VISUAL COHERENCE FOR EVALUATION OF COLOR IMAGE RESTORATION

Thanh Trung DANG, Azeddine BEGHDADI, Chaker LARABI *

L2TI, Institute Galilee, Universite Paris 13, France

*XLIM, Department SIC, Universite de Poitiers, France

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

ABSTRACT

Digital image inpainting technique could provide substantial support for future restoration of images. However, much less effort has been devoted to the development of image inpainting evaluation. In this paper, we proposed a novel objective metric for assessing the quality of color image inpainting which takes into account some constraints and characteristics related to the specific objectives of inpainting approaches. The used characteristics are the visual coherence of the recovered regions and the visual saliency describing the visual importance of the area. A series of psychophysical experiments have been conducted to evaluate the performance of the proposed image quality index. The results show that the proposed image quality metric is well adapted for image quality inpainting evaluation.

I - INTRODUCTION

The goal of digital image inpainting is to fill the lost content automatically according to current image information in a visually plausible way, i.e., the restored parts are not easily detectable by viewers and not annoying. Digital inpainting has found widespread use in many applications such as image editing (removing undesired objects, restoring the scratches), film reproduction (deleting the logos, subtitles, etc), or even creating artistic effects (reorganizing objects, smart resizing of images, blending images), etc. In the field of archaeology, it can be employed for digital restoration of damaged objects such as statues, buildings, walls, etc.

Several image inpainting approaches have been proposed in the literature. These could be categorized into two types based on the objective [1]. The first category consists of geometry-oriented approaches [2, 3, 4]. These methods are mainly designed for filling narrow or small holes but work less well for large missing regions, their drawback consists of introducing blur artifacts that become more visible when inpainting larger holes. Exemplar-oriented algorithms could overcome this drawback by reconstructing large image regions from sample textures [5, 6, 7, 8, 9] .

However, there is an acknowledged lack of quantitative metrics for image inpainting quality evaluation. At our best knowledge the inpainted images are very often evaluated subjectively or by using some objective metrics not well adapted to the specificities of image inpainting criteria. In fact, the best way to evaluate the output quality of image inpainting is to use some perform subjective scoring systems. But, subjective evaluation experiments are time consuming, complex and sometimes unpredictable due to some uncontrolled human factors such as fatigue, visual discomfort and other high level perceptual vision factors. The traditional image quality assessment (IQA) metrics [10, 11, 12] also could not be directly applied for Inpainted Image Quality Assessment (IIQA). Indeed, the specificities and goals of both image quality, in its broad sense, and image inpainting are quite different. For instance, in the case of image inpainting the recovered regions are totally different from the original ones and a reference image might not always be available for comparison. Furthermore, the evaluation is performed not only on the basis of the recovered region but also on the visual coherence with the other surrounding parts in the image.

The purpose of our work is to introduce an objective IIQA that could predict the perceptual quality of the recovered image. The proposed metric is computed by using a coherence map, which refers to the global term, and a saliency map which reflects the local structure continuity. The coherence term, related to the undesired visual artifacts, is computed by the correlation between the inpainted pixels and existing pixels. The structure continuity related to human attention is computed and normalized by the saliency map. By analyzing the experimental results and the comparison with other approaches, our approach provides an impressive objective quality index for image inpainting quality assessment. Our index not only correlates with subjective judgment of observers but also 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: section 2 introduces some the related works. A definition of our index is given in section 3. Section 4 describes some experimental results and comparison with existing approaches. Finally, the paper ends with some conclusions and future works.

II - RELATED WORKS

Inpainting quality evaluation is a challenging issue that has been relatively less investigated compared to classical Image Quality Assessement (IQA) in its broad sense. Currently, a few papers of IIQA have been published but they have some shortcomings. In order to discuss the related works for objective quality evaluation, some notations and conventions are defined in the following. The whole image domain, I, is composed of two disjoint regions: the inpainting region, Ω, and the known region, Φ (Φ=I-Ω). Furthermore, the basic unit of synthesis at pixel p is a patch or window, Ψ_p, centered at pixel p.

In [13], an analysis of gaze pattern ~~which~~ was recorded by using the Seeing Machines face-LAB eye tracker ~~is performed~~ to assess the output quality. This method combines a computational model with collected gaze data for predicting inpainted image quality. But, it suffers the same disadvantages as the subjective evaluation methods.

Wang et al. [14] proposed a full-reference assessment using modified SSIM index composed of three aspects: luminance, definition and gradient similarity to evaluate the blur artifacts produced during inpainting. This approach is suitable only for the first category of inpainting methods when the reference image is available.

In [15] P. Ardis et al. define two other metrics, the *Average Squared Visual Salience* (*ASVS*) and the *Degree of Noticeability* (*DN*), which represent~~s~~ ~~for~~ two types of observable artifacts in an inpainted image, referred to as *in-region* and *out-region*. The former accounts for artifacts inside the hole, Ω, while the out-region artifacts correspond to the outside-hole regions, Ψ. These indexes do not need reference image. Higher scores of *ASVS* and *DN* can be interpreted as an indicator of highly visible artifacts and thus correspond to poor inpainting performance.

Mahalingam [16] also proposed other two visual saliency-based metrics for quantifying inpainting quality, GD_{in} and GD_{out} which define the gaze density within and outside the hole in an inpainted image. The inpainting quality is evaluated based on the change in the saliency maps corresponding to the inpainted and original images.

Based on Mahalingam's framework, A. I. Oncu [17] introduced two improved metrics, the *Border Saliency* (*BorSal*), which evaluates the change of flow of attention at only the Border region that extends three pixels inside and outside the hole, and the *Structural Border Saliency* (*StructBorSal*) which combines the *BorSal* metric with a structure measure, $SSIM\_{IPT}$ to account the structure restoration.

However, these metrics are rarely applied ~~by researches~~ to assess new inpainting techniques because either they do not consider the constraints of inpainting problem, for example there is no original image, or they do not take into account the global term which measures the visual coherence of the inpainted regions with the rest region in image.

III - INPAINTED IMAGE QUALITY INDEX

Currently, most algorithms evaluate the quality of inpainted images based on the artifacts appearance in the holes of pre- and post-inpainting images. ~~Namely,~~ A analysis of saliency map is implemented in order to estimate the change of gaze density within and outside the holes and their results are employed to develop the quality metrics. The saliency map plays an important role in assessing the inpainting quality but it is insufficient. From the observation of many image inpainting results, we found that the compatibility between the reconstructed and the rest regions is an important factor affecting to the inpainting quality. The reconstructed region should not generate new visual artifacts which do not exist in the rest part of image.

Based on these observations, we propose a novel metric for inpainted image quality combining two terms: coherence term which determines the new undesired visual artifacts and a structure term which reflects the local structure continuity in the holes. The inpainted image quality index, *Q*, is defined through equation (1) and it is shown that it is possible to predict reasonably the inpainting image quality in many cases.

$$Q = \frac{\sum_{p \in \Omega} C(p) S(p)}{||\Omega||} \qquad (1)$$

3.1. Coherence term

An inpainted region, Ω, corresponds to a global visual coherence with the rest of the image, Φ, if every new generated pixel is consistent with existing pixels. On the other hand, the local patch, Ψ_p, should be similar to another one, Ψ_q, within Φ. As a result, we define the coherence term for each pixel $p(x, y)$ ($p \in \Omega$) as follows:

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

where Ψ_p, Ψ_q denote small patches around *p* and *q*, respectively. The patches need not necessarily be isotropic and can have difference sizes in the spatial domain. *SIM* is an objective function to evaluate the similarity between two patches. This could be considered as a measure of the goodness degree of an inpainted pixel (*p*) based on existing pixels (Ψ_p) and its neighbors in Ψ_q. A good objective function needs to agree perceptually with a human observer. The MSE or PSNR are used widely for patch similarity but they are insufficient to provide the desired results. The main reason for this is that they do not take into account the human visual features.

In [18], a new similarity function based on the structural information of patches was proposed. This metric accounts only for luminance patches and does not take into account color information. It consists of three terms and is given below.

$$SSIM(\Psi_p, \Psi_q) = [l(p, q)].[c(p, q)].[s(p, q)] \qquad (3)$$

where $l(p, q)$, $c(p, q)$ and $s(p, q)$ denote the mean luminance, contrast and structure in patches Ψ_p and Ψ_q, respectively. To compute the coherence between patches, we exploit the idea developed in [19] where a similarity is adapted to color image by developing the similarity on each channel of the image in color space IPT. The similarity function is then defined as follows:

$$SIM(\Psi_p, \Psi_q) = SSIM_I(\Psi_p, \Psi_q).SSIM_P(\Psi_p, \Psi_q).SSIM_T(\Psi_p, \Psi_q) \quad (4)$$

3.2. Structure term

The second term affecting image inpainting quality is the structure factor. Because human observer would pay more attention to perceptually relevant regions, which usually correspond to contours and details [xxx], but less attention to the rest of the image, 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:

$$S(p) = \frac{SM(p)}{\max_I\{SM\}} \quad \forall p \in \Omega \qquad (5)$$

where *SM* is the saliency map of the inpainted image. Several computational models have been proposed to simulate human's visual attention [20, 21, 22]. However, the high computational cost and variable parameters are still the weaknesses of these models. Authors of [23] 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 equation 6:

$$SM = ||I_\mu - I_G|| \qquad (6)$$

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

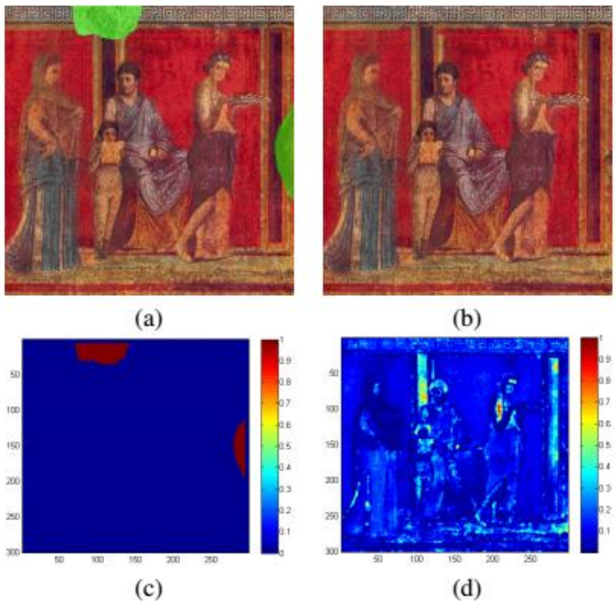

Figure 1. A local quality map. (a) The original image; (b) The inpainted image using method in [8]; (c) Coherence map; (d) Structure map

IV - EXPERIMENTAL RESULTS

In order to evaluate the consistency/performance of the proposed metric, a subjective study has been launched. This experimental study is mandatory for constructing a ground truth for inpainted images for which the literature is almost inexistent. 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 groups: images containing small inpainted regions and images containing large inpainted regions. Since the restoration for small holes aims to make it appear as close to the original image as possible, some full-reference quality metric is used for comparison with our index. However, this is not suitable to the case of large hole. In this situation, a subjective test database using human viewers was built, and MOS (Mean Opinion Score) model is considered as the reference index. Furthermore, the prediction accuracy of all metrics was evaluated using Spearman rank order correlation coefficient, *SCC*, and Pearson product-moment correlation coefficient, *PCC*, [24] in evaluating the performance of the metrics considered.

4.1. Experimental setup

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 eight input images were encoded into PNG format (typically 300 × 200 or similar size) in two cases: small and large inpainting regions. In the first case, we used three images and four different inpainiting methods corresponding to the algorithms in [2, 3, 4]. For the second case, six images are used for inpainting and each of them is restored by five different inpainting methods in [5, 6, 7, 8, 9]. The experiment was carried out as a web-based experiment. All images will be shown on the website for observers.

The subjective test consists in scoring the inpainted images in comparison to the original one. 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 5 ~ 10 subjects and takes about around twenty seconds per image. Figure 2 illustrates the opinion of the ten observers for results given in figure 3. The mean opinion score (*MOS*) is computed for each image/method in order to be used in the performance evaluation step.

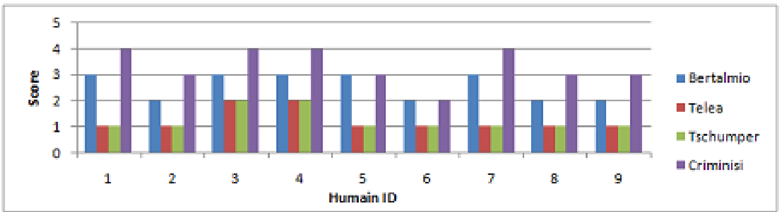

Fig. 2. Subjective score for results of Fig. 3

4.2. Case 1: Small inpainted regions

Figure 3 shows an example of output in [2, 3, 4, 5], respectively. The pseudo-inpainted regions occupy 4.52% of the total image. In order to evaluate the quality metrics, some full-reference is developed in comparison with subjective scores. The obtained results presented in table 1 indicate that our index is the most consistent with *MOS* values and produces the highest mean value for Spearman's rho (*SCC*=0.863) and Pearson product-moment (*PCC*=0.713)

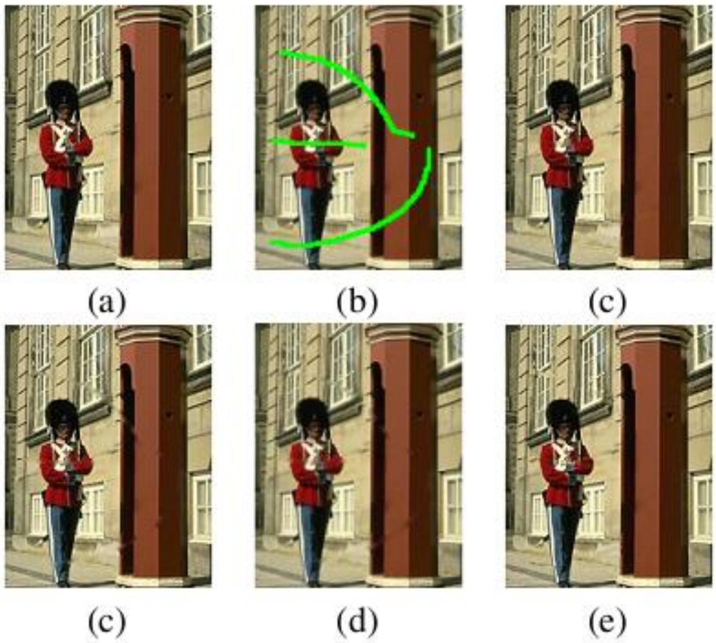

Fig. 3. The inpainting results. (a) The original image; (b) The masked image; (c) Result in [2]; (d) Result in [3]; (e) Result in [4]; (f) Result in [5]

Table 1. The mean scores of SCC and PCC.

	image 1		image 2		image 3	
	SCC	PCC	SCC	PCC	SCC	PCC
MSE	0.14	0.40	-0.13	0.40	-0.52	-0.40
$PSNR$	-0.09	-0.40	0.03	-0.40	0.47	0.40
$MSSIM$	0.24	0 .13	-0.5	-0.4	0.6	0.4
$PWIIQ$	0.85	0.94	0.6	0.2	0.28	0.2
$ASVS$	0.98	-0.94	-0.10	0.4	0.75	-1.0
DN	0.99	-0.94	-0.92	1.0	0.76	-0.94
GD_{in}	-0.22	-0.40	0.38	0.4	0.89	0.8
GD_{out}	0.09	0.4	0.74	1.0	-0.92	-0.72
(Q)	0.99	0.94	0.64	0.4	0.96	0.80

4.3. Case 2: Large inpainted regions

The exemplar based methods achieve impressive results in recovering the large damaged regions and are the most commonly used. In our experiments, we selected the methods in [5, 6, 7, 8, 9], respectively. Figure 4 shows some results of these methods. The original masked image is displayed in figure 4a. The inpainted regions occupy 3.81% of the total image.

The mean values of SCC and PCC in the Table 2 show that our metric has a higher correlation with human visual system than the other considered metrics. It shows that our metric is good enough in evaluation of inpainted quality.

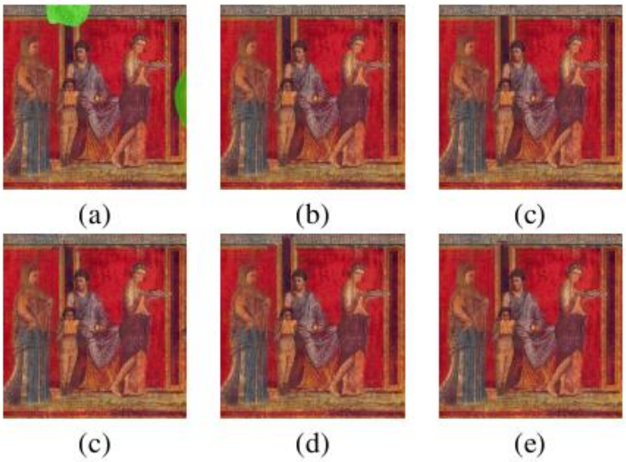

Fig. 4. The inpainting results. (a) The inpainting image; (b) Result in [5]; (c) Result in [6]; (d) Result in [7]; (e) Result in [8]; (f) Result in [9]

Table 2. The mean scores of SCC and PCC.

	image 4		image 5		image 6		image 7		image 8	
	SCC	PCC	SCC	PCC	SCC	PCC	SCC	PCC	SCC	PCC
$ASVS$	-0.50	0.71	-0.00014	0.6	-0.3	0.3	-0.67	0.77	-0.39	0.36
DN	-0.67	-0.88	-0.0004	0.6	-0.22	-0.4	-0.27	0.00	-0.46	0.36
Q	0.88	0.53	0.04	0.9	0.63	0.70	0.97	0.96	0	0

V - CONCLUSION

In this study, 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 objectives of image completion problem and some characteristics of the humane visual system, such as perceptual saliency, an efficient measure could be 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 performed experimental results and comparison with all two approaches for image inpainting methods confirm the efficiency of the proposed index..

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

VI - REFERENCES

[1] P. Arias, G. Facciolo, V. Caselles, and G. Sapiro, “A variational framework for exemplar-based image in-painting,” Journal International Journal of Computer Vision, vol. 93, pp. 319 – 347, 2011.

[2] M. Bertalmo, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting,” Proceedings of SIGGRAPH 2000, pp. 1019–1043, 2000.

[3] A. Telea, “An image inpainting technique based on the fast marching method,” J. Graphics Tools, vol. 9, pp. 23–24, 2004.

[4] D. Tschumperle, “Fast anisotropic smoothing of multi-valued images using curvature-preserving pdes,” Inter-national Journal of Computer Vision, vol. 68, pp. 6582, 2006.

[5] A. Criminisi, P. Perez, and K. Toyama, “Region filling and object removal by exemplar-based image inpaint-ing,” 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 Informa-tion Science and Engineering, 2011.

[7] T. T. Dang, C. Larabi, and A. Beghdadi, “Multi-resolution patch and window-based priority for digital image inpainting problem,” 3rd International Confer-ence on Image Processing Theory, Tools and Applica-tions, 2012.

[8] J. Wu and Q. Ruan, “Object removal by cross isophotes exemplar based image inpainting,” Proceeding of Inter-national Conference of Pattern Recognition, pp. 810 – 813, 2006.

[9] W. Cheng, C. Hsieh, S. Lin, C. Wang, and J. Wu, “Ro-bust algorithm for exemplar-based image inpainting,” Proceeding of International Conference on Computer Graphics, Imaging and Visualization, pp. 64 – 69, 2005

[10] 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.

[11] N. Komodakis and G. Tziritas, “Image completion using global optimization,” IEEE Computer Society Confer-ence on Computer Vision and Pattern Recognition, vol. 1, pp. 442 – 452, 2006.

[12] Y. Pritch, E. Kav-Venaki, and S. Peleg, “Shift-map im-age editing,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 151–158, 2009.

[13] 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.

[14] S. Wang, H. Li, X. Zhu, and P. Li, “An evaluation index based on parameter weight for image inpainting qual-ity,” the 9th International Conference for Young Com-puter Scientists, vol. II, pp. 786–790, 18-21 Nov 2008.

[15] P. Ardis and A. Singhal, “Visual salience metric for im-age inpainting,” Proceedings of the SPIE, vol. 7257, 2009.

[16] V. V. Mahalingam, “Digital inpainting algorithms and evaluation,” Ph.D. dissertation, 2010.

[17] A. I. Oncu, F. Deger, and J. Y. Hardeberg, “Evaluation of digital inpainting quality in the context of artwork restoration,” Proceeding ECCV’12 Proceedings of the 12th international conference on Computer Vision, vol. 1, pp. 561 – 570, 2012.

[18] H. Brettel N. Bonnier, F. Schmitt and S. Berche, “Eval-uation of spatial gamut mapping algorithms,” Proc. Of the 14th Color Imaging Conference, 2006.

[19] L. Itti, C. Koch, and E. Niebur, “A model of saliency-basedvisual attention for rapid scene analysis,” IEEE Transactions on Pattern Analysis and Machine Intelli-gence, vol. 20 (11), pp. 1254–1259, 1998

[20] X. Hou and L. Zhang, “Saliency detection: A spectral residual approach,” IEEE Conference on Computer Vi-sion and Pattern Recognition, pp. 17–22, 2007.

[21] Q. Ma C. Guo and L. Zhang, “Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform,” IEEE Conference on Computer Vision and Pattern Recognition, p. 116, 2008.

[22] R. Achantay, S. Hemamiz, F. Estraday, and S. Susstrunky, “Frequency-tuned salient region detection,” IEEE International Conference on Com-puter Vision and Pattern Recognition, pp. 1597 – 1604, 2009.

[23], A. Beghdadi, C.M. Larabi, A. Bouzerdoum, K. M. Iftekharuddin, « A Survey of Perceptual Image Processing Methods », Signal Processing: Image Communication, Volume 28, Issue 8, September 2013, Pages 811–831