J. Vis. Commun. Image R. 49 (2017) 177–191

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

A critical survey of state-of-the-art image inpainting quality assessment metrics ☆

Muhammad Ali Qureshi [a,b,*], Mohamed Deriche [a], Azeddine Beghdadi [c], Asjad Amin [a,b]

[a] King Fahd University of Petroleum and Minerals (KFUPM), Dhahran 31261, Saudi Arabia
[b] The Islamia University of Bahawalpur, 63100, Pakistan
[c] L2TI, Institut Galilée, Université Paris 13, Sorbonne Paris Cité, France

ARTICLE INFO

Article history:
Received 1 January 2017
Revised 2 August 2017
Accepted 8 September 2017
Available online 9 September 2017

Keywords:
Image inpainting
Image quality assessment
Inpainting quality
Inpainting databases
Image inpainting quality assessment
Survey

ABSTRACT

Image inpainting is the process of restoring missing pixels in digital images in a plausible way. Research in image inpainting has received considerable attention in different areas, including restoration of old and damaged documents, removal of undesirable objects, computational photography, retouching applications, etc. The challenge is that the recovery processes themselves introduce noticeable artifacts within and around the restored image regions. As an alternative to subjective evaluation by humans, a number of approaches have been introduced to quantify inpainting processes objectively. Unfortunately, existing objective metrics have their own strengths and weaknesses as they use different criteria. This paper provides a thorough insight into existing metrics related to image inpainting quality assessment, developed during the last few years. The paper provides, under a new framework, a comprehensive description of existing metrics, their strengths, their weaknesses, and a detailed performance analysis on real images from public image inpainting database. The paper also outlines future research directions and applications of inpainting and inpainting-related quality assessment measures.

1. Introduction

Image inpainting is generally defined as the process of restoring missing pixels and damaged regions, or removing unwanted objects in digital images in a plausible way [1]. Considerable research has been carried in developing inpainting algorithms, and a plethora of image inpainting algorithms have been proposed [2,1,3,4]. Image inpainting has recently received considerable attention in different areas related to image processing. While the applications of image inpainting are countless, we outline below the most common and practical ones.

- *Removing Unwanted Objects:* Unwanted objects can be removed from the image using inpainting techniques. The scenario is seen as a special class of image tampering. Fig. 1 shows a nice example of image inpainting where the cage in the original image is removed in the inpainted image [5].
- *Restoring Photos:* The deterioration in photos with the passage of time can be overcome using inpainting. The scratches in the photos resulting from improper handling can also be removed. This is also the case of restoring images from cultural archives, etc. Fig. 2 shows an example in which the scratches in the old photograph have been removed using inpainting [1].
- *Photo Retouching:* Another widely used application of image inpainting is in the media industry where photos of actors/actresses, models, etc., are manipulated by removing wrinkles, mole marks, or undesirable facial features to make these "more attractive". Fig. 3 shows an example of image inpainting where the face is made more attractive by removing some marks using inpainting [6].
- Text Removal: Image inpainting can also be used for removing unwanted text, stamps, copyright logos, etc., in digital images. Fig. 4 shows an example of a street image with superimposed text, from which the text is removed in the inpainted image.

In a way, image inpainting can be seen as a modified copy-move tampering process which is used to recover or remove some parts of the image without any perceptual loss [4,7]. It is different from copy-move forgery in the sense that different blocks or regions come from different locations of the image (see Fig. 5).

☆ This paper has been recommended for acceptance by Zicheng Liu.

* Corresponding author at: King Fahd University of Petroleum and Minerals (KFUPM), Dhahran 31261, Saudi Arabia.
E-mail addresses: ali.qureshi@iub.edu.pk (M.A. Qureshi), mderiche@kfupm.edu.sa (M. Deriche), azeddine.beghdadi@univ-paris13.fr (A. Beghdadi), asjad.amin@iub.edu.pk (A. Amin).

http://dx.doi.org/10.1016/j.jvcir.2017.09.006
1047-3203/

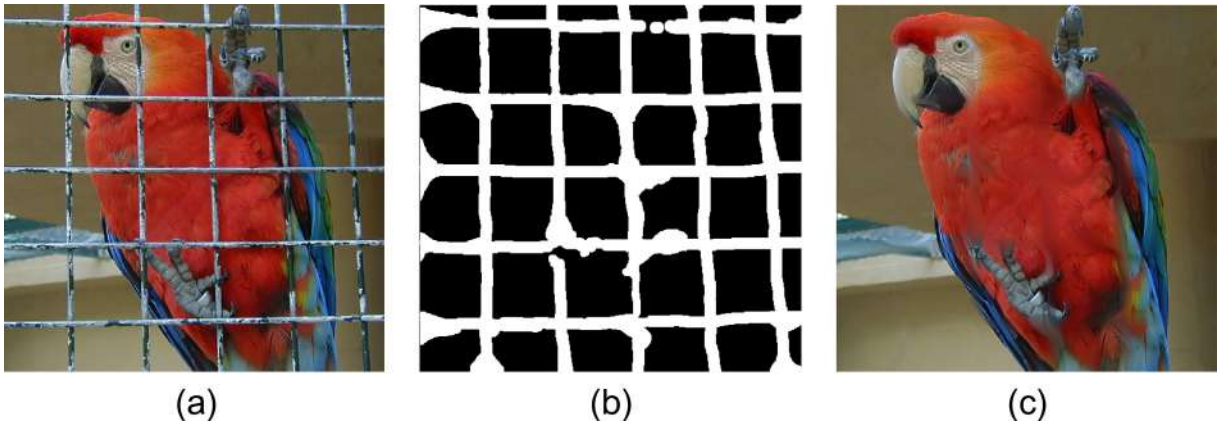

Fig. 1. An example of image inpainting for object removal (a) original image, (b) binary mask, and (c) inpainted image [5].

Fig. 2. An example of image inpainting used in restoration, original image (left), restored image (right) [1].

Fig. 3. An example of inpainting for photo retouching, original image (left), retouched image (right) [6].

Although a substantial amount of research has been carried out in developing robust inpainting algorithms, limited efforts have been put in developing quality assessment metrics to evaluate the performance of image inpainting (restoration) methods. Image Inpainting Quality Assessment (IIQA) is a complex and a challenging problem [2]. The objectives of image inpainting assessment are quite different from those of classical image quality evaluation [9,10]. Here, for inpainting IQA, the goal is to evaluate the quality of the restored images. This task can be performed using either subjective or objective methods. The main goal here is different, traditional IQA fidelity-based metrics which were mainly developed for quantifying *distortions* in degraded images. Hence, traditional IQA metrics are not well suitable for evaluating the quality of *restored* images and cannot directly be used. This is due to the

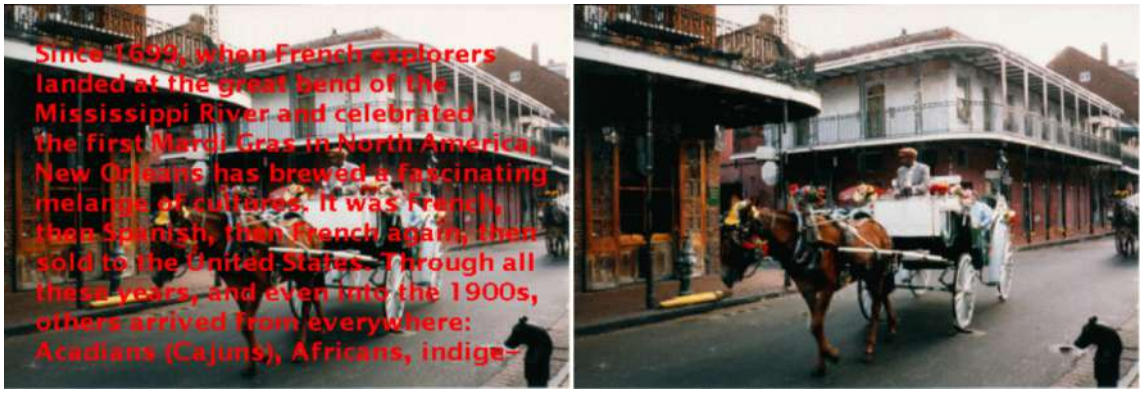

Fig. 4. An example of image inpainting for text removal, original image (left), restored image (right) [1].

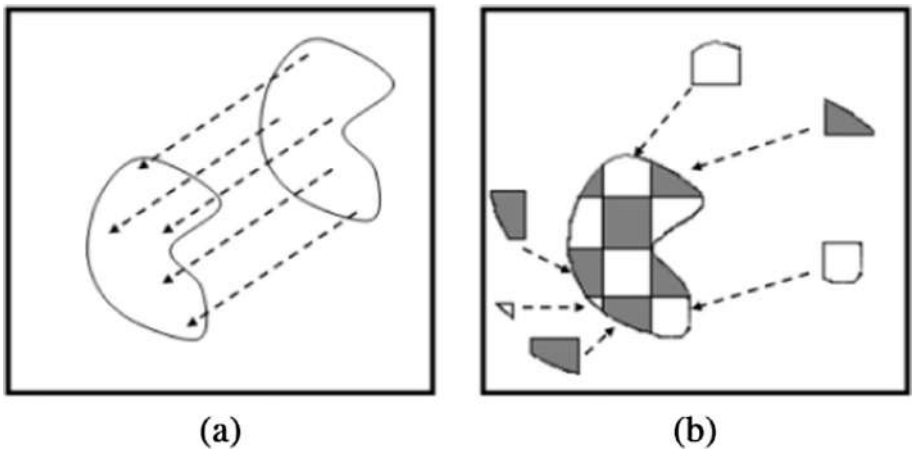

Fig. 5. Difference between two types of tampering (a) copymove forgery and (b) inpainting [8].

Fig. 6. An example of disconnected edges in the inpainted image, original image with mask to be filled (left), inpainted image using [11,2] (right).

fact, that restored images in inpainting are different from their original counterparts. Image inpainting is, to some extent, similar to image enhancement. In image enhancement, we start from an input image with poor quality and try to improve its quality. This process is expected to produce more visible structures and the final images are rather different from the original ones. This process also introduces new types of artifacts which affect the perceived quality in a different manner. Among these artifacts, blur is introduced around edges when restoring large inpainted regions. The curved boundaries are not produced correctly as well. Given the unique nature of artifacts introduced in image inpainting and the inappropriateness of traditional IQA metrics to quantify quality, we have decided to provide a comprehensive and a critical review of different methods developed for quality assessment of inpainted images. We implemented and tested different state-of-the-art IIQA metrics on real images from public image inpainting database and performance analysis is performed in terms of correlation with the subjective perception-based evaluation. This review will be the first of

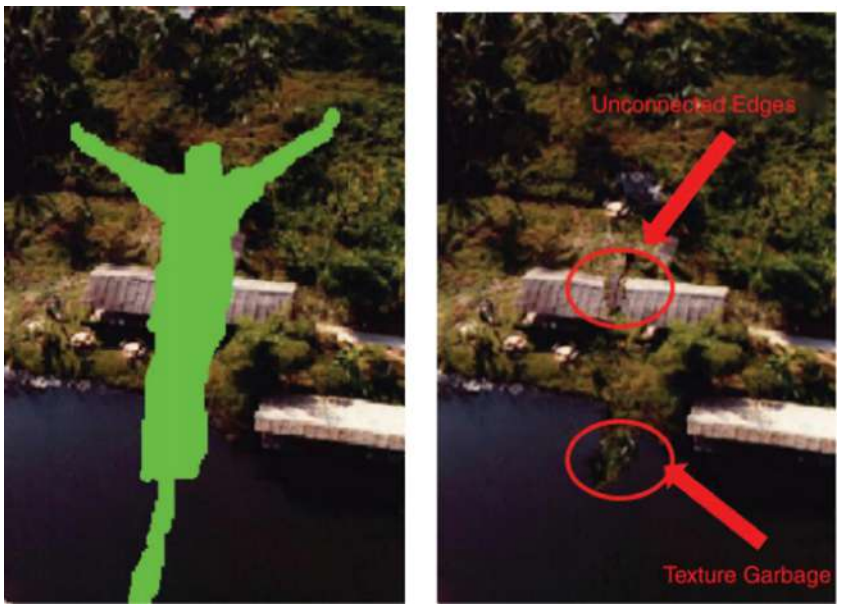

Fig. 7. An example of inconsistent texture regions in exemplar-based methods, original image with mask to be filled (left), inpainted image using [4,2] (right).

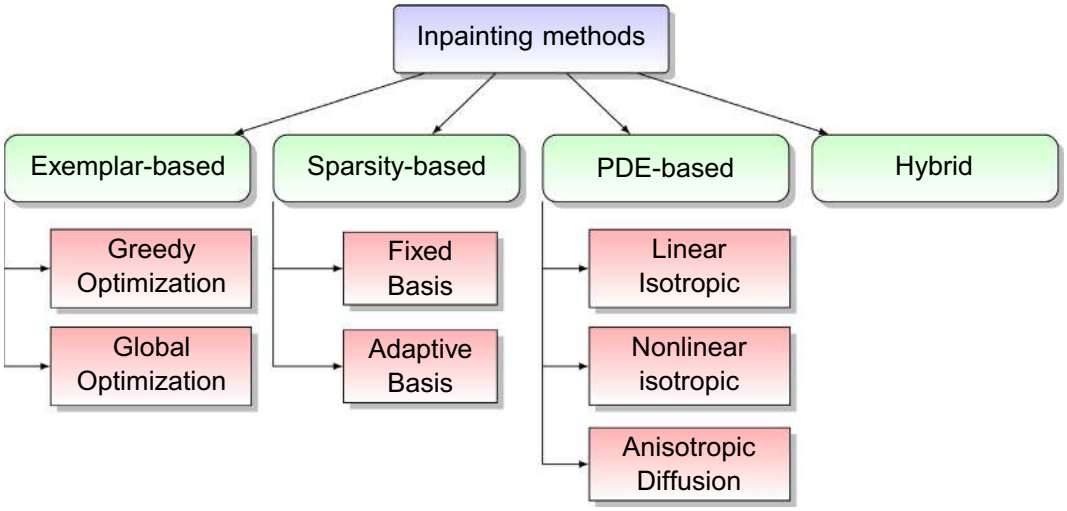

Fig. 8. Tree diagram of different classes of inpainting techniques.

its kind and is expected to help researchers working in this area in benchmarking new inpainting techniques, and in developing more robust methods for inpainting quality assessment, and benchmarking their results.

The rest of the paper is structured as follows: Section 2 provides a discussion of common inpainting algorithms. Section 3 briefly discusses the state-of-the-art IIQA metrics. Different inpainting databases are discussed in Section 4. The experimental results in terms of correlation performance are discussed in Section 5. Finally, the paper is concluded in Section 6.

2. Image inpainting methods

The main objective of inpainting algorithms is to restore the unknown regions to create a more pleasing and realistic feeling about the new (restored) image. Among different types of inpainting artifacts, most commonly observed ones are blurring, disconnected edges, inconsistent pieces of texture, etc. [2,1]. Based on our analysis of the state-of-the-art, we propose here to group inpainting algorithms into four broad categories: Exemplar-based, Partial Differential Equation (PDE)-based, Sparsity-based, and Hybrid (combination of Exemplar-, PDE-, or Sparsity-based) approaches. We display in Fig. 8 a tree diagram of different classes of most commonly used inpainting algorithms. Since, the paper aims to provide a critical review of IIQA metrics instead of the inpainting algorithms themselves, we will only provide a brief discussion of each of these categories. Note that in image inpainting, the unknown pixels are estimated using the known pixels information with the assumption that the pixels in the known and unknown regions share similar geometrical structures and statistical properties.

In PDE-based or diffusion-based inpainting methods, the local structure information is transferred or diffused from the known region to the unknown (target) region [1]. Several variations of PDE-based methods were introduced based on the flow of texture information in linear, nonlinear, isotropic, or anisotropic directions. The PDE-based methods are well-adopted for restoring long narrow regions (cracks, lines). However, they are not suitable for

Fig. 9. An example of blurring artifacts in PDE-based image inpainting method, original image with large hole to be filled (left), inpainted image using [12,2] with introduced blur int the large restored regions (right).

restoring large unknown texture regions, due to the introduction of blur in the textured regions (see Fig. 9). The PDE-based inpainting methods have fast processing time compared.

In Exemplar-based inpainting techniques, the structure completion process is carried out using texture synthesis i.e., the target regions are restored by selecting the patches in the known regions similar (in terms of structure) to the partially unknown patches in the target regions [4]. These techniques use greedy and global optimization cost functions in filling the target regions with the similar known regions/pixels. In greedy methods, the target region is filled in one iteration by copying multiple patches in a greedy manner [4,13]. Whereas, the global optimization and energy-based techniques minimize a certain energy function in several iterations until convergence is reached [14].

In contrast to PDE-based inpainting methods, Exemplar-based inpainting algorithms perform better in restoring large regions at the cost of high computation cost. The time required to restore the unknown patches best matching the known patches, is also high. The choice of patch size is also critical in Exemplar-based inpainting. For small source patches, it is difficult to find the best match and results in poor inpainted image quality. Moreover, some undesirable artifacts in terms of disconnected edges and inconsistent texture regions are visible in the restored images using exemplar-based inpainting methods. Figs. 6 and 7 show examples of these artifacts produced due to exemplar-based inpainting.

Sparse representations of images, over a particular basis (e.g., Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), etc.) has recently attracted a lot of attraction [15]. Sparsity has also been exploited as a powerful criterion in image inpainting [16,17]. In Sparsity-based inpainting methods, the basic idea is representing an image by sparse combination of transformed bases with the assumption that the known and unknown image regions share the same sparse representations. Therefore, the missing pixels are filled by adaptively updating the sparse representations. The choice of the dictionary is very important in sparse representations. The dictionary may be fixed (like DCT and DWT) or adaptive/content dependent. In [16], Chan et al. proposed an inpainting method using the basic Harr-wavelet based fixed dictionary for sparse representation of image. Guleryuz et al. in [17] proposed an image inpainting algorithm based on adaptive sparse reconstruction and iterative denoising.

The Exemplar-based and the Sparsity-based inpainting methods, above, were shown to perform better the traditional PDE-based methods for filling large texture regions. Various Hybrid techniques also exist that combine the strengths and different types of inpainting methods for performance improvement. These combine PDE-, Exemplar-, and Sparsity-based inpainting methods to reconstruct large missing regions first, then reconstruct missing pixels in the thin regions [18–20].

To demonstrate the effect of different inpainting algorithms, Fig. 10 shows a very nice example of inpainting where broken pieces of the kiwi fruit are restored using different inpainting algorithms [21]. The broken area is shown as a green mask. It is clear that Fig. 10(d) and (f) represents more realistic inpainting output compared to the other methods.

After this brief survey on commonly used inpainting algorithms, we now move to the focus of the paper and discuss in more details different types of subjective and objective IIQA metrics, commonly used in the literature.

3. Image inpainting quality assessment measures

Image inpainting methods were initially used for removing missing or damaged areas in an image [22]. The main criterion was that the restored image should be "close" to the original one. The traditional fidelity metrics were used to evaluate the quality of inpainted images. The Mean Squared Error (MSE) and Peak Signal to Noise Ratio (PSNR), which are considered as the most widely used fidelity metrics, were the simplest ones available. Oliveira et al. in [23], for example, used these metrics for quality evaluation of inpainted images. However, both MSE and PSNR are not well correlated with perceptual quality assessment [24]. In inpainting applications, the objective is to restore the original image such that it is more appealing and that the artifacts introduced inside, outside, and around the inpainted regions, are not noticeable/visible.

For performance evaluation of different inpainting algorithms, the metric of choice would be a qualitative judgment averaged over a number of human observers. In this regard, Hays et al. [25] qualitatively evaluated inpainting image quality for the first time using subjective experiments. The purpose of the experiment was just the identification of the original and the tampered (inpainted) images. The proposed method was compared with an exemplar-based approach [4]. Twenty naive observers participated in the subjective tests and were asked to differentiate between the real image and the inpainted image (tampered image). The detection rate for tampered images was achieved as 34%, 64%, and 3% for images in [25,4], and the original images. The purpose of the study was to investigate whether or not the proposed inpainted algorithm produced better perceptual quality image compared to other methods. The results showed that the inpainted images from [25] looked closer to the original images with good perceptual quality. However, the study did not provide any quantitative ratings of inpainted images.

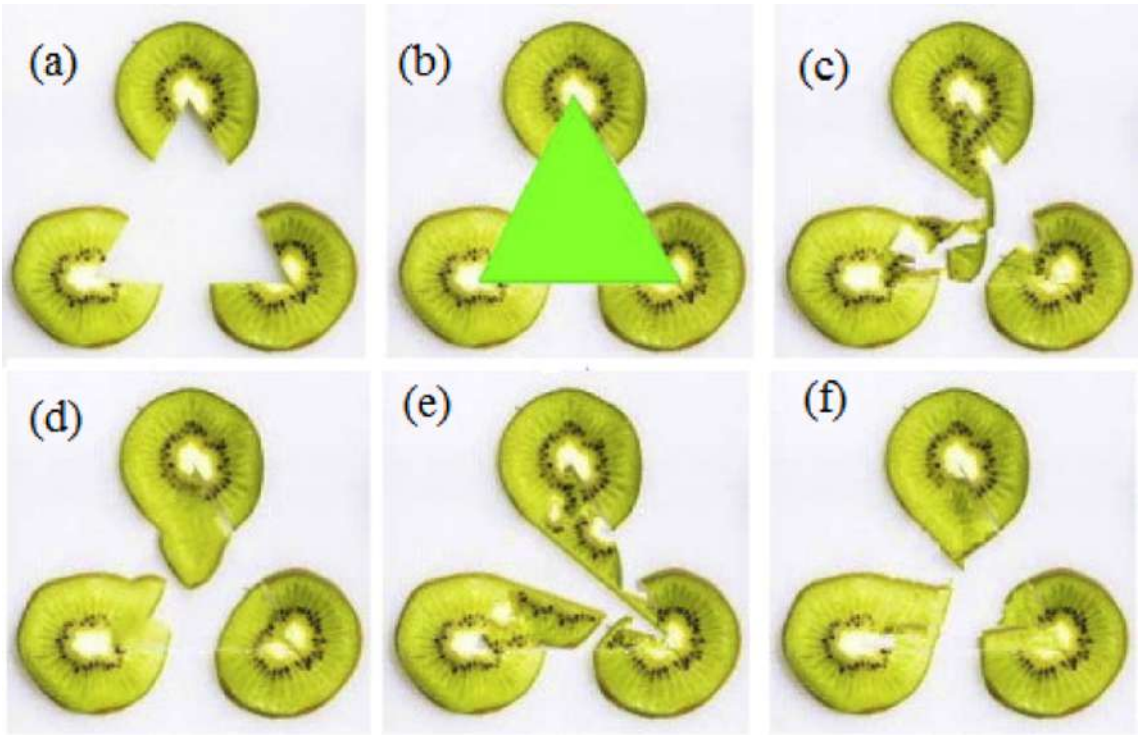

Fig. 10. An example of broken object restoration using inpainting (a) original image, (b) original image with target region in green color, and (c–f) inpainting results from different algorithms [21]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

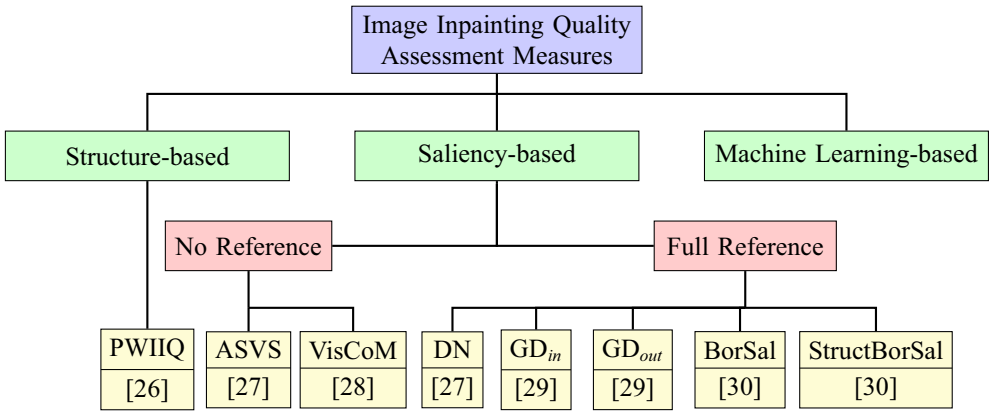

Fig. 11. Proposed framework for grouping the different IIQA metrics.

Subjective assessment methods involve humans and the ratings provided are considered as most reliable and accurate in relation to perceived quality. However, these methods are time-consuming, laborious, and require a significant number of observers to be consistent. They also require a well-controlled environment and lighting conditions. This has motivated researchers in this field to develop alternative objective metrics for inpainting quality assessment without the need of human involvement. Such objective methods use mathematical tools to extract characteristic features from either the reference or the test images or both. These features are then used to get a single quality score for the given image. The aim of objective quality assessment techniques is to predict perceived image quality, the way a human observer perceives it.

Traditional objective quality assessment methods, depending upon the availability of the original image, are grouped into Full Reference (FR), Reduced Reference (RR), and No Reference (NR) methods. In FR methods, the original image is required in addition to the processed image (inpainted image). These are impractical as the original image is usually not available. With NR quality prediction methods, the original image is not available. For RR techniques, partial information about the original images is available in the form of some extracted features. RR techniques are seen as a compromise between FR and NR methods. For the case of inpainting, only NR methods are appropriate due to the unavailability of original image information. Based on our analysis of the literature, we are proposing to group existing IIQA measures into three broad categories, i.e., structure-based, saliency-based, and machine learning-based measures. We display in Fig. 11 a tree diagram listing the different metrics commonly used for IIQA. The different categories are now discussed in more details.

For the sake of completeness, we outline the notation we used in this paper in Table 1. We represent an image to be inpainted

Table 1
Notations used for IIQA metrics.

Notation	Description
I_r	Original image
I_{inp}	Inpainted image
$S(.)$	Saliency map of the original image
$S'(.)$	Saliency map of the inpainted image
Ω	Inpainted region
Φ	Remaining region $I-\Omega$
p	Pixel value under consideration
W	Image width
H	Image height
b	Block size
(i,j)	Pixel index

with I_r and the inpainted image with I_{inp}. The original image to be inpainted is decomposed into three distinct regions, (1) the region to be restored or modified by the inpainting algorithm is represented by Ω, (2) the remaining area is denoted by Φ, and (3) the boundary between the two regions is indicated by $\delta\Omega$. Fig. 12a shows a simple image inpainting model where different regions are clearly labeled.

After outlining our notation, we will now start discussing each of the groups of the IIQA metrics mentioned in Fig. 11.

3.1. Structure-based IIQA measures

In image inpainting, some of the structural details in the original image are either removed or replaced. Inspired by the use of the Structural Similarity Measure (SSIM) [31] in traditional IQA, Wang et al. [26] proposed a FR metric using luminance, definition, and gradient similarities, to determine a quality index for inpainted images. The metric, defined as Parameter Weight Image Inpainting Quality (PWIIQ), is calculated as follows:

$$\mathrm{PWIIQ} = \left[L(I_r, I_{inp})\right]^{\alpha}\left[D(I_r, I_{inp})\right]^{\beta}\left[G(I_r, I_{inp})\right]^{\gamma} \tag{1}$$

where the terms L, D, and G, represent the variances of image luminance, definition, and gradient similarity between the original and inpainted images. The α, β, and γ are positive parameters used to determine the importance of each term in the final quality score.

For implementation purposes, both the original and the inpainted images are first divided into $b \times b$ fixed-size blocks, and the luminance similarity between the corresponding blocks is computed:

$$l(I_r, I_{inp}) = \frac{2\mu_r\mu_{inp} + K_1}{\mu_r^2 + \mu_{inp}^2 + K_1} \tag{2}$$

where μ_r and μ_{inp} represent the mean values of the original and inpainted image blocks respectively, while K_1 is a positive constant with very small value to avoid instability when the denominator is close to zero.

The weighted block means for the original and inpainted images are used. The weights are computed from the symmetrical Gaussian filter window of size 11×11 pixels. The resulting weighted mean is given by:

$$\mu = \sum_{i=1}^{N} \mathrm{w}_i x_i \tag{3}$$

where $\mathrm{w} = \left\{\mathrm{w}_i \text{ such that } \sum_{i=1}^{N}\mathrm{w}_i = 1,\ i = 1,2,\ldots,N\right\}$, where N denotes the number of pixels in the window.

The luminance component, L, in Eq. (1) is computed as the average of the luminance similarities across all blocks:

$$L(I_r, I_{inp}) = \frac{1}{B_1 \times B_2}\sum_{i=1}^{B_1}\sum_{j=1}^{B_2} l(I_r^{ij}, I_{inp}^{ij}) \tag{4}$$

where B_1 and B_2 represent the number of overlapping blocks along the rows and columns of the image.

Secondly, the image definition function, D, is computed as follows:

$$D(I_r, I_{inp}) = \frac{\sum_{i=0}^{W-1}\sum_{j=0}^{H-1}|\mathcal{F}_{inp}^{ij}| - |\mathcal{F}_{inp}^{00}|}{\sum_{i=0}^{W-1}\sum_{j=0}^{H-1}|\mathcal{F}_{r}^{ij}| - |\mathcal{F}_{r}^{00}|} \tag{5}$$

where $\mathcal{F}(.)$ represents the Fourier transform of an image and $\mathcal{F}_{00}$ is the dc component or overall mean value of an image.

Finally, the gradient similarity component is defined as:

$$G(I_r, I_{inp}) = \frac{2\sum_{i=0}^{W-1}\sum_{j=0}^{H-1} G_r^{ij} G_{inp}^{ij} + K_2}{\sum_{i=0}^{W-1}\sum_{j=0}^{H-1}\left[G_r^{ij}\right]^2 + \sum_{i=0}^{W-1}\sum_{j=0}^{H-1}\left[G_{inp}^{ij}\right]^2 + K_2} \tag{6}$$

where the $G(.)$ represents the gradient magnitude computed from the Sobel filter mask of size 3×3 in the vertical and horizontal directions and K_2 is a small positive constant.

Similarly to its IQA counterpart, the structure-based methods suffer from some serious limitations. Since image inpainting operations do not require the original images, the large inpainted regions may be quite different from the actual ones. Consequently, the structural similarity based methods (e.g., [26]) may fail for images with large inpainted regions (see Fig. 9). To overcome the drawbacks of structure-based methods, researchers started introducing image saliency to derive new measures for evaluating

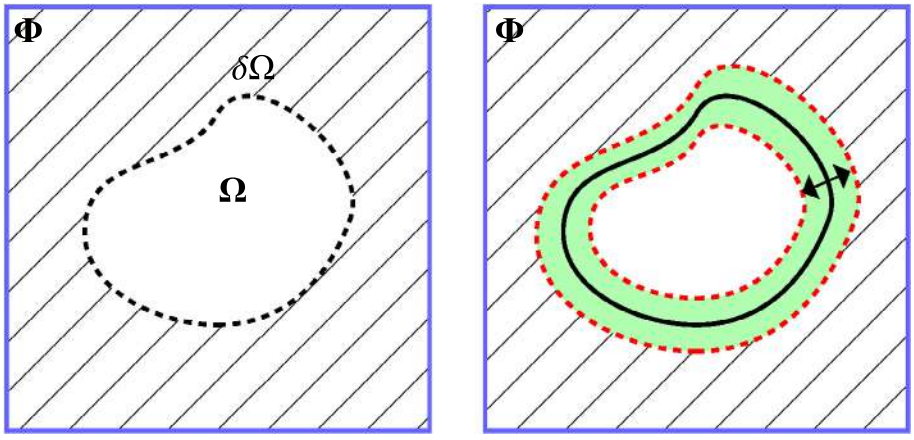

Fig. 12. (a) A simple model used in image inpainting techniques. (b) A typical model used for BorSal and StructBorSal IIQA metrics computations (shaded region is used for IIQA).

quality of inpainted images. Other variants of the SSIM have also been proposed with limited success [32].

3.2. Saliency-based IIQA measures

Visual saliency plays a significant role in image quality assessment applications. Image saliency is used to highlight the areas towards which the human vision system is more sensitive/attracted. Various saliency detection algorithms exist in the literature [33–35]. Given its importance in IQA, saliency has been used in estimating visibility of different artifacts introduced by inpainting. The basic idea is that salient regions change after inpainting. The most prominent IIQA metrics using the concept of saliency are now briefly outlined.

3.2.1. Average Squared Visual Salience (ASVS)

In [27], Ardis et al. proposed two objective metrics for quality assessment of inpainted images. The image saliency was used in capturing the distortions introduced during the restoration process. The first metric is the ASVS, which is represented by the normalized sum of squares of the saliency values within the inpainted region. The ASVS metric relates to the noticeability of the inpainted pixels compared to the overall scene. ASVS is a NR metric as it does not require the original image information. It is calculated as follows:

$$\text{ASVS} = \frac{1}{\|\Omega\|}\sum_{p\in\Omega}|S'(p)|^2 \tag{7}$$

where $S'(p)$ represents the saliency value for the inpainting pixel, p, within to the inpainted region, Ω. High values of the ASVS correspond to more visibility of inpainting related artifacts and reduced perceptual quality [27].

3.2.2. Degree of Noticeability (DN)

Ardis et al. in [36], categorized inpainting artifacts into two broad classes, i.e., in-region and out-region artifacts. During the restoration operation in image inpainting, the in-region artifacts occur due to the introduction of distinct color and structures in the inpainted regions only. These artifacts result in an increased saliency in the inpainted areas and thus disturb attention flow. The ASVS metric relates to the in-region artifacts as it only considers the salient pixels within the inpainted region.

Similarly, the out-region artifacts occur when the local colors or structures are not properly extended to the inpainted region by the inpainted method. These artifacts result in an increase in the saliency of the inpainted region neighborhood and hence decreases attention flow within the inpainted region. The in-region and out-region artifacts are computed as follows:

$$\text{In-region} = \text{ASVS} = \frac{1}{\|\Omega\|}\sum_{p\in\Omega}|S'(p)|^2 \tag{8}$$

$$\text{Out-region} = \frac{1}{\|\Phi\|}\sum_{p\in\Phi}|S'(p)-S(p)|^2 \tag{9}$$

Ardis et al., in [27], took into account both in-region and out-region artifacts and proposed another metric, namely the DN (Degree of Noticeability). The DN measure is intended to identify non-noticeable inpainted regions and indicates the change in attention flow in the neighborhood of the inpainted regions. It is calculated as follows:

$$\text{DN} = \frac{\|\Omega\|}{\|\Omega\|+\|\Phi\|}\text{in-region} + \frac{\|\Phi\|}{\|\Omega\|+\|\Phi\|}\text{out-region} \tag{10}$$

Eq. (10) can further be simplified as follows:

$$\text{DN} = \frac{1}{\|\Omega\|+\|\Phi\|}\left(\sum_{p\in\Omega}|S'(p)|^2 + \sum_{p\in\Phi}|S'(p)-S(p)|^2\right) \tag{11}$$

For both ASVS and DN calculations, the saliency maps are generated using the iLab Neuromorphic Vision Toolkit (iNVT) version 3.1, using scale-4 and discretization of $1:16$. The expected visual cortex stimulation was set with 0.1 ms observation cutoff. Furthermore, four orientation scales, three center scales (2–4), and two center-surround channels $(3,4)$ were considered [27].

Similarly to the ASVS, high values of the DN correspond to more visibility of inpainting related artifacts and reduced perceptual quality. The authors claimed a good correlation for both metrics, with subjective ratings. However, the subjective ratings were not considered reliable as only three observers participated in the psychophysical experiment. Moreover, the overall visual appearance of an image is also ignored while calculating DN and ASVS IIQA metrics.

3.2.3. Gaze density (GD)-based IIQA measures

Following the work in [27], Mahalingam et al. [29] proposed two visual saliency-based metrics for inpainting quality assessment within and outside the inpainted regions. From eye-tracking experimental data, the gaze densities were used to capture the saliency in the original and inpainted images. The motivation was that changes in the saliency map in the inpainted image is related to its perceptual quality.

In their subjective experiments, 45 reference images and 90 modified images were obtained using two different inpainting algorithms. The images were equally distributed into three subsets. Twenty-four naive observers without any prior knowledge of the original and the inpainted images rated the subgroups under ambient lighting conditions and at a distance of 65 cm from the display screen. The average gaze distribution was calculated for each image from the eye-tracking experiment. It was observed that the Human Visual System (HVS) is more attracted towards the regions with more noticeable inpainting artifacts (see Fig. 13). The gaze density was calculated for both inside and outside the inpainted regions using:

$$\text{GD}_{\text{in}} = \frac{1}{\|\Omega\|}\sum_{p\in\Omega}S'(p) \tag{12}$$

$$\text{GD}_{\text{out}} = \frac{1}{\|\Phi\|}\sum_{p\in\Phi}S'(p) \tag{13}$$

The gaze density measures of the inpainted image were normalized by the gaze densities of the original image to account for variations in textures and sizes. The final normalized gaze densities were given by:

$$\overline{\text{GD}}_{\text{in}} = \frac{\sum_{p\in\Omega}S'(p)}{\sum_{p\in\Omega}S(p)}, \tag{14}$$

and

$$\overline{\text{GD}}_{\text{out}} = \frac{\sum_{p\in\Phi}S'(p)}{\sum_{p\in\Phi}S(p)} \tag{15}$$

The experimental results showed a strong correlation between the rankings from the subjective experiments and the gaze density based measures. However, these methods require the original image and are not suited for practical inpainting applications, where the original image is usually unavailable. Similar to the ASVS

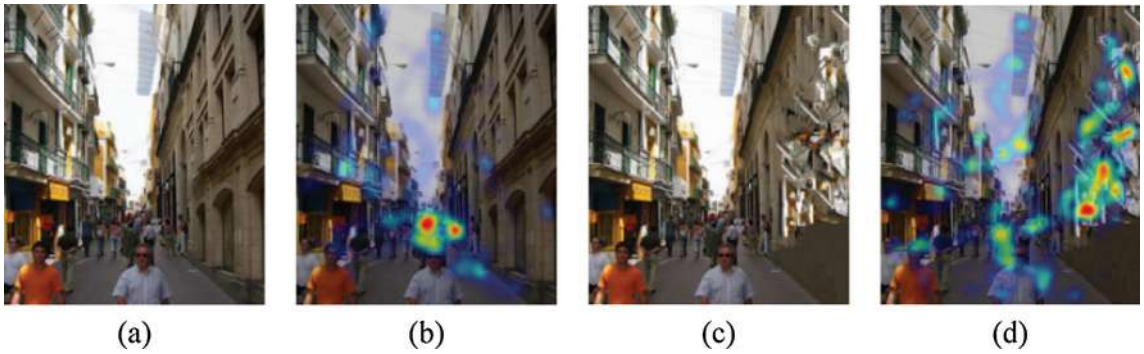

Fig. 13. An example of concentration of gaze distribution or HVS attention around discernible artifacts regions [29] (a) Image inpainted using [25]. (b) Average gaze distribution on (a). (c) Image inpainted using [4] having visible artifacts. (d) Average gaze distribution on (c).

and the DN metrics, both GD_{in} and GD_{out} highlight the change in attention flow within and outside the inpainted regions respectively, and do not consider the global visual appearance of the image. However, these measures are computationally more efficient compared to the ASVS and DN, since summation is carried over the salient pixels in the expressions of the GD-based IIQAs, rather than the square summations in the ASVS and DN IIQA measures.

3.2.4. Border Saliency based measures (BorSal)

The measures proposed in [27,29] considered the in-region and out-region artifacts separately. Oncu et al. [30] showed that saliency map pixels in the neighborhood of the inpainted region are sufficient to capture the changes in saliency due to inpainting. The BorSal metric was proposed to compute the normalized gaze density using the border pixels extended only to three pixels inside and outside the inpainted regions. The six pixels wide border area simultaneously contain information from both in-region and out-region artifacts (see Fig. 12b). The BorSal metric was computed as follows:

$$\text{BorSal} = \frac{\sum_{p\in Border} S'(p)}{\sum_{p\in Border} S(p)} \tag{16}$$

3.2.5. Structural Border Saliency based measures (StructBorSal)

The BorSal IIQA metric accounts for changes in the flow of attention over the inpainted image. Oncu et al. [30] proposed another metric called StructBorSal, to account for the structure information in the whole image and to highlight the artifacts in the restored image. The $SSIM_{IPT}$ based measure [37] was used by taking the geometrical mean of the three SSIM computed for each color channel separately. The StructBorSal metric combines the BorSal metric with the structural measure as follows:

$$\text{StructBorSal} = \text{BorSal} + \text{SSIM}_{IPT} \tag{17}$$

The correlations between the subjective ratings and 14 quality measures (IQA metrics for distortions as well as inpainting) were calculated. The results showed poor performance of existing metrics. The inpainting IQA metrics performed well for images with small and less structured inpainted regions.

The above mentioned saliency-based inpainting IQA metrics, i.e., DN [27], GD_{in} [29], GD_{out} [29], BorSal [30], and StructBorSal [30] require the original image whereas in the restoration process, usually the original image is not available. The overall visual appearance of the image also plays a significant role in the quality perception. These metrics are also lacking in considering the global visual appearance of an image, limiting their use in practical setups.

3.2.6. Visual Coherence Metric (VisCoM)

In image inpainting, the reference image is usually not available, therefore the restored pixels rely solely on the surrounding pixels. The restored pixels in the inpainted regions, however, should exhibit consistency with existing pixels. The coherence of the inpainted regions, which is associated with the degree of annoyance of noticeable distortions, is computed by taking the correlation between the inpainted pixels and the existing ones in the region outside the inpainted region. Similarly, the HVS is more sensitive to the edges and contours in an image. The presence of contours and edge details are more attracted by the HVS compared to the remaining regions. The saliency map of a given image relates to the degree of attention in the image and hence can be used to weight the coherence map in evaluating the final quality index.

Trung et al. [28,38–40], proposed a NR quality metric using visual coherence and visual saliency of restored regions. The index is computed as follows:

$$\text{VisCoM} = \frac{1}{\|\Omega\|}\sum_{p\in\Omega} C(p)^{\alpha} S(p)^{\beta} \tag{18}$$

where $C(p)$ and $S(p)$ represent the coherence term and the saliency or structure term respectively. The exponents α and β control the significance of each term in the final quality score.

The coherence term, C, is basically a similarity index between the inpainted regions and the remaining ones in the inpainted image. It is defined as follows:

$$C(p) = \max\left[SIM(\Psi_p, \Psi_q), \quad \text{all} \quad \Psi_q \in \Phi \quad \text{and} \quad \text{all} \quad \Psi_p \in \Omega\right] \tag{19}$$

where Ψ_p and Ψ_q represent small patches around pixels, p and q, respectively. The SIM is the similarity function between two patches. Instead of the MSE or PSNR, the SSIM index [31], used in classical IQA, is exploited here to find the coherence between patches of size 7×7. The SIM is defined as:

$$SIM(\Psi_p, \Psi_q) = \frac{(2\mu_p\mu_q + K_1)(2\sigma_{pq} + K_2)}{(\mu_p^2 + \mu_q^2 + K_1)(\sigma_p^2 + \sigma_q^2 + K_2)} \tag{20}$$

where μ_p, σ_p and μ_q, σ_q represent the mean and standard deviation of the patches, Ψ_p and Ψ_q, centered at pixels, p and q, respectively, whereas σ_{pq} is the cross correlation between the patches Ψ_p and Ψ_q, and K_1 and K_2 are small positive coefficients to insure stability when denominator is zero.

The local structure term is computed from the saliency values which are further used as weights in the final quality index. Among different saliency detection algorithms, the authors used a simple and computationally efficient method for salient region detection [41]. In [41], color and luminance information were used for sal-

iency detection. For a given image, I, the saliency map was generated using:

$$S' = \|I_\mu - I_G\| \tag{21}$$

where I_μ represents the mean value of the original image and I_G is a Gaussian blurred (5×5 filter mask) version of the original image. The operation is performed in the CIELab color space. The method is simple, computationally efficient, and does not need any down-sampling operation during the estimation of the saliency map. Finally, the saliency map, defined in Eq. (21), is normalized to the range $[0, 1]$:

$$S(p) = \frac{S'(p)}{\max_I(S')} \qquad \forall p \in \Omega \tag{22}$$

The authors in [28,38–40] used visual coherence of recovered regions and visual saliency describing visual importance to develop their index shown in Eq. (18). The proposed approach showed promising results but could only handle a limited number of possible inpainting artifacts.

Based on our study of existing approaches, we present in Table 2 a summary of both structure-based and saliency-based IIQA metrics. It is important to note that there exists only two NR-IIQA metric among these metrics. To overcome this limitation, among others, researchers tried to use advanced machine learning approaches in developing robust quality assessment metrics for practical inpainting applications.

3.3. Machine learning-based IIQA measures

Machine learning-based approaches were originally developed for solving classification and regression problems efficiently and provide good approximation of functional relationships between input features and output classes/scores scores from the training session. In the testing stage, a set of features is extracted from a given image. The trained model and the extracted features are then used for predicting the quality rating of the test image [42].

Among the first approaches using machine learning for IIQA is the metric proposed by Viacheslav et al. in [43]. The method is a NR approach for IIQA based on natural scene statistics and machine learning. First, the saliency map of the inpainted image is calculated to identify most important perceptual information in the inpainted image. The saliency map is then thresholded using average gaze density computed from the outside inpainted regions using Eq. (13) for proto-objects. Then, the DCT is calculated only for the proto-objects and used to train a dictionary of 100 classes, where each word in the dictionary is a DCT coefficient. For each DCT block, the histogram of words is then used as a feature vector. The quality scores collected from the subjective experiments and the extracted features were then used to train a Support Vector Regression (SVR) network, and to predict the quality of inpainted images resulting from different algorithms.

The same authors in [44], replaced the DCT based features with the traditional Local Binary Pattern (LBP) features given their power in describing image structures effectively. The quality scores collected from the subjective experiments and the extracted features are then used to train an SVR for quality prediction. For the subjective experiments, a database consisting of 300 images with different structures and textures was used. The database also included some real images. The images were restored using a mask and using four different inpainted methods. Ten observers participated in the subjective experiments and rated the quality of the inpainted images on a scale 1–5 (5 for excellent quality, 1 for worst quality). The results showed good correlation with human ratings of quality.

More recently, Markio et al. [45] showed that saliency is not an absolute requirement for assessing inpainting quality. They performed an experiment using a learning-to-rank approach. Instead of determining the absolute scores for inpainted images, the preference order is obtained among inpainted images from different inpainting algorithms. They demonstrated that visual saliency map is useful but not a requirement. Rather, they showed that some features can be used to reflect the changes within and outside the modified areas in an inpainted image. Such features are extracted from gaze measurements using a simple Tobii eye-tracker device. From each original image, twelve inpainted images are generated using two inpainted methods, three patch sizes, and two multiscale parameters. A hundred and eleven original images were used in the experiments. The proposed metric was compared to other existing metrics in terms of prediction accuracy in estimating the preferences order ranking. The authors showed that existing saliency-based IIQA metrics fail in ordering the inpainted images correctly due to the small significant difference in the saliency maps in the inpainted regions. The results using the proposed metric showed an improvement of at least 7% over other

Table 2
A summary of structure-based and saliency-based IIQA measures.

Metric	Year	Type	Regions used	Description	Strengths	Weaknesses
PWIIQ [26]	2008	FR	Overall image	statistical features used. A value close to 1 means better quality	Simple and fast	Requires reference image, fails for images with large holes (inpainted regions).
ASVS [27]	2009	NR	In-region	Used when fidelity is not important. High value means more visibility of artifacts, poor quality	Reference image is not needed	Ignore overall visual appearance of an image
DN [27]	2009	FR	In-region, out-region	Used for applications where preservation of original saliency is desired.	Highlights change in attention flow beyond the inpainted regions	Requires reference image, ignore overall visual appearance of an image
GD_{in} [29]	2010	FR	In-region	Value close to 1 means no deviation of attention flow in the inpainted image	Highlights change in attention flow within the inpainted regions	Requires reference image,
GD_{out} [29]	2010	FR	Out-region	Same as GD_{in}	Highlights change in attention flow beyond the inpainted regions	Requires reference image,
BorSal [30]	2012	FR	Border region	A single border region around the hole (inpainted region) is used.	Fast, a single border region around the hole (inpainted region) is used.	Requires reference image,
StructBorSal [30]	2012	FR	Border region	In addition to BorSal, structural information is also used.	Fast, a single border region around the hole (inpainted region) is used, structural artifacts are also captured	Requires reference image, coherence of the inpainted regions with remaining is ignored
VisCoM [28]	2013	NR	Overall image	Uses visual coherence along with structural information	Overall global visual appearance of image is considered	Requires reference image,

metrics with 68.65% prediction accuracy. To summarize the work on IIQA using machine learning techniques, we present in Table 3, a brief summary of most prominent approaches.

Before leaving the above discussion on IIQA, we now discuss another important issue of high relevance to IQA applications; that is the collection of quality ratings averaged over a number of observers over comprehensive databases with inpainted images from different algorithms. These databases are used as a ground truth for validating and testing or both inpainting algorithms and IIQA metrics.

4. Image inpainting quality assessment databases

With the tremendous increase of research activities in image inpainting algorithms and applications, it was crucial to develop comprehensive databases for performance evaluation and benchmarking of different inpainting methods. In the literature, usually, the performance of an inpainting algorithm is evaluated on own local images or using standard databases used for standard IQA problems. Given the importance of image inpainting in multimedia applications, publicly-available databases are needed for unbiased performance comparison. In this regards, Tiefenbacher et al. provided, for the first time, a public database namely the Technische Universitt München Image Inpainting Database (TUM-IID) [47], for objectively estimating quality of inpainted images and performance evaluation of different IIQA metrics. The database contained 17 reference images with diverse texture types and resolution of 640×480 pixels stored in PNG format. Each image in the database is inpainted using four state-of-the-art inpainted methods [48–51] and for four inpainting regions. Then, each inpainted image in the database was rated by 21 observers using a Single Stimulus (SS) subjective experiment protocol, and the ratings from all observers were averaged to get a single score for each image. Some sample images, inpainting masks, and inpainted images from public and private databases are shown in Figs. 14 and 15 respectively.

In an effort to summarize existing work in inpainting using different databases (private and public), we also present in Table 4, the most common experimental setups used in the literature.

5. Experimental results and discussions

The main objective of the paper is to provide a comprehensive performance evaluation and critical review of different state-of-the-art IIQA metrics. Therefore, after providing a detailed description of the existing IIQA metrics, the IIQA metrics are investigated for their consistency with the human visual perception [52]. For this purpose, we have used the psychophysical subjective experimental data provided with the TUMIID [47], a public image inpainting database. For performance evaluation, we have used the Spearman Rank Order Correlation Coefficient (SROCC) measure. The SROCC is widely used as a non-parametric measure to determine the monotonicity between the ranks of two variables and

Table 3
A summary of machine learning-based IIQA measures.

Method	Year	Type	Feature description	Regression
Voronin et al. [43,46]	2014	NR	DCT-based dictionary	SVR, RBF[a] kernel
Voronin et al. [44]	2015	NR	LBP histograms	SVR, EMD[b] kernel
Markio et al. [45]	2016	NR	Gaze features	RankingSVM, RBF kernel

[a] RBF (Radial Basis Function).
[b] EMD (Earth Mover's Distance).

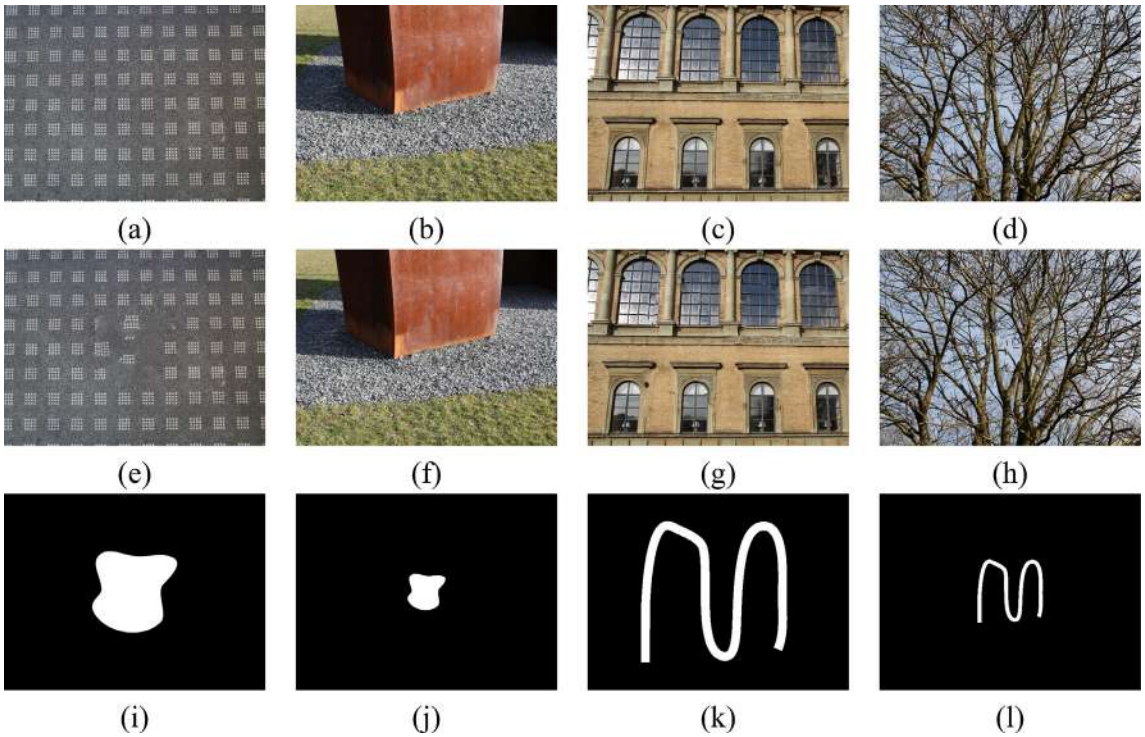

Fig. 14. Sample images from the TUM-IID [47] database: (a–d) Reference images, (e–h) inpainted images using [48], and (i–l) masks used for inpainting [48].

Fig. 15. Sample local images used in [40] for inpainting quality assessment (a) original image with mask, inpainted image using [4] (b), [54] (c), and [55] (d).

Table 4
Experimental setups used for private and public inpainting databases.

Method	[26]	[29]	[30]	[27]	[56]	[47]	[40]	[45]
Year	2008	2009	2010	2012	2013	2015	2015	2016
Original images	3	–	45	6	8	300	7	111
Inpainted images	9	–	90	48	–	1200	56	2466
Image resolution	–	–	–	–	–	–	640×480	–
Subjective method[a]	Rank-order	–	SS	DS	Web	SS	SS	Rank
Score type	Ranking	4-scale	Gaze density	5-scale	5-scale	5-scale	5-scale	5-scale
Screen resolution[a]	–	20 cm ht.	1280×960	–	–	–	–	1280×1080
Observers	–	5	24	69	15–20	10	21	24
Viewing distance[c]	–	0.7 cm	65 cm	Web	–	–	–	60 cm
Inpainting methods	3	–	2	8	–	4	4	2
Evaluation measures	2	2	2	14	–	4	3	6
Access	Private	Private	Private	Private	Private	Private	Public	Private

[b] ht. (height).
[a] SS (Single Stimulus), DS (Double Stimulus).
[c] Web (through web-based interface).

its value ranges from -1 to $+1$. The value is close to $+1$ in the case of strong correlation between the ranks of two variables, and -1 in the case of strong disagreement between the two variables. The SROCC gives zero value when there is no correlation between the ranks. In our study, the aim is to observe how well an IIQA measure is consistent in capturing the ranking for the four inpainted versions of each original image in the database. Therefore, before evaluating the correlations, we must consider, how the change in the magnitude of metric values affects image quality. For some metrics, high values correspond to good quality, whereas for other metrics, the opposite is true (see Section 3). For preference ranking, the highest score is highly ranked. Whereas, the metrics with high/low values corresponding to good quality are also highly ranked. Using the concept of strength of correlation alone without signs does not convey the desired purpose [53].

It has been shown that the performance of an IQA measure is substantially related to certain image characteristics [57]. Therefore, in this work, we calculate the image-wise correlations for the selected IIQA metrics. For each image in the database, we have the ranking scores for its four inpainted versions as well as the quantitative scores.

If I_i represents an original image and $I_{i,j}$, its inpainted version processed by method M_j, for $i = 1, 2, \ldots, n_I$ and $j = 1, 2, \ldots, n_J$, for ($n_I = 7, n_J = 4$). Here n_I and n_J represent the number of original images and the number of inpainting methods respectively. We compute the SROCC for each image using the following relation:

$$\rho_{i,k} = 1 - \frac{6d_{i,j}^2}{n_J(n_J^2 - 1)}, \quad \text{for } i = 1, 2, \ldots, n_I \tag{23}$$

where $d_{i,j}$ represents the difference in the ranks of subjective preferences and objective scores of the kth IIQA measure for the ith image. The image-wise correlation and mean SROCC's for two different types of inpainting masks (i.e., compact and spread-out masks) are reported in Tables 5 and 6.

From the results, it is clear that the performance of the metrics is effected by the choice of the inpainted regions. For compact inpainted regions, it is difficult to restore the missing pixels compared to the spread-out (thin) inpainted regions. Obviously, evaluating quality when reconstructing large regions is more challenging than when only small regions are missing due to long term memory requirements for large regions. From the results, it is clear that IIQA metrics showed better correlation performance in the case of spread-out (thin) inpainted regions compared to the compact (large) inpainted regions. The ASVS and the DN metrics showed strong disagreements with the subjective ratings (negative

Table 5
Image-wise correlation performance of the IIQA metrics using the compact inpainting masks.

Image	2	3	6	7	10	14	17	Mean SROCC
ASVS [27]	−1	−1	−1	−0.8	−1	−0.8	−0.8	−0.9143
DN [27]	−1	−1	−1	−0.8	−1	−0.8	−0.8	−0.9143
GDin [29]	1	1	0.8	0.8	1	0.8	0.8	0.8857
GDout [29]	0.4	0.2	0.4	−0.4	−0.4	0.8	0	0.1429
BorSal [30]	0.8	1	0.8	0.8	0.4	0.8	0.4	0.7143
StructBorSal [30]	0.8	1	1	0.8	0.4	0.8	0.4	0.7429
VisCOM [28]	0.8	1	0.8	0.4	0.4	1	0.4	0.6857
PWIIQ [26]	0.8	1	0.8	0	−0.8	0.8	0.8	0.4857

Table 6
Image-wise correlation performance of the IIQA metrics using the spread-out inpainting masks.

Image	2	3	6	7	10	14	17	Mean SROCC
ASVS [27]	1	−1	−0.8	−0.4	−1	−1	−1	−0.8857
DN [27]	−1	−0.8	−0.8	−0.4	−1	−1	−1	−0.8571
GDin [29]	1	1	0.8	0.4	1	1	0.4	0.8000
GDout [29]	1	1	0.8	0.4	1	1	0.4	0.8000
BorSal [30]	0.8	1	0.8	0.8	0.4	1	0.8	0.8000
StructBorSal [30]	0.8	1	0.6	1	0.4	1	0.4	0.7429
VisCOM [28]	1	1	0.4	0.8	0.8	0.4	0.6	0.7143
PWIIQ [26]	0.8	1	1	0.8	0.4	1	1	0.8571

Table 7
Execution time for different IIQA metrics.

Metric	Time (s)
ASVS [27]	0.3897
DN [27]	0.8471
GDin [29]	0.8943
GDout [29]	0.8657
BorSal [30]	0.8541
StructBorSal [30]	2.845
VisCOM [28]	300
PWIIQ [26]	0.2436

correlation values) for both types of inpainted regions. Whereas, other metrics provide good correlation performance with the subjective judgements. The GDin, BorSal, StructBorSal, and VisCOM IIQA metrics showed consistent correlation performance for both types of inpainting regions.

5.1. Computational time analysis

In real-time applications, low-complexity IIQA metrics are needed. Therefore, in addition to comparing the correlation performance, we also compute the computational time of each IIQA metric. We list in Table 7, the time taken (in s) to compute the metric score for a single image of resolution 512×512. The experiments were performed on a notebook Intel Core i5-24500 M CPU@2.5 GHz and 4G RAM. The software platform is MATLAB R2013b under Windows 8.1. The execution times are only provided to show the relative complexity of the different metrics. The results show that among the NR IIQA metrics, the VisCOM which is good in terms of correlation requires the largest computational time while the ASVS requires second least computational time but has poor correlation performance. Whereas, among various FR-IIQA metrics, the BorSal, and StructBorSal have moderate computational complexity and good correlation performance. The PWIIQ metric provides the least (best) computational time with moderate correlation performance. The results in Table 7 only provide an idea about the comparative complexity of different IIQA metrics. In real-time applications, the algorithms can be substantially optimized.

6. Conclusion and future trends

The quality assessment of inpainted images continues to be a complex and challenging problem. It is substantially different, in nature, from classical IQA, due to the different types of artifacts not commonly observed in other applications, the importance of perceptual clues in scene continuity, and the need for convincing and plausible recovered images. While the advances made in developing robust methods for inpainting are significant, little efforts have been put in developing inpainting-dedicated quality assessment (IIQA) metrics. To bridge this gap, we present this survey paper which summarizes current research efforts carried in the development of robust IIQA measures as well as current challenges and future research directions and applications of inpainting and inpainting-related quality assessment measures. We also provide a performance comparison of different metrics in terms of correlation performance and computational complexity.

Our study revealed that among existing IIQA metrics, many of these require the availability of the original image, whereas image inpainting is usually used in the case of unavailability of the reference or original image. The need for developing robust NR-IIQA metrics is still persistent. From the review of state-of-the-art, we also noticed that saliency-based methods fail for cases where no difference in perceptual saliency can be seen around the damaged/restored regions. Machine learning-based are shown to overcome some of these limitations. Moreover, rather than determining the absolute quality rating scores, the preference order estimation-based methods were shown to provide very convincing results.

We also observed that most of the developed metrics were designed and validated over private databases consisting of a limited number of images and human raters. Currently, only one public database [47] exists which contains a limited number of images. Given the importance of this evolving field in multimedia, it is becoming crucial to develop new public databases consisting of large numbers of inpainted images, generated using various inpainted methods. This will help in providing unbiased comparisons among different IIQA metrics, highlighting their shortcomings, and in introducing new efficient quality measures well-correlated with human perception of quality of inpainting operations.

The overall performance of IIQA measures also depends on the selected features, and how these are exploited in scene continuity and assessing resulting artifacts. The features describing naturalness, colorfulness, continuity of pixels, among others, around the vicinity of damaged regions were shown to facilitate the IIQA task.

Our analysis also showed the importance of understanding image content before selecting the most appropriate IIQA metric. We have seen that while the majority of inpainting techniques focus on small regions continuity (e.g., PDE), others work better with large regions (e.g., Exemplar). Obviously, evaluating quality when reconstructing large regions is more challenging than when only small regions are missing due to long term memory requirements for large regions. As such, IIQA metrics should be designed (or selected) taking into consideration whether the given inpainting image contains large or small reconstructed regions.

Currently, we see a growing number of inpainting applications embedded in new generation mobile devices to remove (and reconstruct) certain objects from captured photos. Hence, computational efficiency of inpainting and IIQA algorithms becomes more relevant to these types of platforms. Other applications of IIQA are the quality prediction of inpainted images over the cloud or in wireless environments. Video inpainting is also a challenging problem when there is a need to remove and track undesired objects in videos or movies. Research in Video Inpainting Quality Assessment (VIQA) is expected to flourish over the next few years as we see more people working in this all-important branch of multimedia.

Acknowledgments

The authors would like to thank the editor and the anonymous reviewers for their valuable comments. The work presented here has been developed in collaboration with L2TI Research Lab, Univ. Paris 13. The research was supported in part by the project GTEC 1401-1402 under the joint Center of Energy and Geoprocessing (CeGP) at King Fahd University of Petroleum & Minerals (KFUPM) and Georgia Tech.

References

[1] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, Image inpainting, in: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '00, ACM Press, New York, USA, 2000, pp. 417–424, http://dx.doi.org/10.1145/344779.344972.

[2] C. Guillemot, O. Le Meur, Image inpainting: overview and recent advances, IEEE Sig. Process. Magaz. 31 (1) (2014) 127–144, http://dx.doi.org/10.1109/MSP.2013.2273004.

[3] Z. Chen, C. Dai, L. Jiang, B. Sheng, J. Zhang, W. Lin, Y. Yuan, Structure-aware image inpainting using patch scale optimization, J. Vis. Commun. Image Represent. 40 (2016) 312–323, http://dx.doi.org/10.1016/j.jvcir.2016.06.029.

[4] A. Criminisi, P. Perez, K. Toyama, Region filling and object removal by exemplar-based image inpainting, IEEE Trans. Image Process. 13 (9) (2004) 1200–1212, http://dx.doi.org/10.1109/TIP.2004.833105.

[5] M.A. Qureshi, M. Deriche, A bibliography of pixel-based blind image forgery detection techniques, Sig. Process.: Image Commun. 39 (2015) 46–74, http://dx.doi.org/10.1016/j.image.2015.08.008.

[6] Inpaint Photo Restoration Software <http://www.theinpaint.com/>.

[7] Z. Liang, G. Yang, X. Ding, L. Li, An efficient forgery detection algorithm for object removal by exemplar-based image inpainting, J. Vis. Commun. Image Represent. 30 (2015) 75–85, http://dx.doi.org/10.1016/j.jvcir.2015.03.004.

[8] D.T. Trung, A. Beghdadi, M.-C. Larabi, Blind inpainting forgery detection, in: 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), IEEE, Atlanta, GA, USA, 2014, pp. 1019–1023, http://dx.doi.org/10.1109/GlobalSIP.2014.7032275.

[9] A. Beghdadi, M.C. Larabi, A. Bouzerdoum, K.M. Iftekharuddin, A survey of perceptual image processing methods, Sig. Process.: Image Commun. 28 (8) (2013) 811–831, http://dx.doi.org/10.1016/j.image.2013.06.003.

[10] M. Qureshi, A. Beghdadi, B. Sdiri, M. Deriche, F.A. Cheikh, A comprehensive performance evaluation of objective quality metrics for contrast enhancement techniques, in: European Workshop on Visual Information Processing (EUVIP), Marseille, France, 2016, pp. 1–5. http://dx.doi.org/10.1109/EUVIP.2016.7764589.

[11] O. Le Meur, C. Guillemot, Super-resolution-based inpainting, in: European Conference on Computer Vision (ECCV), Florence, Italy, 2012, pp. 554–567. doi: http://dx.doi.org/10.1007/978-3-642-33783-3_40.

[12] A. Telea, An image inpainting technique based on the fast marching method, J. Graph. Tools 9 (1) (2004) 23–34, http://dx.doi.org/10.1080/10867651.2004.10487596.

[13] O. Le Meur, J. Gautier, C. Guillemot, Examplar-based inpainting based on local geometry, in: 2011 18th IEEE International Conference on Image Processing (ICIP), IEEE, 2011, pp. 3401–3404.

[14] P. Arias, V. Caselles, G. Facciolo, Analysis of a variational framework for exemplar-based image inpainting, Multisc. Model. Simul. 10 (2) (2012) 473–514.

[15] M.A. Qureshi, M. Deriche, A new wavelet based efficient image compression algorithm using compressive sensing, Multim. Tools Appl. 75 (12) (2016) 6737–6754.

[16] R.H. Chan, S. Setzer, G. Steidl, Inpainting by flexible haar-wavelet shrinkage, SIAM J. Imag. Sci. 1 (3) (2008) 273–293.

[17] O.G. Guleryuz, Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part i: theory, IEEE Trans. Image Process. 15 (3) (2006) 539–554.

[18] N. Komodakis, G. Tziritas, Image completion using efficient belief propagation via priority scheduling and dynamic pruning, IEEE Trans. Image Process. 16 (11) (2007) 2649–2661.

[19] F. Cao, Y. Gousseau, S. Masnou, P. Pérez, Geometrically guided exemplar-based inpainting, SIAM J. Imag. Sci. 4 (4) (2011) 1143–1179.

[20] M. Bertalmio, L. Vese, G. Sapiro, S. Osher, Simultaneous structure and texture image inpainting, IEEE Trans. Image Process. 12 (8) (2003) 882–889.

[21] R. Rodriguez-Sánchez, J.A. García, J. Fdez-Valdivia, Image inpainting with nonsubsampled contourlet transform, Patt. Recog. Lett. 34 (13) (2013) 1508–1518, http://dx.doi.org/10.1016/j.patrec.2013.06.002.

[22] I. Peterson, Filling in blanks, Sci. News 161 (19) (2002) 299–300, http://dx.doi.org/10.2307/4013521.

[23] M.M. Oliveira, B. Bowen, R. McKenna, Y.-S. Chang, Fast digital image inpainting, in: Proceedings of the International Conference on Visualization, Imaging and Image Processing (VIIP), Marbella, Spain, 2001, pp. 261–266.

[24] Zhou Wang, Bovik, Ligang Lu, Why is image quality assessment so difficult?, IIEEE International Conference on Acoustics Speech and Signal Processing, vol IV, IEEE, 2002, pp. 3313–3316, http://dx.doi.org/10.1109/ICASSP.2002.1004620.

[25] J. Hays, A.A. Efros, Scene completion using millions of photographs, Commun. ACM 51 (10) (2008) 87, http://dx.doi.org/10.1145/1400181.1400202.

[26] S. Wang, H. Li, X. Zhu, P. Li, An evaluation index based on parameter weight for image inpainting quality, in: 2008 The 9th International Conference for Young Computer Scientists, IEEE, Hunan, 2008, pp. 786–790, http://dx.doi.org/10.1109/ICYCS.2008.461.

[27] P.A. Ardis, A. Singhal, Visual salience metrics for image inpainting, in: M. Rabbani, R.L. Stevenson (Eds.), Proceedings of SPIE 7257, Visual Communications and Image Processing, San Jose, CA, 2009, p. 72571W, doi: http://dx.doi.org/10.1117/12.808942.

[28] T.T. Dang, A. Beghdadi, M.-C. Larabi, Visual coherence metric for evaluation of color image restoration, in: 2013 Colour and Visual Computing Symposium (CVCS), IEEE, Gjovik, Norway, 2013, pp. 1–6, http://dx.doi.org/10.1109/CVCS.2013.6626268.

[29] V.V. Mahalingam, S.-c. S. Cheung, Eye tracking based perceptual image inpainting quality analysis, in: 2010 IEEE International Conference on Image Processing, IEEE, Hong Kong, 2010, pp. 1109–1112, http://dx.doi.org/10.1109/ICIP.2010.5653640.

[30] A.I. Oncu, F. Deger, J.Y. Hardeberg, Evaluation of digital inpainting quality in the context of artwork restoration, in: Proceeding of the 12th International Conference on Computer Vision, 2012, pp. 561–570, doi:http://dx.doi.org/10.1007/978-3-642-33863-2_58.

[31] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612, http://dx.doi.org/10.1109/TIP.2003.819861.

[32] T.-J. Liu, Y.-C. Lin, W. Lin, C.-C.J. Kuo, Visual quality assessment: recent developments, coding applications and future trends, APSIPA Trans. Sig. Inf. Process. 2 (e4) (2013) 1–20, http://dx.doi.org/10.1017/ATSIP.2013.5.

[33] X. Ma, X. Xie, K.-m. Lam, J. Hu, Y. Zhong, Saliency detection based on singular value decomposition, J. Vis. Commun. Image Represent. 32 (2015) 95–106, http://dx.doi.org/10.1016/j.jvcir.2015.08.003.

[34] X. Ma, X. Xie, K.-M. Lam, Y. Zhong, Efficient saliency analysis based on wavelet transform and entropy theory, J. Vis. Commun. Image Represent. 30 (2015) 201–207, http://dx.doi.org/10.1016/j.jvcir.2015.04.008.

[35] Q. Zhao, C. Koch, Learning saliency-based visual attention: a review, Sig. Process. 93 (6) (2013) 1401–1407.

[36] P.A. Ardis, C.M. Brown, A. Singhal, Inpainting quality assessment, J. Electron. Imag. 19 (1) (2010) 011002, http://dx.doi.org/10.1117/1.3267088.

[37] N. Bonnier, F. Schmitt, H. Brettel, S. Berche, Evaluation of spatial gamut mapping algorithms, in: 14th Color and Imaging Conferenc (CIC), 2006, pp. 56–61, doi:10.1.1.497.9528.

[38] D.T. Trung, A. Beghdadi, C. Larabi, Perceptual quality assessment for color image inpainting, in: 2013 IEEE International Conference on Image Processing (ICIP), IEEE, Melbourne, VIC, 2013, pp. 398–402, http://dx.doi.org/10.1109/ICIP.2013.6738082.

[39] D.T. Trung, A. Beghdadi, M.-c. Larabi, Perceptual evaluation of digital image completion quality, in: 21st European Signal Processing Conference (EUSIPCO 2013), Marrakech, 2013, pp. 1–5.

[40] T.T. Dang, A. Beghdadi, M.C. Larabi, U. Paris, U.D. Poitiers, Inpainted image quality assessment, in: 2013 4th European Workshop on Visual Information Processing (EUVIP), Paris, France, 2013, pp. 76–81.

[41] R. Achanta, S. Hemami, F. Estrada, S. Susstrunk, Frequency-tuned salient region detection, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Miami, USA, 2009, pp. 1597–1604, http://dx.doi.org/10.1109/CVPR.2009.5206596.

[42] M.A. Qureshi, M. Deriche, A fast no reference image quality assessment using laws texture moments, in: 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), IEEE, Atlanta, GA, USA, 2014, pp. 979–983, http://dx.doi.org/10.1109/GlobalSIP.2014.7032267.

[43] V. Viacheslav, F. Vladimir, M. Vladimir, G. Nikolay, S. Roman, F. Valentin, Low-level features for inpainting quality assessment, 2014 12th International Conference on Signal Processing (ICSP), vol. 53, IEEE, Hangzhou, 2014, pp. 643–647, http://dx.doi.org/10.1109/ICOSP.2014.7015082.

[44] V. Voronin, V. Marchuk, E. Semenishchev, S. Maslennikov, I. Svirin, Inpainted image quality assessment based on machine learning, in: WSCG 2015 Conference on Computer Graphics, Visualization and Computer Vision, 2015, pp. 167–172.

[45] M. Isogawa, D. Mikami, K. Takahashi, A. Kojima, Eye gaze analysis and learning-to-rank to obtain the most preferred result in image inpainting, in: 2016 IEEE International Conference on Image Processing (ICIP), IEEE, Phoenix, Arizona, USA, 2016, pp. 3538–3542, http://dx.doi.org/10.1109/ICIP.2016.7533018.

[46] V.A. Frantc, V.V. Voronin, V.I. Marchuk, A.I. Sherstobitov, S. Agaian, K. Egiazarian, Machine learning approach for objective inpainting quality assessment, in: S.S. Agaian, S.A. Jassim, E.Y. Du (Eds.), Proceedings of the. SPIE 9120, Mobile Multimedia/Image Processing, Security, and Applications, 2014, p. 91200S, doi:http://dx.doi.org/10.1117/12.2063664.

[47] P. Tiefenbacher, V. Bogischef, D. Merget, G. Rigoll, Subjective and objective evaluation of image inpainting quality, in: IEEE International Conference on Image Processing (ICIP), IEEE, Quebec City, Canada, 2015, pp. 447–451.

[48] A. Bugeau, M. Bertalmiío, V. Caselles, G. Sapiro, A comprehensive framework for image inpainting, IEEE Trans. Image Process. 19 (10) (2010) 2634–2645, http://dx.doi.org/10.1109/TIP.2010.2049240.

[49] Z. Xu, J. Sun, Image inpainting by patch propagation using patch sparsity, IEEE Trans. Image Process. 19 (5) (2010) 1153–1165.
[50] P. Getreuer, Total variation inpainting using split bregman, Image Process. On Line 2 (2012) 147–157.
[51] J. Herling, W. Broll, Pixmix: a real-time approach to high-quality diminished reality, in: International Symposium on Mixed and Augmented Reality, IEEE, 2012, pp. 141–150.
[52] S. Siegel, Nonparametric Statistics for The Behavioral Sciences, second ed., McGraw-Hill, 1956.
[53] M. Rubinstein, D. Gutierrez, O. Sorkine, A. Shamir, A comparative study of image retargeting, ACM Trans. Graph. 29 (6) (2010) 1, http://dx.doi.org/10.1145/1882261.1866186.
[54] Q. Zhang, J. Lin, Exemplar-based image inpainting using color distribution analysis, J. Inf. Sci. Eng. 28 (4) (2012) 641–654.
[55] T.-T. Dang, C. Larabi, A. Beghdadi, Multi-resolution patch and window-based priority for digital image inpainting problem, in: 3rd International Conference on Image Processing Theory, Tools and Applications (IPTA), IEEE, 2012, pp. 280–284.
[56] V.V. Voronin, V.A. Frantc, V.I. Marchuk, A.I. Sherstobitov, K. Egiazarian, Inpainted image quality assessment based on machine learning (report), in: K.O. Egiazarian, S.S. Agaian, A.P. Gotchev (Eds.), Proceedings of the SPIE, 2015, p. 93990U, doi:http://dx.doi.org/10.1117/12.2076507.
[57] J.Y. Hardeberg, E. Bando, M. Pedersen, Evaluating colour image difference metrics for gamut-mapped images, Color. Technol. 124 (4) (2008) 243–253.