View metadata, citation and similar papers at core.ac.uk brought to you by CORE

Robust object detection in images corrupted by impulse noise

Rykhard Bohush[1] [0000-0002-6609-5810], Sergey Ablameyko[2] [0000-0001-9404-1206] and Yahor Adamovskiy[1] [0000-0003-1044-8741]

[1]Polotsk State University, Blokhina st., 29, Novopolotsk, Republic of Belarus, 211440
{ r.bogush, adamovskiy.y }@pdu.by
[2] Belarusian State University, Nezavisimosti avenue 4, Minsk, Republic of Belarus
ablameyko@bsu.by

Abstract. This paper proposes two effective normalized similarity functions for robust object detection in very high density impulse noisy images. These functions form an integral similarity estimate based on relations of minimum by maximum values for all pairs of analyzed image features. To provide invariance under the constant brightness changes, zero-mean additive modification is used. We explore properties of our functions and compare them with other commonly used for object detection in images corrupted by impulse noise. The efficiency of our approach is illustrated and confirmed by experimental results.

Keywords: similarity functions, object detection, impulse noise

1 Introduction

Impulse noise is frequently encountered in digital image and video due to transmission errors, defective pixels in camera sensors, faulty memory locations, and timing errors during the conversion [1, 2]. An important property for this noise type is that the corrupted pixels can take only the maximum and minimum values from dynamic range and that only part of the pixels is corrupted, and the rest are noise-free. Noise reduction increases the computational cost and it does not allow to get clear image for high noise level. Therefore, it is important to use image processing noise reducing methods for solving the mentioned tasks.

Object detection and precise definition of object location in images and videos are widely used for many applied tasks: industrial inspection and medical diagnostics, stereo vision, reference marks detection and localization in space images of the earth's surface, tracking targets of airborne radar stations, etc [3, 4]. Absolute and relative object position in image or video can be determined after their localization. Camera calibration, image georeferencing, formation of stereo images by computer methods and other applications contain similar tasks. Therefore, many scientific works are devoted to the development of methods for detecting and localizing objects in images and video. Generally, many approaches calculate the similarity of reference object and an input image and compared with a threshold value.

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR-WS.org/Vol-2608/paper83.pdf

Unluckily, there is no a single similarity function that works very well for various images and for all tasks. Different tasks require different measure properties. Therefore, traditional similarity metrics are improved [5, 6] or new functions are provided [7, 8, 9] considering expansion of applications of image processing, and different similarity functions are selected for various applications [10, 11]. For example, image detection in large databases requires high speed and low resistance to deformation of the image, but people tracking in video desires robustness to various types of noise and high accuracy of comparing selected features.

Even in solving only one task uncertainty may also occur. For example, to correctly identity tracked people in a current frame, a maximum accuracy of the estimation for people feature similarity with previous frames is required [12]. However, high accuracy for high density noise or overlapping parts of people by objects will lead to loss their index. It means that a person from previous frames will be identified as another person who has entered into the surveillance area and displayed on the current frame. This problem is also related to the fact that the number of possible grayscale and color images of the same size is huge, their correlation characteristics in practice are not perfect. This leads to false identification or inaccurate object localization when detection is performed in presence of various types of noise components in video or image.

In this paper, we present a new effective function similarity for image and video corrupted by impulsive noise. These functions form an integral normalized similarity estimate based on sequential division of minimum by maximum values for all pairs of features. In Section 2, comprehensive analysis of main similarity functions used in image processing is presented. Section 3 presents our normalized similarity functions and computational costs for them. Section 4 presents experimental results. Finally, the conclusion and feature work details are provided in Section 5.

2 Main similarity functions for image processing

If two images O and B are identical, then value of the normalized metric $M(O,B)=0$, and the value of the normalized similarity function $S(O,B)=1$. For another extreme case, when the differences between the images are maximized, the normalized metric $M(O,B)=1$, and the similarity function usually $S(O,B)=0$ or $S(O,B)=-1$, if the mean value of the brightness levels is not taken into account.

To compare two images $O=\{o_{ij}\}$ and $B=\{b_{ij}\}$, $N \times N$ size, several similarity functions are often used [13]. It is known that linear relation between brightness of images allows to efficiently apply the correlation coefficient for image corrupted by additive noise [14]. The value of the normalized cross-correlation function (NCC) varies from 0 to 1 and is calculated as:

$$S^{NCC}=\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}o_{ij}b_{ij}}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(o_{ij}\right)^2}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(b_{ij}\right)^2}} \quad (1)$$

The normalization embodied into the zero-mean normalized cross-correlation function (ZNCC) allows one to tolerate linear brightness variations. Also, due to the subtraction of the local mean, the ZNCC provides better robustness than the NCC since it tolerates uniform brightness variations as well. In this case, ZNCC takes a value of (-1) to (+1) and is calculated by formula:

$$S^{ZNCC}=\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(o_{ij}-\overline{o}\right)\left(b_{ij}-\overline{b}\right)}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(o_{ij}-\overline{o}\right)^2}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(b_{ij}-\overline{b}\right)^2}} \quad (2)$$

where $\overline{o}$ and $\overline{b}$ – mean values for images O and B (3):

$$\overline{o}=\frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}o_{ij}\ ,\ \overline{b}=\frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}b_{ij} \quad (3)$$

Similarity function based on the Euclidean distance (ED) is characterized by a lower computational cost and is defined as:

$$S^{ED}=1-\sqrt{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left(o_{ij}-b_{ij}\right)^2} \quad (4)$$

Sum of squared differences (SSD) function is robust to Gaussian noise [15] and is calculated as:

$$S^{SSD}=1-\frac{1}{LN^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(o_{ij}-b_{ij}\right)^2 \quad (5)$$

where L – of valid brightness values range.

The function based on the weighted sum of squared differences (SSDW) is more stable to linear distortion of levels of analyzed characteristics compared to previous functions:

$$S^{SSDW} = 1 - \frac{1}{L} \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(o_{ij} - b_{ij}\right)^2}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(o_{ij}\right)^2}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(b_{ij}\right)^2}} \tag{6}$$

The Hausdorff metric-based function (Hd) can only be used for low noise, as increasing noise level quickly reduces the estimation accuracy. This function is determined by the formula:

$$S^{Hd} = 1 - \frac{1}{L}\max_{ij}\left|o_{ij} - b_{ij}\right| \tag{7}$$

In general case, to detect an object in an image, the similarity value is calculated for features of all subimages and object. The decision on object presence is made based on the comparison of obtained values with a threshold. If the condition is satisfied, decision about correspondence on subimage and object is made. For accurate localization similarity function value must exceed threshold only if the object is correctly positioned. However, ensuring the minimum probability of missing object (false-negative error) requires a lower threshold level, but this leads to an increase in the number of subimages, including near the correct location of the object in the image that does not correspond to it, known as false-positive error.

3 New normalized similarity functions

The proposed two similarity functions use relationship calculation between the minimum and maximum values for all pairs of analyzed features. Summation of the calculated values is used to obtain an integral normalized value that characterizes similarity of two images. Relationship calculation between descriptors will better emphasize local differences compared to subtraction. High resistance to noise is achieved through the use of summation when obtaining a complex normalized value. Minimum or maximum attribute is necessary to determine when searching a relationship between them, therefore, the proposed functions are called normalized minimax similarity functions.

To estimate image similarity of object $O = \{o_{ij}\}$, $N \times N$ size, and image $B = \{b_{ij}\}$, $N \times N$ size, functions are described as:

• normalized minimax additive similarity p-function (MMADD^{P}):

$$S^{MMADD^P} = \frac{1}{NN}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\min\left(o_{ij}^p, b_{ij}^p\right)}{\max\left(o_{ij}^p, b_{ij}^p\right)}, p \in Z, p \geq 1 \tag{8}$$

• normalized zero-mean minimax additive similarity p-function (ZMMADD^{P}) :

$$S^{ZMMADD^p} = \begin{cases} \frac{1}{NN}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{b_{ij}-\bar{b}}{o_{ij}-\bar{o}}\right)^2, if\left|o_{ij}-\bar{o}\right| \geq \left|b_{ij}-\bar{b}\right| \\ \frac{1}{NN}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{o_{ij}-\bar{o}}{b_{ij}-\bar{b}}\right)^2, if\left|o_{ij}-\bar{o}\right| < \left|b_{ij}-\bar{b}\right| \end{cases} \tag{9}$$

For both proposed minimax functions, the following basic properties are obvious:

$$\begin{aligned} &S(O,B) = 1, \Leftrightarrow O = B; \\ &S(O,B) = S(B,O); \\ &S(O,B) = 0, \Leftrightarrow o_{ij} = 0, b_{ij} = L \end{aligned} \tag{10}$$

Presented functions are universal, as $MMADD^p$ returns a normalized value from (0) to (1) and $ZMMADD^p$ returns a value from (-1) to (1) for randomly selected features of a pair of images. Table 1 shows a comparison of computational costs for similarity functions for images with size $N \times N$ pixels.

Table 1. Computational costs for similarity functions

Function name	Number of addition / subtraction operations	Number of multiplication / division operations	Number of comparison operations
NCC	3N(N-1)	$3N^2+3$	–
ZNCC	$2N^2+5(N-1)^2$	$3N^2+7$	–
ED	$N^2+N(N-1)+1$	N^2+1	–
SSD	$N^2+N(N-1)+1$	N^2+3	–
SSDV	$N^2+3N(N-1)+1$	$3N^2+4$	–
Hd	N^2+1	1	N^2-1
$MMADD^p$ (p=2)	N(N-1)	$3N^2+2$	N^2
$ZMMADD^p$ (p=1)	$2N^2+3N(N-1)$	N^2+4	N^2

Thus, in comparison with function of normalized correlation, the offered minimax function provides reduction of calculation complexity, as minimum twice. As result, minimax similarity functions can be used for search of objects in a static image and for detecting moving objects in video sequences.

4 Experimental results

4.1 Parameter determination

It is proposed to use the following parameters for the similarity functions analytical assessment when detecting objects in image:

- function value calculated for object and subimage (A – main peak). Value should aim to 1;
- main peak variance (D_A). The value should aim to zero, meaning smaller value deviations A from expected value;
- function maximum value from all side peak (S_L) can be used to determine threshold T , which should be less than T .
- maximum side peak variance (D_{S_L}) should also be as less as possible;
- side peak number at levels higher than 0.95 (N_{S_L}). The parameter can be used to evaluate the possible false-positive detection results number if the threshold value is incorrectly selected.

Research software was developed using the MatLab package for experimental determination of normalized minimax similarity functions qualitative characteristics. The software is based on an object detection algorithm in a grayscale image using pixel brightness as features. It includes the following steps:

1. Selecting a subimage $B_{kl}(k \in 0..M-m, l \in 0..N-n)$, $m \times n$ size, from the upper left of the bitmap image;
2. Similarity function calculation for a reference object O of size $m \times n$ and the selected subimage B_{kl} :

$$S = F(O, B_{kl}) \tag{11}$$

   where F is the mathematical transformation operator;
3. Deciding on object availability by the rule:

$$\begin{cases} if \quad S > T \quad then \quad B_{kl} = O \\ else \quad B_{kl} \neq O \end{cases} \tag{12}$$

   where T is the threshold;
4. If the shift number is less than $(M-m)\times(N-n)$, move right or down one pixel and go to step 1, otherwise the search is finished. The analyzed total number of fragments is defined as $(M-m+1)\times(N-n+1)$.

All parameters were calculated for 20 different images without noise and distortion of size 150 × 150. For each of them 20 reference objects of 15×15 size were used with obtained values averaging. Values $A=1$ and $D_A=1$ are obtained for all similarity functions.

The resulting values of the remaining parameters are given in Table 2. Analysis of the Table 2 shows that the proposed *ZMMADD*^{p} function has the best characteristics for accurately determining object coordinates in images without noise.

Table 2. Function characteristics for noise-free images

Similarity Function	S_L	D_{S_L}	N_{S_L}
NCC	0,97786	0,00008	8568
ZNCC	0,69987	0,01455	3
ED	0,91882	0	481
SSD	0,9918	0	11971
SSDV	0,99948	0	20263
Hd	0,74588	0,00054	43
MMADD^p (p=2)	0,64973	0,00496	263
ZMMADD^p (p=1)	0,47403	0,01489	6

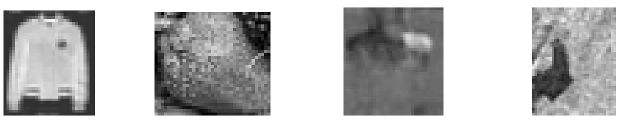
4.2 Noise resistance research

Impulse noise is one of the interference types and it appears as random white or black dots in an image, i.e. corrupted pixels take a maximum or minimum valid value, for example, 255 and 0 for 8-bit images. Interference with an amplitude value greater than the useful signal dynamic range occurs, for example, during a fast transient, and is the cause of the appearance of impulse noise in the image. The impulse noise probability density function of a random variable z is given by:

$$p(z) = \begin{cases} P_a & for \quad z = a \\ P_c & for \quad z = c \\ else & 0 \end{cases} \tag{13}$$

A pixel with a brightness value of $c > a$ looks like a white point when c (?), and a pixel with a brightness value of c looks like a black point in image. Почитай. Здесь совсем непонятно.

Impulse noise is a combination of white and black points (“salt and pepper”). Impulse noise can be characterized by the noise density ρ, which determines the percentage of the corrupted pixels n to a total number, i.e. the probability of distortion for each pixel. For impulse noise, ρ values [0.1, 0.3, 0.5, 0.8, 0.95] was chosen. Objects for detection in Fig. 1. are presented.



a) b) c) d)

Fig. 1. Object tampletes: (a) for Fig. 2a, e, i; (b) for Fig. 2b, f, j; (c) for Fig. 2c, g, k; (d) for Fig. 2d, h, l;

Fig. 2 shows some examples of images with the various noise levels.

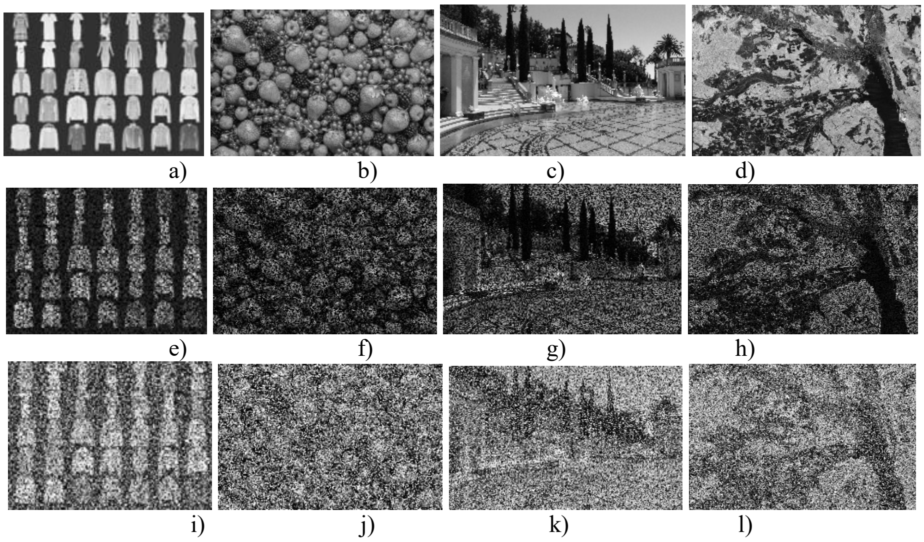


Fig. 2. Some test images: (a-e) noise-free images; (f-j) images corrupted by single-level impulse noise with ρ = 0.5; (k-n) images corrupted by «salt and pepper» noise with ρ = 0.5

The difference values between similarity function main and maximum side peaks show an assessment completeness and comparison of similarity functions in terms of resistance to various noise types (Table 3-4).

Table 3. Experimental results for similarity functions stability to single-level impulse noise

Similarity Function	$(A-S_L)$				
	ρ =0.1	ρ =0.3	ρ =0.5	ρ =0.8	ρ =0.95
NCC	-0.00365	-0.03498	-0.06312	-0.11406	-0.2196
ZNCC	0.06579	-0.00573	-0.07419	-0.18602	-0.21083
ED	0.01497	-0.02903	-0.02102	-0.02727	-0.0044
SSD	-0.0009	-0.00857	-0.01165	-0.01096	-0.01039
SSDV	-0.00001	-0.00037	-0.00083	-0.0038	-0.01349
Hd	-0.13882	-0.05725	-0.05569	-0.04392	-0.03216
MMADD^p (p=2)	0.20776	0.14718	0.1144	0.02539	-0.01688
ZMMADD^p(p=1)	0.07336	0.08461	-0.1426	-0.14419	-0.10195

Correct object detection is possible when the side peak amplitude is below the similarity function main peak level:

$$(A-S_L)>0 \tag{14}$$

Object detection is possible for *MMADD*^{p} function (*p=2*) at $\rho = 0.8$ when the image is distorted by single-level impulse noise and salt and pepper noise. For other functions at this noise level, condition $(A - S_L) > 0$ is not observed.

Table 4. Experimental results for similarity functions stability to salt and pepper noise

Similarity Function	$(A-S_L)$				
	ρ =0.1	ρ =0.3	ρ =0.5	ρ =0.8	ρ =0.95
NCC	-0.01046	-0.05543	-0.09423	-0.08755	-0.09563
ZNCC	0.12058	0.01313	-0.12529	-0.19709	-0.34173
ED	0.01589	0.00848	-0.0138	-0.03309	-0.03137
SSD	0.00663	-0.00103	-0.02537	-0.02423	-0.05641
SSDV	0.00006	-0.00009	-0.00053	-0.00045	-0.00062
Hd	-0.20314	-0.09176	-0.11529	-0.07843	-0.07843
MMADD^p (p=2)	0.22005	0.16443	0.10312	0.01452	-0.0042
ZMMADD^p (p=1)	0.38812	0.16848	0.04697	-0.03497	-0.07924

The *ZMMADD*^{p} (ρ *=1*) provides invariance under the constant change in brightness. Object detection is possible for this function at $\rho = 0.3$ when the image is distorted by single-level impulse noise and $\rho = 0.5$ for salt and pepper noise.

5 Conclusion

Normalized similarity functions for image evaluation are proposed. They are characterized by satisfactory computational costs and can improve the detection and object localization in image and video corrupted by impulse noise. Proposed functions are universal, they return a normalized value from (0) to (1) for randomly selected features. Similarity function resistant to a linear change of analyzed feature values of the compared images is proposed that returns a normalized value from (-1) to (1). Experiments were performed to compare these functions with the known ones for images corrupted by single-level impulse and salt and pepper noise. Pixel brightness values are used as image features for experiments. The obtained results have confirmed efficiency of presented similarity functions.

6 References

1. Gonzalez R., Woods R.: Digital Image Processing, 3rd Edition. Prentice Hall (2008)
2. Pitas I., Venetsanopoulos A.: Order statistics in digital image processing Proceedings of the IEEE 80 (12) pp 1893-1921 (1992)
3. Rezaei M., Fränti P.: Set Matching Measures for External Cluster Validity Transactions on Knowledge and Data Engineering 28 (8) pp 2173–86 (2016)
4. Ges V., Starovoitov V.: Distance-based functions for image comparison Pattern Recogn. Lett. 20 207–14 (1999)

5. Wang L., Zhang Y., Feng J. On the Euclidean Distance of Images IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (8) pp 1334–39 (2005)
6. Nakhmani A., Tannenbaum A.: New Distance Measure Based on Generalized Image Normalized Cross-Correlation for Robust Video Tracking and Image Recognition Pattern Recogn. Lett. 34 (3) pp 315–21 (2013)
7. Zagoruiko N. G., Borisova I. A., Dyubanov V. V., Kutnenko O. A.: A quantitative measure of compactness and similarity in a competitive space J. Appl. Industr. Math. 5 (1) pp 144–54 (2011)
8. Lv G.: A novel similarity measure for matching local image descriptors IEEE Access 6 pp 55315–25 (2018)
9. Kim J., Hyun Ch., Han H., Kim H.: Evaluation of Matching Costs for High-Quality Sea-Ice Surface Reconstruction from Aerial Images Remote Sensing 11 (9) pp 1055-72 (2019)
10. Zitova B., Flusser J.: Image registration methods: a survey Image and vision computing 21 (11) pp 977–1000 (2003)
11. Cha S. H.: Comprehensive survey on distance/similarity measures between probability density functions Int. J. of Mathematical Models and Methods in Applied Sciences 1 (4) p 300 (2007)
12. Bohush R., Zakharava I.: Robust Person Tracking Algorithm Based on Convolutional Neural Network for Indoor Video Surveillance Communications in Computer and Information Science 1055 pp 289–300 (2019)
13. Goshtasby A. A.: Similarity and dissimilarity measures in Image Registration Principles, Tools and Methods. Advances in Computer Vision and Pattern Recognition (London) pp 7–66 (2012)
14. Voronov S. V., Tashlinskii A. G.: Efficiency analysis of information theoretic measures in image registration Pattern Recognit. Image Anal. 26 502–5 (2016)
15. Voronov S. V., Tashlinskii A. G.: Analysis of objective functions in the problem of estimating mutual geometric deformations of images Pattern Recognit. Image Anal. 24 (4) 575–582 (2014)