

PERCEPTUAL QUALITY ASSESSMENT FOR COLOR IMAGE INPAINTING

A. DANG Thanh Trung, B. Azeddine BEGHDAI, C. Chaker LARABI

L2TI, Institut Galilée, Université Paris 13, France A-B
XLIM, Département SIC, Université de Poitiers, France C

Email: {dang.thanhtrung, azeddine.beghdadi}@univ-paris13.fr, chaker.larabi@univ-poitiers.fr

ABSTRACT

A novel objective measure for assessing the quality of image inpainting is proposed. In contrast to standard image quality metrics, the proposed one takes into account some constraints and characteristics related to the specific goals of inpainting techniques. The idea is to combine spatial low-level features and perceptual criteria in the design of the objective Image Inpainting Quality Metric (IIQM). The used characteristics are the visual coherence of the recovered regions and the visual saliency describing the visual importance of an area. Experimental results demonstrate the good performance of the proposed IIQM and its well adaptation to the evaluation of image inpainting results.

Index Terms— Image inpainting, image completion, image quality assessment, inpainting quality assessment.

1. INTRODUCTION

Image inpainting becomes a very active field of research for many real world applications such as digital cinema, computational photography, archaeological image restoration and archival documents restoration [1]. Most inpainting algorithms could be classified into two broad categories depending upon their scope and goals. First, geometry-oriented approaches are mainly designed for filling narrow or small holes because they are less suitable for synthesizing semantic textures or structures [2–4]. Second, exemplar-oriented methods could deal with large holes and they could be further subdivided into two subgroups: greedy strategy [5–7] and global optimization strategy [8–10]. However, when it comes to the evaluation of inpainting results, visual inspection is often used because there is almost no dedicated objective assessment tool.

Image quality assessment (IQA) plays a prominent role for many applications, including video streaming monitoring, medical imaging and lossy compression control among others. IQA in its broad sense refers to the problem of evaluating the level of perceptual quality of an image. Many interesting methods for predicting image quality have been proposed in the literature [11–15]. After all, the subjective evaluation is still the most reliable approach. The latter is not sound for real-time applications because it is tedious and time consuming. It is worth noticing that the goals of IQA in the case of image inpainting are sub-

stantially different from the classical image quality evaluation. Indeed, in the case of inpainting, the intent is to evaluate the quality of the restored image. It is also worth understanding that inpainting could be considered as a specific image restoration. In both problems, inpainting and classical image restoration, the existing IQA metrics could not be directly applied because of the specificity of the aforementioned applications. For example, in the case of image inpainting, the recovered region is totally different from the original one (which is in the major cases unknown because of degradations or occlusions). This operation aims at restoring the missing parts or replacing some parts of the image in a visually plausible way. The intent of Image Inpainting Quality Assessment (IIQA) metric is then to evaluate the visual quality of the inpainted regions in terms of spatial coherence with the existing parts of the image.

Results of image inpainting are very often evaluated subjectively or by using some objective metrics not well adapted to the specificities of their criteria. However, subjective evaluation experiments are known being time consuming, complex and sometimes unpredictable due to uncontrolled human factors such as fatigue, visual discomfort, knowledge, etc.

A few interesting works on objective IIQA have been published recently where the proposed approaches suffer from some shortcomings. For instance, an analysis of gaze patterns obtained thanks to eye-tracking experiments has been described in [16] for predicting inpainted image quality. This method has the same disadvantages as subjective evaluation. Authors of [17] proposed a full-reference quality metric for image inpainting based on a variant of *SSIM* composed of luminance, definition and gradient similarity. Actually, image inpainting is known as a blind image completion where the notion of "reference" is nonexistent. Moreover, this index cannot be applied when inpainted areas are large because original and inpainted images may highly differ from each other. Two other quality indexes are introduced in [18] using saliency maps. The key idea is based on the variation of the saliency map before and after inpainting, expressed by two metrics: average squared visual salience (*ASVS*) and degree of noticeability (*DN*). However, these indexes do not take into consideration the global visual appearance of the image that significantly affects to the restoration quality.

The most contribution of this paper is to propose a global image inpainting quality index. The proposed metric is constructed

by using a coherence map, which refers to the global visual coherence, and a saliency map reflecting the local structure continuity. The combination of these criteria allows achieving results outperforming the state-of-the-art. Subjective experiments have been conducted for the sake of studying the performance of the metric with regards to human judgment. The proposed IQA index can be applied to most of the inpainting image approaches such as geometry-oriented methods, texture-oriented methods and hybrid methods.

The remainder of this paper is structured as follows: the proposed IQA index is described in details in section 2. Section 3 is devoted to experimental results and comparison with existing approaches. Finally, the paper ends with some conclusions and gives ideas about future works.

2. INPAINTED IMAGE QUALITY ASSESSMENT

Through the analysis of image inpainting results, we found that two main aspects need to be considered in image inpainting quality assessment. First, the completion is blindly performed without any cue from the original content of images, that is, the restored regions depend on only the rest of the image. Therefore, the new generated pixels should be consistent with the existing ones. This refers to coherence of inpainted regions that determines the new undesired visual artifacts. Second, the human gaze is more attracted by structures or contours than other parts; hence more importance should be associated to these salient features. In our work, a saliency map, representing an approximation of the human attention, is used as the weighting map for the evaluation. As a result, an efficient quality index for IQA taking into account the mentioned factors is proposed.

In the following, some notations that are similar to those in [7] are adopted. The whole image domain, I , is composed of two disjoint regions: the inpainting region, Ω , and the known region, Φ ($\Phi = I - \Omega$). The inpainted image quality index, Q , is defined as expressed by Eq. 1.

$$Q = \frac{\sum_{\Omega} C(p)^\alpha S(p)^\beta}{\|\Omega\|} \quad (1)$$

where $C(p)$ and $S(p)$ are respectively the coherence and structure terms defined below. Two positive parameters α and β are associated with the aforementioned terms in order to be able to adjust their influence on the quality index (in our implementation, we set $\alpha = \beta = 1$).

2.1. Coherence term

On the one hand, an inpainted region Ω has a global visual coherence with the rest of the image, Φ , if every new generated pixel is consistent with existing ones in terms of a structural information. On the other hand, a local patch Ψ_p , corresponding to a window centered on p , should be similar to the one within Φ . The coherence term for each pixel $p(x, y)$ ($p \in \Omega$) is then

defined as follows:

$$C(p) = \max\{SIM(\Psi_p, \Psi_q), \forall \Psi_q \in \Phi\} \quad (2)$$

where Ψ_p and Ψ_q denote small patches around p and q , respectively. The patch size is a global parameter. SIM is an objective function to evaluate the similarity between two patches that measures an appropriateness degree of an inpainted pixel (p) based on both pixels in Ψ_q and its neighbors in Ψ_p . A new similarity function based on the structural information of patches was proposed in [11] but it works with luminance only. To compute the coherence between patches, we exploit the idea developed in [12] where a similarity measure combining structure and color information is introduced. The similarity function is then defined as follows:

$$SIM(\Psi_p, \Psi_q) = (1 - \theta)SS(\Psi_p, \Psi_q) + \theta HS(\Psi_p, \Psi_q) \quad (3)$$

where, θ is a positive constant within the range $[0, 1]$ defining the relative importance between structure similarity (SS) and hue similarity (HS), corresponding to color information. The structure and hue similarity indexes are defined by Eqs. 4 and 5, respectively:

$$SS(\Psi_p, \Psi_q) = \frac{(\sigma_{pq} + C_1)}{(\sigma_p + \sigma_q + C_1)} \quad (4)$$

$$HS(\Psi_p, \Psi_q) = \frac{(2\lambda_p\lambda_q + C_2)}{(\lambda_p^2 + \lambda_q^2 + C_2)} \quad (5)$$

where (λ_p, σ_p) and (λ_q, σ_q) denote the mean hue and standard deviation set of patches Ψ_p and Ψ_q , respectively, while σ_{pq} denotes their cross correlation. C_1 and C_2 are small positive constant values to avoid computational instability in very dark or homogeneous regions. In our experiment, constants are set as in [11, 12], i.e, $\theta = 0.1667$; $C_1 = C_2 = 6.5025$.

2.2. Structure term

The second term affecting image inpainting quality is the structure factor. Given an image, human observer would pay more attention to perceptually relevant regions, which usually correspond to contours and details, but less attention to the rest of the image. Thus, the contours and other relevant structures in the inpainted regions attract more human gaze than the other components. For that reason, we may identify the structure term using the information provided by a saliency map as follows:

$$S(p) = \frac{SM(p)}{\max_I\{SM\}} \quad \forall p \in \Phi \quad (6)$$

where SM is the saliency map of the inpainted image. Several computational models have been proposed to simulate human's visual attention [19–21]. However, the high computational cost and variable parameters are still the weaknesses of these models. Authors of [22] proposed a simple and efficient method based on the idea that objects attracting the gaze of an

observer should have characteristics that go beyond the average behavior of the image. A simple formulation of the aforementioned saliency map, SM , can be expressed by Eq. 7:

$$SM = ||I_\mu - I_G|| \quad (7)$$

where I_μ and I_G are the arithmetic mean pixel value and the Gaussian blurred version of the inpainted image, respectively. The operation is performed in the $CIEL^*a^*b^*$ color space. Fig. 1 illustrates an example of coherence map and structure map by pseudo-colored mask images where red refers to higher value and blue to lower value.

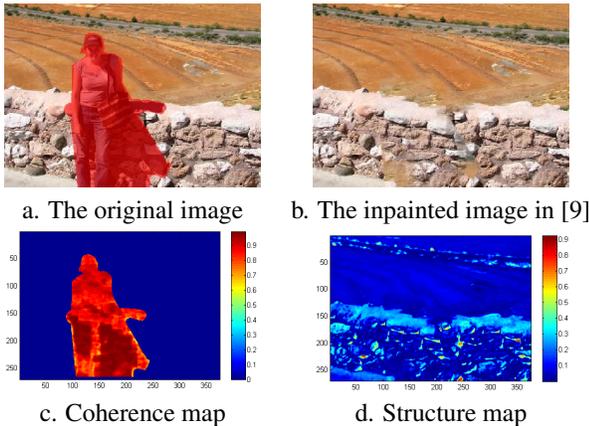


Fig. 1. Illustration of the structure and coherence terms.

3. EXPERIMENTAL RESULTS

In order to evaluate the consistency/performance of the proposed metric, a subjective experiment has been conducted. This experimental study is mandatory for constructing a ground truth for inpainted images for which the literature is almost nonexistent. Therefore, a panel of observers has been invited to assess the quality of several inpainted images. The used test-set of images has been divided into two subsets: images containing small and large inpainted regions. Since the restoration for small areas aims to make it appear as close to the original image as possible, one full-reference quality metric is used for comparison with our index. However, this is not suitable to the case of large holes. In this situation, collected subjective scores from human observers MOS (Mean Opinion Score) are considered as the reference index. Furthermore, the prediction accuracy of all metrics was evaluated using Spearman rank order correlation coefficient (SRCC), r_S , in comparison with the reference indexes ($MSSIM$ and MOS).

3.1. Qualitative scoring of inpainted images

In this section, we describe the experimental setup for the subjective evaluation conducted with the aim of validating the developed metric for inpainting techniques. A set of six input images, commonly used in the inpainting community (typically 300 × 200 or similar size), was used as input of for two subgroups of

inpainting methods: greedy and global optimization strategies. Each image from the set is restored by at least three different inpainting methods producing thus eighteen output results. The subjective test consists in scoring the inpainted images in comparison to the original one (input image). The inpainted images were randomly presented and shown without including the name of inpainting methods to avoid any bias or influence.

Observers participating to the test have a normal vision (good acuity and no color blindness). They were asked to provide their judgment of inpainting quality on a discrete scoring scale of adjectives: *Unacceptable*, *Poor*, *Acceptable*, *Good* and *Perfect*. Each test was viewed by 15 ~ 20 subjects and took around ten seconds per image. Fig. 2 illustrates the opinion of the fifteen observers about resulting images shown in Fig. 5. For the following steps, the mean opinion score (MOS) is computed and normalized in range [0, 1] for each image/method in order to be used in the performance evaluation step.

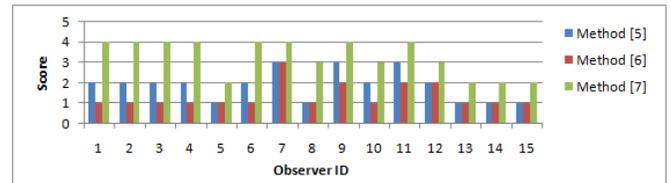


Fig. 2. Subjective scores for resulting images given on Fig. 5.

3.2. Case 1: Small inpainted regions

In order to restore small or thin defects as those shown on Fig. 3b, we selected three approaches described in [3], [4] and [7]. The results of these approaches are shown in Fig. 3. The pseudo-inpainted regions occupy 3.72% of the total image. As mentioned above, since the missing area is relatively small, a well-known full-reference quality measure, $MSSIM$, is used as the reference to evaluate the performance of the proposed index in comparison with the others. Fig. 4 shows $MSSIM$ scores versus inpainted image quality indexes such as $PWIIQ$ [17], $ASVS$ and DN [18] and gives the resulting the Spearman's rho, r_S . As it can be noticed, our index is the most consistent with $MSSIM$ values and produces the highest value for SRCC ($r_S = 1$).

3.3. Case 2: Large inpainted regions

The exemplar-based methods achieve impressive results in recovering large damaged regions and are the most commonly used. In our experiments, we selected methods described in [5], [6] and [7] for greedy strategies and [8], [9] and [10] for global optimization strategies. Fig. 5 shows some results of these methods where the masked image is displayed in Fig. 5a. The remaining images are the outputs of [5], [6] and [7], respectively. Fig. 6 illustrates the results obtained with methods described in [8], [9] and [10]. The inpainted regions occupy 14.87% (Fig. 5) and 21.60% (Fig. 6) of the total image. Since the inpainting regions are large and the main objective of these methods is to restore the damaged images in a visually plausible way without referring to

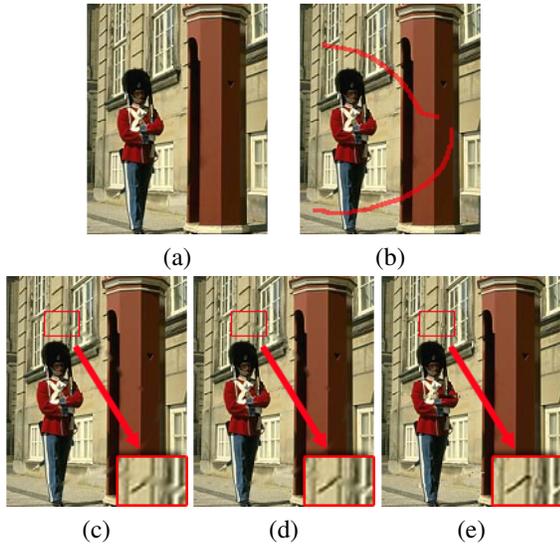


Fig. 3. The inpainting results in case of small inpainted regions. (a) The original image; (b) The image to be inpainted; Outputs when using the methods in (c) [3]; (d) [4]; and (e) [7].

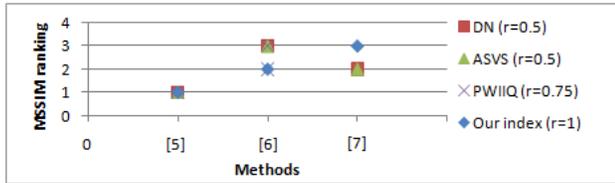


Fig. 4. *MSSIM* ranking versus quality indexes scores.

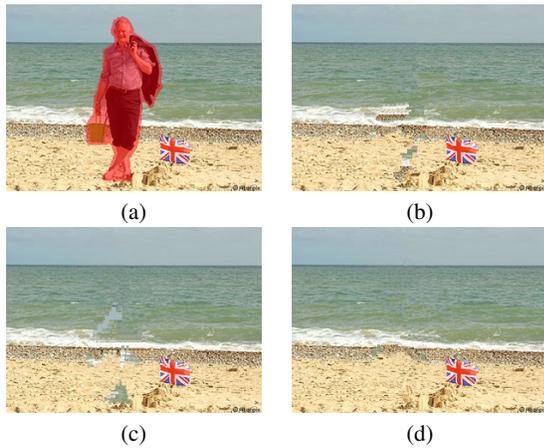


Fig. 5. The inpainting results for greedy strategies. (a) The image to be inpainted; The outputs when using the methods in (b) [5]; (c) [6]; and (d) [7].

the original image, the outputs are subjectively evaluated producing the *MOS* instead of *MSSIM*. From Table 1 and Fig. 7, it is obvious to notice that our index perfectly matches the subjective score (*MOS*) in ranking the quality of output. The SRCC for our index obtains the highest value.

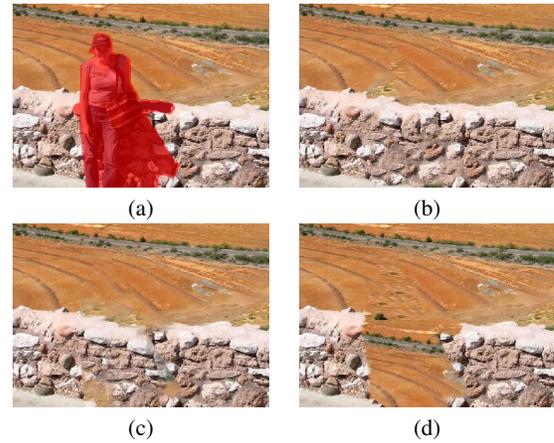


Fig. 6. The inpainting results for global optimization strategies. (a) The image to be inpainted; The outputs when using the methods in (b) [8]; (c) [9]; and (d) [10].

Table 1. The indexes of IIQA.

Strategy	Greedy			Global optimization		
Results in	[5]	[6]	[7]	[8]	[9]	[10]
<i>MOS</i>	0.40	0.34	0.68	0.84	0.56	0.36
<i>ASVS</i>	0.093	0.161	0.114	0.180	0.126	0.101
<i>DN</i>	0.027	0.043	0.034	0.022	0.015	0.012
<i>Ourindex</i>	0.061	0.058	0.063	0.081	0.077	0.056

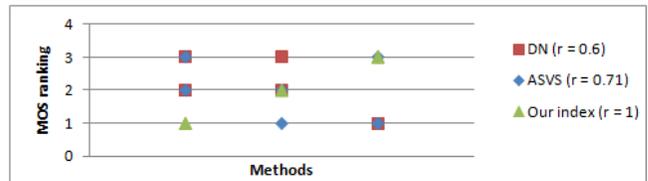


Fig. 7. *MOS* ranking versus quality indexes scores.

4. CONCLUSION

In this paper, a novel approach for inpainted image quality evaluation has been proposed. It is shown that the traditional image quality index could not be used for evaluating the inpainting results. By taking into account the specificities and goals of image completion problem and some characteristics of the humane visual system, such as visual saliency, an efficient measure has been derived. The proposed image inpainting quality index not only correlates with subjective evaluation but also could be applied to most of image inpainting approaches. The experimental results and comparison with all three approaches for image inpainting methods confirm the efficiency of the proposed index.

For future works, a perceptual patch similarity that is more robust and efficient for color images are being further studied and the objective evaluations metrics of video or context-based images will be developed.

5. REFERENCES

- [1] P. Arias, G. Facciolo, V. Caselles, and G. Sapiro, "A variational framework for exemplar-based image inpainting," *Journal International Journal of Computer Vision*, vol. 93, pp. 319–347, 2011.
- [2] J. Shen and T.F. Chan, "Mathematical models for local nontexture inpainting," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1019–1043, 2002.
- [3] D. Tschumperle, "Fast anisotropic smoothing of multi-valued images using curvature-preserving pdes," *International Journal of Computer Vision*, vol. 68, pp. 6582, 2006.
- [4] A. Bugeau and M. Bertalmio, "Combining texture synthesis and diffusion for image inpainting," *Proceeding International Conference Computer Vision Theory and Applications*, p. 2633, 2009.
- [5] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transaction of Image Processing*, vol. 13 (9), pp. 1200–1212, Apr. 2004.
- [6] Q. Zhang and J. Lin, "Exemplar-based image inpainting using color distribution analysis," *Journal of Information Science and Engineering*, vol. 28, pp. 641 – 652, 2012.
- [7] T. T. Dang, C. Larabi, and A. Beghdadi, "Multi-resolution patch and window-based priority for digital image inpainting problem," *3rd International Conference on Image Processing Theory, Tools and Applications*, 2012.
- [8] Y. Wexler, E. Shechtman, and M. Irani, "Space-time video completion," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 29, pp. 463 – 476, March 2007.
- [9] N. Komodakis and G. Tziritas, "Image completion using global optimization," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 442 – 452, 2006.
- [10] Y. Pritch, E. Kav-Venaki, and S. Peleg, "Shift-map image editing," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 151–158, 2009.
- [11] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13 (4), pp. 600–612, Apr. 2004.
- [12] Y. Shi, Y. Ding, R. Zhang, and J. Li, "Structure and hue similarity for color image quality assessment," *International Conference on Electronic Computer Technology*, pp. 329–333, 2009.
- [13] A. Beghdadi and R. Iordache, "Image quality assessment using the joint space/spatial-frequency space," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 40–47, 2006.
- [14] Z. Haddad, A. Beghdadi, A. Serir, and A. Mokraoui, "Image quality assessment based on wave atoms transform," *IEEE International Conference on Image Processing ICIP2010*, pp. 305–308, 2010.
- [15] A. Maalouf and M. C. Larabi, "A grouplet-based reduced reference image quality assessment," *Int. workshop on Quality of Multimedia Experience (QoMEX)*, pp. 59–63, 2009.
- [16] M. V. Venkatesh and S. S. Cheung, "Eye tracking based perceptual image inpainting quality analysis," *17th IEEE International Conference on Image Processing (ICIP)*, pp. 1109–1112, 26-29 September 2010.
- [17] S. Wang, H. Li, X. Zhu, and P. Li, "An evaluation index based on parameter weight for image inpainting quality," *the 9th International Conference for Young Computer Scientists*, vol. II, pp. 786–790, 18-21 Nov 2008.
- [18] P. Ardis and A. Singhal, "Visual salience metric for image inpainting," *Proceedings of the SPIE*, vol. 7257, 2009.
- [19] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20 (11), pp. 1254–1259, 1998.
- [20] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 17–22, 2007.
- [21] C. Guo, Q. Ma, and L. Zhang, "Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 116–123, 2008.
- [22] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 1597 – 1604, 2009.