Perceptual Similarity: A Texture Challenge

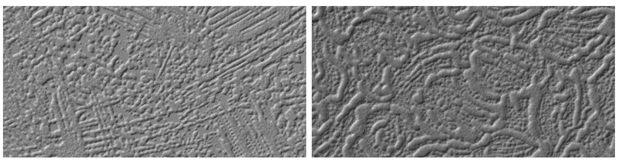
Alasdair D. F. Clarke[1]
http://www.macs.hw.ac.uk/texturelab
Fraser Halley[1]
http://www.macs.hw.ac.uk/texturelab
Andrew J. Newell[2]
emailhomepage
Lewis D. Griffin[2]
emailhomepage
Mike J. Chantler[1]
http://www.macs.hw.ac.uk/~mjc/

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

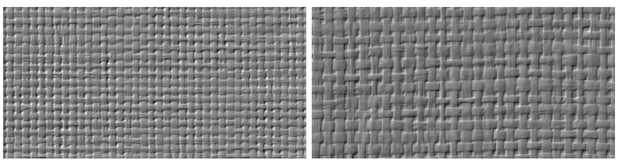
[2] Vision & Image Sciences
Computer Science
University College London, UK

Texture classification and segmentation have been extensively researched over the last thirty years. Early on the Brodatz album[1] quickly became the de facto standard in which a *texture class* comprised a set of non-overlapping sub-images cropped from a *single* photograph. Later, as the focus shifted to investigating illumination- and pose-invariant algorithms, the CUReT database[3] became popular and the *texture class* became the set of photographs of a single physical sample captured under a variety of imaging conditions. While extremely successful algorithms have been developed to address classification problems based on these databases, the challenging problem of measuring perceived *inter-class* texture similarity has rarely been discussed.

This paper makes use of a new texture collection[4]. It comprises 334 texture samples, including examples of embossed vinyl, woven wall coverings, carpets, rugs, window blinds, soft fabrics, building materials, product packaging, etc. Additionally, an associated *perceptual* similarity matrix is provided. This was obtained from a grouping experiment using 30 observers. The similarity scores, $S(I_i, I_j)$, for each texture pair were calculated simply by dividing the number of observers that grouped the pair into the same sub-set by the number of observers that had the opportunity do so. A *dissimilarity matrix* was then defined as $d_{sim}(I_i, I_j) = 1 - S(I_i, I_j)$. Hence $d_{sim}(I_i, I_i) = 0$ for all images I_i, and $d_{sim}(I_i, I_j) = 1$ if none of the participants grouped images I_i together with I_j.



(a) A pair of images with: $d_{sim}(27, 131) = 1$. None of the human observers grouped these textures together



(b) A pair of images which all but one of human observers grouped together: $d_{sim}(168, 176) = 0.07$.

Figure 1: Some examples of textures from the dataset with their similarity matrix results

We used this to perform two experiments testing *classification* and *inter-class similarity* performance. In the two experiments, we evaluated four different texture features (LBP[5], MRV8[6], MRF[7], BIF[2]. We also tested a multi-scale implementation of the BIF algorithm (MS-BIF). These feature sets were chosen as they have all demonstrated excellent results on previous databases and are available to run unchanged as ‘off the shelf algorithms.’ The aim of Experiment 1 (classification) was to test the texture classification algorithms on the new dataset, mimicking the structure of the CUReT database[3] and the protocol commonly used on it (for example, see [2]). Our results are in agreement with previous work[2, 5, 6, 7].

In the second experiment we tested the ability of the four texture algorithms to predict the inter-class similarities derived from our 30 observers. As can be seen inTable 1 the computed distances between histograms do not correlate well with human judgementss, with the best performance giving $R^2 = 0.04$. We also examined Spearman’s rank correlation coefficient, which showed that there was only a weak relationship between perceptual and computational dissimilarity ($\rho = 0.21$).

Feature	R^2	R^2 - log(Feature)	ρ
LBP	0.031	0.025	0.131
MRV8	0.042	0.077	0.180
MRF	0.031	0.071	0.206
BIF	0.009	0.063	0.166
MS-BIF	0.011	0.058	0.176

Table 1: Similarity performance. We can see that the best computational feature explains $\approx 5\%$ of the variation in human responses. ρ is Spearman’s correlation coefficient.

We believe that this set of 334 textures is currently the largest texture database that has been captured under controlled illumination conditions and, perhaps more importantly, is accompanied by an associated perceptual similarity matrix. It is also contains height data allowing illumination-independent generation of features and relighting under arbitrary illumination conditions.

In Experiment 1 we investigated the performance of four state-of-the-art classification schemes and showed that they provide near-ceiling texture classification performance when tested on this new texture database using a protocol similar to that commonly used with the CUReT image set. However, Experiment 2 showed that the perceptual similarity matrix obtained using 30 human observers does not correlate well with machine performance. This is likely to be due to either (a) the perceptual data not being representative of the population or (b) that the algorithms not exploiting all of the texture features that are used by human observers. These include salient, longer range spatial interactions that are not detected by the relatively small spatial neighbourhoods that machine vision features use.

[1] P. Brodatz. *Textures: A photographic album for artists and designers*. Dover Publications, 1966.

[2] M. Crosier and L. D. Griffin. Using basic image features for texture classification. *Internation Journal of Computer Vision*, 88:447–460, 2010.

[3] K.J. Dana, B. Van-Ginneken, S.K. Nayar, and J.J. Koenderink. Reflectance and texture of real world surfaces. *ACM Transactions on Graphics (TOG)*, 18:1–34, 1999.

[4] F. Halley. Pertex v1.0, 2011. URL `http://www.macs.hw.ac.uk/texturelab/resources/databases/pertex/`.

[5] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24: 971–987, 2002.

[6] M. Varma and A. Zisserman. Texture classification: Are filter banks necessary? In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 691–698, 2003.

[7] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1–2):61–81, 2005.