Recovering real-world scene: high-quality image inpainting using multi-exposed references

Z.J. Zhu, Z.G. Li, S. Rahardja and P. Fränti

A novel method of high-quality image inpainting for recovering an original scene from degraded images using reference images of different exposures is proposed. It consists of a new inter-pixel relationship function and the respective refinement to synthesise missing pixels from existing spatially co-related pixels, and a dual patching to minimise the noise caused by dynamic range lost. Experiments on the method have been conducted and the results demonstrate the reliability of the proposed method.

Introduction: Images of the same scene can be captured with different exposures and combined with computing power to synthesise an image that overcomes the limitations of conventional cameras. However, useful data can be lost owing to camera shake, especially when capturing by a hand-held device, which generates noticeable artefacts in the synthesised image. In other words, different from the traditional image inpainting [1, 2] and image completion [3], which generate only photorealistic patches, the degraded image inpainting in digital photography requires true luminance values of the real-world scene. Therefore, the challenge of patching is to find useful relations between missing pixels and the remaining pixels. The camera-response-function (CRF)-based fixing method [4, 5] uses only inter-image relationships. Unfortunately, the patched pixel is just a luminance shift from the reference pixel, and cannot represent the pixel value at the correct exposure of the degraded image. Motivated by these observations, we propose a new method using the refined inter-pixel relationship function (IRF) with both inter-image and intra-image correlations to recover the real scene luminance reliably and a dual patching to reduce the dynamic range lost, which further enhance the inpainting accuracy.

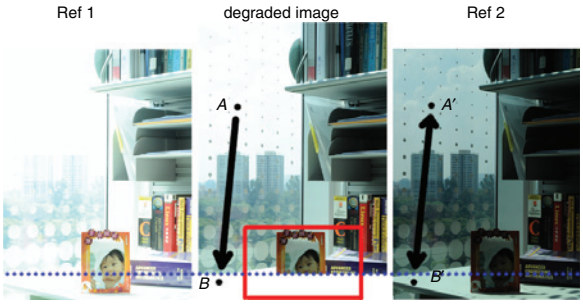

Fig. 1 *Three differently exposed images with degraded image due to camera shake*

Pixel value of B can be copied from A as their co-locations A′ and B′ have same intensity in reference image. Details inside rectangular box displayed in Fig. 3

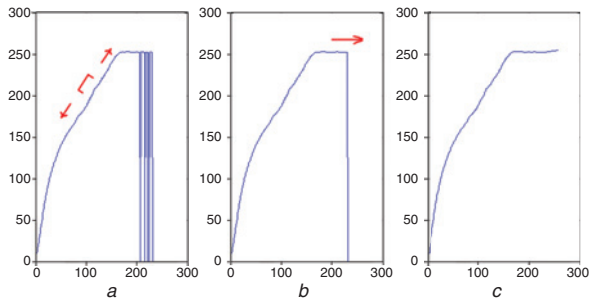

Fig. 2 *IRF curves*

a Initial values
b After smoothing with median filter
c After extending empty values at end
x-axis is intensity of reference image; *y*-axis is calculated co-location intensity in degraded image

Inter-pixel relationship function: As shown in Fig. 1, the pixels A′ and B′ have the same intensity in the reference image. According to photography reciprocity, when the exposure time changes, the pixel values of A′ and B′ change correspondingly. Intuitively, the missing value of B can be copied from A in the degraded image. However, during the image capturing process, sensor noise, sampling noise and compression noise are commonly generated. Thus, it is more accurate to find all the pixels with the same intensity as B′ in the reference image ($\widehat{Z}$) and calculate their co-location values in the degraded image (Z) using mean average. We define the IRF as

$$\psi_c(\widehat{Z}_c(x,y)) = \frac{\sum_{(\acute{x},\acute{y})\in\Omega(\widehat{Z}_c(x,y))} Z_c(\acute{x},\acute{y})}{|\Omega(\widehat{Z}_c(x,y))|} \quad (1)$$

where c is the colour channel, and $|\Omega(\widehat{Z}_c(x,y))|$ is the cardinality of the spatially co-related pixels set $\Omega(\widehat{Z}_c(x,y))$, which is defined as

$$\Omega(\widehat{Z}_c(x,y)) = \{(\acute{x},\acute{y})|\widehat{Z}_c(\acute{x},\acute{y}) = \widehat{Z}_c(\vec{x},\vec{y})\} \quad (2)$$

The IRF has three useful characteristics inherited from the physical camera response: *Char1*: the IRF is a monotonically increasing function: *Char2*: the pixels located at the left end (dark pixels) and right end (bright pixels) are highly compressed owing to dynamic range limit; *Char3*: when choosing different reference images, shorter exposure time leads to a smaller slope at the left end and a bigger slope at the right end.

An 'empty value' problem is raised in the raw IRF when the reference image does not span the whole dynamic range, as shown in Fig. 2. Using *Char1*, the refinement has two steps. First, a median filter is adopted starting from the middle of the valid values towards left and right separately. The median filter corrects the monotonic errors, and recovers the empty values between the valid values. The second step extends two ends of the curve by using the neighbourhood slope. The refined IRF is defined as

$$\Psi_c(z) = \text{extend}(\text{median}(\psi_c(z))), \quad z \in [0, 255] \quad (3)$$

Fig. 3 *Degraded area and patches*

a Original degraded image
b Fixed image using exemplar-based inpainting
c Fixed image using CRF
d Fixed image using our method
e HDR image synthesised with degraded image
f HDR image synthesised with patched image

Dual patching: To increase the accuracy, the reference image is selected to have the smallest exposure difference with the degraded image. However, the dynamic range loss caused by *Char2* is still inevitable. If the reference image has shorter exposure time than the degraded image, then as can be seen from *Char3*, the dark pixels in the reference image are mapped from a big dynamic range to a small one. In other words, it is a compressing process mapping multiple values to one, which in turn makes the IRF in this area reliable. On the other hand, a highly compressed bright pixel in the reference image is mapped into multiple values in the degraded image, which causes inaccuracy owing to the dynamic range lost. Thus, multiple reference images,

with longer and shorter exposure time, respectively, can be adopted to recover the lost dynamic range and enhance the patching accuracy. The missing pixel intensity is calculated by

$$P_c = \frac{\dot{\Psi}_c(\dot{Z})W(\max(\dot{Z}_R, \dot{Z}_G, \dot{Z}_B)) + \ddot{\Psi})_c(\ddot{Z})W(\max(\ddot{Z}_R, \ddot{Z}_G, \ddot{Z}_B))}{W(\max(\dot{Z}_R, \dot{Z}_G, \dot{Z}_B)) + W(\max(\ddot{Z}_R, \ddot{Z}_G, \ddot{Z}_B))} \quad (4)$$

where $\dot{Z}$ and $\ddot{Z}$ are the intensities of two reference images at the same co-location, c is the colour channel ($c = R, G, B$) and W is the weighting function defined as

$$W(z) = \begin{cases} \log(z+1), & EV(\text{ref}) > EV(\text{degrade}) \\ \log(256 - z), & EV(\text{ref}) < EV(\text{degrade}) \end{cases}, z \in [0, 255] \quad (5)$$

Result: As shown in Fig. 3, the degraded area destroys the integrity of the original image composition. Clearly, the exemplar-based inpainting [2] algorithm works well on simple texture, such as the table top. However, obvious errors can be seen for complex content, like the baby's face and the title of the book, because of shortage of reference. The CRF method [4] recovers the content by the luminance shift from the reference pixels, where an obvious artefact can be seen at the border. Our algorithm restores the original scene effectively. Table 1 shows the peak signal-to-noise ratio (PSNR) of each method.

Table 1: PSNR comparison

Method	Exemplar-based inpainting [2]	CRF method [4]	Our method
PSNR (dB)	11.75	20.54	33.66

In addition, we have tested our algorithm in high dynamic range (HDR) image synthesis [6]. The border artefact generated by the degraded image in Fig. 3*e* is completely removed after patching using our algorithm, as shown in Fig. 3*f*.

Conclusions: This Letter describes a new image inpainting method to patch the degraded image in an exposure set. As it uses all the relations of the valid pixels with refined IRF and reconstructs the missing area by dual patching, it demonstrates better quality than other algorithms. Experimental results with the HDR image synthesis further verify the efficiency of the proposed method.

23 September 2009
doi: 10.1049/el.2009.2686

Z.J. Zhu, Z.G. Li and S. Rahardja (*Department of Signal Processing, Institute For Infocomm Research, A*STAR, 1 Fusionopolis Way, 21-01 Connexis, 138632, Singapore*)

E-mail: zhuzj@i2r.a-star.edu.sg

P. Fränti (*Department of Computer Science, University of Joensuu, PO Box 111, Joensuu 80101, Finland*)

References

1 Hung, J.C., Hwang, C.H., Liao, Y.C., Tang, N.C., and Chen, T.J.: 'Exemplar-based image inpainting base on structure construction', *J. Software*, 2008, **3**, (8), pp. 57–64
2 Criminisi, A., Pérez, E., and Toyama, K.: 'Region filling and object removal by exemplar-based image inpainting', *IEEE Trans. Image Process.*, 2004, **13**, (9), pp. 1200–1212
3 James, H., and Alexei, A.E.: 'Scene completion using millions of photographs'. SIGGRAPH, San Diego, CA, USA, August 2007, article No. 4
4 Grosch, T.: 'Fast and robust high dynamic range image generation with camera and object movement'. Proc. of Vision Modeling and Visualization, Aachen, Germany, 2006, pp. 277–284
5 Mann, S.: 'Comparametric equations with practical applications in quantigraphic image processing', *IEEE Trans. Image Process.*, 2000, **9**, (8), pp. 1389–1406
6 Debevec, P., and Malik, J.: 'Recovering high dynamic range radiance maps from photographs'. SIGGRAPH, NJ, USA, 1997, pp. 130–135