Emerging Images Synthesis from Photographs

Mao-Fong Jian[1] Hung-Kuo Chu[1*] Ruen-Rone Lee[1] Chia-Lun Ku[2] Yu-Shuen Wang[2] Chih-Yuan Yao[3]

[1]National Tsing Hua University [2]National Chiao Tung University [3]National Taiwan University of Science and Technology

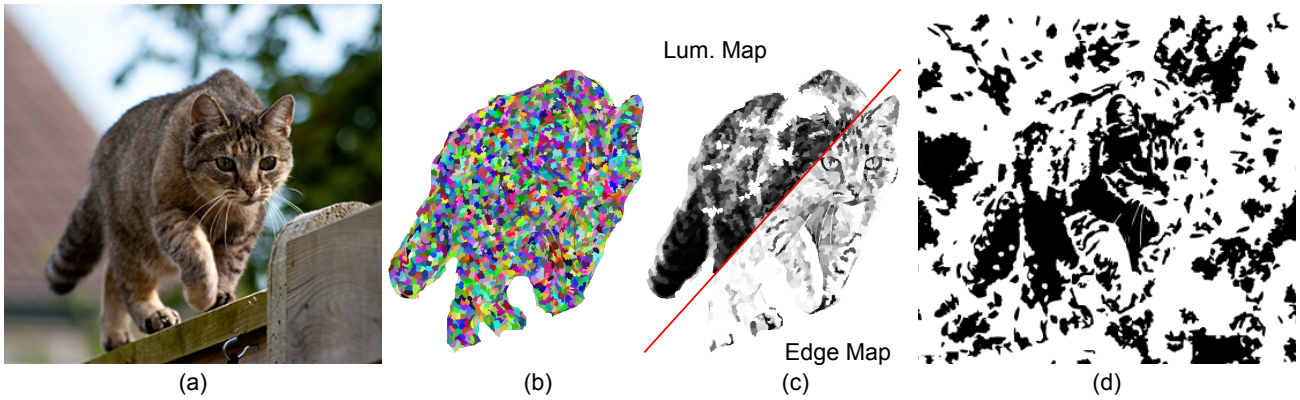

Figure 1: *(a) Input photograph. (b) Over-segmentation using super-pixels. (c) Luminance map and edge map. (d) Our result.*

1 Introduction

Emergence refers to a phenomenon by which human perceives complete objects in a seemingly noisy image not by recognizing local parts of image but viewing the image as a whole. The Dalmatian dog image created by R. C. James is probably the best demonstration of emergence [Bach 2002]. It shows that local windows from the image reveal nothing but meaningless, complex and random black splats. Only when the image is viewed as a whole, a Dalmatian dog suddenly appears. The absence of meaningful information in local image parts largely hinders existing computer vision algorithms from recognizing emerging figures. Therefore, it makes the emergence an new type of CAPTCHA to tell human and machine apart.

Inspired by the Dalmatian dog image, Mitra *et al*. [2009] proposed an automatic algorithm to synthesize emerging images from 3D objects. During the synthesis process, two guiding principles are carefully evaluated to ensure the synthesized images are easy for human to recognize while hard for bots to detect the embedded objects. However, using 3D objects as input brings two drawbacks that make the system an unqualified CAPTCHA. Firstly, available resources of 3D model repository are limited, therefore it is unable to generate a huge database of emerging images. Secondly, operations on 3D space are computational expensive means user cannot obtain or refresh a CAPTCHA puzzle in an interactive rate.

In this work, we propose an automatic algorithm to synthesize emerging images from common photographs. To generate images that are easy for human, Mitra *et al*. [2009] rendered complex splats that capture silhouette and shading information of 3D objects. However, we realize that comparative information could be retrieved from photographs as well, and we try to replace the rendering of black complex splats with super-pixels. It takes two further post processing steps to make segmentation harder for bots, and both of them could find counterpart operations in image domain. Supporting by public image databases such as flickr and Picasa, we can envision a potential CAPTCHA application of our approach to massively and efficiently generate emerging images from photographs.

*Corresponding author:hkchu@cs.nthu.edu.tw

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
SIGGRAPH 2013, July 21 – 25, 2013, Anaheim, California.
2013 Copyright held by the Owner/Author.
ACM 978-1-4503-2261-4/13/07

2 Our Approach

Given a photograph, our system first segments the foreground object using semi-automatic saliency cut [Cheng 2011]. Then it applies an image over-segmentation algorithm [Achanta et al. 2012] to obtain super-pixels of foreground object (see Figure 1(b)). We formulate the problem of rendering an emerging foreground as a binary labeling problem that assigns binary color (black or white) to super-pixels.

To leave necessary cues for human to recognize, we introduce two maps to guide the assignment of binary color as shown in Figure 1(c). One is a luminance map which captures the grayness of each super-pixel while the other is an edge map that captures the relative contrast between adjacent super-pixels. Value of each super-pixel in luminance map and edge map is calculated by averaging the image intensity and gradient of its atomic pixels, respectively. Finally, we employ a graph cut algorithm to efficiently solve the labeling problem consisting of data and smoothness terms with data term measures the distance between new color and original grayness of superpixels while the smoothness term favors the preservation of original contrast. Next, we adopt a similar copy-perturb-paste operation of [Mitra et al. 2009] to duplicate super-pixels from foreground to background such that the object and the rest of image look similar everywhere when observed through small windows. Since our approach operates on 2D super-pixels, the computational complexity is much cheaper than [Mitra et al. 2009] which takes 5 seconds to generate a result while ours take 1 second with image of moderate size using the unoptimized codes.

References

ACHANTA, R., SHAJI, A., SMITH, K., LUCCHI, A., FUA, P., AND SUSSTRUNK, S. 2012. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. PAMI 34*, 11, 2274–2282.

BACH, M., 2002. Hidden Figures – Dalmatian Dog.

CHENG, MING-MING, Z. G.-X. M. N. J. H. X. H. S.-M. 2011. Global contrast based salient region detection. In *IEEE CVPR*.

MITRA, N. J., CHU, H.-K., LEE, T.-Y., WOLF, L., YESHURUN, H., AND COHEN-OR, D. 2009. Emerging images. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 28*, 5, 163:1–163:8.