Texture Similarity Estimation Using Contours

Xinghui Dong
http://www.macs.hw.ac.uk/texturelab

Mike J. Chantler
http://www.macs.hw.ac.uk/~mjc/

The Texture Lab
School of Mathematical and Computer Sciences
Heriot-Watt University
Edinburgh, UK

Although performances in the high nineties are typically obtained for tasks such as texture segmentation and classification the same cannot be said of judging texture similarity where a classifier has to estimate the degree to which pairs of textures appear similar to human observers. In an investigation of 51 computational feature sets Dong et *al*. [1] showed that none of these managed to estimate similarity data derived from a population of human observers better than an average agreement rate of 57.76%. Coincidently, none of these computed higher order statistics (HOS) over large regions (≥ 19×19 pixels).

We have discovered few methods that encode long-range, aperiodic characteristics of texture; however, it is well-known that such data are critical to human perception of imagery [2, 3]. For instance, scrambling phase spectra (while leaving the power spectra intact) will often render imagery unintelligible to the human observer [3]. It is also well-known that humans are extremely adept at exploiting the long-range visual interactions evident in contour information [2, 4]. Therefore, we designed an experiment with human observers in order to determine which of three different types of information (2nd-order statistics, local higher order statistics and contour information, see Figure 1) are more important for the perception of texture.

Ten human observers were used in a 2AFC (two-alternative forced choice) scheme with 334 texture images drawn from the *Pertex* database [5]. In each trial the observer was required to compare an original texture image quarter and one variant image quarter ("variant" being one of either contour, power spectrum or randomized block) and decide whether the variant represented the original texture or not (50% of the time they did not). Different quarters of the same texture sample were used in order to prevent observers from performing pixel-wise comparisons. It was found that contour data is more important than local image patches, or 2nd-order global data, to human observers.

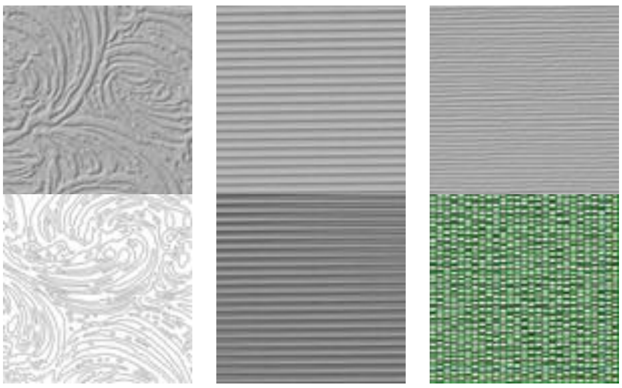
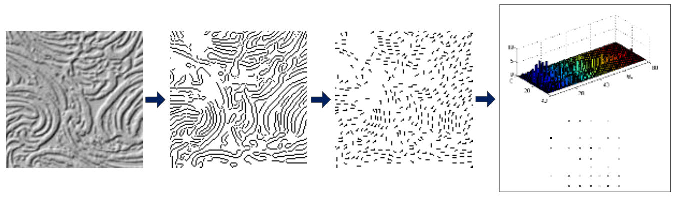


Figure 1: Each of the three columns shows two images derived from the same texture sample (although not the same physical texture area). The upper row shows unprocessed images. The lower row shows, from left to right, the corresponding contour map, power spectrum image and randomized, blocked image.

We therefore developed a contour-based feature set that exploits the long-range HOS encoded in the spatial distribution and orientation of contour segments. A contour is first fragmented into a set of equidistant segments and is then encoded using the spatial distribution and orientation of these segments. Note that images are first processed with the Canny edge detector [6] followed by a morphological erosion operator [7] in order to produce skeleton maps (see Figure 2 (b)). Connected component labelling [7] is performed on skeleton maps. Subsequently, the Moore-neighbour tracing algorithm with Jacob's stopping criteria [7] is applied to each contour and a sequence of points is obtained from each contour. Each contour is then divided into a series of equidistant segments. We represent segments by their mid-point position (on themselves) and chord orientation θ ($\theta \in (0°, 180°]$).

We use these data in two ways as outlined in Figure 2. In the first we encode the average shape of the contours in a segment joint orientation/distance histogram (see Figure 2 (d) upper). This provides data on the long-range higher-order visual interactions. In the second we used basic aura matrices [8] (see Figure 2 (d) lower) to encode the spatial distributions and orientations of the all of the segments within a local window without regard to which contour they belong. These data naturally provide relatively short-range (23×23 or less) HOS. The mean of all segment orientation/distance histograms and each basic aura matrix were concatenated into one feature vector which we refer to as "SDoCS" (spatial distribution of contour segments). We test it with two different segment angle quantization schemes (using A bins, $A \in \{18,36\}$), five different segment lengths ($SL \in \{3,5,7,9,11\}$) and one multi-scale case ($SL =$"MS") which concatenates all five feature vectors derived from the five different segment lengths.



(a) (b) (c) (d)

Figure 2: A representation of the basic information flow: (a) original texture image; (b) skeleton map; (c) segment map. For display purposes, only a part of pixels are shown for each approximate segment; and (d) the joint histogram (upper) and basic aura matrix [8] (lower, only one is shown here).

SDoCS was compared against the 51 feature sets tested by Dong et *al*. [1, 9] and another contour model derived from shape recognition. A pair-of-pairs based evaluation method and a ranking-based evaluation method [1, 9] were applied. The results show that the proposed method outperforms all the other feature sets in the pairs-of-pairs task and all but two feature sets in the ranking task.

We feel that the key point, however, is that we have showed the usefulness of long-range HOS in computing texture similarity and hope that this will inspire other developments of texture features based on such information.

[1] X. Dong and M. J. Chantler. The Importance of Long-Range Interactions to Texture Similarity. Proceedings of the 15th International Conference on Computer Analysis of Images and Patterns, 8047: 425-432, 2013.

[2] D. J. Field, A. Hayes and R. F. Hess. Contour integration by the human visual system: evidence for a local "association field". *Vision Research*. 33: 173-193, 1993.

[3] A. V. Oppenheim and J. S. Lim. The Importance of Phase in Signals. *Proceedings of the IEEE*, 69 (5):529-541, 1991.

[4] L. Spillmann and J. S. Werner. Long-range interactions in visual perception. *Trends in Neurosciences*. 19: 428-434, 1996.

[5] A. D. F. Clarke, F. Halley, A. J. Newell, L. D. Griffin and M. J. Chantler. Perceptual Similarity: A Texture Challenge. *Proceedings of British Machine Vision Conference*, 120.1-120.10, 2011.

[6] J. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6): 679-698, 1986.

[7] R. C. Gonzalez and R. E. Woods. *Digital Image processing*, Prentice Hall Upper Saddle River. NJ, 2002.

[8] X. Qin and Y. Yang. Basic Grey level aura matrices: theory and its application to texture synthesis. *Proceedings of the Tenth IEEE International Conference on Computer Vision*, 1: 128-135, 2005.

[9] X. Dong, T. Methven, and M. J. Chantler. How Well Do Computational Features Perceptually Rank Textures? A Comparative Evaluation. Proceedings of the ACM 2014 International Conference on Multimedia Retrieval, 281-288, 2014.