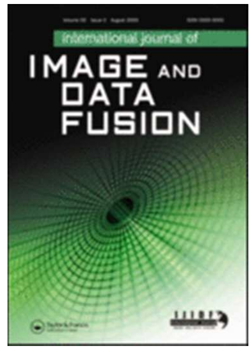

Image Similarity/Distance Measures: What is really behind MSE and SSIM?

Journal:	*International Journal of Image and Data Fusion*
Manuscript ID	TIDF-2016-0046.R1
Manuscript Type:	Review Paper
Keywords:	similarity, distance, composite, correlation coefficient, translation invariant, image processing

Image Similarity/Distance Measures: What is really behind MSE and SSIM?

Gintautas Palubinskas

Remote Sensing Technology Institute, German Aerospace Center (DLR), Oberpfaffenhofen, Germany

German Aerospace Center (DLR), Remote Sensing Technology Institute, Münchener Str. 20, 82234 Wessling, Germany, Tel: +49 8153 281490, Fax: +49 8153 281444, E-mail: Gintautas.Palubinskas@dlr.de

Image Similarity/Distance Measures: What is really behind MSE and SSIM?

Similarity/distance measures play an important role in various signal/image processing applications such as classification, clustering, change detection and matching. In most cases, maybe excluding visual perception, the distance measure should be amplitude/intensity translation invariant what means that it depends only on the relative difference of compared variables/parameters, but not on their absolute values. The two most popular measures: Mean Squared Error (MSE) and Structural SIMilarity (SSIM) index used in image processing have been analysed theoretically and experimentally by showing their origin, similarities/differences, and main properties. Both measures depend on the same parameters: sample means, standard deviations and correlation coefficient. It has been shown that SSIM originates from the two generalized Dice measures and thus inherit their main property scale invariance. Consequently, this property leads to the dependence of the measure on absolute mean and standard deviation values. Similarly, MSE depends on the absolute standard deviation values. A new composite similarity/distance measure based on Means, Standard deviations and Correlation coefficient (CMSC) which has been proposed recently exhibits translation invariance property with respect to means and standard deviations. Experiments on simulated and real data corrupted with various types of distortions: mean shift, contrast stretching, noise (additive/multiplicative, impulsive) and blurring, supported theoretical results.

Keywords: similarity; distance; Euclidian; Dice; composite; correlation coefficient; translation invariant; image processing

1. Introduction

Similarity or distance measures are unavoidable for solving various signal/image processing tasks such as compression, restoration, de-noising, registration, matching, segmentation, classification, detection and recognition, e.g. see references (Jain and Dubes 1988, Fukunaga 1972, Richards and Jia 1999). There exist a lot of such measures proposed in the literature, e.g. see surveys (Cha 2007, Li *et al*. 2010, Goshtasby 2012).

Usually, a broadly recognized and accepted measure MSE is used. Recent investigation (Wang and Bovik 2009) showed its weakness in some applications, e.g. visual perception of images. Thus, a new SSIM index has been proposed in (Wang and Bovik 2002, Wang *et al.* 2004). It is spreading fast, not only in computer vision community, but also within other communities, such as remote sensing, with very different tasks, mostly requiring only a relative comparison of data, such as classification, clustering (Richards and Jia1999), change detection (Alberga 2009), matching (Goshtasby 2012), and image fusion (Palubinskas 2013, Palubinskas 2016). Recently, an enhanced version of SSIM has been introduced (Sampat *et al.* 2009) and some of its properties have been analysed mathematically (Brunet *et al.* 2012) and experimentally (Dosselmann and Yang 2011). Nevertheless, the following questions are arising. Can SSIM be simply transferred to other applications requiring mainly relative comparison of data (amplitude/intensity translation invariance)? Can SSIM replace MSE? Which measure is most suitable for translation invariant applications? To answer these questions a deep theoretical analysis of MSE and SSIM measures is necessary. This analysis allowed identifying different properties of these two measures. Moreover, in (Palubinskas 2014, Palubinskas 2015) a new similarity measure, the composite similarity/distance measure, based on sample moments (means, standard deviations) and the correlation coefficient (CMSC), is proposed which exhibits the translation invariance property with respect to means and standard deviations.

In order to avoid confusion with a usual/standard paper on quality measures for visual image perception application the following should be stated. This paper is not considering this popular and broad application in image processing. Other so-called translation invariant applications such as clustering, classification, change detection and matching, just to list few of them, which require only relative comparison of data, are of interest in this work. Thus, a lot of quality measures developed in visual perception field, see e.g. recent reviews (Liu and Wu 2011, George and Prabavathy 2013, Mohammadi *et al.* 2015, Hanhart *et al.* 2015), are not relevant for this paper. Of course, SSIM is also not relevant for applications aimed in this paper too. But SSIM is used in this paper because of several reasons. First, it is used to show its origin, some properties and its relation to MSE. Second, a new measure CMSC is somehow inspired by SSIM. ~~Third, this study should warn researchers about careful usage of SSIM in their applications.~~ SSIM is designed for a visual image quality assessment and thus should be used only for this application. But this measure is spreading very quickly not only in visual perception applications but also in other applications without careful reasoning if it is an appropriate measure. Thus, the main concern of this paper is not to present or show that a particular measure is the best for a specific application as it is in a usual/standard paper. The main aim is to present a deep understanding of measures, their origin and properties and thus to help a reader to select a right measure or even to create his own measure for his particular application. For example, the application of a new measure CMSC for image fusion quality assessment task is presented in (Palubinskas 2015).

The paper is organized in the following way. First, general definitions for similarity and distance measures, their properties, e.g. metric, normalization/scaling, transformation from a distance to similarity and vice versa, and different ways of similarity combination are introduced. Further, a theoretical analysis of MSE and SSIM measures is performed, which naturally leads to the proposal of a new measure CMSC. Then, the analysed measures are investigated on simulated data and real data (satellite panchromatic image and "Lena" image) corrupted with various types of distortions: mean shift, contrast stretching, different types of noise (additive Gaussian, speckle and impulse) and blurring in order to confirm theoretical results on one hand and maybe to discover new properties or to show suitability for various applications on the other hand, thus enhancing the preliminary conference paper (Palubinskas 2014).

2. Theory

First, some notations used in this paper will be introduced. Let $x = \{x_i \mid i = 1,\dots,N\}$ and $y = \{y_i \mid i = 1,\dots,N\}$ denote two images or image patches or more generally signals to be compared, where x_i , y_i are real numbers in a finite range of values $\min \le x_i , y_i \le \max$ (e.g. min=0 and max=255 for 8bit images), $R = \max - \min$, N is the number of pixels/samples.

2.1 Distance and similarity

Distance *d* is defined as a measure indicating how close/far apart two samples/objects are. It exhibits high values for objects which are far from each other and low values for near objects. Quite often it is also called dissimilarity measure. For example, the Euclidian measure

$$d_E = \sqrt{(1/N) \cdot \sum_{i=1}^{N} (x_i - y_i)^2} \tag{1}$$

is probably the most popular distance measure. The inverse measure to distance *d* is a similarity measure *s*, which exhibits high values of similar objects and low values for different objects. Here, a Correlation Coefficient (CC) as a most popular similarity measure can be mentioned

$$\rho = \sigma_{xy} / (\sigma_x \cdot \sigma_y) , \tag{2}$$

where sample means, standard deviations and covariance are defined as follows:

$$\mu_x = (1/N) \cdot \sum_{i=1}^{N} x_i , \quad \mu_y = (1/N) \cdot \sum_{i=1}^{N} y_i , \tag{3}$$

$$\sigma_x{}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2 , \quad \sigma_y{}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \mu_y)^2 , \qquad (4)$$

$$\sigma_{xy} = \left(1/\left(N-1\right)\right)\cdot\sum_{i=1}^{N}(x_i - \mu_x)\cdot(y_i - \mu_y) . \qquad (5)$$

The correlation coefficient is generally between -1 (negatively correlated objects) and 1 (fully correlated objects). It is 0 for uncorrelated objects. For many tasks, e.g. image processing applications listed in the Introduction section, negative ρ is of no interest, thus it can be assumed that the following relation is valid: $0 \leq \rho \leq 1$, e.g. by setting negative ρ to 0. This assumption simplifies formula and contributes towards easier understanding and interpretation of the results. It does not influence the trends and conclusions observed, because of the constant range factor of 2 which should be used for the scaling of the measures.

Some measures such as Euclidian distance or MSE are defined directly on original data values *x, y* whereas, e.g. the correlation coefficient depends additionally on sample moments (means, standard deviations) and sample covariance of the data.

This study concentrates on similarities involving any two signals/images. In a particular case, e.g. image quality assessment task one of these images can be interpreted as a reference. Thus measures used in this task are quite often called full reference measures. As already stated earlier this paper will not consider a visual image perception task, which requires different similarity or image quality measures based on scale invariance (human perception property). The main interest is in finding out if widely used measures are translation invariant measures. A further aim of this analysis is to educate a user and thus help him with a selection of an appropriate similarity measure for his particular application.

2.2 Measure properties

Usually, it is supposed that measures fulfil some conditions or exhibit some properties. For example, distance measure is called a metric if it fulfils the following four conditions: positive definiteness (non-negativity: $d(x,y) \geq 0$ and coincidence axiom: $d(x,y)=0$ if and only if $x = y$), symmetry $d(x,y)=d(y,x)$ and the triangle inequality $d(x,z) \leq d(x,y)+d(y,z)$. Thus, the Euclidian measure or distance (1) is a metric, whereas the squared Euclidian measure is not a metric as it does not satisfy the triangle inequality. Measures should not be metrics in order to be successful, e.g. in image similarity/quality assessment. Here, the most prominent examples such as squared Euclidian distance or SSIM can be mentioned. Two more important properties of distance measures are: translation invariance

$$d(p_1+c, p_2+c)=d(p_1, p_2) \tag{6}$$

and scale invariance

$$d(c \cdot p_1, c \cdot p_2)=d(p_1, p_2) \tag{7}$$

defined for all variables/parameters p_i and some fixed constant c (Palubinskas 2014, Sect. 2.1). Parameters p_i can be original image values x_i or means or standard deviations in our case. From (6) follows directly (see Appendix 1)

$$d(p, p+c) = \text{constant} \tag{8}$$

for all p and some fixed c, what means that translation invariance implies an independence of the measure of the absolute parameter values or equivalently dependence only on the relative relation, e.g. difference of the parameters (Palubinskas 2014). For example, correlation coefficient is both translation and scale invariant with respect to original data values x, y. Thus, the selection of a particular measure is application dependent, e.g. for image matching, clustering or classification applications translation invariant measures such as MSE can be more suitable whereas for visual perception applications scale invariant measures such as SSIM are preferable (Palubinskas 2014).

Further, most of the measures can be divided into two large groups: geometric and probabilistic. Geometric measures, e.g. MSE, are calculated directly on original data values *x, y*. Probabilistic measures are based on probability density function (pdf), which can be estimated from data histogram or modelled by a distribution assumption. Thus, in the latter case a distance/similarity measure depends on the parameters of the pdf: sample means and standard deviations (e.g. luminance and contrast of SSIM).

2.3 Scaling of measures

For scaled/normalized measures the following relationships hold. For distance *d,* normalized to an interval $0 \le d \le 1$, e.g. using $d_{norm} = (d - d_{\min})/(d_{\max} - d_{\min})$, where $d_{\min} = \min\{d_i\}$ and $d_{\max} = \max\{d_i\}$, the corresponding similarity is simply equal to $s = 1 - d$. Analogously, for similarity *s,* normalized to interval $0 \le s \le 1$, the corresponding distance is simply equal to $d = 1 - s$. Quite often such normalized measure is called index, coefficient or degree of measure. Only such normalized measures are considered in this work.

2.4 Composite similarities

Two or more similarity measures can be combined/composed by averaging, summation and/or multiplication operators. For example, two distance measures d_1 and d_2, each normalized to the interval $0 \le d_1, d_2 \le 1$, are first transformed to individual similarities $s_1 = 1 - d_1$ and $s_2 = 1 - d_2$ and then a composite similarity is calculated by averaging

$$s_{ave} = \frac{s_1 + s_2}{2} = 1 - \frac{d_1}{2} - \frac{d_2}{2} = 1 - \frac{d_1 + d_2}{2} \quad (9)$$

or multiplication

$$s_{mult} = s_1 \cdot s_2 = (1 - d_1) \cdot (1 - d_2) = 1 - d_1 - d_2 + d_1 \cdot d_2 . \quad (10)$$

The following relation holds (see Appendix 2)

$$s_{mult} \le s_{ave} . \quad (11)$$

Of course, mixed composite measures are possible, e.g. $s = 1/2 \cdot (s_1 + s_2) \cdot s_3$. All distances and similarities analysed in this work are summarized in Table 1 and some of their properties in Table 2.

Table 1.

Table 2.

2.5 MSE

The MSE is a very popular distance measure, which is based on original data *x, y*, and is defined as follows

$$d_{MSE} = (1/N) \cdot \sum_{i=1}^{N} (x_i - y_i)^2 . \quad (12)$$

The normalized version of MSE (nMSE) is used in this work

$$d_{nMSE} = \left(1/R^2\right) \cdot d_{MSE} , \quad (13)$$

where $0 \le d_{nMSE} \le 1$. It is inversely related to other known distance measure called peak signal-to-noise measure

$$PSNR = 10 \cdot \log_{10}\left(1/d_{nMSE}\right) . \quad (14)$$

Thus PSNR exhibits no additional information in comparison with MSE and is not considered in this work.

It has been shown experimentally in (Wang and Bovik 2009) that MSE can be rather poor in some cases, especially for visual image quality perception. Some properties and aspects of MSE have been analysed, but still not all reasons are known for such behaviour of MSE. In this paper theoretical and experimental analysis helps to see what is really behind MSE.

Using sampled statistics: $\mu_x, \sigma_x{}^2$ and correlation coefficient distance measure nMSE (13) can be rewritten as already proposed in (Ward and Folland 1991) and (Horé and Ziou 2013)

$$\begin{aligned} d_{nMSE} &= \left(1/R^2\right)\cdot(1/N)\textstyle\sum_{i=1}^{N}(x_i - \mu_x + \mu_x - y_i + \mu_y - \mu_y)^2 \\ &= \left(1/R^2\right)\cdot\left((\mu_x - \mu_y)^2 + \sigma_x{}^2 + \sigma_y{}^2 - 2\cdot\sigma_x\cdot\sigma_y\cdot\rho\right) \end{aligned} . \tag{15}$$

The nMSE is a sum of two distances: d_1 – normalized squared Euclidian measure for means

$$d_1 = \left(1/R^2\right)\cdot(\mu_x - \mu_y)^2 \tag{16}$$

and d_2

$$d_2 = \left(1/R^2\right)\cdot(\sigma_x{}^2 + \sigma_y{}^2 - 2\cdot\sigma_x\cdot\sigma_y\cdot\rho) . \tag{17}$$

Because of $0 \le (d_1 + d_2) \le 1$ the similarity of nMSE can be defined as $s_{nMSE} = 1 - (d_1 + d_2)$ (see Table 1, T2.3). Moreover, the nMSE is translation invariant with respect to variables *x, y* (see Table 2). For the comparison with SSIM and discussion, see sections 2.6 and 2.8, respectively.

2.6 SSIM

The SSIM measure proposed in (Wang and Bovik 2002) and (Wang *et al.* 2004) can be written as a composite measure (multiplication) of tree similarities: luminance (means) s_1, contrast (standard deviations) s_2 and correlation coefficient s_3

$$s_{SSIM} = s_1 \cdot s_2 \cdot s_3 = \frac{2 \cdot \mu_x \cdot \mu_y}{\mu_x{}^2 + \mu_y{}^2} \cdot \frac{2 \cdot \sigma_x \cdot \sigma_y}{\sigma_x{}^2 + \sigma_y{}^2} \cdot \rho \, . \tag{18}$$

Constants for avoiding singular case (zero in the denominator) are omitted for simplicity during theoretical analysis.

The SSIM is based on the same moments and correlation coefficient as MSE (15). So this is the first observation/property of MSE and SSIM: both measures are composed of the same parameters which are only combined in a different way.

It can be seen easily that the first two similarities are generalized Dice measures (Cha 2007) and (Dice 1945), though the authors of SSIM never mentioned it (see Table 1, T1.3).

In order to better understand the behaviour of the composite measures MSE and SSIM a more detailed look at their components (basic measures) is needed.

2.7 Basic measures

2.7.1 Dice measure

Dice similarity measure was independently introduced by botanists Dice (1945) and Sørensen (1948) for measuring similarity of two samples

$$s_{Dice} = (2 \cdot \# C)/(\# A + \# B), \tag{19}$$

where #*A* and #*B* are the number of species in two samples and #*C* is the number of species shared by the two samples. This measure can be viewed as a measure defined over two sets *A* and *B*

$$s_{Dice} = (2 \cdot A \cap B) / (|A| + |B|), \tag{20}$$

where |*A*| is a number of elements in *A*. For example, for the following two sets *A*={*S1, S3*} and *B*={*S1, S2, S4*} the Dice measure is equal to $s_{Dice} = 2/5$. Further, it can be generalized to a vector form of binary and real data (Cha 2007) also known as Hodgkin-Richards index (Hodgkin and Richards 1987)

$$s_{Dice} = (2 \cdot X \cdot Y) / \left(X^2 + Y^2\right). \tag{21}$$

For the above presented example of sets binary vectors are the following: *X*=(1,0,1,0) and *Y*=(1,1,0,1).

It can be seen easily that the first two similarities in SSIM are Dice measures for a special case: for non-negative scalar pairs $X = \mu_x, Y = \mu_y$ and $X = \sigma_x, Y = \sigma_y$, respectively. The authors/inventors of SSIM measure never mentioned this fact in their numerous publications about SSIM, e.g. (Sampat *et al.* 2009) and (Brunet *et al.* 2012) and only one publication is known up to now (Mo *et al.* 2014), where it is observed that the first term in SSIM is a Hodgkin-Richards index for means. So this is the second observation/property of SSIM: SSIM is composed of two generalized Dice measures: one for means and another one for standard deviations.

Dice similarity measure values for real mean values $\mu_x = \{-128,...,127\}$ and $\mu_y = \{-128,...,127\}$ are presented in Figure 1(a). It is easy to see from (21) that this measure is unstable for zero point (0,0) and can be used as similarity measure only if both values are of the same sign.

Figure 1.

So the third observation for SSIM is its instability around the zero point (0,0) and the fourth one – it can only be used for data of the same sign. The authors of SSIM solve these problems by introducing small constants, and restricting the usage to non-negative data only, respectively.

It can be seen directly from (21) that, e.g. for $\mu_x = 0$ and any $\mu_y = \{0,...,255\}$ the Dice similarity measure is equal to 0 (independent on μ_y). Moreover, it depends not only on the relative difference of two values, but also on the absolute values.

The fifth observation for Dice measure is, and this applies for SSIM too, that it depends on the absolute values of input parameters what the consequence of scale invariance property (7) is. First, it is insensitive at all if one of the parameters is equal 0. Second, its sensitivity is decreasing by the increase of absolute parameter values. This can be a severe drawback in applications assuming translation invariance (6).

Similarity measure should be dependent only on the relative difference of the two parameters and independent of the absolute parameter values (6) for many applications assuming translation invariance. Here, it should be noted that SSIM was designed originally for visual image quality assessment by approximately simulating the human eye behaviour (Weber's law or scale invariance). Unfortunately, this measure is spreading into other applications requiring only relative data comparison, e.g. classification, clustering, matching, change detection, and image fusion where it may not be optimal and lead to wrong results.

2.7.2 Squared Euclidian distance

Normalized squared Euclidian distance for means is defined as

$$d_{nSE} = \left(1/R^2\right) \cdot (\mu_x - \mu_y)^2 \tag{22}$$

and fulfils $0 \le d_{nSE} \le 1$. Its corresponding similarity measure $s_{nSE} = 1 - d_{nSE}$ depends only on the relative mean difference. Thus, it fulfils the distance measure requirement for translation invariance (6), which is illustrated in Figure 1(b). Moreover, it is easy to see that it can be defined for real data (independent of the sign).

2.7.3 Correlation coefficient

Correlation coefficient ρ defined in (2) assumes a linear relationship between x and y and is generally between -1 (negatively correlated objects) and 1 (fully correlated objects). It is 0 for uncorrelated objects. For many tasks, e.g. image processing applications listed in the Introduction section, ρ is usually non negative: $0 \le \rho \le 1$. One can set 0 for negative ρ. Thus, the corresponding distance measure in this case is $d_\rho = 1 - \rho$. It depends on the same parameters (sampled moments) as MSE and SSIM except the covariance term. Moreover, it is both translation and scale invariant with respect to *x, y* (see Table 2).

2.8 Discussion

Here the analysis of composite similarity measures s_{SSIM}, s_{nMSE} based on the information presented in Table 1 (T1.3, T2.3) and Table 2 will be performed.

First, it is easy to observe that under the assumptions of $\mu_x = \mu_y,\ \sigma_x = \sigma_y$ the SSIM similarity is simply reduced to the correlation coefficient $s_{SSIM} = \rho$. Second, under the following assumptions of $\mu_x = \mu_y,\ \sigma_x = \sigma_y = 1,\ R = \sqrt{2}$ the nMSE similarity is also reduced to the correlation coefficient $s_{nMSE} = \rho$, what shows an identity of both measures under some (quite similar) conditions.

Analysis of nMSE second similarity term based on distance (17) shows its dependence on the absolute values of two standard deviation values (see Figure 2). Thus, nMSE shares the same property as SSIM. Only for $\rho = 1$ it converges to a squared Euclidian measure as can be seen in Figure 2 and Table I (T2.6).

Figure 2.

This last observation has served as an inspiration for a new similarity measure, the composite image distance/similarity measure based on means, standard deviations and correlation coefficient (CMSC), which exploits the advantages of both MSE and SSIM measures at the same time being translation invariant with respect to means and standard deviations (Palubinskas 2014).

2.9 New similarity measure CMSC

After analysing the two most popular composite similarity measures MSE and SSIM, a new composite image similarity measure based on means, standard deviations and correlation coefficient (CMSC) consisting of the three components: two normalized squared Euclidian measures and one correlation coefficient (Table 1, T3.1) can be proposed. Depending on the way of combination, e.g. these three versions are possible: CMSCam uses averaging and multiplication of individual similarities, CMSCm - only multiplication of similarities (Palubinskas 2015) and CMSCa – only averaging of similarities. The formulas for new measures are

$$s_{CMSCam} = \left(1 - \left(1/2\right) \cdot \left(d_1 + d_2\right)\right) \cdot \rho \,, \tag{23}$$

$$s_{CMSCm} = (1 - d_1) \cdot (1 - d_2) \cdot \rho \,, \tag{24}$$

$$s_{CMSCa} = \left(1/3\right) \cdot \left(2 - \left(d_1 + d_2\right) + \rho\right), \tag{25}$$

where distances are defined in Table 1 (T3.1). For the normalization of the second distance d_2 including standard deviations, a two times smaller constant $R/2$ can be used. For a proof of the statement see Appendix 3.

All proposed measures exhibit translation invariance property and thus are more suitable as similarity/quality measures not only for images but also for signals. From (11) directly follows $s_{CMSCm} \leq s_{CMSCam}$. Here is worth to note that averaging of similarities gives a possibility for weighting.

Generating of the CMSC can be seen as a general way (framework or guide) to produce composite similarity/quality measures, e.g. by introducing additionally gradient, texture, spectral information (similarly as it was already proposed in (Blasch *et al.* 2008) for SSIM), or higher sample moments: skewness and kurtosis to account for pdf form variations.

3. Experiments

Experiments are performed both for simulated and distorted real data in order to verify/validate theoretical results and/or to identify new properties of measures.

3.1 Illustration of theory using simulated data

Simulated data are produced by generating two random normally distributed signals/images x and y with a given means μ_x, μ_y, standard deviations σ_x, σ_y and correlation coefficient ρ between them. Of course, any other distribution could be used with the following calculation of sample statistics. The used image dimensions have a size of 256x256 pixels and all necessary parameters are estimated on the whole image. Several experiments were performed to analyse the behaviour of similarity measures, e.g. dependence on various parameter settings, or illustrate the theoretical results. Here, the following three examples illustrate the main results/trends.

3.1.1 Mean influence

In the first experiment, 155 pairs of images are generated with the following properties: $\mu_x \in \{1,...,155\}$ (in order to stay in 8bit representation), $\mu_y - \mu_x = 100$, $\sigma_x = \sigma_y = 50$ and $\rho = 0.5$. This configuration allows us to investigate the dependence of the similarity measure on absolute mean value. As can be seen in Figure 3, SSIM is dependent on the absolute mean value and its value increases with the increase of the absolute mean value. Other measures exhibit no such dependency, what can be seen directly from formulas. Nevertheless, the values are presented for the completeness sake. Moreover, CMSCm exhibits lower values than CMSCam (11). Both results support well theoretical observations of sect 2. Additionally, CMSCa has similar values to nMSE. Similar trends are observable for other parameter value settings.

Figure 3.

3.1.2 Standard deviation influence

In the second experiment, 76 pairs of images are generated with the following properties: $\sigma_x \in \{1,...,76\}$ (in order to stay in 8bit representation), $\sigma_y - \sigma_x = 50$, $\mu_x = \mu_y = 127$ and $\rho = 0.5$. This configuration allows to investigate the similarity measure dependence on the absolute standard deviation value. As can be seen in Figure 4, SSIM is dependent on the absolute standard deviation value and its value increases with the increase of the absolute standard deviation value (similarly as for the mean value, see previous sub-section). In this case, also the nMSE appears to be dependent on the absolute standard deviation value, but with an inverse trend when compared to SSIM, that is, its value decreases with the increase of the absolute standard deviation value. Again, new measures exhibit no such dependency, what can be seen directly from formulas. A relation between CMSCam and CMSCm remains the same as in the previous experiment. The results support well the theoretical observations of sect 2.

Figure. 4

3.1.3 Correlation coefficient influence

In the third experiment, 11 pairs of images are generated with the following properties: $\rho \in \{0, 0.1, \ldots, 0.9, 1\}$, $\mu_x = \mu_y = 1$ and $\sigma_x = \sigma_y = 127$. This configuration allows to investigate the similarity measure dependence on the correlation coefficient value. The results are presented in Fig. 5. SSIM=CMSCam behaves as expected that is optimal. In this case, CMSCam=CMSCm according to the definition. The nMSE values are shifted to higher values, thus reducing dynamic range (sensitivity) of the measure. The last observation is valid for CMSCa too.

Figure 5.

On one hand, experiments on simulated data show different properties of SSIM (dependence on absolute mean and standard deviation values) and MSE (dependence on absolute standard deviation values) and simultaneously confirm theoretical results. On the other hand, new measures are translation invariant with respect to both means and standard deviations and thus are more suitable for similarity/quality assessment in applications assuming translation invariance.

3.2 Simulation based on real data

Several experiments were performed on real data to investigate the similarity/quality measures. In this paper, WorldView-2 (WV-2) satellite optical remote sensing panchromatic image (512x512 pixel size, 0.5 meter pixel resolution, Figure 6), corrupted with different types of distortions: mean shift and contrast stretching, various types of noise (additive, multiplicative and impulsive) and blurring is used. To avoid stationary problems, similarity measures are calculated locally on small patches 8x8 pixel size, followed by averaging.

Figure 6.

3.2.1 Mean shift

A small part of WV-2 satellite panchromatic image is 9bit and exhibits the following statistics: $\mu = 133.91,\ \sigma = 96.37$ and $\min = 0,\ \max = 511$. In order to avoid clipping of pixel values after a mean shifting, the image is modified/converted into 10bit representation, simultaneously increasing the number of possible experiments. The mean of the image is shifted from 133.91 to 643.91 to remain in 10bit radiometric resolution and this modified image is compared with the original image having mean 133.91. It is obvious, that such mean shift does not influence the standard deviation and the correlation coefficient is always equal to 1. Similarities in dependence of mean difference are presented in Figure 7. SSIM exhibits a slightly different type of slope and is more sensitive (higher dynamic range) than other measures. All measures behave as expected, that is: decreasing with increasing mean difference. In this case, nMSE=CMSCm.

Figure 7.

3.2.2 Contrast stretching

The mean shifted ($\mu = 512$) and 10bit panchromatic image is used in this experiment. The standard deviation of the image is changed from 1 to 129 in order to keep data in 10bit radiometric resolution. Again, the correlation coefficient is equal to 1. Similarities in dependence of standard deviation difference are presented in Figure 8. SSIM again exhibits different types of slope and is much more sensitive than other measures. Nevertheless, all measures behave as expected, thus decrease with an increasing standard deviation difference.

Figure 8.

3.2.3 Additive Gaussian noise

In this experiment an additive Gaussian noise n is added to the modified panchromatic image s (mean shifted and 10bit): $x = s + n$, where $\mu_s = 512$, $\sigma_s = 96.37$ and $n \sim N(0, \sigma^2)$. Standard deviation of noise σ ranges from 1 to 39 for 10bit data. Similarities in dependence of standard deviation of noise are presented in Figure 9. CC denotes the correlation coefficient between the two images compared. In this case, $nMSE \approx 1$ (quite insensitive) and $CMSCam \approx CMSCm$. Moreover, CMSCam curve follows approximately SSIM, CMSCa follows CC, but with a slightly different slope. Additionally, CMSCam exhibits higher sensitivity than SSIM.

Figure 9.

3.2.4 Multiplicative noise (speckle)

In the following experiment a multiplicative noise n (speckle) is added to the modified panchromatic image s (10bit): $x = s \cdot n$,

where $n = (1/L) \cdot \sum_{i=1}^{L} \left| complex(re_i / \sqrt{2}, im_i / \sqrt{2}) \right|^2$, real and imaginary parts of complex variable $re, im \sim N(0,1)$, L – number of looks and $\mu_s = 133.91$ $\sigma_s = 96.37$. The number of looks L ranges from 15 to 100 for 10bit data. Similarities in dependence of the number of looks are presented in Figure 10. CMSCam curve follows approximately SSIM, CMSCa follows CC and nMSE is quite insensitive. Again, in this case $CMSCam \approx CMSCm$.

Figure 10.

3.2.5 Impulsive (salt & pepper) noise

In this experiment an impulsive noise n is added to original panchromatic image s (9bit): $x = s + n$, where $\mu_s = 133.91$ $\sigma_s = 96.37$, for noisy pixels $x = n$, $n = 0$ *or* 511 and K – the number of noisy pixels in percentage %. Similarities in dependence of the number of noisy pixels K are presented in Figure 11. CMSCam curve approximately follows SSIM, CMSC follows CC and nMSE is quite insensitive. In this case, $CMSCam \approx CMSCm$.

Figure 11.

3.2.6 Blurring

In this experiment an original panchromatic image s (9bit) is blurred using a low pass filter: $x = \left| FFT^{-1}(FFT(s) \cdot LPF) \right|$, where $\mu_s = 133.91$ $\sigma_s = 96.37$ and LPF – low pass filter (Gaussian) in the Fourier domain with a desired cutoff frequency (blurring parameter B). Similarities in dependence of blurring parameter B are presented in Figure 12. CMSCam curve approximately follows SSIM, CMSC follows CC and nMSE is quite insensitive. In this case, $CMSCam \approx CMSCm$. Moreover, CMSCam appears to be more sensitive to blurring than SSIM.

Figure 12.

Experiments on real data performed support both theoretical and simulation results. Several new properties can be observed: nMSE is quite insensitive to all types of distortions maybe except mean shift and impulsive noise; CMSCam and CMSCm exhibit different, but at the same time not too diverging behaviour from SSIM for noise and blurring distortions; the correlation coefficient is somewhere between nMSE and SSIM in its sensitivity. It seems that a new measure CMSC exhibits higher sensitivity than SSIM for additive Gaussian and blurring distortions which are the main distortions occurring during the acquisition of optical remote sensing imagery. Similar trends were observed also for other types of images, e.g. "Lena" image. Of course validation of a real application with ground truth data would be desired to confirm the potential of a new measure.

3.3 General application use scenario

Here, a general application use scenario for a new measure CMSC is introduced or in other words, it is shown when and how a new measure can be used. Let's assume that the aim of the application is to compare distance/similarity of two random variables x,y with a set of values as defined at the beginning of the Sect 2. For example, for image processing applications x,y can be two patches of images. For restoration applications the first image can be assumed to be the original image and the second one the distorted or reconstructed image (accuracy assessment task). For filtering/de-noising application variables x,y can be different patches of the same image. For matching or change detection applications, these patches can originate from any two different images. MSE is used quite often for such type of comparison (patch-based), which is known to be translation invariant with respect to variables x,y. In the case, when the comparison should be performed with respect to parameters such as sample means and standard deviations, it is known that MSE is not translation invariant with respect to the standard deviation. In this case, a new measure CMSC, which is translation invariant with respect to both parameters: means and standard deviations, can be more useful/more suitable than MSE. In optical remote sensing data (imagery) processing applications quite often the assumption about translation invariance holds, thus ensuring quite great potential of a new measure. For example, see an application of a CMSCm measure for the image fusion quality assessment task in (Palubinskas 2015). That way, a new measure can be seen as a possible replacement of MSE, where translation invariance is assumed, e.g. in clustering, classification, matching, change detection and many other applications requiring a relative comparison of values/parameters independent of their absolute values.

4. Discussion

Theoretical analysis of MSE and SSIM similarity measures showed that they depend on the same five parameters: two sample means, two standard deviations and correlation coefficient only combined in a different way. SSIM is composed of two generalized Dice measures, one for means and another one for standard deviations thus inheriting several properties of Dice measure. First, instability around the zero point (0,0) and usage restriction only to data of the same sign. Second, its dependence on absolute mean and standard deviation values (scale invariance property). nMSE shares the last part of the property with SSIM, that is it depends also on the absolute values of standard deviation. Though different in definition, the two measures converge to the same correlation coefficient under quite similar conditions (but very special case). Nevertheless, further experiments on simulated and real data show a great difference of these two measures. Finally, three new composite similarity measures CMSC are translation invariant with respect to means and standard deviations.

Experiments on simulated data support the theoretical analysis concerning possible drawbacks of SSIM (dependence on absolute mean and standard deviation values) and MSE (dependence on absolute standard deviation values) for applications assuming translation invariance. On the other hand, new measures are fully translation invariant thus can be considered as potential candidates to replace MSE in many applications.

Experiments on real data covering various image distortions: mean shift, contrast stretching, various types of noise (additive Gaussian, multiplicative speckle and impulsive salt&pepper) and blurring, are performed on satellite panchromatic image. Mean shift and contrast stretching experiments support theoretical and simulation results concerning SSIM dependence on absolute parameter values. Other experiments (noising and blurring) show that two pairs: SSIM and CMSCam/CMSCm, and CMSCa and CC exhibit similar trends. nMSE appears to be very insensitive and less suitable for these applications. In total, it seems that a new measure CMSCam/CMSCm can replace MSE for many applications requiring similarity/quality measure.

This study has not followed a usual/standard way of introducing a new measure, comparing it with several known measures and showing that it is superior under some conditions of a particular application. This paper could be more seen as a guide how to select an appropriate measure for a particular application. A deep understanding of a measure and application is needed in order to be able to select a right measure for a given application.

Further research can be conducted towards introducing an additional gradient, texture, spectral information for CMSC (similarly as it was already proposed in (Blasch *et al.* 2008) for SSIM), or higher sample moments: skewness and kurtosis to account for different forms of a pdf.

Acknowledgments

I would like to thank DigitalGlobe and European Space Imaging (EUSI) for the collection and provision of the WorldView-2 scene over Munich city.

References

Alberga, V., 2009. Similarity measures of remotely sensed multi-sensor images for change detection applications. *Remote Sensing,* 1, 122-143.

Blasch, E., Li, X., Chen, G., and Li, W., 2008. Image quality assessment for performance evaluation of image fusion. *In: Proceedings of International Conference on Information Fusion*, June 2008 Cologne, Germany, 1-6.

Brunet, D., Vrscay, E.R., and Wang, Z., 2012. On the mathematical properties of the structural similarity index. *IEEE Trans. Image Processing,* 21, 1488-1499.

Cha, S.-H., 2007. Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences*, 1, 300-307.

Dice, L.-R., 1945. Measures of the amount of ecologic association between species. *Ecology*, 26, 297-302.

Dosselmann, R., and Yang, X.D., 2011. A comprehensive assessment of the structural similarity index. *SIViP*, 5, 81-91.

Fukunaga, K., 1972. *Introduction to Statistical Pattern Recognition*, New York, USA: Academic Press.

George, A.G., and Prabavathy, A.K., A survey on different approaches used in image quality assessment. *International Journal of Emerging Technology and Advanced Engineering*, 3, 197-203.

Goshtasby, A.A., 2012. Similarity and dissimilarity measures. In: A.A. Goshtasby, ed. *Image Registration: Principles, Tools and Methods.* London, UK: Springer 27-66.

Hanhart, P., Bernardo, V.B., Pereira, M., Pinheiro, A.M.G., and Ebrahimi, T., 2015. Benchmarking of objective quality metrics for HDR image quality assessment. *EURASIP Journal on Image and Video Processing*, 39, 1-18.

Hodgkin, E.E., and Richards, W.G., 1987. Molecular similarity based on electrostatic potential and electric field *International Journal of Quantum Chemistry*, 32, 105–110.

Horé, A., and Ziou, D., 2013. Is there any relationship between peak-signal-to-noise ratio and structural similarity index measure? *IET Image Process.*, 7, 12-24.

Jain, A.K., and Dubes, R.C., 1998. *Algorithms for Clustering Data,* New Jersey, USA: Prentice Hall.

Li , S., Li, Z., and Gong, J., 2010. Multivariate statistical analysis of measures for assessing the quality of image fusion, *International Journal of Image and Data Fusion*, 1(1), 47-66.

Liu. Z., and Wu, W., 2011. The use of the contrast sensitivity function in the perceptual quality assessment of fused image, *International Journal of Image and Data Fusion*, 2(1), 93-103,

Mo, R., Ye, C., and Whitfield, P.H., 2014. Application potential of four nontraditional similarity metrics in hydrometeorology. *Journal of Hydrometeorology*, 15, 1862-1880.

Mohammadi, P., Ebrahimi-Moghadam, A., and Shirani, S., 2015. Subjective and objective quality assessment of image: a survey. *Majlesi Journal of Electrical Engineering*, 9, 55-83.

Palubinskas, G., 2013. Fast, simple and good pan-sharpening method. *Journal of Applied Remote Sensing*, **7**, 1-12.

Palubinskas, G., 2014. Mystery behind similarity measures MSE and SSIM. *In: Proceedings of International Conference on Image Processing*, October 2014 Paris, France: IEEE, 575-579.

Palubinskas, G., 2015. Joint quality measure for evaluation of pansharpening accuracy. *Remote Sensing*, 7, 9292-9310.

Palubinskas, G., 2016. Model-based view at multi-resolution image fusion methods and quality assessment measures, *International Journal of Image and Data Fusion*, **7**(3), 203-218.

Richards, J.A., and Jia, X., 1999. *Remote Sensing Digital Image Analysis*. 3rd ed. Berlin, Germany: Springer.

Sampat, M.P., Wang, Z., Gupta, S., Bovik, A.C., and Markey, M.K., 2009. Complex wavelet structural similarity: a new image similarity index. *IEEE Trans. Image Processing*, 18, 2385-2401.

Sørensen, T., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskab*, 5, 1–34.

Ward, M.N., and Folland, C.K., 1991. Prediction of seasonal rainfall in the north Nordeste of Brasil using eigenvectors of sea-surface temperature. *International Journal of Climatology*, 11, 711-743.

Wang, Z., and Bovik, A.C., 2002. A universal image quality index. *IEEE Signal Processing Letters*, 9, 81–84.

Wang, Z. and Bovik, A.C., 2009, Mean squared error: love it or leave it? A new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26, 98-117.

Wang, Z., Bovik, A.C., Sheikh, H.R., and Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Processing*, 13, 600-612.

Appendix 1. Proof of equation (8).

To prove (8) first I define $p_1 = p_1$ and $p_2 = p_1 + c$. Then by inserting variables I can write $d(p_1, p_2) = d(p_1, p_1 + c)$ and $d(p_1 + c, p_2 + c) = d(p_2, p_2 + c)$. Applying (6) leads directly to $d(p_1, p_1 + c) = d(p_2, p_2 + c)$ for any p_1 and p_2. Thus (8) has been proved.

Appendix 2. Derivation of relation between equations (9) and (10).

To establish a relation between the two combination methods (9, 10) I can rewrite (10)

$$s_{mult} = 1 - \frac{d_1}{2} - \frac{d_2}{2} - \frac{d_1}{2} - \frac{d_2}{2} + d_1 \cdot d_2 = s_{ave} + \frac{d_1 \cdot (d_2 - 1) + d_2 \cdot (d_1 - 1)}{2} \quad . \tag{26}$$

Due to $0 \le d_1, d_2 \le 1$ the last term in (26) is non positive and thus the relation (11) holds.

Appendix 3. Derivation of maximum of variance (4).

In order to derive normalization constant for d_2 of CMSC measure (see Table 1, T3.1), first the maximum of the variance (4) should be found. From the assumption $\min \le x_i \le \max$ follows that the following holds $\min \le \mu \le \max$ and $\mu_{\max} = \max - \min = R$. Looking at (4) I can see that variance is maximized only when $x_i = \min$ or $x_i = \max$ for a fixed μ. Thus I can rewrite (4)

$$\sigma^2 = (1/N) \cdot \left(n \cdot (\min - \mu)^2 + m \cdot (\max - \mu)^2\right), \tag{27}$$

where $n + m = N$. Moreover, it is obvious that in this case

$$\mu = (1/N) \cdot (n \cdot \min + m \cdot \max). \tag{28}$$

After some algebra from (27, 28) follows

$$\sigma^2 = (\max - \mu) \cdot (\mu - \min). \tag{29}$$

Maximization of (29) leads to

$$\mu = (\min + \max)/2 . \tag{30}$$

Inserting (30) into (29) results in

$$\sigma_{\max}{}^2 = ((\max - \min)/2)^2 = (R/2)^2 . \tag{31}$$

Table 1. Summary of distance and similarity measures used in this work. R is a normalization constant, e.g. $R=255$ for 8bit data (Palubinskas 2014).

	Distance d_1	Distance d_2	d_3	Ct [1]	Similarity
T1.1 Dice	$\frac{(\mu_x-\mu_y)^2}{\mu_x^2+\mu_y^2}$	-	-	-	$\frac{2\cdot\mu_x\cdot\mu_y}{\mu_x^2+\mu_y^2}$
T1.2 CC $0\le\rho\le1$	$1-\rho$	-	-	-	ρ
T1.3 SSIM	$\frac{(\mu_x-\mu_y)^2}{\mu_x^2+\mu_y^2}$	$\frac{(\sigma_x-\sigma_y)^2}{\sigma_x^2+\sigma_y^2}$	$1-\rho$	M	$(1-d_1)\cdot(1-d_2)\cdot\rho$
T1.4 SSIM $\mu_x=\mu_y,\ \sigma_x=\sigma_y$	0	0	$1-\rho$	M	ρ
T2.1 Normalized squared Euclidian nSE	$\frac{(\mu_x-\mu_y)^2}{R^2}$	-	-	-	$1-d_1$
T2.2 nMSE (original data)	$\frac{\sum_{i=1}^{N}(x_i-y_i)^2}{N\cdot R^2}$	-	-	-	$1-d_1$
T2.3 nMSE (sample moments)	$\frac{(\mu_x-\mu_y)^2}{R^2}$	$\frac{\sigma_x^2+\sigma_y^2-2\cdot\sigma_x\cdot\sigma_y\cdot\rho}{R^2}$	-	S	$1-(d_1+d_2)$
T2.4 nMSE $\mu_x=\mu_y,\ \sigma_x=\sigma_y=1,\ R=\sqrt{2}$	0	$\frac{2\cdot(1-\rho)}{R^2}$	-	S	ρ
T2.5 nMSE $\rho=0$	$\frac{(\mu_x-\mu_y)^2}{R^2}$	$\frac{\sigma_x^2+\sigma_y^2}{R^2}$	-	S	$1-(d_1+d_2)$
T2.6 nMSE $\rho=1$	$\frac{(\mu_x-\mu_y)^2}{R^2}$	$\frac{(\sigma_x-\sigma_y)^2}{R^2}$	-	S	$1-(d_1+d_2)$
T3.1 CMSCam	$\frac{(\mu_x-\mu_y)^2}{R^2}$	$\frac{(\sigma_x-\sigma_y)^2}{(R/2)^2}$	$1-\rho$	AM	$\left(1-\frac{(d_1+d_2)}{2}\right)\cdot\rho$
T3.2 CMSCm	$\frac{(\mu_x-\mu_y)^2}{R^2}$	$\frac{(\sigma_x-\sigma_y)^2}{(R/2)^2}$	$1-\rho$	M	$(1-d_1)\cdot(1-d_2)\cdot\rho$
T3.3 CMSCa	$\frac{(\mu_x-\mu_y)^2}{R^2}$	$\frac{(\sigma_x-\sigma_y)^2}{(R/2)^2}$	$1-\rho$	A	$\frac{2}{3}-\frac{d_1+d_2}{3}+\frac{\rho}{3}$

[1] Combination type: M – multiplication, S – summation, A – averaging

Table 2. Summary of properties for various distance/similarity measures used in this paper.

Measure	MSE [1]	ρ	MSE sample moments		SSIM		CMSC	
Arguments / Properties	x_i, y_i	x_i, y_i	$\mu_{x,y}$	$\sigma_{x,y}$	$\mu_{x,y}$	$\sigma_{x,y}$	$\mu_{x,y}$	$\sigma_{x,y}$
Translation invariant	+	+	+	-	-	-	+	+
Scale invariant	-	+	-	-	+	+	-	-

[1] MSE is expressed in original data *x*,*y*, + stands for satisfied property for a particular argument, - property is not satisfied.

Figure caption list

Figure 1. (a) Dice similarity measure (21) for $\mu_x, \mu_y \in \{-128,...,127\}$ and (b) Euclidian similarity measure (22) for $\mu_x, \mu_y \in \{0,...,255\}$.

Figure 2. Similarity measure nMSE for $\mu_x = \mu_y, \sigma_x, \sigma_y \in \{0,...,255\}$ in dependence of ρ (Palubinskas 2014) .

Figure 3. Similarity measures of SSIM, nMSE, CMSCam, CMSCm and CMSCa for constant mean difference $\mu_y - \mu_x = 100$, $\sigma_x = \sigma_y = 50$ and $\rho = 0.5$ in dependence of the mean value $\mu_x \in \{1,...,155\}$.

Figure 4. Similarity measures of SSIM, nMSE, CMSCam, CMSCm and CMSCa for standard deviation difference $\sigma_y - \sigma_x = 50$, $\mu_x = \mu_y = 127$ and $\rho = 0.5$ in dependence of the standard deviation value $\sigma_x \in \{1,...,76\}$.

Figure 5. Similarity measures of nMSE, CMSCam and CMSCa for $\mu_y = \mu_x = 1$ and $\sigma_x = \sigma_y = 127$ in dependence of the correlation coefficient ρ. In this case SSIM=CMSCam=CMSCm.

Figure 6. WorldView-2 satellite panchromatic image of Munich center (Frauenkirche).

Figure 7. Similarity measures of nMSE, SSIM, CMSCam and CMSCa for $\mu_x = 134$, $\mu_y \in \{134,135,...,510\}$, $\sigma_x = \sigma_y = 96.37$ and $\rho = 1$ in dependence of mean difference $\mu_y - \mu_x$. In this case nMSE=CMSCm.

Figure 8. Similarity measures of nMSE, SSIM, CMSCam, CMSCm and CMSCa for $\sigma_x = 1$, $\sigma_y \in \{1,2,...,129\}$, $\mu_x = \mu_y = 512$ and $\rho = 1$ in dependence of standard deviation difference $\sigma_y - \sigma_x$ (a) and different y-axis scaling (b).

Figure 9. Similarity measures of nMSE, SSIM, CMSCam, CMSCm and CMSCa in dependence of standard deviation of noise $\sigma \in \{1,...,39\}$(a) and different y-axis scaling (b). CC - correlation coefficient. In this case $nMSE \approx 1$, $CMSCam \approx CMSCm$.

Figure 10. Similarity measures nMSE, SSIM, CMSCam, CMSCm and CMSCa: (a) in dependence of number of looks $L \in \{15,...,100\}$ and (b) same as (a) but with different y-axis scaling. CC - correlation coefficient. In this case, $CMSCam \approx CMSCm$.

Figure 11. Similarity measures of SSIM, nMSE, CMSCam, CMSCm and CMSCa in dependence of the number of noisy pixels $K \in \{0\%,...,90\%\}$. CC - correlation coefficient. In this case $CMSCam \approx CMSCm$.

Figure 12. Similarities of nMSE, SSIM, CMSCam, CMSCm and CMSCa measures in dependence of blurring parameter (a) and different y-axis scaling (b). CC - correlation coefficient. In this case $CMSCam \approx CMSCm$.

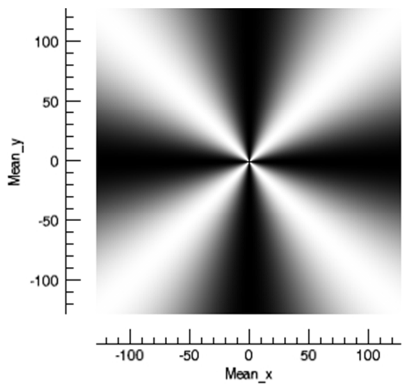

Figure 1. (a) Dice similarity measure (21) for   and (b) Euclidian similarity measure (22) for .

213x152mm (100 x 100 DPI)

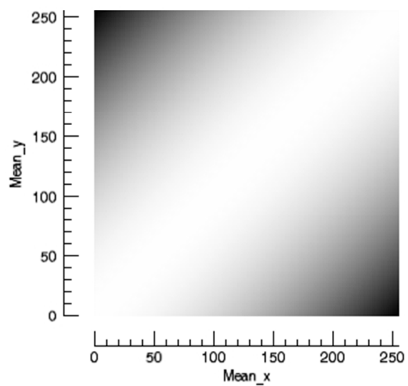

Figure 1. (a) Dice similarity measure (21) for   and (b) Euclidian similarity measure (22) for .

213x152mm (100 x 100 DPI)

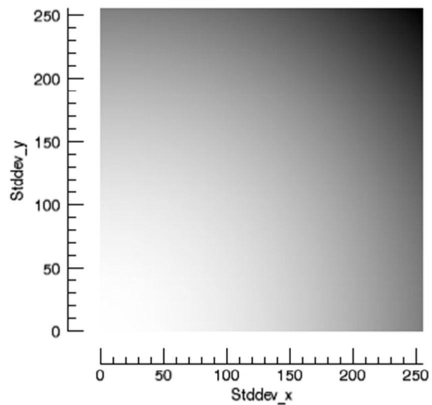

Figure 2. Similarity measure nMSE for , in dependence of (Palubinskas 2014) .

201x130mm (100 x 100 DPI)

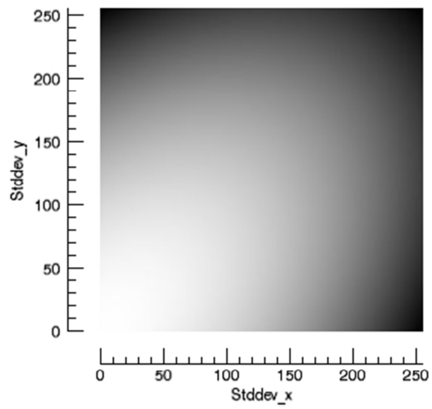

Figure 2. Similarity measure nMSE for , in dependence of (Palubinskas 2014) .

201x130mm (100 x 100 DPI)

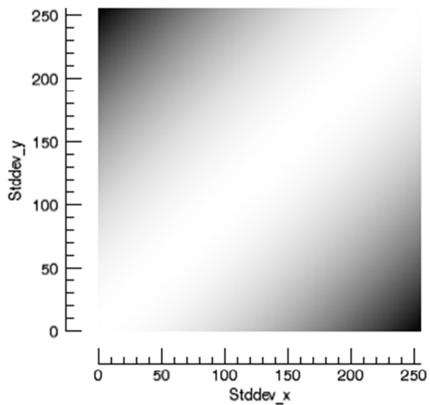

Figure 2. Similarity measure nMSE for , in dependence of (Palubinskas 2014) .

201x130mm (100 x 100 DPI)

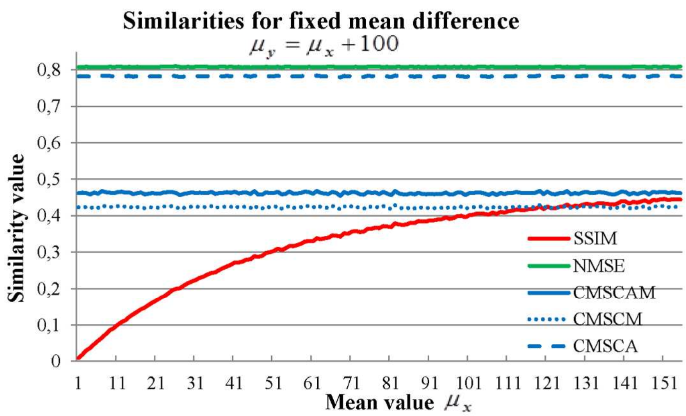

Figure 3. Similarity measures of SSIM, nMSE, CMSCam, CMSCm and CMSCa for constant mean difference , and in dependence of the mean value .

221x312mm (200 x 200 DPI)

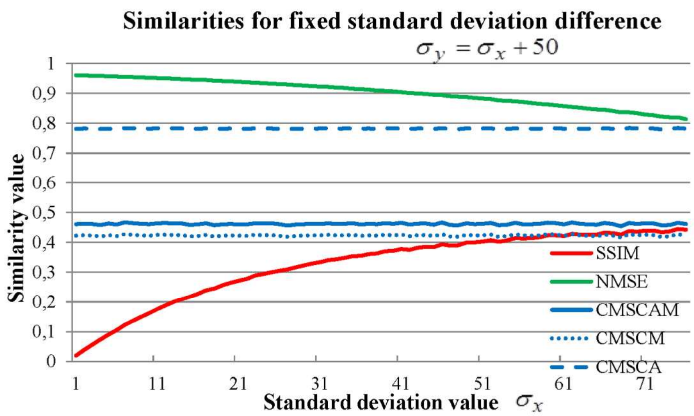

Figure 4. Similarity measures of SSIM, nMSE, CMSCam, CMSCm and CMSCa for standard deviation difference , and in dependence of the standard deviation value .

221x312mm (200 x 200 DPI)

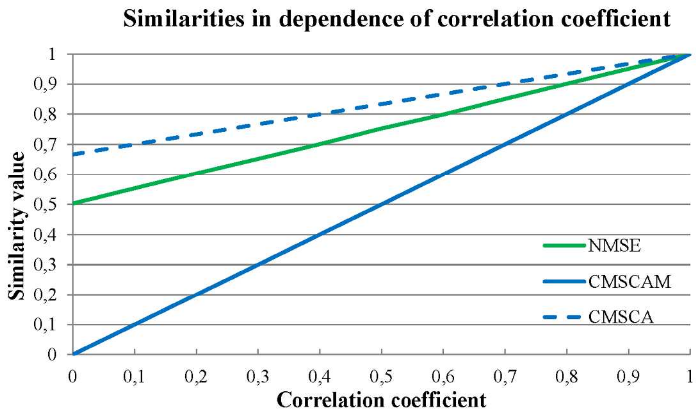

Figure 5. Similarity measures of nMSE, CMSCam and CMSCa for and in dependence of the correlation coefficient . In this case SSIM=CMSCam=CMSCm.

221x312mm (200 x 200 DPI)

Figure 6. WorldView-2 satellite panchromatic image of Munich center (Frauenkirche).

130x130mm (100 x 100 DPI)

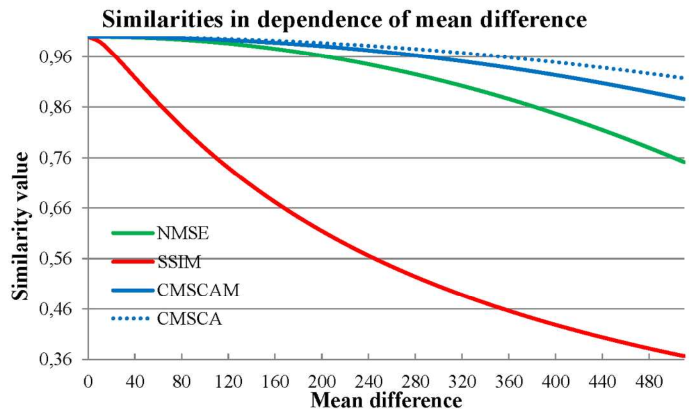

Figure 7. Similarity measures of nMSE, SSIM, CMSCam and CMSCa for , , and in dependence of mean difference . In this case nMSE=CMSCm.

221x312mm (200 x 200 DPI)

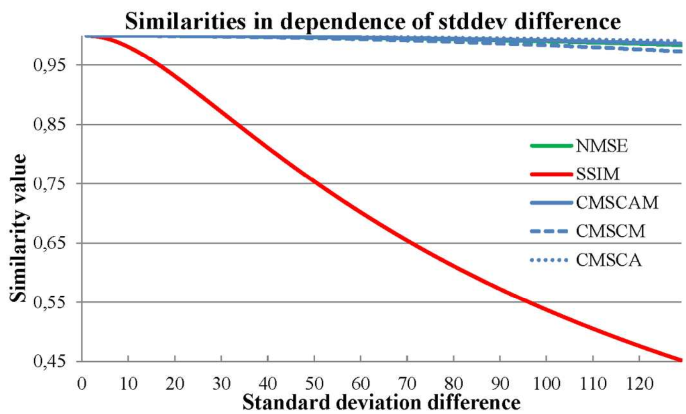

Figure 8. Similarity measures of nMSE, SSIM, CMSCam, CMSCm and CMSCa for , , and in dependence of standard deviation difference (a) and different y-axis scaling (b).

221x312mm (200 x 200 DPI)

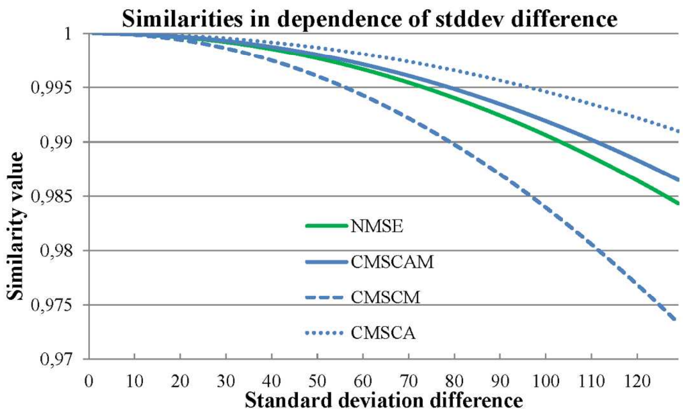

Figure 8. Similarity measures of nMSE, SSIM, CMSCam, CMSCm and CMSCa for , , and in dependence of standard deviation difference (a) and different y-axis scaling (b).

221x312mm (200 x 200 DPI)

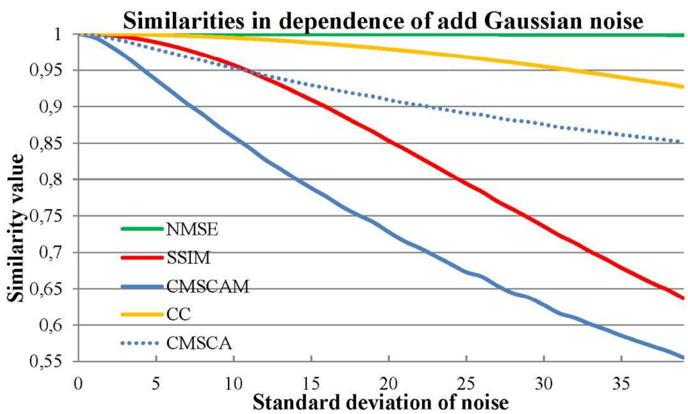

Figure 9. Similarity measures of nMSE, SSIM, CMSCam, CMSCm and CMSCa in dependence of standard deviation of noise (a) and different y-axis scaling (b). CC - correlation coefficient. In this case .

221x312mm (200 x 200 DPI)

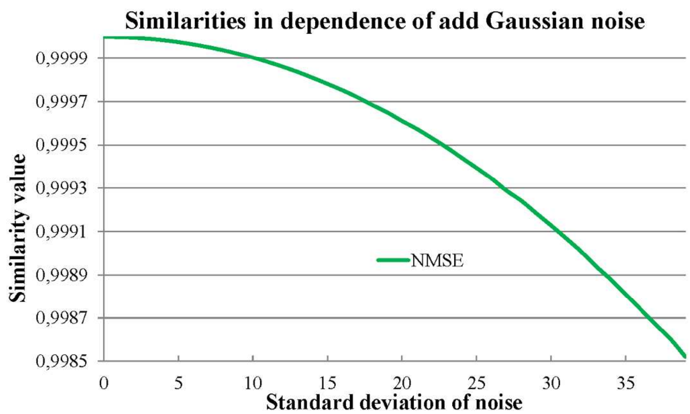

Figure 9. Similarity measures of nMSE, SSIM, CMSCam, CMSCm and CMSCa in dependence of standard deviation of noise (a) and different y-axis scaling (b). CC - correlation coefficient. In this case .

221x312mm (200 x 200 DPI)

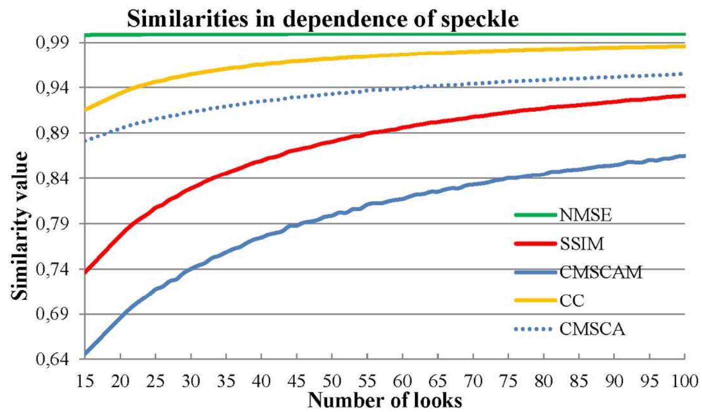

Figure 10. Similarity measures nMSE, SSIM, CMSCam, CMSCm and CMSCa: (a) in dependence of number of looks and (b) same as (a) but with different y-axis scaling. CC - correlation coefficient. In this case, .

221x312mm (200 x 200 DPI)

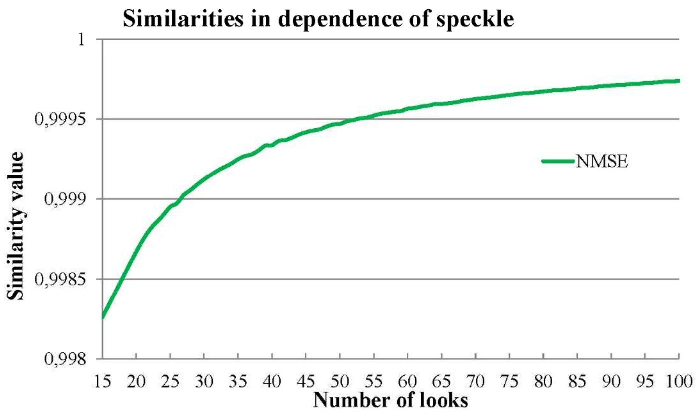

Figure 10. Similarity measures nMSE, SSIM, CMSCam, CMSCm and CMSCa: (a) in dependence of number of looks and (b) same as (a) but with different y-axis scaling. CC - correlation coefficient. In this case, .

221x312mm (200 x 200 DPI)

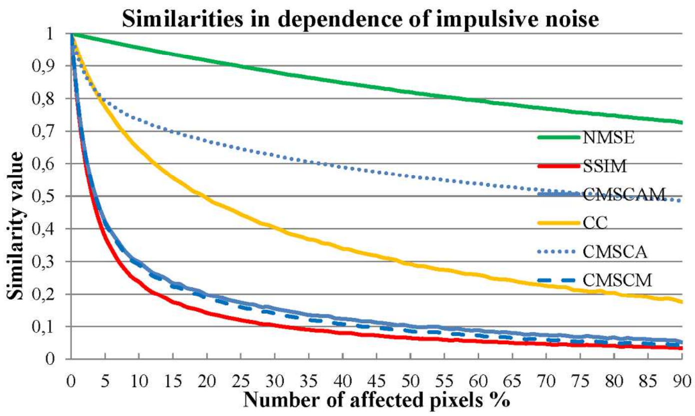

Figure 11. Similarity measures of SSIM, nMSE, CMSCam, CMSCm and CMSCa in dependence of the number of noisy pixels . CC - correlation coefficient. In this case .

221x312mm (200 x 200 DPI)

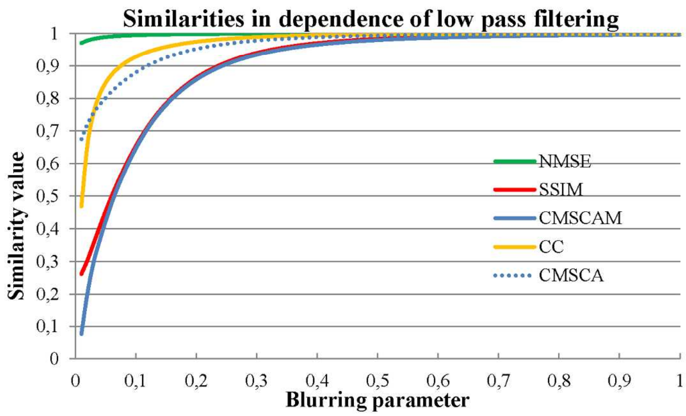

Figure 12. Similarities of nMSE, SSIM, CMSCam, CMSCm and CMSCa measures in dependence of blurring parameter (a) and different y-axis scaling (b). CC - correlation coefficient. In this case .

221x312mm (200 x 200 DPI)

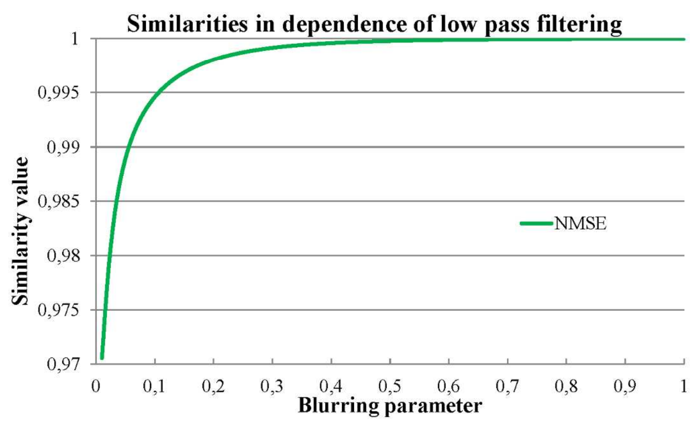

Figure 12. Similarities of nMSE, SSIM, CMSCam, CMSCm and CMSCa measures in dependence of blurring parameter (a) and different y-axis scaling (b). CC - correlation coefficient. In this case .

221x312mm (200 x 200 DPI)