Feature based prediction of the perceived and aesthetic properties of visual textures

Stefan Thumfart, Richard H.A.H. Jacobs, Koen V. Haak, Frans W. Cornelissen, Josef Scharinger and Christian Eitzinger

***Abstract*—Texture is an essential factor for human aesthetic perception. We investigate the relationship between computational texture features and the perceived and aesthetic properties of visual textures, obtained from a psychological experiment. We use linear feature selection and neural networks to extract a set of 8 texture features which are capable of predicting 6 perceived and aesthetic properties.**

***Index Terms*—texture, perception, aesthetics, feature selection.**

I. Introduction

TEXTURE is a key factor for human perception and thus widely used in fields such as product design. Whereas issues related to texture segregation, classification and retrieval received much attention in the scientific community during the last decades, high level perceptual and aesthetic content was not investigated systematically. Based on the results of a psychological experiment we investigate the relationship between a set of texture features and perceived and aesthetic properties of visual textures. To facilitate readability, we refer to perceptual and aesthetic properties of visual textures as texture aesthetics or aesthetic properties in the following sections. We show that there is a significant relationship between texture feature subsets and aesthetic properties. In addition a robust interpretation method, to assess the relevance of single features for the prediction of human judgments, is presented.

In section IV we outline a hybrid feature selection method combining linear feature selection and a nonlinear subset selection method. This evaluation is based on two major data sources obtained for a set of 70 visual textures. The aesthetic properties of these textures are assessed in a psychological experiment as discussed in section II. The corresponding feature values are computed using 6 different texture analysis approaches, as outlined in section III.

II. Experimental design

We conducted a semantic differential study, similar to [1], to collect the judgments on 9 aesthetic properties for a set of 70 textures. The rating was done by 20 subjects.

TABLE I
THE ASETHETIC PROPERTIES USED WITHIN THE SEMANTIC DIFFERENTIAL STUDY.

feeling	naturalness	roughness
elegance	complexity	warmth
beauty	colourfulness	hardness

A. Stimuli

The texture set (available at [2]) consists of 65 original texture samples, which are selected based on their score with respect to computational texture features or previous psychophysical experiments. The remaining 5 textures are generated by fusing 5 texture pairs from the already selected ones, using a texture mixing algorithm.

B. Procedure

The experiment is composed of two parts. First a *practice session* is conducted, to show all stimuli to the subjects for reference. The participants are informed that the textures are covering the whole range of perceptual qualities and instructed to assess them quietly (without an evaluation).

Before the *evaluation session*, all aesthetic properties (see table I) are presented to the participants. During the evaluation the subjects are instructed to assess the presented images by moving a sliding bar towards a word that describes the image best. Every experimental run addresses only a single aesthetic property. The antonyms describing this property (e.g. naturalness: natural, artificial) are shown on the left and right side of the sliding bar. For half of the participants the words were mirrored. The sequence of runs (aesthetic properties) was presented to the subjects in pseudorandom order.

Within a run, the texture images are shown until the participant fixes his evaluation by clicking the mouse. Before the next stimuli, a gray screen is shown for one second. Furthermore each texture image is masked by a Gaussian transparency mask to provide a smooth transition between image and background. The sliding bar position is linearly rescaled to the range $[-100; 100]$.

C. Results

The semantic differential study provides us with the ratings of 9 aesthetic properties for 70 textures. We use the average

This work was funded by the EC under grant no. 043157, project SynTex. It reflects only the authors' view.

S. Thumfart and C. Eitzinger are with Profactor GmbH, Im Stadtgut A2, 4407 Steyr-Gleink, Austria, email {stefan.thumfart, christian.eitzinger}@profactor.at

R. H.A.H. Jacobs, K. V. Haak and F. W. Cornelissen are with the Laboratory of Experimental Ophthalmology & BCN NeuroImaging Centre, School of Behavioural and Cognitive Neurosciences, University Medical Centre Groningen, University of Groningen, PO Box 30.001, Groningen 9700 RB, The Netherlands.

J. Scharinger is with the Department of Computational Perception, Johannes Kepler University Linz, Altenberger Str. 69, 4040 Linz, Austria.

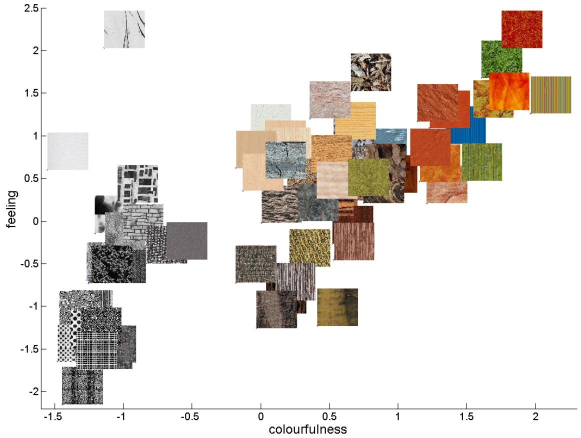

Fig. 1. Plot of texture thumbnails, rated against colourfulness and feeling (i.e. positive vs. negative) . The judgments are standardized to $\mu = 0, \sigma = 1$.

ratings (over subjects) to evaluate the relevance of computational texture features for the prediction of aesthetic properties (see section IV). See figure 1 for a plot of all textures for the judgments feeling and colourfulness.

III. Texture Features

This section introduces the texture analysis methods we used. Please refer to [2] for an exhaustive list of all features (and the corresponding parameters).

A. Gray Level Co-occurrence Matrix

The use of a Gray Level Co-occurrence Matrix (GLCM) to analyze the 2^{nd} order statistical properties of textures was first proposed by Haralick et al. in [3]. A matrix entry at index i, j describes the frequency of pixel pairs with intensity i and j, separated by a displacement vector d. Based on this two dimensional histogram, several statistical measures such as contrast, entropy or correlation are computed.

We use 4 different angles $(0°, 45°, 90°, 135°)$ for a fixed length of d to obtain results that are insensitive to rotation, as proposed in [3]. Next, we compute 11 statistical measures for the 4 resulting GLCMs. These measures are combined using mean and range, forming a feature vector of length 22. We perform these steps for 4 different vector lengths $(|d| = 1, 2, 4, 8)$ to capture both, large scale structures and fine details of the texture.

B. Neighbourhood Gray-Tone Difference Matrix

The Neighbourhood Gray-Tone Difference Matrix (NGTDM) describes the intensity difference between pixels with a specific intensity i and their local neighbourhood. In [4] Amadasun and King propose 5 perceptual texture features that are extracted from the NGTDM. We compute these features for two neighbourhood sizes $(d = 1, 2)$.

C. Tamura

The 6 texture features proposed in [5] are not based on a specific intermediate texture representation, but aim only at high correlation with human perception of textural properties such as coarseness, contrast, and directionality. Hence, this category of features is widely used in content based image retrieval applications [6].

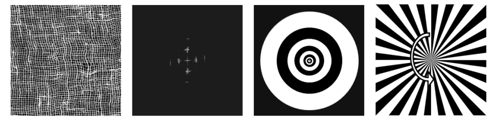

Fig. 2. Partition of the Fourier power spectrum into circular rings and wedges for texture D103 from the Brodatz texture album [9].

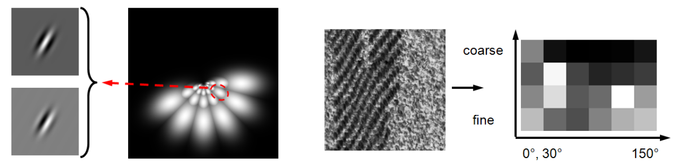

Fig. 3. F.l.t.r. real and imaginary component of a Gabor filter; Gabor filter bank in Fourier domain; a visual texture; Gabor energy map computed for the visual texture using the specified filter bank.

D. Fourier spectrum

We use the Fourier power spectrum to compute texture features by dividing it into circular rings and wedges as depicted in figure 2. The energy distribution between these partitions can be used to assess properties like coarseness or directionality [7]. We refer the reader to [8] for an exhaustive discussion of the Fourier transform and the properties of the Fourier power spectrum.

E. Gabor

Gabor filter based approaches to model visual perception are an active research area since Daugman's finding that the receptive field of cortical cells can be modelled best by Gabor functions [10]. A two-dimensional Gabor function consists of a sinusoidal plane wave of a certain frequency and orientation modulated by a Gaussian envelope. It is given by:

$$f(x, y) = \exp\left(-\frac{1}{2} \cdot \left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right]\right) \cos(2u_0\pi x + \phi)$$

where u_0 and ϕ are the frequency and phase of the sinusoidal wave. The values σ_x and σ_y are the sizes of the Gaussian envelope in the x and y directions, respectively. The Gabor function at an arbitrary orientation can be obtained by a rigid rotation of the x-y plane [7], [11].

We use a Gabor filter bank, containing filters at 4 scales and 6 orientations, as depicted in figure 3. This filter bank is convolved with the texture, resulting in 24 filtered images. We use a Gabor energy map (see figure 3) as an intermediate representation of the filter response magnitudes per image [12]. Apart from the 24 energy map entries, we compute features such as SGOED (sum of Gabor orientation energy differences) as proposed by Kim et al. in [12]. Following the ideas of Kim et al. we further extend our Gabor feature set by measures such as SGSED (sum of Gabor scale energy differences), MGOED (mean of Gabor orientation energy difference) and MGSED

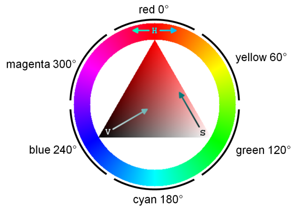

Fig. 4. Partitioning of the hue component into 6 equally sized sectors.

(mean of Gabor scale energy difference). The final set of 31 Gabor features is completed by mean, standard deviation and entropy of the energy map entries.

F. Colour

Human aesthetic perception is highly depending on colour properties. Hence we include colour measures into our feature set. Datta et al. recently presented a work [13], dealing with the aesthetic properties of photographic images. Among others, Datta et al. propose three colour related features. Whereas the average intensity (*f1*) and average saturation (*f3*) features are based on the HSV colour space representation of the image, the colourfulness *f2* is assessed using LUV colour space and earth mover's distance (EMD).

To gain further insight into the colour distribution of the texture image, we compute features based on the HSV colour space. We partition the hue component into 6 sectors (each spanning over 60°, see figure 4) and compute the relative frequency of pixels within each sector. After normalizing this sector frequency by the average image value and saturation, we obtain 6 features representing e.g. the greenness of an image.

IV. Feature Relevance Evaluation

In this section we use linear feature selection, followed by a neural network based subset evaluation method to show that there is a significant relationship between 6 aesthetic properties and the features outlined in section III. Furthermore we analyse the relevance of single features for specific aesthetic properties, to facilitate future modelling approaches.

A. Feature Selection

Feature selection is essential for most machine learning applications that deal with a high dimensional feature space [14]. Also our feature set, containing 188 different features, needs to be reduced before we can apply our neural network based subset evaluation. We use *sequential forward selection* (see [15]), that starts with an empty feature set. Features are included one by one, to minimize the prediction error of a linear regression model. We select the group of four most linearly predictive features per aesthetic judgment, resulting in a reduced set containing 31 unique features.

B. Subset Evaluation

One major drawback of sequential forward selection is that it does not account for statistical dependencies within the set of available features, as they are added one by one. In addition only linear dependencies between features and judgments are considered during the forward selection process.

Artificial neural networks are well suited to model scenarios with nonlinear correlations, even if the input and output data tend to be noisy or contain outliers. Hence we conduct feature subset evaluation using ANNs to assess the relevance of feature groups for an aesthetic judgment. The basic processing steps can be summed up as follows:

(1) select a group of features of size S (e.g. Tamura-coarseness, average saturation, SGSED)
(2) select one aesthetic property (human judgments e.g. beauty)
(3) perform 5-fold cross validation:
    a) separate the selected feature data and judgments into 5 folds
    b) use 4 folds for ANN training
    c) use the remaining fold for testing
    d) measure the similarity of the ANN output (predictions) and the original human judgments using the Pearson correlation coefficient R
    e) repeat steps b)-d) until all fold combinations are used
    f) store the average correlation
(4) goto (1) until all feature combination of size S are evaluated

We have 70 samples available to train and test our ANNs. To avoid overfitting, we have to restrict ourselves to a small feed-forward neural network (1 hidden layer consisting of 3 neurons with a $\tanh$ shaped transfer function). Together with a feature group size $S = 3$, we get a reasonably low number of network weights that need to be trained.

C. Nomination Analysis

After performing the sequence of processing steps outlined above we obtain an average correlation between every possible feature group (4495) and every aesthetic property (9). This information can be used to retrieve the feature group, having the highest correlation with a specific adjective. However, as randomness is involved in the neural network training, only considering single correlation results may possibly be misleading. Hence we answer the question, whether there is a relationship between computational texture features and aesthetic properties by comparing two correlation histograms $h_{original}$ and h_{random} (see figure 5).

The histogram $h_{original}$ shows the distribution of correlation coefficients computed for every combination of feature triplet and aesthetic property. Before any conclusions can be drawn from $h_{original}$ we have to assess the effect of neural network training (particularly the randomness involved) on the results.

Inspired by the parallel analysis factor retention method described in [16], we compute the random correlation histogram h_{random} which contains correlation coefficients resulting from the randomness involved in network training and the distribution of feature and experimental data. We compute this baseline histogram h_{random} by repeating all processing steps outlined in section IV-B for a random dataset, distributed like the original feature and experimental data. The random

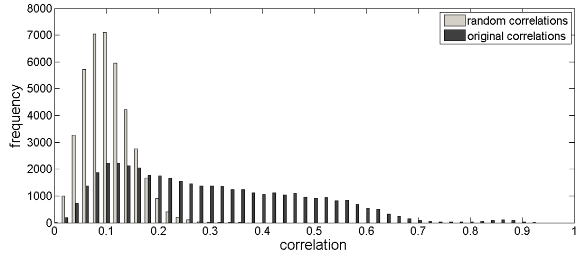

Fig. 5. Correlation histogram of the original and the randomly permuted data. The maximum correlation value of the randomly permuted data is 0.37. 32% of the correlation values for the original data are larger than 0.37.

dataset is generated by random permutation of the samples for the feature matrix, hence destroying the relationship between features and judgments. In figure 5 we can easily see that there is a significant difference between the two distributions and therefore a relationship between texture features and human aesthetic judgments.

We already mentioned, that it is dangerous to focus on single correlation values to analyse the feature relevance. Hence we assess the number of feature nominations to achieve robust results. We define that a feature is nominated for a specific aesthetic property, if it is element of a feature group that reached a correlation larger than a threshold t. The lower bound for t is the maximum correlation obtained for the random data, i.e. 0.37. To increase the robustness of the following interpretation we decided to choose $t = 0.5$.

We use the number of nominations per aesthetic property to assess whether this property can be sufficiently predicted using the current feature set. The number of nominations a feature receives for a particular aesthetic property indicates its relevance for the prediction of this specific property.

D. Results

It is particularly interesting that 3 out of 9 aesthetic properties did not receive a single nomination, indicating that the reduced feature set is not suited to predict these properties, encouraging further investigations related to the initial feature set, as well as the feature selection step discussed in section IV-A. The dominating aesthetic properties are *elegance, feeling* and *colourfulness*, receiving 74% of the total nominations (see table II).

Furthermore, the colourfulness feature, designed particularly to measure the colour distribution in an image, is not relevant for the judgment colourfulness, which is dominated by the average saturation measure.

The extension of the feature set, described in [1], by Gabor features clearly provides relevant information for the prediction. We can see that 5 of 6 feature categories are contributing to the set of relevant features (table II), supporting our approach to use a variety of different analysis methods.

TABLE II
THE TABLE CONTAINS ALL AESTHETIC PROPERTIES WHICH RECEIVED NOMINATIONS. THE COLUMN 'MAX. R' CONTAINS THE MAXIMUM CORRELATION OF A FEATURE TRIPLET WITH AN AESTHETIC PROPERTY. THE COLUMN 'RELEVANT FEATURES' CONTAINS FEATURES THAT RECEIVED THE HIGHEST NUMBER OF NOMINATIONS FOR AN AESTHETIC PROPERTY.

aesthetic property	number of nominations	relevant features (sorted by nominations)	max. R
elegance	1695	SGSED, std(eMap), Tamura-contrast, avgInt, MGSED	0.78
feeling	1530	std(eMap), SGSED, avgInt, MGSED, Tamura-contrast, avgSat	0.78
colourfulness	1299	avgSat	0.91
beauty	588	std(eMap), SGSED, NGTDM-strength	0.68
complexity	569	m-sumEntropy	0.72
hardness	427	Tamura-contrast	0.68

V. Conclusions

We could show that there is a relationship between computational texture features and human aesthetic judgments. Furthermore we could predict 6 out of 9 aesthetic properties with a correlation $R > 0.68$. The aesthetics *naturalness, roughness* and *warmth* could not be predicted with reasonable accuracy using our current feature set, encouraging further experiments and investigations focused on these properties.

References

[1] S. Thumfart, R. H. Jacobs, F. W. Cornelissen, and C. Eitzinger, "A feature based approach towards predicting the perceived and aesthetic properties of visual textures," in *Proceedings of the Austrian Computer Society OCG, 32nd Workshop of the Austrian Association for Patttern Recognition (AAPR/OAGM)*. Austrian Computer Society, 2008, pp. 211–222.

[2] Profactor, "Syntex official web page," http://www.syntex.or.at, section 'public material', 2006.

[3] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics.*, vol. 3, no. 6, pp. 610–621, 1973.

[4] M. Amadasun and R. King, "Textural features corresponding to textural properties," in *IEEE Transactions on Systems, Man and Cybernetics.*, vol. 19, 1989, pp. 1246–1274.

[5] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," in *IEEE Transactions on Systems, Man and Cybernetics.*, vol. 8, 1978, pp. 460–473.

[6] T. Deselaers, D. Keysersand, and H. Ney, "Features for image retrieval: A quantitative comparison," in *LNCS: Pattern Recognition*. Springer-Verlag, 2004, pp. 228–236.

[7] M. Tuceryan, *Texture Analysis*, 2nd ed. World Scientific Publishing Co., 1998, pp. 207–248.

[8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Prentice Hall, 2002, pp. 150–156.

[9] P. Brodatz, *A Photographic Album for Artists and Designers*. New York: Dover Publications, 1966.

[10] J. G. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," *Vis. Res.*, vol. 20, pp. 847–856, 1980.

[11] N. Sebe and M. S. Lew, *Robust Computer Vision: Theory and Applications*. Norwell, MA, USA: Kluwer Academic Publishers, 2003.

[12] M. Kim, C. Park, and K. Koo, "Natural / man-made object classification based on gabor characteristics," *Image and Video Retrieval*, pp. 550–559, 2005.

[13] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Studying aesthetics in photographic images using a computational approach," in *Proceedings of the European Conference on Computer Vision.*, vol. 3953, 2006, pp. 288–301.

[14] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2003.

[15] J. Kittler, "Feature set search algorithms," *Pattern Recognition and Signal Processing*, pp. 41–60, 1978.

[16] J. C. Hayton, D. G. Allen, and V. Scarpello, "Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis," in *Organizational Research Methods*, vol. 7, 2004, pp. 191–205.