2ax\),
- \(Y = -2by\),
- \(Z = a^2 + b^2 - r^2\).

Figure 10 presents images containing a designated approximated circle and coordinates of the central moment of a set of points (calculation based on equation (7 – 10)).

![Images with selected approximation circle (red) set of points and coordinates of the central moment (blue): (a) primary image in the RGB colour palette, (b) image with the use of the modified Spearman version.](image)

The presented methods characterizing the set of points can be parameters in the recognition of images. In the case of circle approximation, it can be both the coordinates of the centre of the circle and the radius of the circle. The coordinates of the central moment can be the element connecting this parameter with the circle approximation, e.g. by determining the vector determining the distance between the centre of the circle and the coordinates of the central moment.

### 5. Conclusion

This article presents a modified method determining the Spearman's rank correlation coefficient. The first part of the article presents an algorithm for ranking a picture with known Spearman rho relationships. The simplistic Spearman rho dependence was compared with the relation taking into account the repetition of ranks in a given set. From the comparison of both, one can conclude that the method including the repeatability of ranks returns a higher number of points, which can be more accurate. By conducting a critical analysis of these methods, the modified Spearman rho dependence for digital image processing was developed. The method is more effective than the traditional one because it returns the maximum values if the potential image edge occurs. From the comparison of the modified to the traditional one, one can conclude that it is more accurate because it eliminates potential interference.

The last chapter presents examples of mathematical instruments for a set of points that can characterize the obtained results. In the further part of the work an algorithm for object recognition will be made using the modified version of the Spearman rank correlation coefficient. The next stages of the research will be to improve the functionality of the algorithm as well as to develop a comparative method of the proposed algorithm with commonly known methods.

### 6. References

[1] Lysek K, Gwiazda A and Herbus K 2018 Application of CAM systems to simulate of a milling machine work *IOP Conf. MSaE* 400 042037

[2] Lysek K, Gwiazda A and Herbus K 2018 Modelling of CAM systems to simulate of a milling machine work *IOP Conf. MSaE* 400 042036
[3] Mansour G, Mitsi S, Bouzakis K D, Sagris D and Varitis E 2010 Developed hybrid genetic algorithm for optimizing reverse engineering methods *International Journal of Modern Manufacturing Technologies* II(1) 43-48

[4] Cohal V and Cohal A Shape recognition of various parts produced within a manufacturing process using images from different angles *International Journal of Modern Manufacturing Technologies* VII(2) 32-37

[5] Canny J 1986 A Computational Approach to Edge Detection *IEEE Trans. Pattern Anal. Mach. Intell* PAMI-8 6 679–698

[6] New version of Canny edge detection algorithm Research Gate(Online), https://www.researchgate.net/publication/236024099_New_version_of_Canny_edge_detection_algorithm. accessed: 21.11.2018

[7] Heath M, Sarkar S, Sanocki T and Bowyer K 1996 Comparison of edge detectors: a methodology and initial study *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (San Francisco, USA: CA) 143–148

[8] Manousaki A G 2006 A simple digital image processing system to aid in melanoma diagnosis in an everyday melanocytic skin lesion unit. A preliminary report *Int. J. Dermatol.* 45(4) 402–410

[9] Calautti C 2007 The relationship between motor deficit and hemisphere activation balance after stroke: A 3T fMRI study *Neuro Image* 34(1) 322–331

[10] Pan B, Qian K, Xie H and Asundi A 2009 Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review *Meas. Sci. Technol.* 20(6) 062001

[11] Wanghua H, Wang Q, Xu W and Zhou Y 2016 Robustness analysis of digital image spearman’s rho correlation AEST DOI: 10.2991/aest-16.2016.29

[12] Onwuegbuzie A J, Daniel L and Leech N L 2007 Spearman’S Rho *Encyclopedia of Measurement and Statistics* Thousand Oaks: SAGE Publications Inc. 3 928–933

[13] Faherty V 2008 Correlation: Spearman’s rho and Pearson’s r *Compassionate Statistics: Applied Quantitative Analysis for Social Services, with exercises and instructions in SPSS, Thousand Oaks: SAGE Publications, Inc.* 183–200

[14] Yue S, Pilon P and Cavadias G 2002 Power of the Mann–Kendall and Spearman’s rho tests for detecting monotonic trends in hydrological series *J. Hydrol* 259(1–4) 254–271

[15] Gibbons J D and Chakraborti S 2003 *Nonparametric statistical inference* (New York: Marcel Dekker)

[16] Xu W, Hou Y, Hung Y S and Zou Y 2013 A comparative analysis of Spearman’s rho and Kendall’s tau in normal and contaminated normal models *Signal Process* 93b 261–276

[17] Zezula P, Amato G, Dohnal V and Batko M 2006 *Similarity Search: The Metric Space Approach* (US: Springer)