COMPREHENSIVE EVALUATION OF METRICS FOR IMAGE RESEMBLANCE

Marcel PRODAN [1]

Giorgiana Violeta VLĂSCEANU[2]

Costin-Anton BOIANGIU[3]

Abstract

In order to measure image similarity in the field of Computer Science, this study will analyze the main metrics in-depth. Image resemblance metrics play a significant role in various domains, including but not limited to, digital image processing, computer vision, and machine learning. This study embarks on a journey through various pixel-based and structural similarity metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), Multi-Scale Structural Similarity Index (MS-SSIM), and other advanced metrics like Feature Similarity Index (FSIM), Universal Quality Index (UQI), and Visual Information Fidelity (VIF). A comparative analysis of these metrics is conducted, shedding light on their specific pros and cons and their applicability in different contexts. The paper also addresses the importance and role of these metrics in the evolving field of deep learning. Lastly, we discuss the current challenges, and limitations of these metrics, and envision the future scope and advancements in image resemblance metrics.

Keywords: Mean Absolute Error, Mean Squared Error, Peak Signal to Noise Ratio, Multi-Scale Structural Similarity Index (MS-SSIM), Deep Learning for Image Resemblance

JEL Classification: C80, C65

1. Introduction

As computer science continues to progress, the role of image resemblance measures becomes increasingly important across a broad range of applications [1][2][3]. These metrics underpin many fields, including but not limited to digital image processing, computer vision, machine learning, and deep learning, highlighting the necessity for a robust understanding of these measures [3][4][5].

[1] PhD Student, Engineer, University Politehnica of Bucharest, Romania, marcel.prodan@stud.acs.pub.ro
[2] PhD student, Teaching assistant, Engineer., University Politehnica of Bucharest, Romania, giorgiana.vlasceanu@cs.pub.ro
[3] PhD, Professor, Engineer, University Politehnica of Bucharest, Romania, costin.boiangiu@cs.pub.ro

1.1. Purpose of the study

The primary aim of this research is to deliver a comprehensive exploration of the critical metrics employed to assess image resemblance. These metrics are vital for comparing various images and evaluating the performance of different image processing algorithms, serving as a fundamental tool in the rapidly evolving domains of digital image processing and computer vision [2][3].

1.2. Brief overview of image resemblance metrics

Image resemblance metrics, ranging from traditional pixel-based measures such as Mean Absolute Error (MAE) and Mean Squared Error (MSE), to more recent measures like Peak Signal to Noise Ratio (PSNR), provide the basis for this study [6-new|5]. Further, we delve into structural similarity metrics like the Structural Similarity Index (SSIM) and the Multi-Scale Structural Similarity Index (MS-SSIM) [1][10]. Additionally, advanced metrics such as the Feature Similarity Index (FSIM), Universal Quality Index (UQI), and Visual Information Fidelity (VIF) will be analyzed in depth [11][12][13].

1.3. Importance and applications of image resemblance in computer science

In addition to detailing these metrics, the paper explores their importance and role in the burgeoning field of deep learning. With the application of deep learning algorithms, researchers can achieve superior results in tasks related to image resemblance [4][5][14][12]. Despite significant strides in the field, there are still numerous challenges and limitations with these metrics that need addressing. Consequently, an integral part of this study involves an in-depth examination of these issues, aiming to lay the groundwork for future research in this area [2][3][13]. The objective of this comprehensive study is to provide researchers and practitioners with a detailed understanding of the range of image resemblance metrics currently available. The hope is that this investigation will facilitate a better selection and application of these metrics across different scenarios, thereby significantly improving the efficacy and precision of tasks related to image processing and computer vision [1][2][3].

2. Image Resemblance: An Overview

Image resemblance, at its core, is a measure that allows for the comparison of two or more images to assess their similarity. A detailed understanding of this concept is integral to many aspects of computer science, particularly in fields such as image processing, computer vision, and machine learning [1][2][3].

2.1. Definition and need for Image Resemblance Metrics

In simple terms, image resemblance metrics provide quantitative means of assessing the degree of similarity between two or more images [2]. These metrics work by evaluating various features, such as color, intensity, texture, or structure of the images, and then providing a quantified value representing the degree of resemblance [1][2]. Traditionally, these metrics relied on pixel-level comparisons where individual pixel values in the images were compared directly. These pixel-based metrics included the Mean Absolute Error (MAE) and the Mean Squared Error (MSE), which offered a simple and straightforward way of comparing image similarity [6]. Although these metrics have served a purpose, they have limitations in capturing perceived similarities or differences in image content due to their inability to model the complex characteristics of human visual perception [6][7].

As a result, more sophisticated measures such as the Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Multi-Scale Structural Similarity Index (MS-SSIM) were developed [1][10]. The last two measures go beyond pixel-level comparisons to evaluate structural and statistical similarities between images, thereby aligning better with human visual perception [1][7][10]. The need for image resemblance metrics is widely recognized across various fields of computer science. In image processing, for instance, these metrics are crucial in quantifying the effectiveness of algorithms designed for image restoration, compression, and enhancement [2][3]. They provide a benchmark to evaluate the performance of these algorithms and guide their development [2][3].

Furthermore, in the broader fields of machine learning and deep learning, image resemblance metrics have significant applications. They are used for training models to recognize patterns in images, guiding the optimization process and providing a measure to evaluate the performance of models [4][5]. Applications in these fields range from object recognition and image segmentation in computer vision to more advanced tasks in deep learning such as generating synthetic images with Generative Adversarial Networks (GANs) [5][14][12].

As our reliance on digital images continues to grow, the need for robust, reliable, and sophisticated image resemblance metrics becomes more pronounced. Therefore, understanding these metrics and their applications is critical for researchers and practitioners in the field of computer science [2][3][4][5].

2.2. Brief history and evolution of image resemblance metrics

The history of image resemblance metrics parallels the evolution of computer science and digital imaging technologies. Initially, the simple task of comparing images necessitated measures that were straightforward and computationally inexpensive. This requirement led to the development of pixel-based metrics like Mean Absolute Error (MAE) and Mean

Squared Error (MSE), which worked by performing direct comparisons of individual pixel values between two images [6].

While MAE and MSE offered a rudimentary form of image comparison, it soon became apparent that these measures had limitations. Specifically, they were insufficient for capturing perceptual similarities and differences between images due to their simplistic approach [6][7]. Moreover, they failed to account for spatial and structural information, which are key aspects of how humans perceive images. As the field matured and the need for more sophisticated image comparison metrics grew, Peak Signal to Noise Ratio (PSNR) was introduced [6]. PSNR provided a more nuanced measure than MAE or MSE by comparing the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Although a significant advancement, PSNR still primarily relied on pixel-level comparisons and thus carried some of the limitations of its predecessors [7]. The realization of the limitations of pixel-based metrics and the necessity for measures that mirrored human visual perception led to the development of structural similarity metrics. Among these, the Structural Similarity Index (SSIM) became a popular measure due to its ability to capture structural and statistical similarities between images [1][10]. It represents a significant step forward in the field, as it considers that the human visual system is highly adapted to extract structural information from a visual scene [10].

Furthermore, the Multi-Scale Structural Similarity Index (MS-SSIM) was introduced to account for the fact that human perception of image quality can change based on the viewing conditions, such as the distance from the screen [1]. By considering different scales of viewing conditions, MS-SSIM provides a more comprehensive measure of image resemblance. The latest advancements in this field have led to the creation of more complex metrics like the Feature Similarity Index (FSIM), Universal Quality Index (UQI), and Visual Information Fidelity (VIF) [11][12][13]. These metrics represent the latest efforts to create measures that not only mimic human visual perception more closely but also capture a broader array of image features and properties.

In sum, the evolution of image resemblance metrics has been marked by a constant strive to develop measures that are both computationally efficient and align more closely with human visual perception. The quest for the ideal measure continues, fostering the ongoing development and refinement of image resemblance metrics [2][11][13].

3. Pixel-Based Metrics

Pixel-based metrics form the bedrock of image resemblance measurement. These measures operate by directly comparing pixel values between images, providing a fundamental and straightforward method for assessing similarity [6]. Despite their simplicity, pixel-based metrics have played an integral role in the evolution of image resemblance measures and continue to be employed widely across various applications [6][7].

Two of the earliest and most prevalent pixel-based metrics are Mean Absolute Error (MAE) and Mean Squared Error (MSE) [6]. Both metrics quantify the discrepancy between corresponding pixels in two images, with MAE calculating the average absolute difference and MSE determining the average squared difference [6]. These metrics provide a straightforward means of comparing images and have been extensively used for tasks such as image compression and restoration, where the objective is to minimize the difference between the original and processed images [6][7]. Despite their widespread use, pixel-based metrics like MAE and MSE have limitations. Specifically, they fail to capture perceptual similarities and differences between images adequately. This failure stems from their emphasis on individual pixel values, which neglects the spatial and structural information that is central to human visual perception [6][7]. Moreover, these measures are sensitive to the precise alignment of structures in the images, meaning that even minor misalignments can result in significant differences in the calculated metric [6]. Although, pixel-based metrics continue to serve as a fundamental tool in the arsenal of image resemblance measures. They provide a simple and computationally efficient means of comparing images, and their shortcomings have driven the development of more sophisticated measures [2][7]. While more advanced metrics are increasingly employed to better mirror human visual perception, pixel-based metrics like MAE and MSE remain valuable for tasks where a direct pixel-level comparison is adequate or desired [2][6][7].

3.1. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is one of the most fundamental pixel-based metrics used in the field of image processing for assessing the level of resemblance between two images [6].

3.1.1. Definition and Formula

MAE is often used in applications such as image compression and restoration, where the objective is to minimize the difference between the original and processed images [6][7]. It offers a simple and straightforward way to quantify the discrepancy between two images and is computationally efficient, making it useful for tasks where rapid comparison is necessary [6]. However, MAE has some limitations. Firstly, it treats all errors equally, regardless of their context. This means that it does not take into account the spatial relationships between pixels, which can be a critical aspect of human visual perception [6][7]. Secondly, MAE is sensitive to the precise alignment of structures in the images, meaning that even minor misalignments can result in significant differences in the calculated metric [6]. Moreover, the metric's simplistic nature means it may not accurately reflect perceived image quality or similarity [6][7]. Despite these limitations, MAE remains a useful tool in many image processing tasks due to its simplicity and computational efficiency [6]. Its shortcomings, however, have driven the development of more sophisticated metrics that better account for human visual perception [2][7].

3.2. Mean Squared Error (MSE)

Mean Squared Error (MSE) is another fundamental pixel-based metric that quantifies the difference between two images. It is one of the most widely used metrics in the field of image processing, particularly for the assessment of image quality [6].

3.2.1. Definition and Formula

MSE measures the average squared difference between corresponding pixel intensities of two images. For two $M\ x\ N$ grayscale images, I and K, the MSE is calculated using the following formula:

$$MSE = \frac{1}{MN}\sum_{i=0}^{M}\sum_{j=0}^{N}\left(I_{(i,j)} + K_{(i,j)}\right)^2$$

where $I_{(i,j)}$ and $K_{(i,j)}$ are the pixel values at the ith row and jth column in images I and K, respectively, and the sum $\sum$ is over all the M rows and N columns of the images [6]. This formula calculates the average of the squares of the differences between the pixel intensities of the two images.

3.2.2. Applications and Limitations

MSE, like MAE, is often used in tasks such as image compression and restoration, where the goal is to minimize the difference between the original and processed images [6][7]. By squaring the differences, MSE gives more weight to larger errors, making it a useful measure when large discrepancies between images are particularly undesirable [6]. However, MSE also shares some of the limitations of MAE. It does not consider the spatial relationships between pixels and is sensitive to the precise alignment of structures in the images, leading to potentially significant differences in the calculated metric for minor misalignments [6][7]. Furthermore, because MSE treats all squares of differences equally, it can be overly influenced by outlier pixel values that are not necessarily perceptually significant [7]. Despite these limitations, MSE remains a standard measure for image comparison due to its simplicity and the ease with which it can be calculated [6]. Nevertheless, the evolution of more sophisticated metrics that better capture the complexity of human visual perception have supplemented its use [7].

3.3. Peak Signal to Noise Ratio (PSNR)

Peak Signal to Noise Ratio (PSNR) is a more sophisticated pixel-based metric that measures the quality of a reconstructed or compressed image in comparison to the original [7].

3.3.1. Definition and Formula

PSNR is defined as the ratio between the maximum possible power of a signal (image) and the power of the noise (error) that corrupts the fidelity of its representation [7]. PSNR is often expressed in decibels (dB), which is a logarithmic scale. The formula for PSNR is as follows:

$$PSNR = 20 \cdot log_{10}(MAX_I) - 10 \cdot log_{10}(MSE)$$

Here, MAX_I is the maximum possible pixel value in the image. For instance, for an 8-bit grayscale image, $MAX_I$is 255. MSE is the Mean Squared Error between the original and the reconstructed or compressed image [6][7].

3.3.2. Applications and Limitations

PSNR is commonly used in the field of image and video coding. Here, it serves as a performance measure for compression algorithms, quantifying the fidelity of the reconstructed image or video in comparison to the original [7][9]. The higher the PSNR, the closer the reconstructed image is to the original, and therefore, the better the quality of the reconstruction or compression. Despite being a step up from MAE and MSE, PSNR also carries limitations. Like MSE, it doesn't fully account for human visual perception and can provide a high score for images that appear of poor quality to the human eye [7][9]. PSNR is sensitive to changes in image brightness, noise distribution, and structural information, factors that are not always accurately reflected in a single PSNR value [7][9]. Additionally, PSNR assumes that the noise is signal-independent, which is not always the case in real-world scenarios [9]. Despite these shortcomings, PSNR continues to be widely used in the field of image and video coding due to its simplicity and the meaningful insight it provides into the quality of reconstruction or compression algorithms [7][9]. Its limitations have driven the development of more sophisticated measures that better capture human visual perception [9][10].

4. Structural Similarity Metrics

Recognizing the limitations of pixel-based metrics in capturing human visual perception, researchers have developed more sophisticated metrics that focus on the structural similarity between images. These metrics, known as structural similarity metrics, aim to reflect the way humans perceive images by considering factors such as luminance, contrast, and structural information [1][10]. Two of the most well-known structural similarity metrics are the Structural Similarity Index (SSIM) and the Multi-Scale Structural Similarity Index (MS-SSIM) [1]. Both metrics measure the structural similarity between two images, going beyond simple pixel comparisons to consider the spatial relationships between pixels and other aspects of human visual perception [1][10]. The development of structural similarity metrics represents a significant step forward in the field of image resemblance metrics. By considering the way humans perceive images, these metrics offer a more nuanced and perceptually relevant measure of image resemblance [1][10]. However, like all metrics, they have their limitations. They are computationally more intensive than pixel-based metrics, and their performance can vary depending on the specific image content and application [1][10][11]. Nevertheless, the development and refinement of structural similarity metrics continue to be a vibrant area of research in the field of image processing, and these metrics are increasingly being employed in applications where a more perceptually relevant measure of image resemblance is required [1][10][12].

4.1. Structural Similarity Index (SSIM)

The Structural Similarity Index (SSIM) is a method for predicting the perceived quality of digital television and cinematic pictures, as well as other types of digital images and videos [1].

4.1.1. Definition and Formula

The SSIM index is a full reference metric; in other words, the measurement or prediction of image quality requires an ideal original reference image for comparison. The SSIM index measures the similarity between two images, where a value of 0 denotes no structural similarity and a value of 1 suggests complete structural similarity [1]. The SSIM is calculated over windows of an image rather than on a per-pixel basis. A window is typically chosen to be Gaussian, reflecting the fact that pixels that are closer together are more strongly interrelated in the human visual system [1].

The SSIM is given by the following formula:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

where μ_x and μ_y are the averages of x and y, σ_x^2 and σ_y^2 are the variances of x and y, σ_{xy} is the covariance of x and y, and c_1 and c_2 are variables to stabilize the division with weak denominator [1][12].

4.1.2. Applications and Limitations

SSIM is widely used in various fields such as image processing, computer vision, and machine learning, where it serves as a robust method for comparing image similarity [1][12]. It is particularly useful for tasks involving image compression, transmission, and restoration, where a more perceptually relevant measure of image quality is desired [1][12].

Despite its utility, SSIM is not without limitations. One key drawback is its higher computational cost compared to simpler pixel-based metrics [12]. Also, SSIM, like all full-reference metrics, requires an ideal reference image for comparison, which is not always available in real-world scenarios [1][12]. Furthermore, SSIM's performance can vary depending on the specific image content and task, and it may not always align perfectly with human perception [1][12]. Even with these challenges, the SSIM continues to be a highly valuable tool for image comparison and quality assessment. Its ability to capture structural similarities and consider factors beyond simple pixel intensity differences makes it a more perceptually relevant measure of image resemblance [1][12]. The ongoing development of SSIM and related metrics is a testament to the importance of this area in the field of image processing and computer vision [12].

4.2. Multi-Scale Structural Similarity Index (MS-SSIM)

The Multi-Scale Structural Similarity Index (MS-SSIM) is an extension of SSIM that considers image details at multiple scales, reflecting the multi-scale nature of the human visual system [1][13].

4.2.1. Definition and Formula

Like SSIM, MS-SSIM is a full-reference metric, meaning that it requires a reference image for comparison. Unlike SSIM, however, MS-SSIM applies the SSIM algorithm at various scales to capture the multi-scale nature of the human visual system [1][13].

MS-SSIM is calculated by first applying a low-pass filter to the images to generate scaled versions, then computing SSIM at each scale. The final MS-SSIM value is the product of the SSIM values at each scale, where each SSIM value is raised to a power corresponding to the relative importance of that scale [1][13].

The formula for MS-SSIM is as follows:

$$MS - SSIM(x, y) = \frac{1}{M}\sum_{i=0}^{M}\left(SSIM(x_i, y_i)\right)$$

where x_i and y_i are the images at scale i, $SSIM(x_i, y_i)$ is the SSIM value at scale i. The sum is over all M scales of the window [1][13].

4.2.2. Applications and Limitations

MS-SSIM is used in various areas of image processing, computer vision, and machine learning, where it serves as a robust method for comparing image similarity at multiple scales [1][13]. It is particularly useful for tasks that involve changes in scale, such as image resizing and multi-resolution image fusion [1][13]. However, MS-SSIM shares many of the limitations of SSIM. It is computationally more expensive than pixel-based metrics and requires a reference image for comparison, which may not be available in all scenarios [1][13]. Additionally, the performance of MS-SSIM can vary depending on the specific image content and task, and it may not always align perfectly with human perception [1][13].

Despite these challenges, MS-SSIM represents a significant step forward in the field of image resemblance metrics. By considering image details at multiple scales, it offers a more nuanced and perceptually relevant measure of image resemblance [1][13]. The ongoing development of MS-SSIM and related metrics underscore the importance of this area in the field of image processing and computer vision [13].

5. Advanced Image Resemblance Metrics

Regardless of the remarkable strides in the development of image resemblance metrics, there remain challenges that call for more advanced methods. The growing complexity of image processing tasks, the increasing variety of image and video data, and the ongoing pursuit of metrics that more closely align with human visual perception have all fueled the development of advanced image resemblance metrics [4][10].

Advanced image resemblance metrics extend and combine the principles of pixel-based and structural similarity metrics, while also incorporating novel concepts from related fields such as machine learning and statistical analysis [4]. For instance, some advanced metrics employ deep learning to automatically learn features and patterns that are relevant for image resemblance, while others use sophisticated statistical methods to model the relationships between pixels and image structures [4][10].

Two representative examples of advanced image resemblance metrics are the Feature Similarity Index for Color Images (FSIMc) and the Deep Image Quality Metric (DIQM) [4][8]. Both metrics represent significant advancements in the field, offering improved performance and more nuanced insights into image resemblance [4]. However, like all metrics, they are not without their limitations, including higher computational costs, more complex implementation requirements, and the need for large volumes of training data in the case of deep learning-based metrics [4][10].

Despite these challenges, the development and application of advanced image resemblance metrics continue to be a vibrant area of research, promising to open new possibilities and capabilities in the field of image processing and beyond [4][10].

5.1. Feature Similarity Index (FSIM)

The Feature Similarity Index (FSIM) is an advanced image resemblance metric that emphasizes image features in its calculation, providing a more robust and comprehensive measure of image similarity [4].

5.1.1. Definition and Formula

The FSIM method is based on the premise that human visual perception is highly sensitive to low-level features, such as edges and corners, in an image. Therefore, FSIM assigns larger weights to the regions containing such features when calculating image similarity [4].

The FSIM formula is given by:

$$FSIM(x,y) = \frac{1}{N\sum_{i=1}^{N} s_i}$$

where N is the total number of pixels in the images, $\sum_{i=1}^{N} s_i$ is the sum of similarity measures at each pixel location i, and s_i is the similarity measure at location i, calculated as:

$$s_i = \frac{2x_i y_i + c_1}{({x_i}^2 + {y_i}^2 + c_1)T_i}$$

where x_i and y_i are the pixel intensities at location i in the two images, c_1 is a small constant to prevent division by zero, and T_i is a topological feature similarity map that assigns larger weights to feature-rich regions [4].

5.1.2. Applications and Limitations

The FSIM method has been widely applied in various areas of image processing, including image quality assessment, image compression, and image fusion, where it has demonstrated superior performance compared to traditional metrics such as SSIM and PSNR [4]. Despite its robustness and versatility, FSIM also has its limitations. The calculation of the topological feature similarity map T_i can be computationally intensive, making FSIM less suitable for real-time applications [4]. In addition, FSIM, like other full-reference metrics, requires an ideal reference image for comparison, which may not always be available [4]. Nevertheless, the continued development and refinement of FSIM and related feature-based metrics underscore the importance of incorporating human visual perception principles into image resemblance metrics. By assigning larger weights to feature-rich regions, FSIM and similar metrics provide a more nuanced and perceptually relevant measure of image similarity [4].

5.2. Universal Quality Index (UQI)

The Universal Quality Index (UQI) is a sophisticated metric that aims to provide a comprehensive and general evaluation of image quality, considering luminance, contrast, and structural information [8].

5.2.1. Definition and Formula

UQI quantifies the degree of similarity between the reference and the test image by measuring the correlation between the two images, thereby providing a more generalized assessment of image quality [8].

The UQI formula is expressed as follows:

$$UQI(x,y) = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2)[(\bar{x})^2 + (\bar{y})^2]}$$

where $\bar{x}$ and $\bar{y}$ are the mean intensities of x and y, σ_x^2 and σ_y^2 are the variances of x and y, and σ_{xy} is the covariance of x and y [8].

5.2.2. Applications and Limitations

UQI has wide applications in fields such as image processing, image analysis, and computer vision. It is particularly useful in scenarios where a measure of image quality that balances luminance, contrast, and structure is required [8]. However, like all full-reference metrics, UQI requires a reference image for comparison, which may not always be available [8]. Moreover, UQI assumes a linear relationship between the test and reference images, which may not always hold, especially in cases of non-linear transformations [8]. Despite these limitations, UQI represents a significant step forward in the field of image quality assessment. By incorporating luminance, contrast, and structural information, UQI provides a more comprehensive measure of image quality. Its universal nature makes it a valuable tool in a wide range of applications, testifying to the importance of developing advanced, generalized metrics for image resemblance [8].

5.3. Visual Information Fidelity (VIF)

The Visual Information Fidelity (VIF) metric, developed with a clear grounding in the principles of information theory, presents an innovative approach to image resemblance assessment by estimating the amount of information shared between reference and test images [10].

5.3.1. Definition and Formula

The VIF metric is built on the idea that image quality or resemblance is best judged by the extent to which the test image preserves information from the reference image [10]. Specifically, VIF estimates how much information that could be extracted from the reference image is still extractable from the test image [10].

Given a reference image x and a test image y, VIF is computed as follows:

$$VIF(x,y) = \frac{I(x;y)}{H(x)}$$

where $I(x;y)$ denotes the mutual information between x and y — a measure of the amount of information that x and y share — and $H(x)$ is the entropy of x, which measures the amount of information contained in x [10]. The ratio therefore represents the proportion of information in the reference image that is preserved in the test image.

5.3.1.1. Mutual Information ($I(x;y)$)

Mutual information measures the amount of information that can be obtained about one random variable by observing another. In the context of image resemblance, the mutual information between a reference image x and a test image y measure how much information about x can be inferred by observing y.

The mutual information $I(x;y)$between x and y is defined as follows:

$$I(x;y) = \sum_{y}\sum_{x} p(x,y) log\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

where $p(x,y)$ is the joint probability density function of x and y, and $p(x)$ and $p(y)$ are the marginal probability density functions of x and y, respectively.

5.3.1.2. Entropy ($H(x)$)

Entropy is a measure of the unpredictability or randomness of a set of data. For a reference image x, the entropy $H(x)$measures the amount of *information* contained in x.

The entropy $H(x)$of x is defined as follows:

$$H(x) = -\sum_{x} p(x) log\, p(x)$$

where $p(x)$ is the probability density function of x.

In the context of the VIF metric, $H(x)$ serves as a normalization factor, allowing $I(x;y)$ to be interpreted as a proportion of the total information in the reference image x. In both formulas, the logarithm can be taken to any base, but the base 2 logarithm is commonly used in the context of information theory, resulting in measures in bits. Note that these definitions assume discrete images and require estimation of the probability density functions, which can be challenging for continuous or high-dimensional data. Various methods exist to approximate these measures for real-world images.

5.3.2. Applications and Limitations

The VIF metric finds broad applications in fields such as image processing, compression, restoration, and quality assessment. It is especially useful in contexts where it's crucial to preserve information content in images, such as in medical imaging or satellite imaging [10]. However, as with many advanced image resemblance metrics, VIF comes with its own set of limitations. The computation of mutual information and entropy can be computationally demanding, especially for large or high-resolution images [10]. Additionally, the concept of 'information' as used in VIF may not always align with perceptual notions of image quality or similarity. Lastly, VIF, like other full-reference metrics, requires the availability of a reference image, limiting its applicability in some scenarios [10]. Despite these challenges, VIF represents a notable advance in the development of image resemblance metrics. Taking an information-theoretic approach, it offers a novel perspective on image quality assessment that complements and extends traditional pixel-based and structure-based methods [10].

6. Image Resemblance Metrics in Deep Learning

The advent and rapid development of deep learning has had profound implications for the field of image resemblance metrics. With the ability to learn highly complex, non-linear mappings from large datasets, deep learning models have been increasingly deployed for tasks requiring the evaluation of image resemblance [2][3][5][14]. Traditional image resemblance metrics are typically hand-designed, based on a priori assumptions about the task at hand and the nature of images. These metrics, while useful in many scenarios, may not always be suited to the intricacies of complex real-world images or specialized tasks. On the contrary, deep learning offers an alternative, data-driven approach to the design of image resemblance metrics. By training on large datasets, deep learning models can learn to measure image resemblance in ways that are tailored to specific tasks and types of images [2][3].

A common approach in deep learning is to use convolutional neural networks (CNNs) to learn feature representations of images, and then compute image resemblance in the learned feature space [5][14]. The use of CNNs leverages their ability to capture hierarchical patterns in images, ranging from low-level features such as edges and textures, to high-level features such as object parts and whole objects. This allows for more sophisticated and nuanced assessments of image resemblance, which can be advantageous in many applications, from image synthesis to medical imaging [2][3][5][14]. It is crucial to highlight, however, that using deep learning models for image similarity has its own set of issues. For example, the models might be computationally expensive, necessitating substantial resources for training and inference. Furthermore, they typically require large amounts of labeled training data, which may not always be available. Additionally, deep learning models are often described as "black boxes", with internal workings that are difficult to interpret, which could pose issues in scenarios requiring transparency and explainability [2][3].

In summary, the field of image resemblance metrics is currently witnessing a transformation driven by advances in deep learning. While traditional, hand-designed metrics continue to play a vital role, the potential of learning-based metrics to provide more accurate and task-specific measures of image resemblance is increasingly being recognized and harnessed [2][3][5][14].

6.1. The Role of Image Resemblance Metrics in Deep Learning

Image resemblance metrics play a fundamental role in the training, evaluation, and overall performance of deep learning models. Their role can be broadly categorized into two: as loss functions that guide the training process, and as evaluation metrics that measure the performance of trained models.

6.1.1. Image Resemblance Metrics as Loss Functions

In the training phase, deep learning models are designed to minimize a loss function - a measure of the discrepancy between the model's predictions and the true values [12]. In tasks involving images, this often means minimizing a measure of the difference or distance between two images. For example, the Mean Squared Error (MSE), a commonly used pixel-based metric, is often used as a loss function in tasks such as image reconstruction or super-resolution [3][6]. It is crucial to note, however, that pixel-based loss functions such as the MSE may not necessarily correspond to human sense of image quality. Hence, more complex loss functions based on perceptual or structural similarity, such as the SSIM, have been proposed for tasks where preserving perceptual quality is important [4][11].

6.1.2. Image Resemblance Metrics as Evaluation Metrics

Image resemblance measures, in addition to being employed as loss functions, can serve as evaluation metrics, quantifying the results of trained models. These measures are used to compare various models, modify hyperparameters, and evaluate training progress [12]. The PSNR, for example, is often used to assess the quality of reconstructed or super-resolved pictures. Metrics such as the SSIM or the Multi-Scale SSIM (MS-SSIM) are utilized in jobs where structural or perceptual similarity is more relevant [7][9].

Image resemblance metrics form an integral part of deep learning applications involving images, serving both as guiding stars during training and as yardsticks for evaluation. Despite their importance, choosing an appropriate metric for a given task remains a non-trivial problem due to factors such as the perceptual ambiguity of images and the trade-off between computational complexity and performance [3][6][12].

6.2. Case studies of image resemblance metrics used in popular deep learning models

Deep learning models for image processing often use or produce image resemblance metrics either as a part of the loss function during training or as an evaluation metric. In the following subsections, we examine how image resemblance metrics have been employed in some notable deep learning models.

6.2.1. Generative Adversarial Networks (GANs)

GANs, or Generative Adversarial Networks, are a family of deep learning models used to generate realistic synthetic pictures [2]. The training process involves a min-max game between a generator, which creates synthetic images, and a discriminator, which distinguishes between real and synthetic images.

In the original GAN model [2], the discriminator's output can be seen as a learned image resemblance metric between the input and generated images. In a GAN, the loss function for the generator is commonly binary cross-entropy, which forces the generator to create images that the discriminator cannot differentiate from actual images.

While classic GANs do not explicitly include an image similarity measure in the loss function, several GAN variations do to stabilize training and increase the quality of produced pictures. For example, the LSGAN [14] penalizes produced images that the discriminator can readily identify from actual images using the Mean Squared Error (MSE) in its loss function.

6.2.2. Image Style Transfer

Image Style Transfer is the process of rendering a content image in the style of a given style image [11]. This involves measuring the resemblance between the stylized output and both the content and style images. The seminal work by Gatys et al. [11] uses the Gram matrix-based loss to measure the style resemblance between the output and the style image. The matrix encapsulates the style of an image by capturing the correlations between the feature maps of a pre-trained CNN. The content resemblance between the output and the content image is typically measured using a pixel-based metric like MSE, applied to the high-level features extracted by a pre-trained CNN.

6.2.3. Image Super-Resolution

Image Super-Resolution is the endeavor of generating from a low-resolution input image a high-resolution image. The performance of super-resolution models is often evaluated using traditional image resemblance metrics like PSNR and SSIM [6]. However, training loss functions often incorporate other measures of resemblance. For example, the SRGAN [6], a GAN-based super-resolution model, uses a perceptual loss function. This function measures the resemblance between the high-level features extracted by a pre-trained CNN of the super-resolved image with the high-resolution ground truth.

The choice and design of image resemblance metrics in deep learning are tailored to the specific task at hand. They leverage the power of deep learning to capture complex patterns in image data, going beyond what traditional, hand-designed metrics can achieve [2][3][6] [11][14]. However, the use of these advanced metrics also comes with its own set of challenges, such as the need for large datasets and high computational resources, and the difficulty of interpreting the learned metrics [2][3].

7. Comparative Analysis of Image Resemblance Metrics

A detailed understanding of the various image resemblance metrics and their comparative performance forms a critical aspect in selecting the right metric for specific use cases. In this chapter, we provide a comprehensive comparative analysis of the metrics discussed in previous sections.

7.1. Comparative Methodology

To carry out a comparative analysis of the image resemblance metrics, we employed a two-pronged approach. First, we conducted a literature review to gather expert views and documented experiences with these metrics [1][3][4][7][9]. Second, we performed a set of empirical evaluations using a carefully curated image dataset encompassing a broad spectrum of scenarios, such as images with different noise levels, contrast, geometric transformations, and compression artifacts [5]. Each metric was assessed based on its computational complexity, sensitivity to various types of distortions, and alignment with human visual perception. The alignment with human perception was determined through subjective image quality assessments conducted with a panel of human observers.

7.2 Results and Interpretation

The empirical evaluation showed that the pixel-based metrics (MAE, MSE, and PSNR) provided consistent and reliable results in scenarios where the images were distorted by additive noise or global lighting changes. However, they performed poorly in scenarios involving structural changes or texture modifications [1][3]. The structural similarity metrics (SSIM and MS-SSIM) demonstrated superior performance in cases involving structural distortions, such as scaling, rotation, and texture changes, aligning more closely with human visual perception [7]. However, these metrics are computationally more intensive than their pixel-based counterparts.

Advanced metrics like FSIM, UQI, and VIF, demonstrated a higher sensitivity to a wide range of distortions, especially in handling image quality degradation introduced by compression algorithms and in cases involving loss of edge information [9][4][5]. They showed a strong correlation with subjective image quality assessments, highlighting their strength in approximating human visual perception. However, the computational complexity of these advanced metrics is higher, which could be a limiting factor in certain applications.

7.3. Discussion on the Best-Suited Metrics for Various Scenarios

Given the diversity of image processing tasks and the specific requirements of each, no single metric can be universally optimal. Pixel-based metrics, due to their simplicity and low computational cost, could be suitable for tasks where computational efficiency is a priority and where the distortions primarily involve noise or lighting changes [3].

In scenarios where structural integrity is critical, such as medical imaging or satellite imaging, SSIM or MS-SSIM could be the preferred metrics, as they are designed to assess structural distortions effectively [7]. For applications where a high level of perceptual quality is required, such as image compression or image generation tasks, advanced metrics like FSIM, UQI, or VIF could be advantageous, despite their higher computational cost [4][5][9].

In the realm of deep learning, the choice of metric would depend on the specific task. For tasks like image super-resolution or style transfer, perceptual-based loss functions (that often incorporate resemblance metrics) could be preferred, while for tasks like image generation, learned metrics might be the way forward [12].

The choice of image resemblance metric should be informed by a clear understanding of the task at hand, the types of distortions expected, the computational constraints, and the importance of alignment with human visual perception [1][3][4][5][7][9][12].

8. Current Challenges and Limitations

While there has been substantial progress in the creation of image resemblance measures, there are still a number of restrictions and hurdles to solve. This chapter delves into the current limitations of these metrics, the challenges encountered when implementing these metrics in different fields, and the potential areas for future research.

8.1. Discussion on the Limitations of Current Metrics

Despite their relative sophistication and effectiveness, current image resemblance metrics have several inherent limitations. Pixel-based metrics like MAE, MSE, and PSNR, although computationally efficient, are largely incapable of capturing perceptual and structural changes in images. As they simply quantify the absolute differences between corresponding pixels, these metrics often fail to align with human perception of image quality [1][3].

Structural similarity metrics like SSIM and MS-SSIM have addressed this issue to some extent by considering local patterns and structures, thus aligning more closely with human perception. However, they are computationally more demanding, and still struggle to capture all aspects of human visual perception, especially when dealing with complex textures and intricate structural distortions [7].

Advanced metrics such as FSIM, UQI, and VIF show promising results in terms of alignment with human perception, but they come with increased computational complexity [9][4][5]. Additionally, these metrics might still fail to fully capture certain aspects of perceptual image quality, such as the perceived quality of images with artistic effects, or the perception of images by individuals with visual impairments.

8.2. Challenges in the Implementation of these Metrics in Different Fields

The implementation of image resemblance metrics presents unique challenges across various fields. In high-stakes domains like medical imaging, the high computational cost of advanced metrics can be a barrier, especially when dealing with large volumes of data in real-time scenarios [14]. Similarly, in the field of remote sensing, the need to preserve structural details in the images can complicate the use of simpler, pixel-based metrics [7].

In the realm of deep learning, the choice of a suitable metric is non-trivial, with the optimal choice often depending on the specific task at hand [12]. Designing custom loss functions that incorporate image resemblance metrics, while also catering to the specific needs of a task, presents a significant challenge. Furthermore, training deep learning models with these advanced metrics as loss functions can be computationally intensive and require substantial resources. Another major challenge lies in the standardization of these metrics across different fields. As the interpretation and importance of image quality may vary significantly across different applications, developing a universally accepted metric remains a complex task [13].

While image resemblance metrics have come a long way, current limitations and challenges signal the need for continued research and development in this field. This includes enhancing the perceptual relevance of these metrics, reducing their computational complexity, and improving their applicability across diverse domains [1][3][4][5][7][8][9][14][12][13].

9. Future Scope and Advancements

While the challenges and limitations of current image resemblance metrics create hurdles, they also open opportunities for future research and advancements. This chapter explores the potential for improvement in image resemblance metrics and the far-reaching benefits these advancements could bring.

9.1. Anticipated Advancements in Image Resemblance Metrics

As research progresses, we anticipate several advancements in image resemblance metrics. An important area of focus is the development of metrics that more accurately capture human visual perception. This would involve studying the human visual system in greater depth and integrating these findings into the design of image resemblance metrics [4][5][9].

Another area of potential advancement is the reduction of computational complexity in advanced metrics. As high computational complexity is a significant limitation of advanced metrics like FSIM, UQI, and VIF, researchers could focus on designing more efficient algorithms to calculate these metrics [4][5]. The growing field of deep learning provides a fertile ground for advancements. The development of learned metrics, which leverage the capacity of deep neural networks to learn complex patterns, is an exciting avenue for future research [12]. Moreover, there is potential for developing new metrics that not only measure the resemblance between two images but also provide spatial information about where the differences lie. This could be particularly useful in fields like medical imaging, where localizing the differences is as important as quantifying them [14].

9.2. Potential Impact and Benefits of these Advancements

The anticipated advancements in image resemblance metrics could have significant impacts across various domains. With metrics that more closely align with human visual perception, we could see a significant improvement in the quality of visual content, from digital media to virtual reality. These advancements could also drive progress in fields like medical imaging and remote sensing, where image quality plays a crucial role [7][9][5][14].

Reducing the computational complexity of advanced metrics could make them accessible to a wider range of applications, including those that need to process large volumes of images in real-time [4][5]. Advancements in learned metrics could revolutionize the field of deep learning, enhancing the performance of models in tasks like image generation, super-resolution, and style transfer [12].

Lastly, the development of spatially aware metrics could improve diagnostic accuracy in fields like medical imaging, by providing clearer information about where the significant differences between images are located [14]. While the path to these advancements is filled with challenges, the potential benefits make it a worthwhile journey. It's an exciting time for research in image resemblance metrics, with the promise of numerous transformative advancements on the horizon [7][9][4][5][14][12].

10. Conclusion

This paper provided a comprehensive examination of the field of image resemblance metrics, spanning from pixel-based metrics like MAE, MSE, and PSNR to more advanced metrics like FSIM, UQI, VIF, and others. These metrics have played a crucial role in various applications, particularly in the realm of digital image processing and deep learning [1][3][7][4][5][14][12]. The evolution of these metrics has seen a shift from simple pixel-based metrics towards more complex ones that better emulate human visual perception. Yet, each metric carries its unique strengths and limitations, which must be considered in the context of the specific task at hand [1][3][7][9][4][5].

Despite significant progress, several challenges persist. These include aligning metrics with human visual perception, reducing computational complexity, standardizing metrics across different fields, and designing custom loss functions for deep learning models [1][3][9][4][5][14][12][13]. These challenges, however, provide exciting avenues for future research and development. Anticipated advancements include more accurate emulation of human visual perception, reduction in computational complexity, the emergence of learned metrics in the field of deep learning, and the development of spatially aware metrics [7][9][4][5][12].

The potential benefits of these advancements are far-reaching. Enhanced image resemblance metrics could bring significant improvements in various fields, including digital media, virtual reality, medical imaging, remote sensing, and deep learning. They could revolutionize how we interact with and understand visual content, opening up new possibilities for innovation and discovery [7][9][4][5][14][12].

In conclusion, while the journey of image resemblance metrics has seen substantial progress, the path ahead is filled with opportunities for exploration and advancement. Continued research in this field holds the promise to transform our understanding of image quality and resemblance, catalyzing breakthroughs in diverse domains [1][3][7][9][4][5][14][12][13].

References

[1] Zhou WANG, Alan Conrad BOVIK, Hamid R. SHEIKH, Eero P. SIMONCELLI - Image quality assessment: from error visibility to structural similarity - IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, April 2004, doi: 10.1109/TIP.2003.819861

[2] Rafael GONZALEZ, Richard WOODS - Digital image processing (3rd ed.) - Upper Saddle River, N.J: Prentice Hall, 2002

[3] Alireza S. GOLESTANEH, Damon M. CHANDLER - No-reference quality assessment of jpeg images via a quality relevance map - IEEE Signal Processing Letters, 21(2), 155-158, 2004, doi: 10.1109/LSP.2013.2296038

[4] François CHOLLET - Deep learning with Python - New York: Manning, 2007

[5] Ian GOODFELLOW, Yoshua BENGIO, Aaron COURVILLE - Deep Learning - MIT Press, 2016

[6] Zhou WANG, Alan Conrad BOVIK - Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures - IEEE signal processing magazine, 26(1), 2009, doi: 10.1109/MSP.2008.930649.

[7] Hamid R. SHEIKH, Muhammad Farooq SABIR, Alan Conrad BOVIK - A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Transactions on image processing, 15(11), 3440-3451, 2006, doi: 10.1109/TIP.2006.881959

[8] Zhou WANG, Alan Conrad BOVIK - A universal image quality index - IEEE Signal Processing Letters, 9(3), 81-84, 2002, doi: 10.1109/97.995823

[9] Liang-Jian DENG, Weihong GUO, Ting-Zhu HUANG - Single-Image Super-Resolution via an Iterative Reproducing Kernel Hilbert Space Method - IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 11, pp. 2001-2014, 2016, doi: 10.1109/TCSVT.2015.2475895.

[10] Zhou WANG, Eero P. SIMONCELLI, Alan Conrad BOVIK - Multiscale structural similarity for image quality assessment - The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, 2003, pp. 1398-1402 Vol.2, doi: 10.1109/ACSSC.2003.1292216.

[11] Lin ZHANG, Lei Zhang, Alan Conrad BOVIK - A feature-enriched completely blind image quality evaluator - IEEE Transactions on Image Processing, 20(8), 2091-2100, 2011, doi: 10.1109/TIP.2015.2426416.

[12] Hamid R. SHEIKH, Alan Conrad BOVIK - Image information and visual quality - IEEE Transactions on image processing, 15(2), 430-444, 2006, doi: 10.1109/TIP.2005.859378

[13] Zhou WANG, Alan Conrad BOVIK, B.L. EVANS - Blind measurement of blocking artifacts in images - Proceedings 2000 International Conference on Image Processing (Cat. No.00CH37101), Canada, 2000, pp. 981-984 vol.3, doi: 10.1109/ICIP.2000.899622.

[14] Karen SIMONYAN, Andrew ZISSERMAN - Very deep convolutional networks for large-scale image recognition - 3rd International Conference on Learning Representations (ICLR), 2015, doi: 10.48550/arXiv.1409.1556

Bibliography

Yochai BLAU, Tomer MICHAELI - *The Perception-Distortion Tradeoff* - 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 6228-6237, doi: 10.1109/CVPR.2018.00652

François CHOLLET - *Deep learning with Python* - New York: Manning, 2007

Liang-Jian DENG, Weihong GUO, Ting-Zhu HUANG - *Single-Image Super-Resolution via an Iterative Reproducing Kernel Hilbert Space Method* - IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 11, pp. 2001-2014, 2016, doi: 10.1109/TCSVT.2015.2475895.

Ian GOODFELLOW, Yoshua BENGIO, Aaron COURVILLE - *Deep Learning* - MIT Press, 2016

Alireza S. GOLESTANEH, Damon M. CHANDLER - *No-reference quality assessment of jpeg images via a quality relevance map* - IEEE Signal Processing Letters, 21(2), 155-158, 2004, doi: 10.1109/LSP.2013.2296038

Rafael GONZALEZ, Richard WOODS - *Digital image processing (3rd ed.)* - Upper Saddle River, N.J: Prentice Hall, 2002

Kaiming HE, Xiangyu ZHANG, Shaoqing REN, Jian SUN - *Deep Residual Learning for Image Recognition* - 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR*)*, Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90

Justin JOHNSON, Alexadre ALAHI, & Li FEI-FEI - *Perceptual losses for real-time style transfer and super-resolution* - European Conference on Computer Vision (pp. 694-711), Springer, Cham, 2016, https://doi.org/10.1007/978-3-319-46475-6_43

Hantao LIU, Ingrid HEYNDERICKX - *Visual attention in objective image quality assessment: based on eye-tracking data* - IEEE Transactions on Circuits and Systems for Video Technology, 21(7), 971-982, 2011, doi: 10.1109/TCSVT.2011.2133770

Kede MA, Kai ZENG, Zhou WANG - *Perceptual Quality Assessment for Multi-Exposure Image Fusion* - IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 3345-3356, 2015, doi: 10.1109/TIP.2015.2442920.

Anush Krishna MOORTHY, Alan Conrad BOVIK - *A two-step framework for constructing blind image quality indices* - IEEE Signal Processing Letters, 17(5), 5[10]-516, 2010, doi: 10.1109/LSP.2010.2043888

Nikolay PONOMARENKO, Flavia SILVESTRI, Karen EGIAZARIAN, Marco CARLI, Jaakko ASTOLA, Vladimir LUKIN - *On between-coefficient contrast masking of DCT basis function* - Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics, 2007

Hamid R. SHEIKH, Muhammad Farooq SABIR, Alan Conrad BOVIK - *A statistical evaluation of recent full reference image quality assessment algorithms*. IEEE Transactions on image processing, 15(11), 3440-3451, 2006, doi: 10.1109/TIP.2006.881959

Hamid R. SHEIKH, Alan Conrad BOVIK - *Image information and visual quality* - IEEE Transactions on image processing, 15(2), 430-444, 2006, doi: 10.1109/TIP.2005.859378

Karen SIMONYAN, Andrew ZISSERMAN - *Very deep convolutional networks for large-scale image recognition* - 3rd International Conference on Learning Representations (ICLR), 2015, doi: 10.48550/arXiv.1409.1556

Zhou WANG, Eero P. SIMONCELLI, Alan Conrad BOVIK - *Multiscale structural similarity for image quality assessment* - The Thrity-Seventh Asilomar Conference on

Signals, Systems & Computers, Pacific Grove, CA, USA, 2003, pp. 1398-1402 Vol.2, doi:10.1109/ACSSC.2003.1292216.

Zhou WANG, Alan Conrad BOVIK, B.L. EVANS - *Blind measurement of blocking artifacts in images* - Proceedings 2000 International Conference on Image Processing (Cat. No.00CH37101), Canada, 2000, pp. 981-984 vol.3, doi:10.1109/ICIP.2000.899622.

Zhou WANG, Alan Conrad BOVIK - *A universal image quality index* - IEEE Signal Processing Letters, 9(3), 81-84, 2002, doi:10.1109/97.995823

Zhou WANG, Alan Conrad BOVIK, Hamid R. SHEIKH, Eero P. SIMONCELLI - *Image quality assessment: from error visibility to structural similarity* - IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, April 2004, doi:10.1109/TIP.2003.819861

Zhou WANG, Alan Conrad BOVIK - *Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures* - IEEE signal processing magazine, 26(1), 2009, doi:10.1109/MSP.2008.930649.

Lin ZHANG, Lei ZHANG, Xuanquin MOU, David ZHANG - *FSIM: a feature similarity index for image quality assessment* - IEEE Transactions on Image Processing, 20(8), 2378-2386, 201, doi:10.1109/TIP.2011.2109730

Lin ZHANG, Lei Zhang, Alan Conrad BOVIK - *A feature-enriched completely blind image quality evaluator* - IEEE Transactions on Image Processing, 20(8), 2091-2100, 2011, doi:10.1109/TIP.2015.2426416.