XIX IMEKO World Congress
Fundamental and Applied Metrology
September, 6 – 11, 2009, Lisbon, Portugal

MODEL DEVELOPMENT TO PREDICT PERCEIVED DEGREE OF NATURALNESS

Bialek Agnieszka [1], *Forbes Alistair* [1], *Goodman Teresa*[1], *Montgomery Ruth*[1], *Rides Martin* [1], *van der Heijden Gerie*[2], *van der Voet Hilko*[2], *Polder Gerrit*[2], *Overvliet Krista* [3]

[1]National Physical Laboratory (NPL), London, UK, teresa.goodman@npl.co.uk
[2]Biometris, Wageningen University, The Netherlands, gerie.vanderheijden@wur.nl
[3]Parc Científic de Barcelona, Universitat de Barcelona, Spain, krista.overvliet@gmail.com

Abstract – This paper presents the development of a mathematical model to predict the perception of naturalness for a range of materials, based on an understanding of the relationship between the physical attributes of the material and the human sensory inputs. The work is being carried out under an European Union project called 'Measurement of Naturalness' (MONAT), which focuses on understanding the relationships between the physical properties of natural and synthetic materials and the visual and tactile sensory processes that lead to perceptual judgments of naturalness. Integral to the project is the development of novel measurement facilities with dynamic ranges and sensitivities that are relevant for the human sensory systems. The input data to the model are derived from psychophysical and physical studies on pre-selected wood, textile and stone samples.

Keywords: regression, sensory metrology, perception

1. INTRODUCTION

Natural materials such as silk, cashmere and rosewood are generally perceived as being highly desirable and can command high prices Moreover, although we instinctively know whether something is natural or synthetic, the processes involved in this decision are complex and not well understood.

The physical properties of a material or object, e.g. the roughness of the surface, its colour and texture, are generally assessed first by looking at it, and then reinforced or changed by touching it. Interactions between the material and the sensory transducers in our skin and eyes generate sensory impulses, which then pass along nerve fibres to the brain. The strength of these signals depends on factors such as the sensitivity of human sensors, the physical properties of the material and the environmental conditions. Once they reach the brain, the nerve impulses are combined and interpreted to generate a percept; in our case, whether or not the material is natural. But this perception also depends on factors such as memory, expectation and emotional state, and these factors can be just as important as the raw information transmitted by the nerve cells in our eyes and skin. Thus, although we have the feeling that we are in direct contact with our environment, and make decisions based solely on this information, this feeling is generally an illusion. Everything we perceive is determined indirectly, through transformation of physical stimuli into electrical signals and the transformation of these signals into conscious experience. By studying the complete sensory chain, from the properties of the material right through to what happens in the brain, this project is unravelling some of the workings of the perceptual process.

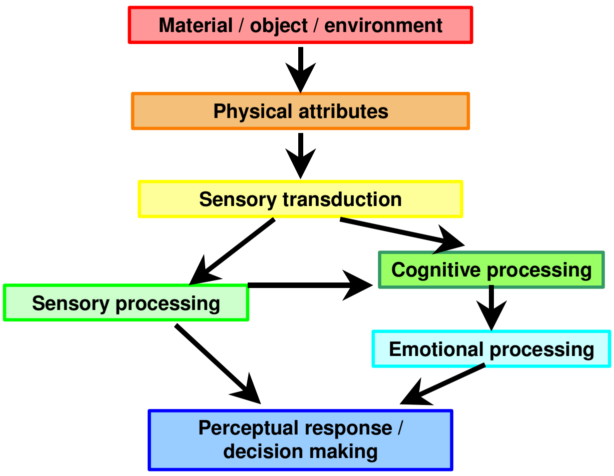

Fig. 1. The perceptual process.

Establishing the chain of perception for naturalness based on sensory inputs enables relationships between the physical attributes of the material and the neural and cognitive process to be identified. A major aim of the project is to develop mathematical models that can predict the perception of naturalness for a range of materials.

2. UNDERSTANDING PERCEPTION IN PRODUCT DEVELOPMENT

On completion of the project, manufacturers may use the information to develop improved materials that appear more natural, leading to products that are more desirable than those made from current synthetics, yet cheaper and more durable than those made from natural materials. These improvements in replica natural products do not only offer economic benefits, but will also help reduce the need to exploit the Earth's dwindling supply of natural resources and may find application in other areas, such as improved virtual reality systems for surgical training.

3. MATHEMATICAL MODELLING

A detailed description of the experimental methods and sample preparation can be found in [1, 2, 3 and 4]. The first material studied was wood and its commercially available synthetic mimics: vinyl, laminate etc. A selected set of 30 samples was measured in the materials and appearance laboratories at the National Physical Laboratory (NPL) to obtain physical characterisation data relevant to the visual and tactile human sensory systems. Psychophysical tests using a nominally identical set of samples were performed at the University of Barcelona. The psychophysical responses during the experiments were recorded in four ways: assignment to finite set of numerical scores (labelled scaling), magnitude estimation, ranked ordering and binary decision [4]. High correlations were found between each of the methods, providing evidence that the psychophysical responses represented perceptual characteristics. In the model below, the responses are represented by the data vector $\mathbf{y}$.

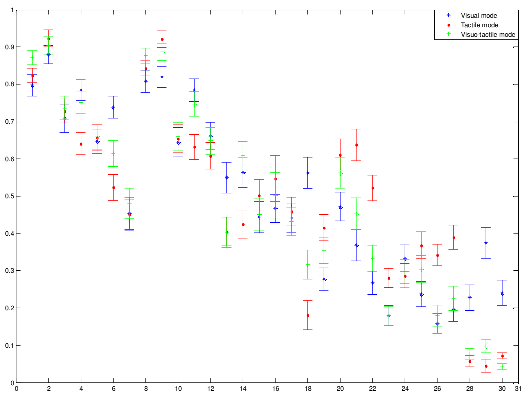

Fig. 2. Average psychophysical tests results using all methods and all participants for each modality.

The results from the psychophysical experiments depend on which sense was used in assessing naturalness. For example, the naturalness of any individual wood sample might be scored differently in the visual mode and in the tactile mode. Moreover when both senses are involved in the decision making process the result is not always an average response from the two modalities. Sometimes the multimodal decision is dominated by one modality, so the result is almost the same as the tactile or visual mode alone.

The physical data creates a set of feature vectors $\mathbf{x}_j$ for each sample, where *i*th element in the *j*th vector represents the measured value of the *j*th feature for the *i*th sample. The physical feature vectors can be very long, as a large number of physical features derived from the physical properties have been measured. In the case of physical measurements relevant to tactile sensation, for example, 81 features are determined. However only 30 different wood samples have been measured. This means that when modelling the relationship between the physical properties for these samples and the perceptual responses, it is necessary to reduce the number of the physical features to ensure that the model uses only the salient ones, to avoid over fitting of the data.

3.1 Classification

The first approach implemented in this project to characterise perceived naturalness was based on classification [1]. The challenge was to separate the samples into defined classes using their physical properties, represented by the feature vectors $\mathbf{x}_j$, as the basis for the discrimination. It is important to remember that the aim was not to classify the samples in terms of the actual type of material (i.e. truly natural or truly synthetic), but in terms of subjective human observations of the degree of naturalness, based on the psychophysical studies. The classes used were therefore defined as:

(a) usually perceived as synthetic;
(b) likely to be perceived as synthetic;
(c) likely to be perceived as natural; and
(d) usually perceived as natural.

A linear discrimination algorithm was able to achieve a useful classification but only at the expense of poor predictive performance on new samples. Potentially a more sophisticated classification algorithm, based on support vector machines for example, might achieve better results. However, it was decided that regression techniques, described below, offered a better approach to predicting the degree of naturalness (interpreted as the probability that a human subject would classify a material as natural).

3.2 Least squares regression

Least squares regression techniques determine the estimate $\hat{\boldsymbol{\beta}}$ of the linear combinations $X\boldsymbol{\beta} = \sum_j \beta_j \mathbf{x}_j$ that best match the response vector $\mathbf{y}$ by minimising

$$F(\boldsymbol{\beta}) = (\mathbf{y} - X\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y} - X\boldsymbol{\beta}) \tag{1}$$

Because for our problem there are more feature vectors than observations it is possible to match the response vector exactly. For this reason the least squares approach has to be modified in order to reduce the number of feature vectors used.

3.3 Partial least squares

The Partial Least Squares (PLS) approach attempts to explain the behavioural response in terms of a set of linear combinations of the feature vectors in an iterative scheme. At the *k*th step in the algorithm, a new linear combination is added to the set in order to reduce the sum of squares as quickly as possible. The application of PLS gave good results in that it was possible to model the psychophysical responses in terms of a small number of combinations of the feature vectors. However, the PLS method does not easily indicate which feature vectors are important for modelling the response.

3.4 LASSO algorithm

The LASSO ('least absolute shrinkage and selection operator') approach was also implemented [5]. This regression method combines the advantages of subset selection regression, i.e., explaining the response in terms of a subset of features, and ridge regression methods, by setting an upper limit:

$$\sum_{j=1}^{p} \left| \beta_j \right| \leq t \tag{2}$$

on the sum of the absolute values of the regression coefficients, with the result that the number of nonzero coefficients is kept small. The algorithm also employs a tuning parameter, s, given by:

$$s = \frac{t}{\sum_{1}^{p} \left| \hat{\beta}_j \right|} \tag{3}$$

where $\hat{\beta}_j$ is the least squares estimate of the jth coefficient.

For $s = 1$, the model is not penalized and least square estimates are obtained; as s decreases towards 0 some coefficients become zero, hence effectively dropping features from the model. Note that this is different to ridge regression, where the coefficients generally do not hit zero. The output is a model with only a small number of non-zero coefficients. This means the model is relatively easy to interpret: if the predictive power of the model using the selected features is found to be acceptable, these features can then be assumed to be the most important ones.

Leave-one-out cross-validation was applied to find the model with the minimal error of prediction (MSEP). Next, predictive accuracy randomisation tests were applied to find the model with the optimal number of physical variables [6]. The optimal model is defined here as the model with as few variables as possible, and an error of prediction not significantly different from the model with minimal error of prediction. Single mode models (visual-only, tactile-only), were created first to identify the most important physical features that determine the perceived naturalness.

3.5 Leave-two-out double cross validation

The set of wood samples used in this study included real wood samples were paired according to the processing and treatments applied to them. In other words, the full set of 30 samples contained 14 real wood samples that had been derived from two different types of oak but then processed using just seven different methods. This led to pairs of samples that had been subjected to the same treatment: raw, weathered, sanded, waxed, oiled, varnished and manufactured (surface treated by a manufacturer to be suitable for wooden flooring)

To ensure that the model is not sensitive to this pairing of the real wood samples, leave-two-out double cross validation tests were conducted. This involved creating models using 28 out of the 30 samples and calculating the MSEP using each model for the two remaining samples. This step was repeated for all possible combinations of two missing samples in the data set. This approach allows observation of the model performance when both of, for example, the raw samples were removed from the data set. Moreover double cross validation produces more rigorous results than single cross validation.

4. RESULTS

A complete tactile model for the wood samples has been developed. The mathematical techniques developed as part of this tactile-mode model are transferable to models for different materials such as textiles and stones.

4.1 Tactile modality model for wood

Eighty-one physical features related to tactile perception of 30 wood samples were used in the first stage of the modelling. The physical measurements covered: 3D texture; friction coefficients in four directions, including circular movements that mimic finger exploration of the samples for the psychophysical experiments; hardness; and thermal properties. Leave-one-out cross-validation for this model gave a minimum value for the mean square error of prediction (MSEP) when 23 out of 81 variables were used to build the model. This number was considered too high, meaning the model would not be easy to interpret, so a randomisation test was performed to justify the prediction accuracy when less than 23 variables are used in the model. The results show that the number of variables in the model could be reduced from 23 to 6 without a significant change in prediction accuracy.

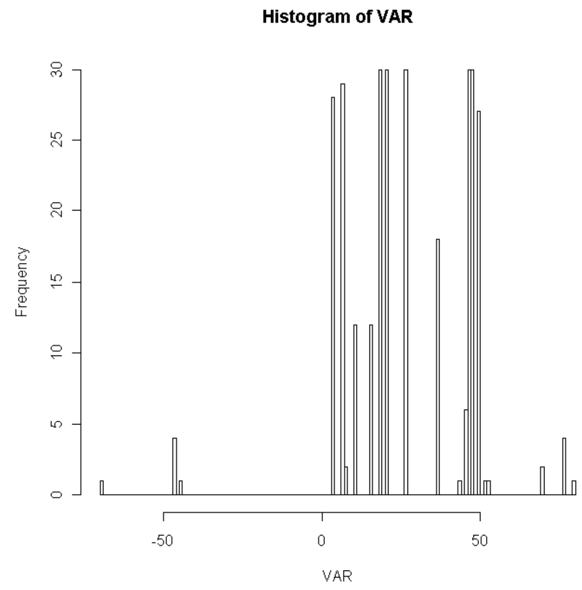

Fig. 3. Histogram presenting frequency of selected variables used in 30 leave-one-out models. The negative values on the x-axis indicate variables that have been dropped from the model.

To ensure that the first variables selected in the model are consistently chosen, the frequency of each individual variable has been checked for 30 models in which one sample (i.e. one feature vector) was missing. The histogram shown in figure 3 presents the results of this test. In all 30 models, each based on data from 29 samples, at least 5 of the first steps in the LASSO variable selection were almost the same.

Figure 3 confirms that even when the data relating to one of the samples is missing in the model, the same physical features are selected as being important. To assess the robustness of the selection of the six features another test was performed. Six new models were created, each with 80 physical variables instead of 81 with one of the six "important" features deleted from the feature set. This test showed that the five remaining features were always selected from the 80 features with the missing feature replaced with another one. Moreover, the new feature was always very highly correlated with the feature that had been replaced. The model performance measured by the R^2 (degree of correlation) value and MSEP did not show a large difference except when the first variable (the most highly correlated with $\boldsymbol{y}$) was missing. This test suggested that the set of 81 features could be reduced by removing highly correlated features.

4.2 Feature reduction using expert opinion

The feature reduction process was conducted using two different methods. In the first, an experienced researcher in material science at NPL selected a set of 16 primary features that were considered to describe fully the key properties of the materials under study. For example, the aforementioned friction coefficient measurements were performed in four different directions (x, y, x-y and circular) but for the purposes of the reduced model only the x and y directions were used. The results for the x-y and circular directions can be determined from a combination of the forces in the x and y directions and these can therefore be considered as redundant data. Similarly, the friction coefficient in any one direction can be characterised using 11 different variables that are all derived from the same raw data and are therefore highly correlated. Therefore only the average and standard deviation of this parameter were chosen in the reduced dataset, and ultimately the high correlation between the x and y directions allowed these too to be combined, thus enabling the number of variables describing the friction coefficient characteristics of each sample to be reduced from 44 to just 4. Similar rules were applied to the rest of the physical measurements.

4.2 Feature reduction using SVD

The second method used the singular value decomposition (SVD) method. The aim was to find the subset of the feature vectors that produced the least mutual correlation. This analysis was done purely on the basis of the feature vectors $\mathbf{x}_j$. (By contrast the PLS and LASSO algorithms use a subset selection approach targeted at approximating the response vector $\mathbf{y}$.)

Both the expert and SVD methods selected similar features. The final set of reduced features had 16 variables that have been chosen by the specialist scientist and then improved by feature selection using SVD. This set was used as the input data to Lasso regression and found to give the best regression performance and has the minimum MSEP for 5 variables. Using this combination of methods, the most salient features for the tactile perception of the naturalness of wood samples were found to be:

(a) Friction coefficient y average
(b) Valley void volume of the surface
(c) Fractal dimension
(d) Texture direction index
(e) Texture aspect ratio

Figure 4 presents the results of the model fitting and model prediction. The model is built with the five features listed above.

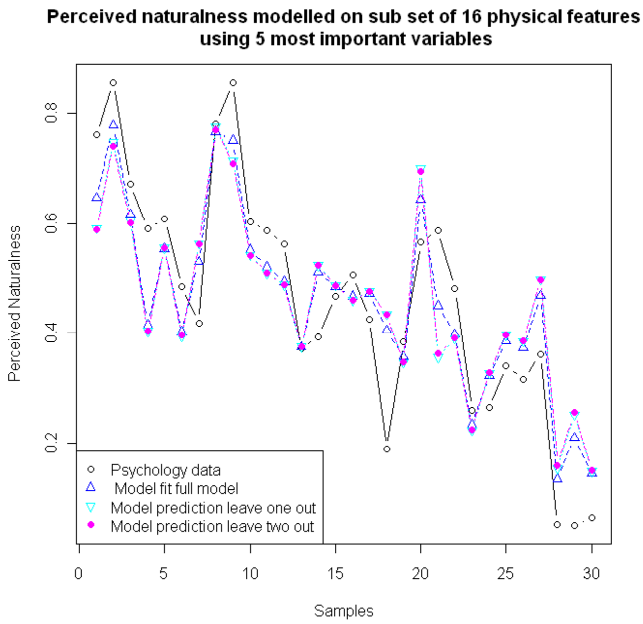

Fig. 4. The performance of the final model for perceived naturalness of wood samples using the tactile modality. R^2=0.815.

The MSEP calculated for leave-one-out validation was equal 0.0134, increasing to 0.0135 for leave-two-out validation, where the optimal number of variables used in the model was still 5. The values plotted in figure 4 as the 'Model prediction leave two out' are the averaged responses over 29 possible combinations of missing pairs for each sample number.

4.3 Leave two out – extreme case

The pairing issue is described in more detail below and the results are plotted in figure 5, where only the models using the pair of same-treatment specimens were removed from the dataset. Therefore the perceived naturalness value was predicted for a completely unknown observation. A significant difference in model prediction was obtained for samples number 2 and 9; these are the weathered pair of specimens. The weathered woods were perceived as the most natural (maximum on the perceived naturalness scale), but when neither of them was included in the model-building process the perceived naturalness scale was reduced. The model was then attempting to perform a challenging task, to predict a response value *y* that was beyond fitting range. Actually this was not achieved with a high degree of accuracy (see figure 5). The explanation of the poor model prediction in this case lies in the salient features that were selected in the first five steps of the Lasso regression. One of the five most important physical features from the full model (fractal dimension) was replaced by the parameter 'ten points height of selected area', which is not well-correlated with fractal dimension.

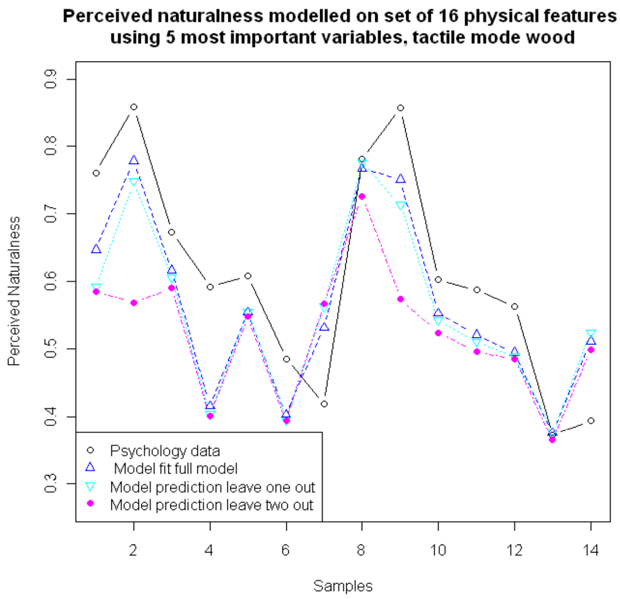

Fig. 5. The performance of the model for perceived naturalness for 14 real wood samples; leave two out present prediction for the missing pairs.

Figure 6 shows the feature space using parallel coordinates, where the first axis on the left hand side is the sample axis, next are the five most salient physical features and the last vertical axis is the new variable that was used in the model when both weathered specimens were missing. The black lines correspond to the missing samples and the fourth axis from the left represents fractal dimension. The noticeable extremes in the fractal dimension values for this pair and the lack of this physical property in the leave-two-out model explain the inaccurate prediction in this particular case. In the following, sixth, step of the Lasso algorithm the added variable was fractal dimension and the prediction was significantly better. The assumption can be made that fractal dimension is a key feature to express the uniqueness of the weathered wood samples.

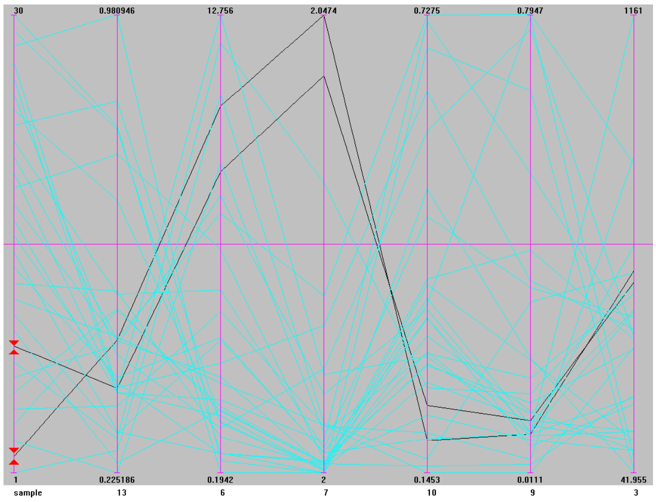

Fig. 6. Parallel coordinates to visualise correlation between salient physical features in the model. The two black lines correspond to two weathered samples.

5. FUTURE WORK

The project is currently focusing on identifying the samples that best encapsulate the extremes of the physical parameters as identified using the perceptual model. These samples will be used in neuroimaging (fMRI) investigations in order to map the neural (sensory and cognitive) responses to different physical variables. These results will be incorporated in the model that links psychophysics with the physical characteristics, with the expectation that this will enhance its validity and accuracy of prediction when applied to new samples.

A new set of 20 wood and wood effect samples is being prepared in order to validate the final model more fully. Like the original sample set, these include a range of natural and synthetic types, but all the samples are different from those used in the original set.

Work is also underway, using the methods described above, to create models for:

(a) Visual and bimodal perception of the naturalness of the wood samples; and

(b) Perception of the naturalness of the remaining types of material (textile and stone) in all 3 modalities.

ACKNOWLEDGEMENTS

This research is part funded by the EU under the Framework 6 Programme (NEST-2004-Path-IMP) and part funded by the National Measurement System Policy Unit of the Department for Innovation, Universities and Skills of the UK government.

REFERENCES

[1] T. Goodman, R. Montgomery, A. Bialek, A. Forbes, M. Rides, A.Whitaker, K. Overliet, F. McGlone, G. van der Heijden. “The Measurement of Naturalness (MONAT)”. IMEKO TCI and TC7 2008, Annecy France. In press

[2] R. Montgomery, T. Goodman, A. Bialek, A. Forbes, M. Rides, A. Whitaker, K. Overliet, F. McGlone, G. van der Heijden “Measuring physical and visual material properties to determine their perceived degree of naturalness” Materials and Sensations 2008, Pau, France

[3] T.A. Whitaker, C. Simões-Franklin, F.N. Newell “The Natural Truth: The contribution of vision and touch in the categorisation of “Naturalness,” EuroHaptics 2008.

[4] K.E. Overvliet, S. Soto-Faraco, T.A. Whitaker, L. Johnstone Sorensen, F. McGlone, G. van der Heijden “Measuring perception of naturalness” Measuring Behaviour 2008, Maastricht, The Netherlands

[5] R. Tibshirani “ Regression Shrinkage and Selection via the Lasso” J.R. Statist. Soc B (1996) 58, No 1, pp. 267-288

[6] H van der Voet “ Comparing the predictive accuracy of models using a simple randomization test” Chemometrics and Intelligent Laboratory Systems 25 (1994) 313-323