Image Editing With Texture Power Maps

Sara L. Su
MIT CSAIL

Frédo Durand
MIT CSAIL

Maneesh Agrawala
Microsoft Research

1   Introduction

A major obstacle in photography is the presence of distracting elements that pull attention away from the main subject and clutter the composition. Photographers have developed post-processing techniques to reduce the *salience* of distractors by altering low-level features to which the visual system is particularly attuned: sharpness, brightness, chromaticity, or saturation. Biologically-inspired models of attention identify salient regions as statistical outliers in these feature distributions. One low-level feature that cannot be directly manipulated with existing image-editing software is *texture variation*. Psychophysical studies have shown that discontinuities in texture can elicit an edge perception similar to that triggered by color discontinuities.

2   Power maps

First-order computational models of saliency measure the response to filter banks that extract contrast and orientation in the image. Various non-linearities can then be used to extract and combine maxima of the response to each feature. A recently-introduced *second-order* model performs additional image processing on the response to a first-order filter bank, effectively performing the same computation as first-order models but on what we term *power maps* rather than on image intensity. Higher-order features describing local frequency content, power maps have been used previously in image analysis; e.g. response to multiscale oriented filters can be used for texture discrimination. We show that power maps are also a powerful representation for *manipulating* frequency content in an image.

3   Texture equalization

We introduce an image-processing technique for selectively reducing spatial variation of texture to reduce the salience of distracting regions. In a nutshell, our *texture equalization* technique modifies distracting regions to make them look more like uniform textures.

We illustrate our technique with a 1D example. The input signal (Fig. 1(a)) is first band-pass filtered (b) and rectified with an absolute value non-linearity (c). (For the 2D case, we use *steerable pyramid* filters to compute frequency content because they permit straightforward analysis, processing, and reconstruction of images.) Pooling the rectified response by applying a low-pass filter with a Gaussian kernel captures the local frequency content. We call the resulting image the *power map* (d).

To reduce texture variation in the image, some portion of the high frequencies of the power maps must be removed, a seemingly trivial image-processing operation. However, we must define how a modification of the power map translates into a modification of pyramid coefficients. The exponent of the high-pass response (e) is used to scale the bandpass response. Because the goal is to reduce variation, a negative multiple of the high-pass is used as the scale factor. Note how the scaled signal (f) has been 'flattened' compared to the input. In the 2D case, the scaled subbands are then recombined to produce the final texture-equalized image.

(a) Input   (b) Band-pass   (c) Rectified   (d) Power map   (e) Scale   (f) Output
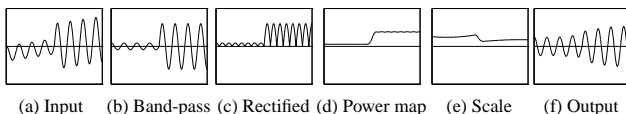
Figure 1: Texture discrimination and manipulation in 1D.

Figure 2: Specular highlights in the leaves and other distractors prevent clear foreground/background separation in the original (top) photograph. Reducing texture variation in the masked region of the bottom image de-emphasizes these distractors, thereby increasing salience of the intended subject, the tiger.

4   Psychophysical study

To validate our technique's effectiveness, qualitative changes in user fixations on original and modified images were recorded using an eye tracker. Emphasized regions attracted and held fixations longer than de-emphasized ones. Results of a search experiment quantified the effect of our technique on response time. Subjects were asked to find a target object in a series of images, some unmodified and some in which distractors had been de-emphasized. Texture equalization resulted in a search speedup of more than 20%.

5   Discussion

Texture equalization is complementary to de-emphasis methods such as Gaussian blur, which increases depth-of-field effects. Reduced sharpness can be undesirable, particularly if distractors are at the same distance as the main subject. Blur removes high frequencies and emphasizes medium ones, possibly resulting in a more distracting object, while our technique makes high-frequencies more uniform, "camouflaging" medium-frequency content. Our technique is most effective for textured image regions, while Gaussian blur works best when small depth-of-field effects are already present and when medium-frequency content is not distracting.

Please see additional images from the poster at
`http://csail.mit.edu/˜sarasu/pub/scpv05`