EUSIPCO 2013 1569735049

PERCEPTUAL EVALUATION OF DIGITAL IMAGE COMPLETION QUALITY

DANG Thanh Trung[1], *Azeddine BEGHDADI*[2], *Mohamed-Chaker LARABI*[3]

[1,2] Lab. L2TI, Institut Galilée, Université Paris 13, France

[3] Lab. XLIM, Dept. SIC, Université de Poitiers, France

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

ABSTRACT

Recently, image completion or inpainting has become a hot research topic because of its extensive applications in various real-world applications. Many papers have been published on this topic, but there are very few studies devoted to objective image inpainting evaluation. The intend of this work is to propose a solution based on Image Quality Metrics (IQMs). The idea is to adapt some IQMs by taking into account some specificities and constraints related to image inpainting objectives. The main characteristics used in this approach are the visual coherence of the recovered regions and the visual saliency. In this paper the notion of Inpainted Image Quality Assessment is introduced and discussed trough some experimental results. The obtained results demonstrate that the proposed image quality metric is well adapted to image inpainting evaluation.

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

1. INTRODUCTION

Digital inpainting refers to techniques used to fill in missing or deteriorated parts in an image in a visually plausible way. These techniques can be applied in many real-world applications such as digital film restoration, post-production of special effects shots for digital cinema, removing undesired or masking objects from photographs for police investigation, etc. Several image inpainting approaches have been proposed in the literature. Most of them could be categorized into two types based on the objectives. The first category consists of diffusion-based approaches [1, 2, 3] in which the missing regions are filled by diffusing the information from the known region into the missing region. These methods are suitable for narrow or small area but are less efficient for large areas. The second category of approaches is the exemplar-based approach. They could be further divided into subgroups: greedy [4, 5, 6] and global optimization strategies [7, 8, 9]. Stemmed from texture synthesis methods, these approaches produced an impressive output in recovering large damaged regions.

However, 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 scoring subjectively by human observers. But, the subjective evaluation experiments are time consuming, complex and unpredictable due to some non controlled human factors such as fatigue, visual discomfort, etc. The purpose of our work is to introduce an objective IIQA that could predict the perceptual quality of the recovered image.

Many image quality assessment (IQA) metrics have been introduced to identify and quantify image quality degradations [10, 11]. Nonetheless, the existing IQA metrics could not be directly applied for IIQA since the specificities and goals of both image quality, in its broad sense, and image inpainting are different. For example, in the case of image inpainting the recovered region is totally different from the original one. Furthermore, the evaluation is performed not only the basis of the recovered region but also on the visual coherence with the other surrounding parts in the image.

Currently, there are a few interesting works on objective IIQA. But they suffer from some shortcomings. M. M. Venkatesh et al [12] proposed two visual saliency-based metrics for predicting inpainted image quality using gaze pattern recorded by the faceLAB eye tracker. This method has the same disadvantages as the subjective evaluation methods. Another full-reference measure based on a modified $SSIM$ metric composed of three aspects: luminance, definition and gradient similarities is introduced in [13] to evaluate blurring effect of the restored regions. Actually, the inpainting problem is known as a blind image completion, where the reference image is unavailable. Moreover, this index cannot be applied when inpainted regions are large and may highly differ from the original ones. P. Ardis et al [14] defined two types of observable artifacts of inpainted image referred to variation of the saliency map before and after inpainting corresponding to two criteria: 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 main contribution of this paper is proposing a complete image inpainting quality index. The proposed metric is computed by combining 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 the known 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. The proposed measure is consistent with subjective judgment of observers. It is worth noticing that it does not depend on any assumption on the considered inpainting method. Thus, it could be used to evaluate any inpainting approach.

The paper is organized as follows: Section 2 is dedicated to the subjective quality protocol used in our experiment. The computation of the proposed metric is described in section 3. Section 4 is devoted to some experimental results and comparison with some metrics and MOS values. Finally, the paper ends by conclusions and future works.

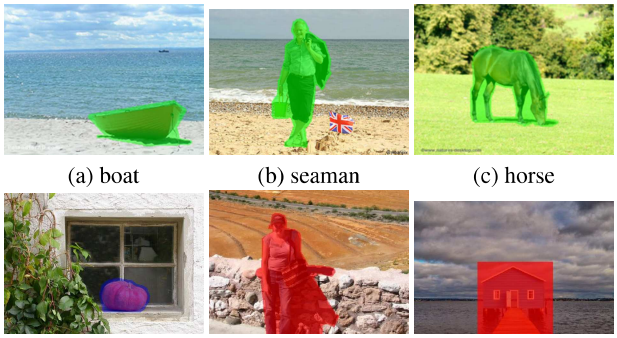
2. SUBJECTIVE QUALITY CONSTRUCTION

As mentioned previously, quality assessment dedicated to inpainting is a relatively new field of research. The performance of the proposed methods is often evaluated subjectively from the end-user including on the basis of some global perceptual criteria such as the global homogeneity of the spatial content and the coherence of added/suppressed regions. Therefore, the performance of a given metric, whatever its scheme is, should be assessed with regards to human judgment. However, to the best of our knowledge, there is no exhaustive inpainting database gathering state-of-the-art methods and their associated subjective scores for a significant set of images. Based on this observation and for the sake of proposing an objective metric for the evaluation of digital image completion quality, we decided to run psychophysical experiments to construct our ground truth. The latter will help the performance of the proposal in comparison to literature.

2.1. Test setup

Two groups of three images each have been constructed depending on the category of inpaining i.e. greedy strategy or global optimization, to be applied on it. Fig. 1 shows a snapshot of the used images. For each set, three representative methods have been selected to perform the inpainting process on images; thus generating nine results for each set.

A panel of seventeen observers, having a normal acuity and no color blindness, participated to the subjective experiment. They were asked to score an inpainted image in presence of the original image and those obtained with the remaining methods. Scores were given on discrete scale ranging from *1-unacceptable* to *5-perfect*. Observers were allowed to change their vote within a same series of images until validation. This has been done to account for the unavailability of a real original image and availability of different versions of results.



(a) boat (b) seaman (c) horse

(d) pumpkin (e) lady (f) house

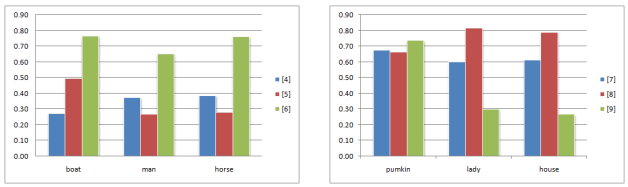
Fig. 1. Sets of images used for the subjective evaluation. Images (a), (b), (c) are used for greedy strategy; Images (d), (e) , (f) are used for gloabl optimization strategy.

2.2. Subjective results

Once the score are obtained from observers, a Mean Opinion Score (MOS), $\overline{u}_{i,k}$ and associated Confidence Interval (CI) is computed thanks to the following equation giving the average opinion of the observers regarding the submitted question.

$$\overline{u}_{i,k} = \frac{\sum_j u_{i,j,k}}{5N} \tag{1}$$

where $\overline{u}_{i,k}$ is the score of the observer i for the inpainting method j of the image k. N represents the number of observer. From the fig. 2, it is clearly possible to conclude about the visual performance of the greedy strategy inpainting method defined in [6] while the remaining methods are less performant and image-dependant. However, for global optimization methods, instead of pumpkin, three decreasing levels of inpainting quality are delivered corresponding to the three considered methods [7], [8] and [9]. The MOSs obtained from this study are of a high importance since it will allow to study the performance of the proposed inpainting quality metric.



(a) (b)

Fig. 2. Mean Opinion Score for (a) greedy strategy group and (b) global optimization group.

3. THE PROPOSED METRIC

The key objective of inpainting is filling in the damaged region without any clue of the reference image. That is the inpainted

region depends only on the rest of image. Therefore, the global coherence term is related to the consistency between the new generated pixels and the pixels in the known region. On the other hand, the coherence term determines the new undesired visual artifacts. Whereas, the local structure term is related to the local continuity of the edges or contours which often attract more attention of observers than other parts. Based on this observation, we proposed an efficient IIQA where both factors are taken into account. The inpainted image quality index, Q, is then defined as follows:

$$Q = \frac{\sum_{\Omega} C(p)S(p)}{||\Omega||} \tag{2}$$

where $C(p)$ and $S(p)$ represent the coherence and structure terms, respectively. The whole image domain, I, is composed of two disjoint regions: the inpainting region (or target region), Ω, and the known region, Φ ($\Phi = I - \Omega$). The basic unit of synthesis at pixel p is a patch, Ψ_p, centered at pixel p. The function sum ($\sum$) indicates that the index, Q, provides a score based on all pixels in inpainted regions instead of restricting the computation to pixels corresponding to maximum global visual saliency (as would be the function max).

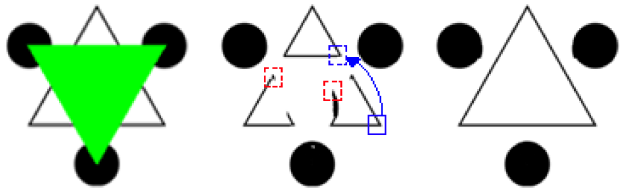
The coherence term C(p) is a measure of the visual coherence between an inpainted region, Ω, and the rest of the image. This could be expressed through a measure of the similarity between the candidate patch Ψ_q and the local patch Ψ_p. The coherence term is then defined as follows:

$$C(p) = max\{SIM(\Psi_p, \Psi_q), \forall \Psi_q \in \Phi\} \quad \forall p \in \Omega \tag{3}$$

where Ψ_p and 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. A good objective function needs to agree perceptually with the human observer. The MSE and $PSNR$ are widely used for patch similarity but they are insufficient. This is mainly due to the fact that they do not take into account perceptual features. To overcome limitations of MSE-based measures, a new similarity measure based on the structural information of patches was proposed in [10]. Following this idea, we propose a new formulation of the similarity measure by incorporating a coherence term. The proposed measure is defined as follows:

$$SIM(\Psi_p, \Psi_q) = \frac{(2\mu_p\mu_q + C_1)(2\sigma_{pq} + C_2)}{(\mu_p^2 + \mu_q^2 + C_1)(\sigma_p^2 + \sigma_q^2 + C_2)} \tag{4}$$

where the parameters μ_p, σ_p and μ_q, σ_q are the mean intensity and standard deviation set of patches Ψ_p and Ψ_q, respectively, while σ_{pq} denotes their cross correlation. The two positive constants C_1 and C_2 are introduce to avoid instability in uniform regions. Fig. 3 illustrates the spatial coherence notion used in the design of the IIQA. Note that the output shown in fig. 3c is more spatially coherent to that of fig. 3b. Indeed, some pixels



(a) (b) (c)

Fig. 3. The global coherence term. (a) An image to be inpainted. The image in (b) is less coherent than one in (c).

in inpainted regions (the red patches) do not look coherent with the original patterns. In other terms, there is no similar patches in in the source region. The blue dashed patch representing for a restored pixel is more reasonable than the red ones because it is similar to a blue solid patch in the known region.

The second factor affecting inpainting quality is continuity of edges or contours. Given an image, human vision would pay more attention to perceptually relevant regions corresponding to contours and details than the rest of the image. Therefore, the contours and other relevant structures in the inpainted regions attract human attention more than the other components. From these observations, the structure term is defined using the information provided by a saliency map in which region with high values corresponds to image components attracting more visual attention.

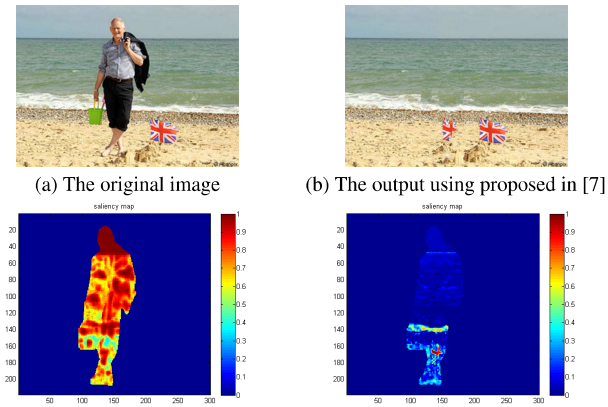
Many models have been proposed to simulate human's visual attention [15]. However, the high computational cost and uncontrolled parameters are still the weaknesses of these models. The authors in [16] proposed a simple and efficient method based on features of color and luminance to extract the salient regions. Moreover, when the inpainted region is manually restricted by user before inpainting, the observers often concentrated only on the inpainted region rather than other parts. Therefore, the saliency map should be locally calculated in this region. A full extension of saliency map SM based on approach in [16] is described as follows:

$$SM(p) = ||I_\mu - I_G(p)|| \quad \forall p \in \Omega \tag{5}$$

where I_μ and I_G are the local mean pixel value and the Gaussian blurred pixel value of the inpainted region. The $CIEL^*a^*b^*$ color space is used in order to correctly estimate the color difference appearance. The structure term based on the saliency map SM in eq. (6) would be normalized within the range $[0, 1]$ to compare in varying situations as follows:

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

Fig. 4 illustrates an image and the associated coherence and structure maps where higher values are represented in red and lower values are in blue.



(a) The original image (b) The output using proposed in [7]

(c) Coherence map (d) Structure map

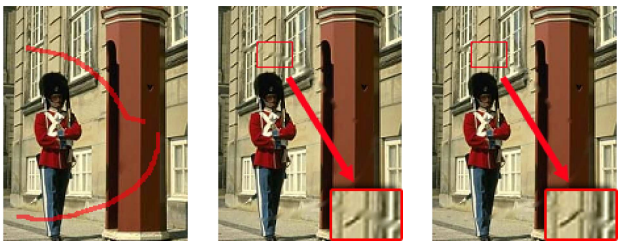
Fig. 4. Local quality maps.

4. EXPERIMENTAL RESULTS

In order to validate the performance of the objective quality metric presented in the previous section, the proposed metric needs to be evaluated with subjective evaluation obtained from psychophysical experiments. Nine inpainting methods of three groups are selected for psychophysical experiments. The size of patch is set to $\{7 \times 7\}$ for the coherence map computation. The size of the convolution mask representing the Gaussian filter used in the saliency map computation is $\{5 \times 5\}$.

4.1. Small inpainted region

Because the inpainted region is small or thin, the goal of inpainting was to make it appearing as close to the original image as possible. In this case, some full-reference image quality assessment metrics are used to evaluate the performance of our metric such as $MSSIM$ and $PSNR$.



(a) (b) (c)

Fig. 5. The inpainted images. (a) The pseudo-inpainting image; The outputs using approach in (b) [2] and (c) [3].

In our experiments, four approaches proposed in [2], [3], [1], and [6] representing two groups of inpainting methods namely geometry oriented and hybrid methods are considered. The inpainting region occupies 3.72% of the total image. Table 1 shows some quality indexes for image inpainting quality assessment. The $PWIIQ$ is the index defined in [13]. The $ASVS$ and DN are two indexes introduced in [14]. From table 1, we can observe that our index provides the values consistent with $MSSIM$ and $PSNR$ values. The higher values in bold represent better quality. It could be noticed that $ASVS$ and DN indexes determine the presence of inpainting artifacts, that is, it intends to produce lower value for better result. This shows that these indexes do not work well in this situation. The values in table 1 show that the approach in [1] gives the best output in term of $MSSIM$ and $PSNR$ values, and our index gives the same ranking.

Table 1. The IIQA indexes in case of thin damaged region.

Output by	[2]	[3]	[1]	[6]
$PSNR$	32.250	31.939	**33.233**	31.935
$MSSIM$	0.980	0.979	**0.985**	0.982
$PWIIQ$	0.981	0.983	0.984	**0.989**
$ASVS$	0.136	**0.134**	0.150	0.149
DN	**0.006**	0.006	0.007	0.007
$Ourmetric$	0.240	0.245	**0.273**	0.262

4.2. Large inpainted region

For the large inpainted regions, the exemplar-based methods are the most common use and achieve impressive results. A large number and variety of inpainting algorithms have been published. To cover all these methods would be infeasible. Here, the experimental comparison study is restricted to some methods of the state-of-the-art: greedy strategy based methods [4, 5, 6] and optimization based methods [7, 8, 9].

The main objective is to restore the deteriorated regions in a visually plausible way without refereeing to the original image. Therefore, the $PSNR$ and $MSSIM$ could not be applied. The output results are then subjectively evaluated using the MOS.

By investigating the values in two tables, it is obvious that our index is the most consistent with the subjective score in categorizing the quality of outputs respected by mean value of Spearman rank order correlation coefficient (SCC_{TB}).

5. CONCLUSION

In this paper a novel approach for inpainting image quality evaluation is proposed. It is shown that the classical image quality measures could not be used for evaluating the inpainting results. By taking into account the specificities and objectives of inpainting problem and some characteristics of the human 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 could be also applied to most of inpainting image approaches. The performed experimental results and comparison with all three approaches for inpainting image methods confirm the efficiency of the proposed index.

For future works, the computational time is too high because of similar patch searching for coherence map evalua-

Table 2. The IIQA of greedy methods.

Images	seaman			boat			horse			SCC_{TB}
Result by	[4]	[5]	[6]	[4]	[5]	[6]	[4]	[5]	[6]	
MOS	0.370	0.267	0.653	0.271	0.494	0.765	0.387	0.280	0.760	*
$ASVS$	0.093	0.1614	0.1125	0.3115	0.2260	0.3097	0.225	0.213	0.186	0.5
DN	0.027	0.043	0.039	0.0556	0.0545	0.0644	0.048	0.049	0.043	0.33
$Ourmetric$	0.073	0.065	0.070	0.145	0.150	0.180	0.077	0.089	0.089	**0.67**

Table 3. The IIQA metrics of global optimization methods.

Images	lady			house			pumpkin			SCC_{TB}
Result by	[7]	[8]	[9]	[7]	[8]	[9]	[7]	[8]	[9]	
MOS	0.814	0.600	0.300	0.613	0.720	0.240	0.675	0.663	0.7375	*
$ASVS$	0.180	0.126	0.101	0.0700	0.0855	0.1593	0.157	0.151	0.236	-0.33
DN	0.022	0.015	0.012	0.0343	0.0392	0.0498	0.014	0.015	0.018	-0.17
$Ourmetric$	0.055	0.059	0.016	0.063	0.061	0.042	0.095	0.102	0.124	**0.67**

tion. Moreover, the objective evaluations metrics of video or context-based image will be also investigated.

6. REFERENCES

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," *Proceedings of ACM Conference Computer Graphics*, pp. 417–424, 2000.

[2] D. Tschumperle, "Fast anisotropic smoothing of multi-valued images using curvature-preserving pdes," *International Journal of Computer Vision*, vol. 68, pp. 65–82, 2006.

[3] A. Bugeau and M. Bertalmio, "Combining texture synthesis and diffusion for image inpainting," *Proceeding International Conference Computer Vision Theory and Applications*, pp. 26–33, 2009.

[4] 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.

[5] Q. Zhang and J. Lin, "Exemplar-based image inpainting using color distribution analysis," *Journal of Information Science and Engineering*, pp. 1–12, 2011.

[6] 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*, pp. 280–284, 2012.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] P. Ardis and A. Singhal, "Visual salience metric for image inpainting," *Proceedings of the SPIE*, vol. 7257, 2009.

[15] 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. 1–8, 2008.

[16] 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.