Comparing computational and human measures of visual similarity

Tim M. Gale
University of Hertfordshire
t.gale@herts.ac.uk

Yi Sun
University of Hertfordshire
Y.2.Sun@herts.ac.uk

Rod Adams
University of Hertfordshire

Neil Davey
University of Hertfordshire

June 10, 2004

There have been many attempts to quantify visual similarity within different categories of objects, with a view to using such measures to predict impaired recognition performance. Although many studies have linked measures of visual similarity to behavioural outcomes associated with object recognition, there has been little research on whether these measures are associated with human ratings of perceived similarity. In a recent study [1], similarity measures extracted from self-organising map[2] representations of greyscale pictures compared favourably with ratings of perceived visual similarity provided by human subjects for the same set of pictures. In the current study, we turn our attention to other computational measures of visual similarity to investigate whether these approaches can provide even stronger associations with human ratings.

In this work, we compare similarity measures extracted from principal components analysis (PCA), isometric feature mapping (Isomap)[3], and wavelets[4] representations with ratings of human subjects. Isomap is an approach that preserves the geodesic distances between all pairs of data points from the high-dimensional data space to the low-dimensional space. A wavelet representation is produced by a 2-D haar-wavelet transformation, decomposing each image into different directional, i.e., vertical, horizontal and diagonal detail images of different resolutions. For each different level in each different direction, we compute the standard deviation, which is used to compose a new feature vector.

Our results show that features extracted by calculating the standard deviation of wavelet coefficients provides the closest fit to the human rating data of all the methods we applied here. The correlation between the wavelet representation and human ratings is 25% higher than that of the raw data and human ratings.

References

[1] T.M. Gale, N. Davey, K.R. Laws, M. Looms, R.J.Frank: Self-organising map represeatitions of greyscale images reflect human similarity judgements. In: Proceedings IEEE IS, 2004.

[2] T. Kohonen: Self-organized formation of topologically coorect feature maps. Biology Cybernetics, 43, 59-69, 1982a.

[3] J.B. Tenenbaum, V.de. Silva, J.C. Langford: A global geometric Framework for nonlinear dimensionality reduction. Science, Vol. 290, 2319–2323, 2000.

[4] B. Vidakovic: Statistical Modeling by Wavelets. John Wiley & Sons, Inc. 1999.